On Advantages of Non-Classical Resources in Information Theoretic Tasks: Communication and Correlations

By

Sumit Rout

Supervisor: Prof. dr . hab. Paweł Horodecki

Faculty of Mathematics, Physics and Informatics
University of Gdańsk

Acknowledgement

This doctoral journey would not have been possible without the incredible people who walked alongside me, challenged me, and believed in me even during the toughest times. First and foremost, I would like to thank my supervisor, Prof. Dr. hab. Paweł Horodecki, for his guidance, patience, support, and deep insights throughout my PhD. His valuable feedback has shaped not only this thesis but also my growth as a researcher.

I am also grateful to my collaborators, who shared their skills, intuition, and passion for the subject. It was a great privilege to work with Dr. Some Sankar Bhattacharya, Dr. Nitica Sakharwade, Dr. Ravishankar Ramanathan, and Dr. Anubhav Chaturvedi. Their insights across different topics of research enriched my work, and their passion for science was truly contagious. I would like to thank Dr. Markus Grassl for the many long and stimulating discussions. I am also grateful to my colleagues Dr. Marco Erba and Dr. Michał Banacki for their valuable input, and to Dr. Mir Alimuddin for several thought-provoking conversations. My sincere thanks also go to my colleagues at ICTQT and IFTiA for their numerous discussions, seminar presentations, and thoughtful questions, which played role in shaping this PhD.

I would like to thank the administrative team at ICTQT for their constant support, and the doctoral school staff for helping me meet the university requirements. In particular, I am grateful to Mrs. Ewa Kaszewska, Mrs. Marta Krzyżykowska, and Mrs. Małgorzata Szczekocka. Their assistance eased the administrative load and allowed me to focus on research.

On a personal note, I would like to begin with my girlfriend Ruchi, who has been extremely supportive throughout this journey. She has been the pillar I could lean on at any time, my anchor through every storm, and my biggest cheerleader in every triumph.

A lifeline for every PhD student is their circle of friends, and I have been nothing less than fortunate in this regard. I wish to thank my collegues and friends Paulo J. Cavalcanti, Abyoudai S. S., Aravinth B. Ravichandran, Paweł Mazurek, Gerardo Suarez, Tanmoy Biswas, Borhan Ahmadi, Tushita Prasad, Ekta Panwar, and Chithra Raj. Thank you for listening to me, for offering perspectives when I felt lost in my own sphere, and for the laughter, banter, smiles, and even the occasional arguments that became the non-academic components of my PhD journey.

I would also like to Aditya, who has been an incredible source of support over the years. Finally, I wish to thank my family, without whom none of this would even have begun. Their unconditional love, encouragement, and countless sacrifices made this achievement possible.

I would also like to acknowledge the technical tools that supported me during this PhD, including Wolfram Mathematica and language models such as ChatGPT.

This work was supported by the Foundation for Polish Science (IRAP project, ICTQT, contract no. MAB/2018/5, co-financed by EU within Smart Growth Operational Programme).

Dedication

To my loving family,

whose endless love, patience, and constant belief gave me strength and inspiration throughout this journey.

Abstract

Non-classical features of quantum theory fuel a quantum over classical advantage by improving task performance or resource efficiency. The operational scenario underlying a task sets constraints within which some quantum resources resist efficient classical simulation and effectively lead to an advantage. Additionally, advantages in information processing tasks serve as tools to certify non-classical behaviour. This thesis investigates the advantages offered by non-classical resources in communication and non-local scenarios, and develops methods to detect non-classical quantum resources.

The first part introduces the *distributed clique labelling problem* for orthogonality graphs and studies the classical and quantum communication required for one-way zero-error *distributed computation* and *relation reconstruction* without shared resources. While distributed computation shows no quantum advantage for some graphs, relation reconstruction reveals an instance of *unbounded* quantum advantage. The classical communication cost without shared and private randomness grows as $\log_2 n$ bit with the order of the graph n and the classical cost without shared randomness is at least $\log_2 K(\mathcal{G})$ bit for graph $\mathcal{G}$ with disjointness number $K(\mathcal{G})$. $K(\mathcal{G})$ is lower bounded by $\max\{\omega(\mathcal{G}),\ \log_2 n + \frac{1}{2}\log_2\log_2 n + O(1)\}$. In contrast, the quantum cost is $\log_2 \omega(\mathcal{G})$ qubit whenever $\mathcal{G}$ admits a faithful orthogonal representation in dimension $\omega(\mathcal{G})$, where $\omega(\mathcal{G})$ is the clique number (size of the maximum clique in the graph). We identify families of graphs where the order n grows without bound, while the faithful orthogonal range and clique number remain constant. For these graphs, the quantum cost remains fixed while the classical cost grows without bound, establishing an unbounded quantum advantage in relation reconstruction. We also show that the necessary shared randomness for bounded classical communication scales as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit with the number of maximum cliques α. For a specific family of graphs, this leads to an unbounded separation between entanglement assistance and shared randomness assistance required for a one-bit classical channel.

The second part presents *randomness-free* schemes for detecting *non-projective-simulable measurements* in scenarios with separated parties sharing systems with bounded local operational dimension. It presents detection schemes for three- and four-outcome qubit non-projective measurements in a bipartite scenario, some of which are robust against arbitrary depolarising noise. It also proposes schemes for detecting five-outcome qutrit non-projective measurements. In a bipartite case with identical devices, the work presents a scheme and provides evidence of its robustness against arbitrary depolarising noise. Relaxing the identical-device assumption, detection schemes for qutrit non-projective simulable measurements are discussed in both bipartite and tripartite scenarios. The work extends the notion of non-projective simulable measurements for General Probabilistic Theories and presents a randomness-free test to show that the square-bit model or box world theory is unphysical.

The third part studies correlation-assisted bounded classical communication tasks in a one-way prepare and measure scenario when the receiver is not given an input. It constructs a Bell inequality tailored to a correlation-assisted classical communication task with a linear payoff using the *wire-cutting* technique. The violation of this inequality implies an advantage in the corresponding task, and vice versa. It then introduces the concept of *wire-reading*, which leverages the readability of the classical messages, and uses it to present two families of tasks where shared randomness assistance gives a strictly suboptimal payoff. In the first, any correlation from a non-local facet of the no-signalling polytope achieves the optimal payoff, with explicit quantum advantage shown for some cases. In the second family, each task is tailored to a specific non-local facet, and correlations on that facet achieve the maximum payoff. Considering tasks tailored to non-local extremal correlations with dichotomic outputs, where the shared randomness-assisted payoff is at most 0.75, it shows any correlation on the isotropic line connecting the extremal point to white noise leads to an advantage as long as the noise fraction is below 0.5. It introduces a third family of tasks characterised by two parameters, demonstrating advantageous assistance from Hardy-type correlations with dichotomic inputs to a one-bit channel. While the payoff with shared randomness is bounded by zero from above, Hardy correlations yield a positive payoff. The maximum quantum payoff using Hardy-type correlations grows with a task parameter. Finally, it shows an instance when a two-qutrit entangled state achieves maximum quantum payoff. In contrast, two-qubit entangled states achieve payoffs higher than the local bound, but strictly lower than the quantum maximum, demonstrating a qutrit over qubit advantage. This provides an operational method to witness the local dimension of the shared entangled system.

Abstrakt

Nieklasyczne cechy teorii kwantowej umożliwiają jej przewagę nad teorią klasyczną poprzez poprawienie wydajności wykonywania zadań lub efektywność zasobów. Scenariusz operacyjny leżący u podstaw danego zadania nakłada ograniczenia, w ramach których pewne zasoby kwantowe opierają się efektywnej klasycznej symulacji i w praktyce prowadzą do przewagi. Dodatkowo przewagi w zadaniach przetwarzania informacji służą jako narzędzia do certyfikacji nieklasycznego zachowania. Niniejsza rozprawa bada przewagi oferowane przez nieklasyczne zasoby w scenariuszach komunikacyjnych i nielokalnych oraz rozwija metody wykrywania nieklasycznych zasobów kwantowych.

Pierwsza jej część wprowadza problem rozproszonego etykietowania klik dla grafów ortogonalności i bada klasyczną oraz kwantową komunikację wymaganą do jednostronnego bezbłędnego rozproszonego obliczania i rekonstrukcji relacji bez wspólnych zasobów. Podczas gdy rozproszone obliczanie nie wykazuje przewagi kwantowej dla niektórych grafów, rekonstrukcja relacji wykazuje przypadek nieograniczonej przewagi kwantowej. Koszt komunikacji klasycznej bez wspólnej i prywatnej losowości rośnie jako $\log_2 n$ bitów wraz z rzędem grafu n, a koszt klasyczny bez wspólnej losowości wynosi co najmniej $\log_2 K(\mathcal{G})$ bitów dla grafu $\mathcal{G}$ o liczbie rozłączności $K(\mathcal{G})$. $K(\mathcal{G})$ jest ograniczone od dołu przez $\max \omega(\mathcal{G})$, $\log_2 n + \frac{1}{2}\log_2\log_2 n + O(1)$. W kontraście do powyższego, koszt kwantowy wynosi $\log_2 \omega(\mathcal{G})$ kubitów, ilekroć $\mathcal{G}$ dopuszcza wierną ortogonalną reprezentację w wymiarze $\omega(\mathcal{G})$, który jest klikową liczbą grafu (równa rozmiarowi największej kliki w grafie). Identyfikujemy rodziny grafów, w których rząd n rośnie bez ograniczeń, podczas gdy wierny zakres ortogonalności i klikowa liczba grafu pozostają stałe. Dla tych grafów koszt kwantowy pozostaje stały, podczas gdy koszt klasyczny rośnie bez ograniczeń, co ustanawia nieograniczoną przewagę kwantową w rekonstrukcji relacji. Pokazujemy również, że konieczna wspólna losowość dla ograniczonej komunikacji klasycznej rośnie jako $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bitów wraz z liczbą α największych klik. Dla pewnej rodziny grafów prowadzi to do nieograniczonej separacji między wspomaganiem splątaniem a wspomaganiem współdzieloną losowością wymaganymi dla jednobitowego kanału klasycznego.

Druga część pracy przedstawia schematy wolne od losowości do wykrywania pomiarów niesymulowalnych projekcyjnie w scenariuszach z rozdzielonymi stronami, które współdzielą systemy o ograniczonym lokalnym wymiarze operacyjnym. Prezentuje schematy wykrywania nieprojekcyjnych pomiarów kubitowych o trzech i czterech wynikach w scenariuszu dwucząstkowym, z których niektóre są odporne na dowolny szum depolaryzacyjny. Proponuje również schematy wykrywania nieprojekcyjnych pomiarów qutritowych o pięciu wynikach. W przypadku dwucząstkowym z identycznymi urządzeniami praca przedstawia schemat i dostarcza argumentów za jego odpornością na dowolny szum depolaryzacyjny. Rozluźniając założenie identycznych urządzeń, omówiono schematy wykrywania nieprojekcyjnych pomiarów symulowalnych qutritowych w scenariuszach dwu- i trójcząstkowych. Praca rozszerza pojęcie pomiarów nieprojekcyjnych symulowalnych w ramach Ogólnych Teorii Probabilistycznych i prezentuje test wolny

od losowości, aby pokazać, że model square-bit lub teoria świata skrzynkowego są niefizyczne. Trzecia część pracy bada zadania klasycznej komunikacji z ograniczeniami wspomaganej korelacjami w jednostronnym scenariuszu typu przygotuj-i-mierz gdy odbiorca nie otrzymuje wejścia. Konstruuje ona nierówności Bella dopasowane do zadań klasycznej komunikacji wspomaganej korelacjami z liniową nagrodą, wykorzystując technikę cięcia przewodów. Naruszenie tej nierówności implikuje przewagę w odpowiadającym zadaniu, i vic versa. Następnie wprowadza pojęcie odczytywania przewodu, które wykorzystuje czytelność klasycznych wiadomości, i używa go do przedstawienia dwóch rodzin zadań, w których współdzielona losowość daje ściśle suboptymalną nagrodę. W pierwszej rodzinie dowolna korelacja z nielokalnej ściany wieloboku bezsygnałowego osiąga optymalną nagrodę, przy czym w niektórych przypadkach wykazuje się wyraźną przewagę kwantową. W drugiej rodzinie każde zadanie jest dopasowane do określonej nielokalnej ściany, a korelacje na tej ścianie osiągają maksymalną wypłatę. Rozważając zadania dostosowane do nielokalnych ekstremalnych korelacji z dychotomicznymi wyjściami, gdzie nagroda wspomagana losowością wspólną wynosi co najwyżej 0.75, pokazano, że dowolna korelacja na linii izotropowej łączącej punkt ekstremalny z szumem białym prowadzi do przewagi, o ile udział szumu jest poniżej 0.5. Wprowadza się trzecią rodzinę zadań charakteryzowanych przez dwa parametry, demonstrując przydatność korelacji typu Hardy'ego z dychotomicznymi wejściami do jednobitowego kanału. Podczas gdy nagroda przy wspólnej losowości jest z góry ograniczona przez zero, korelacje Hardy'ego dają dodatnią nagrodę. Maksymalna kwantowa nagroda przy użyciu korelacji typu Hardy'ego rośnie wraz z parametrem zadania. Na koniec pokazano przypadek, dla którego gdy splątany stan dwóch qutritów osiąga maksymalną kwantową nagrodę. Warto tutaj zaznaczyć że splątane stany dwóch kubitów osiągają nagrodę wyższą niż granica lokalna, ale ściśle niższą niż maksimum kwantowe, co pokazuje przewagę qutritów nad kubitami. Dostarcza to operacyjnej metody doświadczalnej certyfikacji lokalnego wymiaru współdzielonego układu splątanego.

List of Articles During Doctoral Studies

[A] Rout, S., Sakharwade, N., Bhattacharya, S. S., Ramanathan, R., & Horodecki, P. (2025). Unbounded quantum advantage in communication with minimal input scaling. *Physical Review Research*, 7(2), 023104. https://doi.org/10.1103/PhysRevResearch.7.023104

[B] Rout, S., Bhattacharya, S.S., & Horodecki, P. (2025). Randomness-free detection of non-projective measurements: qubits & beyond. *New Journal of Physics*, 27(3), 033024. https://doi.org/10.1088/1367-2630/adc0b4

[C] Rout, S., Chaturvedi, A., Bhattacharya, S. S., & Horodecki, P. (2025). Facets of Non-locality and Advantage in Entanglement-Assisted Classical Communication Tasks. arXiv: 2507.10830 [quant-ph]. https://doi.org/10.48550/arXiv.2507.10830 [**Pre-print**]

[D] Guha, T., Alimuddin, M., Rout, S., Mukherjee, A., Bhattacharya, S. S., & Banik, M. (2021). Quantum Advantage for Shared Randomness Generation. *Quantum*, 5, 569. https://doi.org/10.22331/q-2021-10-27-569

[E] Rout, S., Maity, A. G., Mukherjee, A., Halder, S., & Banik, M. (2021). Multiparty orthogonal product states with minimal genuine nonlocality. *Physical Review A*, 104(5), 052433. https://doi.org/10.1103/PhysRevA.104.052433

Other Articles:

[1] Rout, S., Maity, A. G., Mukherjee, A., Halder, S., Banik, M. (2019). Genuinely nonlocal product bases: Classification and entanglement-assisted discrimination, *Physical Review A*, 100(3), 032321. https://doi.org/10.1103/PhysRevA.100.032321

Table of Contents

List of Abbreviations

FOR	Faithful Orthogonal Representation
OR	Orthogonal Representation
i.e.	id est or that is
iff	if and only if
wlog	without loss of generality
wrt	with respect to
CLP	Clique Labelling Problem
CPTP	Completely Positive Trace Preserving
GPTs	General Probabilistic Theories
nPS	Non-projective-Simulable measurement
PM	Prepare-and-Measure scenario
POVM	Positive Operator Valued Measure
PS	Projective Simulable measurement
PVM	Projective Measurement

List of Symbols

$A_{\mathcal{G}}$	Adjacency matrix of graph $\mathcal{G}$
$\mathbf{B}$	Bell inequality
β_L	Local bound for Bell inequality
$\beta_{\mathcal{Q}}$	Maximum quantum violation of a Bell inequality
$\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$	Bell scenario with input set X,Y, output set A,B
$\mathcal{B}(\mathcal{H})$	Bounded linear operators acting on $\mathcal{H}$
$\mathfrak{B}$	Box world theory
Cl	Classical theory
g_{C}	A clique label for maximum clique C
$\omega(\mathcal{G})$	Clique number of graph $\mathcal{G}$
$\bar{\mathcal{G}}$	Complement of graph $\mathcal{G}$
$\mathbb{C}$	The set of complex numbers
$\cong$	Isomorphic to
$:=$	Definition
$D(\mathcal{H})$	Set of density operators on Hilbert space $\mathcal{H}$
V^*	Dual space of real vector space V
e_i, f_i	Effects in GPTs
$\mathcal{E}_A$	Effect space of system A
ϕ	Faithful orthogonal representation of graph $\mathcal{G}$
$\forall$	For all
$\xi(\mathcal{G})$	Faithful orthogonal range of $\mathcal{G}$
$\mathcal{G}$	Graph with vetex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$
$\mathcal{H}$	Hilbert space
$\dagger$	Hermitian conjugate
$Herm(\mathbb{C}^d)$	Heritian operators on $\mathbb{C}^d$
$\mathbf{I}_d$	Identity map from $\mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$
I_d	$d \times d$ Identity operator
$\iff$	If and only if

$\implies$	Implies
$\in$	Belongs to
K_i	Kraus operator $K_i : \mathbb{C}^d \to \mathbb{C}^{d'}$
$\delta_{i,j}$	Kronecker delta
λ	Local hidden variable
Λ	Local hidden variable theory
$\mathbf{\Phi}$	Completely positive trace-preserving map
P_L	Local extremal point
$\mathcal{L}$	Local polytope
$\mathcal{M}$	Measurment in GPTs
$\oplus_k$	Sum modulo k
$\mathbb{N}$	The set of natural numbers
P_{NL}	Non-local extremal point
$\lvert\vec{n}\rvert$	Norm of vector $\vec{n}$
$\mathcal{NS}$	No-signalling polytope
$\mathbf{d}_A$	Operational dimension of type A
Tr_A	Partial trace operation over some subsystem A
σ_i	Pauli operators $i \in \{1 \leftrightarrow x, 2 \leftrightarrow y, 3 \leftrightarrow z\}$
$\vec{\sigma}$	$(\sigma_1, \sigma_2, \sigma_3)$
$\times$	Product or cartesian product
$\mathbf{TH}$	Physical theory
V_+	Positive cone of real vector space V
$E \geq 0$	Positive semi-definite operator
Tr	Trace operation
E_i, F_i, G_i	Elements of POVM
$\subset$	Proper subset
$\lvert\psi\rangle, \lvert\phi\rangle$	Pure quantum state
QT	Quantum theory
$\mathcal{Q}$	Quantum realisable correlations in $\mathcal{NS}$
$\mathbb{R}$	The set of real numbers
ρ	Density operator
$\mathsf{F}_{\mathcal{G}}$	Set of binary colourings f of $\mathcal{G}$
$\setminus$	Set difference
Ω_A	State space of system A
$\subseteq$	Subset
A, B, A_i	Subsystems of a composite system
Sys_A	System of type A
$\otimes$	Tensor product

$\mathcal{T}$	Transformation in GPTs
$\rho_{A,B}^{T_B}$	Partial transpose operation with respect to system B
ρ^T	Transpose operation on ρ
U	Unitary operator
u_A	Unit effect on system A

List of Tables

Table **Page**

List of Figures

Figure **Page**

Chapter 1

Introduction

"What we observe is not nature itself, but nature exposed to our method of questioning."

—Werner Heisenberg

Over the past several decades, progress in our understanding of quantum theory has inspired research that goes beyond confirming its predictions or examining the foundational implications of the theory. Recent efforts also increasingly aim to employ the non-trivial features of the theory to drive technological applications. This shift in focus following the "second quantum revolution" has led to the proposition of quantum computers [3–5] and the development of several innovative quantum protocols [6–17] whose performance surpasses classical limitations, thereby demonstrating a clear quantum advantage. These works highlight the potential applications of quantum resources in various areas, including communication [18–48] and computation [49–75]. Theoretical advances, together with recent experimental breakthroughs [61, 63, 76–93], have brought us closer to the practical realisation of these quantum advantages. Such developments have also paved the way for commercially available products such as quantum random number generators [94–99], quantum cryptography [100–110], enabled secure quantum communications over relatively large distances [111–116], and contributed towards architectures for noisy intermediate-scale quantum computing [117–119], to name a few. Future advancements in quantum technology are expected to accelerate progress across various domains of science and technology [61, 118, 120–132], including the simulation of molecules and chemical reactions, drug discovery, materials design, machine learning, solving hard combinatorial problems, and metrology, among others.

Quantum technologies leverage quantum behaviours to achieve advantages in information processing. The various manifestations of non-classicality, such as randomness [133], entanglement

[134–136], steering [134, 135, 137], non-locality [133, 138], and contextuality [139, 140], usually fuel such operational advantages. The existence of such features first became apparent in the context of the EPR paradox [138]. According to EPR, quantum entanglement allows for correlations, which suggests that the quantum mechanical description of reality must be incomplete if the theory satisfies the principle of locality. This indicates that the inherent randomness in quantum theory is a consequence of a coarse-grained description of some underlying classical model. The resolution of the EPR paradox came with Bell's theorem [133]. An attempt to model quantum theory using *hidden variables* [141] showed that randomness in quantum theory cannot be simulated using any classical description satisfying the principle of locality. This result cemented the uncertainty inherent in the quantum description as a fundamental feature of nature. These theoretical findings, validated by subsequent experimental observations [142–145] of correlations beyond local realistic theories, inspired the exploration of quantum features as resources in various domains such as quantum communication, cryptography, algorithms, etc.

Broadly, *quantum advantage* often refers to an improvement in the performance of an information processing task when using quantum resources over their classical counterparts. Such an advantage is typically established by either demonstrating a lower quantum resource overhead in accomplishing the task or a higher efficiency when using quantum resources over their classical analogue. In each instance of such an advantage, the relevant quantum resources cannot be efficiently simulated using classical means. The "non-classical" resources leading to the quantum advantage are usually specific to the operational scenario underlying the task.

The associated physical setup imposes natural constraints, which effectively determine the availability or limit on certain resources. These constraints also specify the class of allowed operations using the accessible resources. For instance, in scenarios involving space-like separated parties, communication is forbidden. Consequently, the parties may only access pre-shared resources and perform operations that cannot influence distant systems. In contrast, in scenarios where communication is allowed, the parties may perform local operations, exchange physical systems, and use shared resources. In some operational settings, the effects of quantum resources and operations may be unattainable within classical theory [133], or classical simulation of these effects would incur a significantly higher cost than their realisation within quantum theory [16, 17]. For example, when communication is allowed but considered costly, encoding classical information into quantum systems can significantly reduce the amount of communication required to complete a task. In particular, the quantum communication needed can be exponentially less than the required classical communication [17]. Similarly, when computation time is the key resource, certain problems can be efficiently solved using quantum systems [11], *i.e.*, in polynomial time. In contrast, their classical counterparts require computational time that scales exponentially with input size while using the best-known classical algorithms.

Quantum advantage in information processing identifies quantum states and operations that outperform classical resources in specific tasks.

Besides demonstrating quantum advantage, such tasks also highlight a deeper operational distinction between classical and quantum theories (or even post-quantum theories). Specifically, they show how certain features of quantum systems resist classical simulation under particular physical constraints. In contrast, certain features of a physical theory may be efficiently reproduced using classical resources within the given operational scenario. In such cases, resources in the physical theory exhibiting these features can be effectively treated as classical resources relative to the scenario. For example, in the Bell scenario, the parties are space-like separated. They have access to shared systems and can perform local operations. Some shared quantum states and measurements give rise to correlations that can also be reproduced using shared classical systems. In particular, correlations obtained from separable quantum states, as well as from certain entangled states [146], can always be simulated using shared classical randomness. These states, therefore, behave *classically* within the Bell scenario, despite being quantum in origin. Therefore, tasks that exhibit quantum advantage under such constraints do more than demonstrate the superior efficiency of quantum resources. They can also serve as operational witnesses of non-classical behaviour. More specifically, the inability of classical resources to reproduce the observed effects within the given scenario can be regarded as evidence that some genuinely non-classical feature of the underlying physical system is in use.

Experiments based on tasks in which the presence of non-classical resources can be inferred solely from the observed data are referred to as device-independent detection schemes for non-classical resources. Such schemes do not require any knowledge of the internal workings of the physical devices. For example, under minimal assumptions such as no-signalling and measurement independence, a violation of a Bell inequality using the observed statistics enables the detection of non-locality in a manner that is independent of the specific physical implementation [142–145]. In some cases, additional assumptions, such as the validity of a quantum description for the shared system, allow further conclusions to be drawn. For instance, particular violations of a Bell inequality may permit the certification of the shared quantum state [147, 148]. Other forms of partial characterisation, such as bounding the dimension of the system or limiting the amount of classical communication, give rise to semi-device-independent frameworks. Within this framework, information-theoretic tasks can be tailored to detect specific non-classical resources. Semi-device-independent detection schemes for entanglement [149, 150], non-projective measurements [151, 152], dimension [153–156], etc., have been widely explored. These approaches often offer robust methods for certifying quantum resources and reinforce the operational significance of quantum advantage, particularly in realistic and resource-constrained settings.

Information-theoretic tasks can therefore serve a dual purpose. They are useful to demonstrate the advantage of quantum resources over their classical counterparts. Additionally, they function as practical tools for detecting non-classical features of physical systems under some minimal assumptions. Among various paradigms used to highlight quantum advantage, communication-based tasks, especially in the prepare-and-measure (PM) scenario, have been widely studied. Such scenarios usually involve two distant parties receiving randomly chosen inputs. They are required to produce outputs following some specified rule. The objective of the task is defined through constraints on the input-output statistics that the parties must satisfy. For example, in communication complexity, the goal is to compute a function or a relation while using minimal communication when given randomly chosen inputs. This has been extensively investigated when classical communication [157, 158] is accessible. Although the classical communication capacity of quantum communication channels is bounded by Holevo's theorem [159], several results have shown that encoding classical information into quantum systems and using them for communication can reduce the communication cost for performing some computational task. A quadratic quantum advantage in communication was first demonstrated in [16] when some error in computation was allowed. Remarkably, an exponential quantum advantage in communication complexity was also shown when the computation had to be zero-error [16]. Further works have reported exponential separations between classical and quantum communication complexity for specific functions and relations, even when some error in computation is allowed [17–20, 160]. Besides communication complexity, communication tasks involving the simulation of certain correlations in the PM scenario have also been explored [21–27].

A natural question that follows from these results is about the maximum separation between classical and quantum communication achievable in such prepare-and-measure scenarios. Some earlier studies considered this problem without any shared resources (public randomness) between the parties, Alice and Bob. In one-way PM scenarios without shared resources, Alice's use of quantum communication is characterised by a set of quantum preparations conditioned on her input. Similarly, Bob's decoding is implemented via a measurement procedure based on his input. If the resulting input-output statistics can be simulated by a classical protocol with the same communication cost, then the given preparations and measurements are effectively classical. For example, when Alice's preparations are diagonal states and Bob's measurements commute, the strategy admits a classical description. In this sense, such preparation and measurement pairs correspond to classical strategies within this framework. Massar *et al.* [24] demonstrated a task that can be completed with a single qubit of communication, whereas the best classical protocol requires $\log_2 n$ bit. The input size is $2n$ bit, which grows exponentially in comparison to the classical communication cost. This motivates the search for other tasks where the separation between classical and quantum communication grows arbitrarily large. One question is whether such an unbounded separation can be shown when the required classical communication exceeds $\log_2 n$, if n is the input size. Another question is whether there exist tasks that require a d-dimensional quantum system (beyond a qubit), while the corresponding classical communication cost grows arbitrarily large. We explore these questions in Chapter 3.

Within PM scenarios, one can also consider the role of shared correlation-assistance to classical communication. These previously shared correlations are required to satisfy the principle of *no-signalling*, which prohibits instantaneous communication. Different kinds of no-signalling correlations can be shared between the communicating parties as assistance to classical communication. Some no-signalling correlations are realisable within classical theory and are referred to as shared randomness. Other no-signalling correlations, termed non-local correlations, go beyond classical realisability and can arise from entangled quantum systems or even from hypothetical post-quantum theories [161]. Although non-local correlations cannot be used for communication, several results have demonstrated that they can nonetheless enhance classical communication tasks: either by reducing the amount of classical communication required, or by increasing the success probability when communication is bounded [28, 32, 32–44]. Notably, even in the minimal one-way PM scenario where Bob has no additional input, certain non-local correlations have been shown to outperform shared randomness in assisting classical communication. For instance, shared entanglement and more general non-local correlations can increase the effective zero-error capacity of a noisy classical channel [45, 46, 162], enabling more alphabets to be transmitted noiselessly. Furthermore, both quantum and post-quantum correlations can reduce the classical communication cost needed to simulate a given noisy classical channel [45, 46, 163]. In particular, an interesting communication task in the one-way PM scenario demonstrated an advantage of entanglement-assisted communication over shared randomness, when only a single classical bit was allowed to be transmitted [163]. In that task, any correlation that violates the CHSH inequality [164] was shown to yield an advantage. Subsequently, it was shown that Hardy-type non-local correlations with dichotomic inputs and outputs can assist classical bit communication better than shared randomness [165]. Such results naturally motivate the question of a deeper connection between non-locality and advantage in correlation-assisted classical communication, particularly within minimal one-way PM scenarios. In Chapter 5, we explore the interplay between non-locality and advantageous assistance from some non-local correlations in classical communication tasks.

Quantum non-locality in the Bell scenario has been extensively studied through the violation of Bell inequalities. In a Bell experiment, multiple space-like separated parties receive randomly chosen inputs and generate outputs based on local operations and pre-shared resources. A Bell inequality is defined as a linear constraint on the observed input-output statistics in this scenario. Such inequalities are satisfied by any correlation that can be simulated using shared randomness. However, quantum and post-quantum non-local correlations are known to violate these inequalities. Quantum non-local correlations require both a shared entangled state and the use of measurements that are not jointly measurable. Besides non-joint measurability, quantum theory allows for measurements on a d-dimensional system that yield more than d outcomes. Measurements on a d-dimensional system that cannot be simulated by d outcome projective measurements on the same system are referred to as non-projective measurements. Such measurements have no classical counterpart. Non-projective measurements are instrumental in numerous quantum information processing tasks, including quantum state discrimination [166–169], entanglement detection [170], quantum tomography [171–175], cryptographic protocols [176], port-based teleportation [177–179], quantum metrology [180–182], and randomness certification [183]. While Bell experiments are traditionally employed to witness entanglement or non-locality, they can also be sensitive to other quantum features, including non-projective measurements. In particular, assuming a bound on the dimension of the shared entangled state, Vertesi *et al.* [151] demonstrated that certain Bell inequalities can exhibit a larger quantum violation only when the measurements used are non-projective. This implies that such inequalities can serve as witnesses of non-projective measurements. However, a fundamental assumption in the standard Bell scenario is measurement independence, *i.e.*, the inputs to each party are assumed to be freely and randomly chosen. From an operational standpoint, sources of true randomness require non-classical resources and may be considered costly or non-trivial to implement. This raises a natural question: can measurements on a d-dimensional quantum system which cannot be simulated using a projective measurement followed by post-processing be detected under relaxed assumptions, possibly without requiring input randomness? We explore this question in Chapter 4.

The aim of this work is to explore and demonstrate the advantages of quantum non-classicality in information-theoretic tasks defined in different operational scenarios using communication and correlation as potential resources. We now summarise the main results of this thesis, which address the questions outlined above. The Chapter 3 is based on the article "Unbounded quantum advantage in communication with minimal input scaling" [184]. We introduce the *distributed clique labelling problem*, defined using orthogonality graphs. This problem naturally specifies a relation, and we study the distributed computation of this relation for some graphs in a one-way PM scenario. We show that both classical and quantum one-way zero-error communication complexity without shared resources are equal and scale as $\log_2 \omega(\mathcal{G})$ bit or qubit, where $\omega(\mathcal{G})$ is the clique number of the graph $\mathcal{G}$. We then consider a stricter version of the distributed computation task, referred to as *relation reconstruction*. The classical communication cost without shared and private randomness grows as $\log_2 n$ bit with the order of the graph n. The classical cost without shared randomness is at least $\log_2 K(\mathcal{G})$ bit, where $K(\mathcal{G})$ is the disjointness number of $\mathcal{G}$ (see Definition 3.4). This quantity $K(\mathcal{G})$ is at least

$\max\{\omega(\mathcal{G}),\ \log_2 n + \frac{1}{2}\log_2\log_2 n + O(1)\}$. On the other hand, the quantum communication cost is $\log_2 \omega(\mathcal{G})$ qubit whenever the graph admits a faithful orthogonal representation in dimension $\omega(\mathcal{G})$. Thus, there is a quantum advantage whenever the graph has a faithful orthogonal representation in dimension $\omega(\mathcal{G})$ and $K(\mathcal{G}) > \omega(\mathcal{G})$. We also identify families of graphs for which the associated task can be performed with a constant amount of quantum communication, $\log_2 d$ qubit, independent of the order of the graph. The classical cost in these cases grows arbitrarily large with n and is at least $\max\{\log_2 d, \log_2(\log_2 n + \frac{1}{2}\log_2\log_2 n + c)\}$ bit, where n is the order of the graph and c is a constant. This establishes an unbounded quantum advantage for relation reconstruction. The input size in this task grows logarithmically with n, namely $2\log_2 n + c'$ bit, where c' is a constant. This unbounded quantum advantage disappears in the presence of shared resources such as shared randomness. We therefore investigate how much shared randomness assistance is required to accomplish the relation reconstruction when the classical communication is bounded. For certain types of graphs, we show that the required amount of shared randomness scales as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit, where α is the number of maximum cliques. Moreover, in relation reconstruction for some families of graphs using 1 bit communication, we show that the separation between the amount of shared randomness and entanglement assistance required can be arbitrarily large. In this case, while assistance from one e-bit entanglement suffices, the shared randomness required scales as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit with the number of maximum cliques α.

The Chapter 4 is based on the article titled "Randomness-free detection of non-projective measurements: qubits & beyond" [185]. In this chapter, we present *randomness-free* method for detecting *non-projective-simulable measurements*. The physical scenario we consider involves multiple space-like separated parties who have a shared system where the local operational dimension of each sub-system is upper bounded by, say, d. The parties receive no inputs and simply generate outcomes by performing local measurements on their respective subsystems. Since the setup requires no external inputs, no seed randomness is needed to certify the measurement structure, making the approach randomness-free. We begin by proving that all correlations achievable in this scenario via local projective measurements on a d-dimensional quantum system, followed by classical post-processing of the outcomes, can also be generated using shared classical systems with the same local dimension, and vice versa. Therefore, such measurements on a d-dimensional system can be regarded as effectively classical in this framework. It follows that any observed quantum advantage in an information processing task defined in this setting must stem from the use of non-projective-simulable measurements, *i.e.* those not simulable by a projective measurement followed by post-processing. Consequently, quantum advantage in such tasks serves as an operational witness for the presence of non-projective-simulable measurements. We present some detection schemes in the bipartite scenario for three and four-outcome qubit non-projective-simulable measurements, where the goal is to produce certain target correlations.

Our detection scheme is robust against arbitrary depolarising noise, except in the limit of completely depolarising noise. We further present a detection scheme for five-outcome qutrit non-projective-simulable measurements. Specifically, we show that some correlations cannot be reproduced when both parties perform identical five-outcome qutrit projective-simulable measurements on pre-shared two-qutrit states. However, these correlations can be realised using identical qutrit non-projective-simulable measurements. We provide numerical evidence supporting the robustness of this scheme under an arbitrary amount of depolarising noise. We also present a task in bipartite and tripartite scenarios for cases when five-outcome qutrit measurements could be non-identical. We provide a numerically obtained upper bound for the reward function achievable using qutrit projective-simulable strategies. We show that this bound is violated when using measurements on a shared two-qutrit entangled state. From a foundational standpoint, we generalise the notion of non-projective-simulable measurements to general probabilistic theories (GPTs). Using a randomness-free task, we demonstrate that certain GPTs, such as square-bit or box-world theories, are not physical.

The Chapter 5 is based on the pre-print titled *"Facets of Non-locality and Advantage in Entanglement-Assisted Classical Communication Tasks"* [2] and some unpublished results. In this chapter, we introduce families of tasks in a minimal PM scenario where non-local correlations enhance bounded classical communication. We begin by presenting a Bell inequality tailored to a correlation-assisted classical communication task with a linear payoff function. We obtain the Bell inequality for a communication task using a technique we call "wire-cutting". A violation of this Bell inequality implies an advantage in the corresponding communication task. Conversely, any non-local correlation that yields an advantage violates the Bell inequality. The maximum quantum violation of the Bell inequality corresponds to the maximum entanglement-assisted payoff achievable in the communication task. We then present "wire-reading", based on the observation that classical messages can be accessed without disturbing the communication. Since these messages can be recorded, they may be treated as observed variables influencing the payoff. This additional freedom enriches the task structure. We demonstrate that the addition of wire-reading can be used to exhibit the advantage of quantum non-locality in scenarios where no such advantage exists without it. Using wire-reading, we construct a family of communication tasks in the minimal PM scenario where assistance from shared randomness yields a strictly suboptimal payoff, while assistance from any correlation from a non-local facet of the no-signalling polytope leads to the optimal (maximum) payoff. For some of these tasks, we explicitly demonstrate quantum advantage. While quantum strategies outperform classical ones in these tasks, we observe that the local bound on the payoff approaches the maximum quite rapidly as the task parameters increase. This behaviour motivates the construction of a second family of tasks, each tailored to a specific non-local facet. In these tasks as well, shared randomness assistance leads to suboptimal payoffs. Assistance from any correlation on this non-local facet achieves the optimal payoff. As a concrete example, we consider tasks tailored to non-local extremal correlations with dichotomic outputs, where the shared randomness-assisted payoff is upper bounded by 0.75. However, any correlation on the isotropic line joining the corresponding non-local extremal correlation and white noise provides an advantage as long as the white noise fraction is strictly less than 0.5. We then introduce a third family of tasks with one-bit communication that demonstrate advantageous assistance from Hardy-type non-local correlations with dichotomic inputs and d-outputs [186]. In this case, the payoff with shared randomness is bounded from above by zero, while the payoff with Hardy correlation is positive and increases with a parameter defining the task. Finally, we consider a particular communication task from the first family discussed earlier, where any correlation from non-local correlation is optimal. For this task, the maximum quantum achievable payoff is attained by sharing a two-qutrit entangled state. In contrast, two-qubit entangled states yield payoffs that exceed the classical bound but fall short of the quantum maximum. This gap offers an operational method to witness the local dimension of the shared entangled system.

The thesis is structured in the following manner. Chapter 2 introduces the necessary preliminaries for the subsequent discussion. The main contributions are presented in Chapters 3, 4, and 5. Finally, Chapter 6 summarises the results and concludes with a discussion of open questions for future research.

CHAPTER 2

Preliminaries

In this chapter, we revisit some foundational concepts and mathematical preliminaries. We also present some results relevant to the discussions in later chapters. We begin with a brief overview of quantum theory. This is followed by a discussion on General Probabilistic Theories (GPTs), which offers a unified framework for describing a broad class of physical theories. Classical and quantum mechanics appear as special cases within this framework. The notions presented here will be useful in comparing resources, specifically classical and quantum, in various information-theoretic tasks. We then discuss the no-signalling correlations and present the mathematical tools pertinent to the study of non-locality. Next, we provide an overview of selected topics in graph theory, focusing on graph colouring and orthogonal representations. We conclude with a discussion of communication-based information-processing tasks in a Prepare-and-Measure scenario (PM).

2.1 Quantum Theory

Quantum mechanics explains a wide range of physical phenomena. In this section, we revisit the mathematical formalism of this physical theory and present some definitions. For an in-depth treatment of the topic, we recommend any standard textbook [187–189].

(i) **State space:** The state space $D(\mathcal{H})$ of a physical system with associated Hilbert space $\mathcal{H}$ consists of bounded, positive semi-definite operators with a unit trace. Such operators are called density operators. We will consider only finite-dimensional systems, where $\mathcal{H} \cong \mathbb{C}^d$. The state of the system is denoted by a density operator $\rho \in D(\mathbb{C}^d)$.

A convex mixture $\rho'' = p\rho + (1-p)\rho'$, with $\rho, \rho' \in D(\mathbb{C}^d)$ and $p \in [0,1]$, is also a valid quantum state. The convexity of the state space is important, as any statistical mixture of valid preparations must be a valid preparation as well. Pure states correspond to the extremal points of the convex set $D(\mathbb{C}^d)$ and can be equivalently represented by unit vectors $|\psi\rangle \in \mathbb{C}^d$ where $\rho = |\psi\rangle\langle\psi|$. If $\mathrm{Tr}[\rho^2] = 1$ then the state ρ is *pure*. Otherwise, the state is *mixed*.

(ii) **Composition rule:** A composite quantum system consisting of subsystems $A_1, A_2, \cdots, A_n$ is associated with the Hilbert space $\bigotimes_{j=1}^{n} \mathbb{C}^{d_j}$, where for $j \in [n]$ the subsystem A_j has an associated Hilbert space $\mathbb{C}^{d_j}$.

Given the state $\rho_{A_1A_2\cdots A_n} \in D(\bigotimes_{j=1}^{n} \mathbb{C}^{d_j})$ of the composite system, the reduced state $\rho_{A_j} \in D(\mathbb{C}^{d_j})$ of subsystem A_j is obtained by taking the partial trace of the state $\rho_{A_1A_2\cdots A_n}$ over the complementary subsystems, *i.e.*,

$$\rho_{A_j} = \mathrm{Tr}_{\bar{A}_j}[\rho_{A_1A_2\cdots A_n}] \text{ where } \bar{A}_j = \{A_1, A_2, \cdots, A_n\} \backslash \{A_j\} \tag{2.1}$$

Note that any mixed state $\rho \in D(\mathbb{C}^d)$ can be realised as the reduced state of a pure state $\rho' = |\psi\rangle\langle\psi| \in D(\mathbb{C}^d \otimes \mathbb{C}^{d'})$ where $\mathbb{C}^{d'}$ is the Hilbert space associated with the ancilla. This procedure is called *purification*. All purifications of a mixed state are equivalent up to some local isometry acting on the ancilla [190].

(iii) **Measurement:** A measurement on a quantum system with Hilbert space $\mathbb{C}^d$ is described by a set of positive semi-definite operators $\{E_i \in \mathcal{B}(\mathbb{C}^d) : E_i \geq 0\}_{i=1}^{N}$ which satisfy $\sum_{i=1}^{N} E_i = I_d$. Here, I_d denotes the identity operator on $\mathbb{C}^d$. Such a set is called a positive operator-valued measure (POVM). A projective measurement (PVM) is a POVM where the operators are additionally orthogonal projectors, *i.e.*, $E_iE_j = \delta_{i,j}E_i$ for all $i, j \in [N]$.

(iv) **Probability rule:** The probability of obtaining an outcome $i \in [N]$ for a POVM $\{E_i\}_{i=1}^{N}$ on the quantum system prepared in the state $\rho \in D(\mathbb{C}^d)$ is given by Born's rule. *i.e.*, $p(i) = \mathrm{Tr}[\rho E_i]$. The post-measurement state, conditioned on the observed outcome i, is given by $\rho' = \frac{K_i \rho K_i^\dagger}{\mathrm{Tr}\left[K_i \rho K_i^\dagger\right]}$ where the operators K_i satisfy $E_i = K_i^\dagger K_i$ for $i \in [N]$.

Analogous to the state space, any convex mixture of POVMs is also a valid POVM. Consequently, the set of all measurements on a quantum system forms a convex set. An extremal POVM cannot be realised as a convex mixture of other POVMs. For a d-dimensional system, any extremal POVM has at most $N = d^2$ outcomes [191, 192] and such a POVM with d^2 outcomes always exist [191]. It is also worth noting that the post-measurement state is not uniquely determined by the POVM element E_i.

By Naimark's dilation theorem [193–196], any POVM $\{E_i\}_{i=1}^{N}$ on a d-dimensional quantum system can be realised as a PVM $\{F_i\}_{i=1}^{N}$ on an extended system. Specifically, this composite system consists of the original system and an ancilla of dimension at most N. Importantly, on a system of dimension d, it is still possible to implement a POVM with more than d outcomes. We now define a particular class of such measurements, which we refer to as *projective simulable* (PS) measurements.

Definition 2.1. An N-outcome POVM $\{E_i \in \mathcal{B}(\mathbb{C}^d) : E_i \geq 0\}_{i=1}^N$, with $\sum_{i=1}^N E_i = I_d$, is called projective simulable if there exist a projective measurement $\{F_j \in \mathcal{B}(\mathbb{C}^d) : F_j \geq 0\}_{j=1}^d$ on d dimensional system, *i.e.* $F_j F_k = \delta_{j,k}$ and $\sum_{j=1}^d F_j = I_d$, and a conditional probability distribution $\{p(i|j)\}_{i=1}^N$ for all $j \in [d]$ such that $E_i = \sum_{j=1}^d p(i|j) F_j$.

Such a measurement can be realised by performing a projective measurement on the d-dimensional quantum system and post-processing its outcomes.

Definition 2.2. An N-outcome POVM $\{E_i \in \mathcal{B}(\mathbb{C}^d) : E_i \geq 0\}_{i=1}^N$, with $\sum_{i=1}^N E_i = I_d$, is called *non-projective simulable* (nPS) if it is not projective simulable, *i.e.* if it cannot be expressed in the form specified in Definition 2.1.

This notion differs from the definition introduced in [197–199]. In [197], an N-outcome measurement on a d-dimensional system is called projective measurement simulable if it can be realised by randomising over different PVMs performed on the d-dimensional system, followed by post-processing of the outcomes. The work [197] provides a generalisation of Naimark's theorem, proving that any N-outcome measurement on a d-dimensional system is projective measurement simulable using a d-dimensional ancilla. Considering probabilistic simulation, later it was shown that any measurement on a d-dimensional system is projective measurement simulable with a success probability of $\frac{1}{d}$ without any ancilla by additionally post-selecting measurement outcomes [198]. Subsequently, it was shown that using similar resources and a single qubit ancilla, some measurements can be simulated with success probability independent of the system dimension d [200]. More recently, it was proven that measurements on a d-dimensional system are projective measurement simulable with a constant success probability of 0.125 using only a single qubit ancilla [199]. In contrast, here in Definition 2.1, we focus on the simulability of measurements, particularly using a single projective measurement and classical post-processing, without any ancilla or randomisation over multiple projective measurements and post-selection. We will discuss detection schemes for qubit and qutrit nPS measurements in Chapter 4.

(v) **Transformation:** In quantum theory, the evolution of a state $\rho \in D(\mathbb{C}^d)$ is given by a completely positive and trace preserving (CPTP) linear map $\mathbf{\Phi} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^{d'})$. These conditions are described below:

(i) $\forall \rho \in \mathcal{B}(\mathbb{C}^d),\ \mathrm{Tr}[\rho] = \mathrm{Tr}[\mathbf{\Phi}(\rho)]$ (trace preserving)

(ii) $\forall \rho \in \mathcal{B}(\mathbb{C}^n \otimes \mathbb{C}^d)$ with $\rho \geq 0, n \in \mathbb{N}$, $(\mathbf{I}_n \otimes \mathbf{\Phi})(\rho) \geq 0$ (completely positive)

where $\mathbf{I}_n$ represents the identity map on $\mathcal{B}(\mathbb{C}^n)$.

Alternately, $\mathbf{\Phi} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^{d'})$ is CPTP *iff* there is a set of Kraus operators $\{K_i : \mathbb{C}^d \to \mathbb{C}^{d'}\}_{i=1}^n$ that satisfy $\sum_{i=1}^n K_i^\dagger K_i = I_d$. For $\rho \in D(\mathbb{C}^d)$, $\mathbf{\Phi}(\rho) = \sum_{i=1}^n K_i \rho K_i^\dagger$. In general, the Kraus representations of a CPTP map are non-unique. Moreover, the evolution of a closed system is described by a unitary operator $U : \mathbb{C}^d \to \mathbb{C}^d$, and the transformed state is given by $\rho' = U\rho U^\dagger$. There are linear maps $\tilde{\mathbf{\Phi}} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^{d'})$ which are positive but not completely positive. While such maps transform any positive operator in $\mathcal{B}(\mathbb{C}^d)$ to some positive operator in $\mathcal{B}(\mathbb{C}^{d'})$, they do not satisfy the condition of complete positivity. Such positive but non-completely positive maps are useful for detecting quantum entanglement [201, 202], which is a non-classical feature.

Let us consider a simple example of a two-dimensional quantum system with Hilbert space $\mathbb{C}^2$, also referred to as a qubit. The state of a qubit $\rho \in D(\mathbb{C}^2)$ can be represented as: $\rho = \frac{1}{2}(I_2 + \vec{n} \cdot \vec{\sigma})$ where $\vec{n} \in \mathbb{R}^3$, $|\vec{n}| \leq 1$, $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. In this case, each vector $\vec{n}$ corresponds to a state, and the unit sphere representing the qubit state space is called the *Bloch sphere*. For qubit pure states ($|\vec{n}| = 1$), the state can equivalently be represented as a unit vector $|\psi\rangle \in \mathbb{C}^2$. The eigenvectors corresponding to the $+1$ and -1 eigenvalues of the operator σ_3 are represented as $|0\rangle$ and $|1\rangle$, respectively.

A general form of N-outcome POVM for qubits can be given as $\{E_i = \kappa_i(I_2 + \vec{n_i} \cdot \vec{\sigma})\}_{i=1}^N$ where $\kappa_i \geq 0$, $\vec{n_i} \in \mathbb{R}^3$, $|\vec{n_i}| \leq 1$, $\sum_{i=1}^N \kappa_i = 1$ and $\sum_{i=1}^N \kappa_i \vec{n_i} = \vec{0}$. Analogous to pure states, one can consider extremal POVMs for qubits. For $N = 2$, qubit extremal POVMs correspond to projective measurements of the form $\{E_1 = \frac{1}{2}(I_2 + \vec{n} \cdot \vec{\sigma}), E_2 = \frac{1}{2}(I_2 - \vec{n} \cdot \vec{\sigma})\}$, where $\vec{n} \in \mathbb{R}^3$ is a unit vector. For $N = 3$ or 4, a qubit extremal POVM is composed of elements with $\vec{n}_i \in \mathbb{R}^3$ being linearly independent unit vectors and $\kappa_i > 0$ for all $i \in [N]$. These measurements cannot be realised by post-processing the outcomes of some 2-outcome qubit projective measurements and are examples of qubit non-projective simulable measurements. Note that there exists no extremal POVM on a qubit with more than four outcomes [191]. For a closed qubit system, the evolution is given by some unitary $U = e^{-i\frac{\theta}{2}(\vec{n} \cdot \vec{\sigma})} = \cos\frac{\theta}{2}\ I_2 - i \sin\frac{\theta}{2}\ (\vec{n} \cdot \vec{\sigma})$ where $\vec{n} \in \mathbb{R}^3$ is a unit vector. More general evolutions for qubits are given by CPTP maps. A typical example is the depolarising map $\mathbf{\Phi}_p : \mathcal{B}(\mathbb{C}^2) \to \mathcal{B}(\mathbb{C}^2)$ defined as $\mathbf{\Phi}_p(\rho) = p\,\rho + (1-p)\frac{I_2}{2}$ where $0 \leq p \leq 1$ and $\rho \in D(\mathbb{C}^2)$.

Now, we will briefly discuss a feature of quantum states called *entanglement*, which essentially sparked the debate on the incompleteness of quantum description of reality and hidden variable models for quantum theory [138].

Definition 2.3. A state $\rho_{A_1A_2\cdots A_n} \in D(\bigotimes_{j=1}^n \mathbb{C}^{d_j})$ of a composite quantum system $A_1, A_2 \cdots A_n$ is separable if it can be expressed as $\rho_{A_1A_2\cdots A_n} = \sum_k p_k \bigotimes_{j=1}^n \rho_{A_j}^{(k)}$ where $\forall k\ \rho_{A_j}^{(k)} \in D(\mathbb{C}^{d_j})$, $0 \leq p_k \leq 1$, and $\sum_k p_k = 1$.

Definition 2.4. A state $\rho_{A_1 A_2 \cdots A_n} \in D(\bigotimes_{j=1}^{n} \mathbb{C}^{d_j})$ of a composite quantum system $A_1, A_2 \cdots A_n$ is entangled if it is not separable.

Some examples of entangled states in $\mathbb{C}^2 \otimes \mathbb{C}^2$ are the two-qubit Bell states (see Equation (2.2)). The entanglement for a pure state can be easily determined. A pure state is entangled if the reduced state of any subsystem is mixed. In general, there is no such easy way to check the entanglement of a mixed state. Interestingly, when $\rho_{A,B} \in D(\mathbb{C}^d \otimes \mathbb{C}^{d'})$, the Peres-Horodecki criterion [202] provides a test for entanglement of mixed states. A state $\rho_{AB} \in D(\mathbb{C}^d \otimes \mathbb{C}^{d'})$ is entangled if it is not positive under partial transposition, i.e., $\rho_{AB}^{T_B} \ngeq 0$. This is both necessary and sufficient for entanglement when $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ or $D(\mathbb{C}^2 \otimes \mathbb{C}^3)$. Subsequently, several works have further investigated such a notion of entanglement witness [203–205].

$$|\Phi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|00\rangle_{AB} \pm |11\rangle_{AB}\Big) \tag{2.2}$$

$$|\Psi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|01\rangle_{AB} \pm |10\rangle_{AB}\Big) \tag{2.3}$$

These are commonly referred to as two-qubit maximally entangled states, as the reduced state $\rho_A = \rho_B = \frac{I_2}{2}$ in each case. The state $|\Psi^-\rangle_{AB}$ is also called a singlet. It can be equivalently expressed as $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(\left|\psi\psi^{\perp}\right\rangle_{AB} - \left|\psi^{\perp}\psi\right\rangle_{AB}\Big)$ or any orthonormal basis $\{|\psi\rangle, \left|\psi^{\perp}\right\rangle\}$ of $\mathbb{C}^2$.

2.2 General Probabilistic Theories

Quantum theory exhibits several features with no classical counterpart, such as entanglement, superposition, and non-joint measurability. Such non-classical features, along with the current formalism of the theory, motivate some foundational questions. Can the theory be derived from physical axioms? Could there be a theory more fundamental than quantum mechanics? These philosophical questions also have physical significance. Modifying the physical axioms while not violating constraints like no-signalling may help in developing a theory of quantum gravity. The framework of General Probabilistic Theories (GPTs) helps explore wider physical theories with minimal assumptions [161, 206]. This framework is useful in detecting signatures of non-classicality [207, 208].

These assumptions are motivated by scenarios that can arise in a laboratory. A physical theory **TH** specifies a set of system types and the rules for the composition of multiple systems. The key ingredients associated with any physical theory within the framework include preparation procedures that prepare a system in a state, transformation, measurement, and a probability rule for the outcomes of a measurement. Consider a system $Sys_A(\Omega_A, \mathcal{E}_A)$ of type A, where V_A is a finite-dimensional real vector space associated with the system and V_A^* is the dual space.

(i) **State space:** There exists a unique linear functional $u_A : \Omega_A \to [0,1]$, where $\Omega_A \subset V_A$, called the deterministic (or unit) effect, which maps every $\omega \in \Omega_A$ to 1, *i.e.*, $u_A(\omega) = 1 \; \forall \; \omega \in \Omega_A$. $\Omega_A \subset V_A$ is then the normalised state space of the system $Sys_A(\Omega_A, \mathcal{E}_A)$, which is a convex and compact set of normalised states ω.

$V_{A+} = \{c\omega : \omega \in \Omega_A, c \geq 0\}$ is the convex positive cone containing unnormalised states. From the convexity of the state space, any statistical mixture $p\omega + (1-p)\omega' \in \Omega_A$, where $0 \leq p \leq 1$, of preparations $\omega, \omega' \in \Omega_A$ is also a valid preparation. Additionally, the state space Ω_A must be closed for a physical theory because of the assumption that a state $\omega \in V_{A+}$ can be prepared if there are preparation procedures for states arbitrarily close to ω [206].

Definition 2.5. A state $\omega \in \Omega_A$ is *pure* if for all $\omega', \omega'' \in \Omega_A$ and $0 \leq p \leq 1$, $\omega = p\omega' + (1-p)\omega'' \implies \omega = \omega' = \omega''$, *i.e.*, ω is an extremal point of the convex set Ω_A. Otherwise, ω is a *mixed state*.

In the physical theory, a measurement can be described by a certain collection of elementary effects or linear functionals $e \in V_A^*$ defined over the state space.

(ii) **Effect space:** Let $\Omega_A^* \subset V_{A+}^*$ denote the set of all linear functionals $e : \Omega_A \to [0,1]$ defined over the state space Ω_A. $V_{A+}^* = \{f \in V_A^* : f(\omega) \geq 0 \; \forall \; \omega \in V_{A+}\}$ is the set of positive linear functionals on the state cone. The effect space $\mathcal{E}_A$ for the system $Sys_A(\Omega_A, \mathcal{E}_A)$ is a subset of Ω_A^*. An element $e \in \mathcal{E}_A$ is called an effect. Although $\mathcal{E}_A \subseteq \Omega_A^*$ in general, we will use the *no-restriction hypothesis* [209, 210] and consider theories for which all the linear functionals $e \in \Omega_A^*$ are valid effects, *i.e.*, $\mathcal{E}_A = \Omega_A^*$.

(iii) **Measurement and probability rule:** A measurement $\mathcal{M}_A = \{e_i \in \mathcal{E}_A\}_{i=1}^N$ is a set of effects such that $\sum_{i=1}^N e_i = u_A$. The probability of obtaining outcome associated with effect e_i when performing the measurement $\mathcal{M}_A$ on a state $\omega \in \Omega_A$ is given by $p(e_i|\omega) = \mathrm{Tr}\left[e_i^T \cdot \omega\right]$.

(iv) **Transformation:** While we are interested in the transformation of a single system, its linearity depends on a property of composite systems in the theory, namely local tomography (discussed later in this section). Thus, any transformation is a linear map $\mathcal{T} : V_A \to V_B$ that maps every unnormalised element $\omega \in V_{A+}$ to an unnormalised state $\omega' \in V_{B+}$. Also, these maps either preserve or decrease total probability, *i.e.* $u_B(\mathcal{T}(\omega)) \leq u_A(\omega) \; \forall \; \omega \; \in \Omega_A$. u_B is the unit effect of the system of type B. Deterministic transformations (channels) preserve normalisation for all the states, *i.e.*, $u_B \circ \mathcal{T} = u_A$ [206, 211].

Measurements can be used to distinguish states. It is also possible to identify each state in some ensemble using a single measurement, making them operationally distinguishable.

Definition 2.6. A set of states $\{\omega_i \in \Omega_A\}_{i=1}^N$ of a system $Sys_A(\Omega_A, \mathcal{E}_A)$ are jointly perfectly distinguishable if there exists a measurement $\mathcal{M}_A = \{e_i \in \mathcal{E}_A\}_{i=1}^N$ such that $p(e_j|\omega_i) = \delta_{i,j}$ for all $i, j \in [N]$.

Using ensembles of distinguishable states, it is possible to generalise the notion of *dimension* for GPTs. As communication or sharing correlation requires physical systems, this notion will be particularly useful for comparing communication and shared resources across different theories, especially quantum and classical theories.

Definition 2.7. Operational dimension $\mathbf{d}_A$ of a $Sys_A(\Omega_A, \mathcal{E}_A)$ is the largest cardinality of a subset of states in Ω_A that are jointly perfectly distinguishable by a single measurement.

Now we will discuss systems in GPTs, which consist of multiple subsystems of possibly different types. In the theory **TH**, consider the simplest scenario where a system consists of sub-systems $Sys_A(\Omega_A, \mathcal{E}_A)$ and $Sys_B(\Omega_B, \mathcal{E}_B)$ of type A and B, respectively. The joint system AB is associated with a finite-dimensional real vector space V_{AB}.

(ii) **Composition rule:** The theory **TH** prescribes the composition rule for a joint system $Sys_{AB}(\Omega_{AB}, \mathcal{E}_{AB})$. This relates the state space Ω_{AB} as well as the effect space $\mathcal{E}_{AB}$ of $Sys_{AB}(\Omega_{AB}, \mathcal{E}_{AB})$ respectively with the state spaces Ω_A, Ω_B and effect spaces $\mathcal{E}_A$, $\mathcal{E}_B$ of the subsystems. The normalised state space Ω_{AB} is a convex compact subset of the positive cone V_{AB+}. There must be $\omega_{AB} \in \Omega_{AB}$ corresponding to every $\omega_A \in \Omega_A$ and $\omega_B \in \Omega_B$. Similarly, there also exists an effect $e_{AB} = (e_A, e_B) \in \mathcal{E}_{AB}$ corresponding to each pair of local effects $e_A \in \mathcal{E}_A$ and $e_B \in \mathcal{E}_B$.

The identification of the bipartite state (or effects) with local states (effects) is linear [206]. For a joint state $\omega_{AB} \in \Omega_{AB}$, the probability of obtaining an outcome corresponding to an effect e_{AB} is $p(e_{AB}|\omega_{AB}) = \mathrm{Tr}\left[e_{AB}^T \cdot \omega_{AB}\right] \in [0, 1]$. Although Ω_{AB} is not uniquely defined based on Ω_A and Ω_B, for theories satisfying the following, it is possible to define $\Omega_{AB} = \Omega_A \otimes \Omega_B \subset V_{A+} \otimes V_{B+}$:

(a) For $(\omega_1)_{AB}, (\omega_2)_{AB} \in \Omega_{AB}$, if $\mathrm{Tr}\left[e_{AB}^T \cdot (\omega_1)_{AB}\right] = \mathrm{Tr}\left[e_{AB}^T \cdot (\omega_2)_{AB}\right]$ for $e_{AB} = (e_A, e_B) \in \mathcal{E}_{AB}$ corresponding to each pair of effects $e_A \in \mathcal{E}_A$ and $e_B \in \mathcal{E}_B$, then $(\omega_1)_{AB} = (\omega_2)_{AB}$. This is referred to as *local tomography*. In other words, joint states can be fully characterised using statistics from local measurements.

(b) For any state $\omega_{AB} \in \Omega_{AB}$ of the composite system, the marginal probabilities of outcomes obtained for measurement on system $Sys_A(\Omega_A, \mathcal{E}_A)$ should not depend on the choice of measurement on the system $Sys_B(\Omega_B, \mathcal{E}_B)$ and vice versa.

In this case, the effect space is $\mathcal{E}_{AB} = \Omega^*_{AB}$. The unit effect on the composite system is $u_{AB} = u_A \otimes u_B$. The product effects are defined as $e_{AB} = e_A \otimes e_B \in \mathcal{E}_{AB}$, where $e_A \in \mathcal{E}_A$ and $e_B \in \mathcal{E}_B$. Similarly, the product states are given by $\omega_{AB} = \omega_A \otimes \omega_B \in \Omega_{AB}$, where $\omega_A \in \Omega_A$ and $\omega_B \in \Omega_B$. The probability of obtaining an outcome corresponding to effect $e_{AB} = e_A \otimes e_B$ is $p(e_{AB} = e_A \otimes e_B|\omega_A \otimes \omega_B) = p(e_A|\omega_A)p(e_B|\omega_B)$ for all effects $e_A \in \mathcal{E}_A$ and $e_B \in \mathcal{E}_B$. In general, the choice of the tensor product in $\Omega_A \otimes \Omega_B$ is not unique but lies between two extreme types [161]:

Definition 2.8. The maximal tensor product $\Omega_A \otimes_{max} \Omega_B := \{\omega_{AB} \in V_A \otimes V_B : (e_A \otimes e_B)(\omega_{AB}) \geq 0 \; \forall \; e_A \in \mathcal{E}_A, e_B \in \mathcal{E}_B$ and $(u_A \otimes u_B)(\omega_{AB}) = 1\}$.

Definition 2.9. The minimal tensor product $\Omega_A \otimes_{min} \Omega_B$ is the convex hull of all product states $\omega_{AB} = \omega_A \otimes \omega_B \in \Omega_{AB}$, where $\omega_A \in \Omega_A$ and $\omega_B \in \Omega_B$.

Now, we consider a few examples like classical theory, quantum theory and some post-quantum theories like box world.

2.2.1 Classical Theory

In the case of classical theory $\mathbf{TH} = Cl$, the state space $(\Omega_d)_A$ of a d-level classical system $Sys_A((\Omega_d)_A, (\mathcal{E}_d)_A)$ is defined as $(\Omega_d)_A = \{(p_0, p_1, \ldots, p_{d-1}) \in [0,1]^d : \sum_{i=0}^{d-1} p_i = 1\}$. Each state $\omega \in (\Omega_d)_A$ is a probability distribution over d possible outcomes. Pure states are deterministic distributions $\omega_i = (p_0, p_1, \ldots, p_{d-1})$, where $p_i = 1$ for $i \in \{0, 1, \cdots, d-1\}$. The effect space $(\mathcal{E}_d)_A = [0,1]^d$. The unit effect $u_A = (1, 1, \cdots, 1)$. Measurement is defined by a set of effects $(\mathcal{M}_A^{d\to N})_{Cl} = \{e_i \in (\mathcal{E}_d)_A\}_{i=1}^N$ satisfying $\sum_{i=1}^N e_i = u_A$. Equivalently, with $e_i = (w_{i,1}, w_{i,2}, \cdots, w_{i,d})$, the measurement can be represented as a column stochastic matrix $W_A^{d\to N}$ with element $(W_A^{d\to N})_{i,j} = w_{i,j}$ and $\sum_{i=1}^N w_{i,j} = 1 \; \forall j \in [d]$. The probability of obtaining an outcome corresponding to the effect $e_i = (w_{i,1}, w_{i,2}, \cdots, w_{i,d})$ for a state $\omega = (p_0, p_1, \ldots, p_{d-1})$ is given by $p(e_i|\omega) = \sum_{j=1}^d w_{i,j} p_{j-1}$. In other words, the probability of obtaining the i^{th} outcome is given by the i^{th} element of $(W_A^{d\to N})(\omega)^T$. The operational dimension of the d-level classical system $Sys_A((\Omega_d)_A, (\mathcal{E}_d)_A)$ is d.

Note that in Chapter 4, we will use the convention where N outcome measurement $(\mathcal{M}_A^{d\to N})_{Cl} = \{e_i \in (\mathcal{E}_d)_A\}_{i=0}^{N-1}$ satisfying $\sum_{i=0}^{N-1} e_i = u_A$. In this case, the measurement is represented as a column stochastic matrix $W_A^{d\to N}$ with element $(W_A^{d\to N})_{i,j} = w_{i,j}$ where $i \in \{0, 1, \cdots, N-1\}, j \in \{0, 1, \cdots, d-1\}$ and $\sum_{i=0}^{N-1} w_{i,j} = 1 \; \forall j \in \{0, 1, \cdots, d-1\}$. The probability of obtaining an outcome corresponding to the effect $e_i = (w_{i,0}, w_{i,1}, \cdots, w_{i,d-1})$, where $i \in \{0, 1, \cdots, N-1\}$, for a state $\omega = (p_0, p_1, \ldots, p_{d-1})$ is given by $p(e_i|\omega) = \sum_{j=0}^{d-1} w_{i,j} p_j$.

For a composite classical system $Sys_{AB}((\Omega_{d''})_{AB}, (\mathcal{E}_{d''})_{AB})$ consisting of two subsystems $Sys_A((\Omega_d)_A, (\mathcal{E}_d)_A)$ and $Sys_B((\Omega_{d'})_B, (\mathcal{E}_{d'})_B)$, the state space

$(\Omega_{d''})_{AB} = \{(p_{0,0}, p_{0,1}, \ldots, p_{d-1,d'-1}) \in [0,1]^{d''} : \sum_{i=0}^{d-1}\sum_{j=0}^{d'-1} p_{i,j} = 1\}$ with $d'' = dd'$. The effect space $(\mathcal{E}_{d''})_{AB} = [0,1]^{dd'}$.

2.2.2 Quantum Theory

We will denote quantum theory by $\mathbf{TH} = Q$. As discussed in Section 2.1, a system $Sys_A(\Omega_A, \mathcal{E}_A)$ is associated with Hilbert space $\mathbb{C}^d$. If we denote $V_A = Herm(\mathbb{C}^d) = \mathbb{R}^{d^2}$ as the space of hermitian operators on $\mathbb{C}^d$ and $V_{A+} = \{O \in V_A : O \geq 0\}$ as its positive cone, then it can be seen that the state space $\Omega_A = D(\mathbb{C}^d) \subset V_{A+}$. The unit effect in quantum theory is $u_A = I_d$. The effect space is $\mathcal{E}_A = \{E \in \mathcal{B}(\mathbb{C}^d) : E \geq 0\}$. Each POVM defines a measurement. The probability of obtaining an outcome associated with the POVM element $E \in \mathcal{E}_A$ conditioned on the system's state $\rho \in \Omega_A$ is given by Born rule, *i.e.*, $p(E|\rho) = \mathrm{Tr}[E\rho]$. The composition rule is given using the tensor product. In quantum theory, only mutually orthogonal states can be distinguished perfectly using a single measurement. Thus, the operational dimension of a quantum system associated with $\mathbb{C}^d$ is d.

2.2.3 Box world

Beyond standard quantum mechanics, there are alternative models, often grouped under the term "post-quantum theories". Now, we will provide an example of a particular post-quantum theory called the box-world $\mathbf{TH} = \mathfrak{B}$ [212]. In the Bell scenario, the boxworld states can generate correlations that go beyond those allowed in quantum theory [161, 213]. The state space Ω_A of an elementary system $Sys_A(\Omega_A, \mathcal{E}_A)$ in the box-world is $\Omega_A = Conv\{\omega_1, \omega_2, \omega_3, \omega_4\} \subset \mathbb{R}^3$, the convex hull of four extremal points. Here, we denote the convex hull of x and y by $Conv\{x, y\} = \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$. These four states are the pure states, and they form the vertices of a square in $\mathbb{R}^3$, *i.e.*

$$\omega_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \omega_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}, \omega_3 = \begin{pmatrix}-1\\0\\1\end{pmatrix}, \omega_4 = \begin{pmatrix}0\\-1\\1\end{pmatrix} \tag{2.4}$$

In the box-world, the probability of observing the outcome associated with an effect $e \in \mathcal{E}_A$ on $\omega \in \Omega_A$ is $p(e|\omega) = \mathrm{Tr}\left[e^T \cdot \omega\right]$. The zero effect and unit effect are two special effects. The unit effect plays the role of normalisation, and the zero effect describes an impossible event. They are given below.

$$\mathbb{O} = \begin{pmatrix}0\\0\\0\end{pmatrix}, u_A = \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{2.5}$$

The effect space can be written as $\mathcal{E}_A = Conv\{\mathbb{O}, u_A, e_1, e_2, e_3, e_4\} \subset \mathbb{R}^3$ with the extremal effects given below

$$e_1 = \frac{1}{2}\begin{pmatrix}1\\1\\1\end{pmatrix}, e_2 = \frac{1}{2}\begin{pmatrix}-1\\1\\1\end{pmatrix}, e_3 = \frac{1}{2}\begin{pmatrix}-1\\-1\\1\end{pmatrix}, e_4 = \frac{1}{2}\begin{pmatrix}1\\-1\\1\end{pmatrix} \tag{2.6}$$

A measurement is a collection of effects $\{f_i \in \mathcal{E}_A\}_{i=1}^N$ satisfying $\sum_{i=1}^N f_i = u_A$. Note that the set of states $\{\omega_1, \omega_2\}$ are perfectly distinguishable using the measurement defined by the effects $\{f_1 = e_4, f_2 = e_2\}$, *i.e.* $p(f_i|\omega_j) = \delta_{i,j}$ for $i, j \in \{1, 2\}$. Additionally, any set of three or more states in Ω_A is not perfectly distinguishable using a single measurement in the theory. Thus, the operational dimension of the system $Sys_A(\Omega_A, \mathcal{E}_A)$ is 2.

Consider a composite system $Sys_{AB}(\Omega_{AB}, \mathcal{E}_{AB})$ consisting of two elementary sub-systems $Sys_A(\Omega_A, \mathcal{E}_A)$ and $Sys_B(\Omega_B, \mathcal{E}_B)$ of the same type. Here, we will discuss a specific composition rule for the Box world that gives rise to the PR-model, which in the bipartite scenario allows realisation of a correlation called the PR-box [213]. The PR-box represents the strongest non-local correlations consistent with the no-signalling principle (refer to Section 2.3). The state space $\Omega_{AB} = Conv\{(\omega_1)_{AB}, (\omega_2)_{AB}, \cdots, (\omega_{24})_{AB}\}$ where

$$(\omega_{4(i-1)+j})_{AB} = (\omega_i)_A \otimes (\omega_j)_B^T \quad \text{for } i, j \in [4],\ (\omega_i)_A \in \Omega_A,\ (\omega_j)_B \in \Omega_B \tag{2.7}$$

are extremal product states. The extremal entangled states [212] in the PR-model are

$$(\omega_{17})_{AB} = \frac{1}{2}\left((\omega_2)_A \otimes (\omega_2)_B^T - (\omega_3)_A \otimes (\omega_3)_B^T + (\omega_3)_A \otimes (\omega_4)_B^T + (\omega_4)_A \otimes (\omega_3)_B^T\right) \tag{2.8a}$$

$$(\omega_{18})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_4)_B^T - (\omega_1)_A \otimes (\omega_1)_B^T + (\omega_2)_A \otimes (\omega_2)_B^T + (\omega_4)_A \otimes (\omega_1)_B^T\right) \tag{2.8b}$$

$$(\omega_{19})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_1)_B^T - (\omega_2)_A \otimes (\omega_2)_B^T + (\omega_2)_A \otimes (\omega_3)_B^T + (\omega_3)_A \otimes (\omega_2)_B^T\right) \tag{2.8c}$$

$$(\omega_{20})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_1)_B^T - (\omega_1)_A \otimes (\omega_4)_B^T + (\omega_2)_A \otimes (\omega_4)_B^T + (\omega_4)_A \otimes (\omega_3)_B^T\right) \tag{2.8d}$$

$$(\omega_{21})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_4)_B^T - (\omega_1)_A \otimes (\omega_1)_B^T + (\omega_2)_A \otimes (\omega_1)_B^T + (\omega_4)_A \otimes (\omega_2)_B^T\right) \tag{2.8e}$$

$$(\omega_{22})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_1)_B^T - (\omega_1)_A \otimes (\omega_2)_B^T + (\omega_2)_A \otimes (\omega_2)_B^T + (\omega_4)_A \otimes (\omega_3)_B^T\right) \tag{2.8f}$$

$$(\omega_{23})_{AB} = \frac{1}{2}\left((\omega_2)_A \otimes (\omega_2)_B^T - (\omega_3)_A \otimes (\omega_2)_B^T + (\omega_3)_A \otimes (\omega_3)_B^T + (\omega_4)_A \otimes (\omega_1)_B^T\right) \tag{2.8g}$$

$$(\omega_{24})_{AB} = \frac{1}{2}\left((\omega_1)_A \otimes (\omega_2)_B^T - (\omega_2)_A \otimes (\omega_2)_B^T + (\omega_2)_A \otimes (\omega_3)_B^T + (\omega_3)_A \otimes (\omega_1)_B^T\right) \tag{2.8h}$$

The effect space of the composite system is $\mathcal{E}_{AB} = \{(e)_{AB} = (e)_A \otimes (e)_B^T : (e)_A \in \mathcal{E}_A, (e)_B \in \mathcal{E}_B\}$. It contains all extremal product effects $(e_{4(i-1)+j})_{AB} = (e_i)_A \otimes (e_j)_B^T$ where $(e_i)_A \in \mathcal{E}_A, (e_j)_B \in \mathcal{E}_B$ and $i, j \in [4]$. Next, we will discuss the no-signalling principle and the correlation that can arise only using local operations on a shared state in certain theories.

2.3 No-signalling Correlations and Bell Inequality

The study of non-locality in the foundations of quantum mechanics was initiated as a consequence of the seminal paper by EPR [138], which argued on the incompleteness of the quantum description of reality. They considered a bipartite quantum system in an entangled state. The outcomes corresponding to two non-commuting observables for one of the subsystems can be guessed perfectly without directly interacting with it [214]. However, the quantum description of the state does not account for these realities and, according to EPR, is therefore incomplete. The implication that quantum theory could be completed using some local hidden variable description was later disproven [133, 215]. The principle of no-signalling and non-existence of local hidden variable models using the violation of Bell inequalities has been extensively studied. Here we will briefly overview some of the notions and refer the reader to books such as [216] for a detailed introduction to the topic.

Consider an experiment involving two space-like separated parties, say Alice and Bob, who receive randomly chosen inputs $x \in \mathsf{X}$ and $y \in \mathsf{Y}$ and output an outcome $a \in \mathsf{A}$ and $b \in \mathsf{B}$ respectively. This setup is referred to as the Bell scenario or the non-local scenario, and we denote it by $\mathscr{B}^{\mathsf{X},\mathsf{Y}}_{\mathsf{A},\mathsf{B}}$. $\mathsf{X}, \mathsf{Y}, \mathsf{A}, \mathsf{B}$ are finite sets. The parties can access shared states $\omega_{AB} \in \Omega_{AB}$ allowed in some theory **TH**. The subsystems of Alice and Bob are of type A and B, respectively. In the experiment, based on their received input $x \in \mathsf{X}$ and $y \in \mathsf{Y}$, Alice and Bob select a state $\omega_{AB} \in \Omega_{AB}$ according to a probability distribution $\{p(\omega_{AB}|x,y)\}_{\omega_{AB},x,y}$. Alice performs an $|\mathsf{A}|$-outcome measurement on sub-system A based on the input x and outputs the outcome $a \in \mathsf{A}$ of the measurement. Similarly, Bob performs a $|\mathsf{B}|$-outcome measurement on the sub-system B based on his input y and outputs the outcome $b \in \mathsf{B}$ of the measurement. Using the shared state and measurements, they obtain the conditional probability distribution $\{P(a,b|x,y,\omega_{AB})\}_{a,b,x,y}$. Here, $P(a,b|x,y,\omega_{AB})$ denotes the conditional probability of Alice and Bob getting outcomes a and b when performing measurements labelled x and y, respectively, on the shared state ω_{AB}. The conditional probability distribution $P = \{P(a,b|x,y)\}_{a,b,x,y}$ is obtained by averaging over the state space and can be expressed as

$$P(a,b|x,y) = \int_{\omega_{AB}\in\Omega_{AB}} p(\omega_{AB}|x,y)P(a,b|x,y,\omega_{AB})\, d\omega_{AB} \tag{2.9}$$

Let the state space of a local hidden variable theory be $\Omega_{AB} = \Lambda$. Say, Alice and Bob share states $\lambda \in \Lambda$ from the hidden variable theory distributed according to the probability $\{p(\lambda)\}_{\lambda\in\Lambda}$. In such a scenario, we use the assumption of *measurement independence* or *freedom of choice*, according to which the $p(\lambda|x,y) = p(\lambda)\ \forall x,y,\lambda$. In other words, using Bayes' theorem, the measurement setting is randomly chosen and is independent of the shared state. Using this in Equation (2.9),

$$P(a, b|x, y) = \int_{\lambda \in \Lambda} p(\lambda) P(a, b|x, y, \lambda)\ d\lambda \tag{2.10}$$

For local hidden variable theory, the conditional probability distribution $\{P(a, b|x, y, \lambda)\}_{a,b,x,y}$ factorise or satisfy the property of *locality* for all $\lambda \in \Lambda$, *i.e.*,

$$P(a, b|x, y, \lambda) = P(a|x, \lambda)P(b|y, \lambda) \tag{2.11}$$

The outcomes a or b, respectively, should not depend on the choice of measurement y and x. In this case, the correlation $P = \{P(a, b|x, y)\}_{a,b,x,y}$ obtained in the Bell scenario using shared state and measurements described by a local hidden variable model takes the form:

$$P(a, b|x, y) = \int_{\lambda \in \Lambda} p(\lambda) P(a|x, \lambda)P(b|y, \lambda)\ d\lambda \tag{2.12}$$

For a deterministic local hidden variable model, the shared state λ completely determines the outcome of measurements indexed by $x \in \mathsf{X}$ for Alice and $y \in \mathsf{Y}$ for Bob. Given such λ, the outcomes of measurement x and y are determined by some function $\mathbb{LA} : \mathsf{X} \to \mathsf{A}$ and $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$, respectively. Thus, $P(a|x, \lambda) = \delta_{a,\mathbb{LA}(x)}$ and $P(b|y, \lambda) = \delta_{b,\mathbb{LB}(y)}$. The cardinality of the deterministic local hidden variable states λ in the state space Λ is equal to the number of such pairs of deterministic functions, *i.e.* $|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}$.

More generally, assume Alice and Bob share a black-box having a measurement setting corresponding to inputs $x \in \mathsf{X}$ and $y \in \mathsf{Y}$, respectively. Additionally, the box outputs outcomes $a \in \mathsf{A}$ and $b \in \mathsf{B}$ for Alice and Bob, respectively, based on their local inputs. Say they obtain a conditional probability distribution $P = \{P(a, b|x, y)\}_{a,b,x,y}$ in the experiment. This correlation could, in principle, be realised using systems and measurements in some physical theories, such as classical, quantum, or post-quantum theory. Since Alice and Bob are spatially separated, this correlation must satisfy the no-signalling principle.

Definition 2.10. A correlation $P = \{P(a, b|x, y)\}_{a,b,x,y}$ in the Bell scenario is *no-signalling* if it satisfies the following constraints:

$$\sum_b P(a, b|x, y) = \sum_b P(a, b|x, y') = P_A(a|x)\ \forall\ a, x, y, y' \tag{2.13}$$

$$\sum_a P(a, b|x, y) = \sum_a P(a, b|x', y) = P_B(b|y)\ \forall\ b, x, x', y \tag{2.14}$$

$\{P_A(a|x)\}_{a,x}$ and $\{P_B(b|y)\}_{b,y}$ are marginal distributions of Alice and Bob, respectively. The Equations (2.13) and (2.14) state that Bob's measurement choice does not affect Alice's marginal conditional probabilities, and vice versa. Thus, neither party can infer the other's measurement choice from their outcome statistics, and the correlation cannot be used for signalling. Any correlation P in the Bell scenario that admits a local hidden variable description as in Equation (2.12) is necessarily no-signalling, but not all no-signalling correlations admit such a model. Additionally, any no-signalling correlations P that admit a local hidden variable description can also be realised by a deterministic local hidden variable model [217]. The reverse implication is trivial. In what follows, we use the term local hidden variable model to mean the deterministic local hidden variable model, given their equivalence in the Bell scenario. The local hidden variable models correspond to the classical description and are often referred to as *shared randomness*.

Let $\mathcal{L}$ be the set of no-signalling correlations P in the Bell scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$ that have a local hidden variable realisation. This set is convex as $P = pP' + (1-p)P'' \in \mathcal{L}$ for all $P', P'' \in \mathcal{L}$ and $p \in [0,1]$. Its extremal points correspond to correlations obtained by fixing a shared state $\lambda \in \Lambda$, *i.e.* setting $p(\lambda) = 1$ in Equation (2.12). Thus, $\mathcal{L}$ is a polytope with $|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}$ local extremal points $\{P_L^{(i)}\}_{i=1}^{|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}}$. It is referred to as the local polytope. Similarly, the set of all no-signalling correlations forms a polytope $\mathcal{NS}$ referred to as the no-signalling polytope. This convex set lies in an D-dimensional real vector space $\mathbb{R}^D$ [218]. The subspace $\mathcal{NS}$ is specified by elements satisfying the positivity and normalisation constraints for the conditional probability as well as the no-signalling constraints given in Equations (2.13) and (2.14). Each constraint defines a hyperplane, and since their number is finite, $\mathcal{NS}$ is a polytope. The extremal points $\{P_L^{(i)}\}_{i=1}^{|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}}$ of local polytope $\mathcal{L}$ are also extremal points of $\mathcal{NS}$. An extremal point $P_{NL} = \{P_{NL}(a,b|x,y)\}_{a,b,x,y}$ of $\mathcal{NS}$ such that $P_{NL} \notin \{P_L^{(i)}\}_{i=1}^{|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}}$ is a non-local extremal point. Let $\{P_{NL}^{(i)}\}_i$ denote the set of all non-local extremal points.

Definition 2.11. A non-local facet[1] $Face\{P_{NL}^{(i)}\}_{i\in J}$ is the convex hull $Conv\{P_{NL}^{(i)}\}_{i\in J}$ of non-local extremal points with indices in J such that for all $P \in Face\{P_{NL}^{(i)}\}_{i\in J}$, $P \neq pP' + (1-p)P_L^{(j)}$ where $P_L^{(j)} \in \{P_L^{(i)}\}_{i=1}^{|\mathsf{A}|^{|\mathsf{X}|}|\mathsf{B}|^{|\mathsf{Y}|}}$, $P' \in \mathcal{NS}$ and $p \in [0,1)$.

A no-signalling correlation $P \in \mathcal{NS}$ can become noisy. A particular kind of noise is the white noise correlation $P_{WN} = \{P_{WN}(a,b|x,y) = \frac{1}{|\mathsf{A}||\mathsf{B}|}\}_{a,b,x,y} \in \mathcal{NS}$.

Definition 2.12. A correlation $P' \in \mathcal{NS}$ lies on the isotropic line connecting $P \in \mathcal{NS}$ and white noise $P_{WN} = \{P_{WN}(a,b|x,y) = \frac{1}{|\mathsf{A}||\mathsf{B}|}\}_{a,b,x,y}$ if there is $p \in [0,1]$ such that $P' = pP + (1-p)P_{WN}$.

A correlation $P \in \mathcal{NS}$ is quantum realisable if there is a shared quantum state $\rho_{AB} \in D(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ and POVMs $\mathcal{M}_A^x = \{E_a^x \in \mathcal{B}(\mathbb{C}^{d_A}) : E_a^x \geq 0, \sum_{a\in\mathsf{A}} E_a^x = I_{d_A}\}$ $\forall x \in \mathsf{X}$ for Alice and POVMs $\mathcal{M}_B^y = \{E_b^y \in \mathcal{B}(\mathbb{C}^{d_B}) : E_b^y \geq 0, \sum_{b\in\mathsf{B}} E_b^y = I_{d_B}\}$ $\forall y \in \mathsf{Y}$ for Bob such that

$$P(a,b|x,y) = \mathrm{Tr}[\rho_{AB} E_a^x \otimes E_b^y] \quad \forall \quad a,b,x,y \tag{2.15}$$

[1]Throughout this thesis, we use the term "non-local facet" in a broader sense, referring to any non-local *face* of the no-signalling polytope, not only facets in the strict polyhedral sense of dimension $\dim(\mathcal{NS}) - 1$.

We denote using $\mathcal{Q} \subset \mathcal{NS}$, the set of all no-signalling correlations $P \in \mathcal{NS}$ which are quantum realisable. In general, it is a hard problem to determine if a correlation $P \in \mathcal{NS}$ is quantum realisable or not [219, 220]. In Appendix B, we discuss some known methods to bound the set of quantum realisable correlations $\mathcal{Q}$ from the outside.

Consider a hyperplane in $\mathcal{NS}$ such that the local polytope $\mathcal{L}$ lies entirely on one side. Positivity, normalisation, and no-signalling constraints on the conditional probabilities define trivial hyperplanes. It is possible to define some non-trivial hyperplanes, or Bell functionals.

Definition 2.13. In the Bell scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$, the Bell functional is a linear map $\mathbf{B} : \mathcal{NS} \to \mathbb{R}$ acting on the space of no-signalling correlations, and can be written in the form

$$\mathbf{B}(P) = \sum_{a,b,x,y} c_{a,b}^{x,y} P(a,b|x,y) \tag{2.16}$$

where $P \in \mathcal{NS}$ and $c_{a,b}^{x,y} \in \mathbb{R}$.

Although one can also consider non-linear functionals in general, here we will restrict ourselves to a linear Bell functional. The local bound β_L for the Bell functional is given by

$$\beta_L = \max_{P \in \mathcal{L}} \mathbf{B}(P) \tag{2.17}$$

Since $\mathbf{B}(P) \leq \beta_L$ for all $P \in \mathcal{L}$, any violation of this inequality $\mathbf{B}(P') > \beta_L$ for some correlation $P' \in \mathcal{NS}$ implies that it is non-local, *i.e.* $P' \in \mathcal{NS} \backslash \mathcal{L}$. A standard example is the CHSH inequality [164]. In the Bell scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$ where $\mathsf{X} = \mathsf{Y} = \mathsf{A} = \mathsf{B} = [2]$ and the Bell functional

$$\mathbf{B}(P) = \sum_{a,b,x,y} c_{a,b}^{x,y} P(a,b|x,y) \tag{2.18}$$
$$\text{where } c_{a,b}^{x,y} = \begin{cases} (-1)^{(a-1)\oplus_2(b-1)} & \text{if } (x,y) \in \{(1,1),(1,2),(2,1)\} \\ (-1)^{(a-1)\oplus_2(b-1)\oplus_2 1} & \text{if } x=2, y=2 \end{cases}$$

The local bound for the Bell functional is $\beta_L = 2$. Interestingly, some quantum realisable correlations violate the local bound. The maximum quantum violation for CHSH is $2\sqrt{2}$ [221]. This can be obtained using a shared maximally entangled state $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|00\rangle_{AB} + |11\rangle_{AB}\Big)$. Alice's observables are σ_3 and σ_1, respectively, for $x = 1$ and $x = 2$. Bob's observables are $\frac{\sigma_3+\sigma_1}{\sqrt{2}}$ and $\frac{\sigma_3-\sigma_1}{\sqrt{2}}$, respectively, for $y = 1$ and $y = 2$. This establishes that quantum theory cannot be described by any local hidden variable model. However, quantum theory does not achieve the maximum no-signalling value of this Bell function. This maximum violation $\max_{P\in\mathcal{NS}} \mathbf{B}(P) = 4$ and can be achieved by a non-local extremal point $P_{NL}^{(0,0,0)}$ of $\mathcal{NS}$, also called a PR-box [213] as defined below. In this scenario, there are eight non-local extremal points or PR-boxes denoted by $P_{NL}^{(\alpha,\beta,\gamma)} = \{P_{NL}^{(\alpha,\beta,\gamma)}(a,b|x,y)\}_{a,b,x,y}$, $(\alpha,\beta,\gamma) \in \{0,1\}^3$, where

$$P_{NL}^{(\alpha,\beta,\gamma)}(a,b|x,y) = \frac{\delta_{(a-1)\oplus_2(b-1),f(x,y,\alpha,\beta,\gamma)}}{2} \tag{2.19}$$
$$\text{and } f(x,y,\alpha,\beta,\gamma) = (x-1)(y-1) \oplus_2 \alpha(x-1) \oplus_2 \beta(y-1) \oplus_2 \gamma$$

$x\backslash y$	1	2	3	$\cdots$	$Y-g$	$Y-g+1$	$\cdots$	Y
1	$\frac{1}{2}I_2$	$\frac{1}{2}I_2$	$\frac{1}{2}I_2$	$\cdots$	$\frac{1}{2}I_2$	L	$\cdots$	L
2	$\frac{1}{2}I_2$	$\frac{1}{2}\sigma_1$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	$\cdots$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	L	$\cdots$	L
3	$\frac{1}{2}I_2$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	$\cdots$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	L	$\cdots$	L
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
$X-h$	$\frac{1}{2}I_2$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	$\cdots$	$\frac{1}{2}I_2$ or $\frac{1}{2}\sigma_1$	L	$\cdots$	L
$X-h+1$	K	K	K	$\cdots$	K	M	$\cdots$	M
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
X	K	K	K	$\cdots$	K	M	$\cdots$	M

Table 2.1: Representation of a non-local extremal point in the Bell scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$ with $\mathsf{A} = \mathsf{B} = 2$ and $\mathsf{X} = [X], \mathsf{Y} = [Y]$. Here, $g, h \in \mathbb{N}, 0 \leq g \leq Y-2, 0 \leq h \leq X-2$ and entries of each block matrix corresponding to an x, y pair represent $P(a, b|x, y)$. The rows of the block matrix are indexed by $a \in \{1, 2\}$ and columns are indexed by $b \in \{1, 2\}$.

These correlations can be realised in the boxworld GPTs [212]. More generally, in the Bell scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$ with $\mathsf{A} = \mathsf{B} = [2]$ and $\mathsf{X} = [X], \mathsf{Y} = [Y]$, the extremal non-local points of the $\mathcal{NS}$ polytope take the form (up to local relabelling) as shown in Table 2.1 [222]. Here $g, h \in \mathbb{N}, 0 \leq g \leq Y-2, 0 \leq h \leq X-2$ and

$$\mathrm{K} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \end{pmatrix}, \mathrm{L} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{pmatrix}, \mathrm{M} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.20}$$

Each entry in Table 2.1, corresponding to $x \in \mathsf{X} = [X]$ and $y \in \mathsf{Y} = [Y]$, is a block matrix. This matrix gives values of $P(a, b|x, y)$ where rows index $a \in \{1, 2\}$ and columns index $b \in \{1, 2\}$.

Non-locality can also be revealed in Bell scenarios without Bell inequalities. It can be witnessed through some logical contradictions. Some popular examples are the GHZ paradox [223] and the Hardy paradox [224, 225]. The Hardy paradox was originally formulated in the bipartite scenario $\mathscr{B}_{\mathsf{A},\mathsf{B}}^{\mathsf{X},\mathsf{Y}}$ with $\mathsf{X} = \mathsf{Y} = \mathsf{A} = \mathsf{B} = [2]$, and later extended to multipartite scenarios as well as settings with more inputs and outputs [186, 226–233]. Here we will discuss the Hardy type correlations introduced in [186]. The authors considered a Bell scenario with $\mathsf{X} = \mathsf{Y} = [2]$ and $\mathsf{A} = \mathsf{B} = [d]$, where $d \in \mathbb{N} \setminus \{0, 1\}$. For a no-signalling correlation $P \in \mathcal{NS}$ we introduce two shorthand representations: $P(a < b|x, y) = \sum_{\substack{a,b: \\ a<b}} P(a, b|x, y)$ and $P(a > b|x, y) = \sum_{\substack{a,b: \\ a>b}} P(a, b|x, y)$. The Hardy paradox is expressed through the following conditions on P:

$$P(a < b|2, 1) = 0 \tag{2.21a}$$

$$P(a > b|1, 1) = 0 \tag{2.21b}$$

$$P(a < b|1, 2) = 0 \tag{2.21c}$$

$$P(a < b|2, 2) > 0 \tag{2.21d}$$

The final inequality indicates a non-zero probability of an outcome that cannot be explained classically if the other conditions are satisfied. For quantum realisable no-signalling correlations $P \in \mathcal{Q}$ satisfying these conditions, we define the Hardy payoff, which quantifies the non-classicality of a correlation.

$$H^{d}_{Hardy}(P) = P(a < b|2, 2) \tag{2.22}$$

For $d = 2$, constraints in Equation (2.21) reduce to Hardy's paradox as in [224]. Every local correlation $P \in \mathcal{L}$ violates at least one of the constraints given in Equation (2.21). Some quantum correlations satisfy all constraints. However, correlations from bipartite maximally entangled states fail to satisfy them for any $d \in \mathbb{N}\backslash\{0, 1\}$ [186]. Correlations obtained from mixed two-qubit entangled states also fail to satisfy these constraints for $d = 2$. In quantum theory, a maximum value of $H^{d=2}_{Hardy} \approx 0.09$ can be achieved for $d = 2$ [224]. This is obtained using two-qubit entangled states. Interestingly, the maximum value of H^{d}_{Hardy} achievable in quantum theory increases with d and is bounded from above by 0.5 [186].

The violation of Bell inequality and the logical contradictions highlight the presence of non-classical correlation. To formalise this perspective, one can adopt a resource-theoretic approach. Now, we briefly outline the resource theory of local operations and shared randomness (LOSR) [234, 235]. In the Bell scenario, shared randomness can be taken as a free resource. Any local operation on the shared state by spatially separated parties is a free operation. Using only free resources and free operations, Alice and Bob can generate correlations within the local polytope $\mathcal{L}$. Furthermore, some shared quantum states also yield correlations in $\mathcal{L}$ for all choices of local operations in quantum theory. These states are equivalent to free resources and are considered "classical" in this framework. Examples include separable states and certain entangled states [146]. This notion of classical resources can be extended to any physical theory $\mathbf{TH}$. Violation of a Bell inequality in this framework implies the presence of non-classical resources. Since Alice and Bob treat their system as a black box, assuming only no-signalling, such detection of non-classical resources using Bell-inequality violation is called *device independent* [216].

In contrast to the device-independent scenario, additional assumptions about the inner workings of the shared apparatus define a semi-device-independent framework. For example, assuming the shared system is quantum, the maximum quantum violation of some Bell inequality can self-test the shared state and measurements up to local isometry [147, 148]. Similarly, assuming a bound on the local dimension of the shared state can enable certification of certain non-classical measurements. In [151], the authors considered a Bell scenario $\mathscr{B}^{\mathsf{X,Y}}_{\mathsf{A,B}}$ where $\mathsf{X} = [3], \mathsf{Y} = [2], \mathsf{A} = [3], \mathsf{B} = [2]$. For $x = 1, 2$, the outcome $a = 3$ does not effectively contribute to the Bell functional $\mathbf{B}$ they studied. Assuming the shared state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ and general qubit POVMs, they obtained correlations P for which $\mathbf{B}(P)$ exceeded the value possible with any projective measurement on a two-qubit state. This enables semi-device-independent detection of non-projective measurements. These scenarios typically assume random inputs, *i.e.*, measurement independence. A natural extension is to lift this assumption to detect measurements on d-dimensional system that are neither projective nor post-processing of outcomes obtained using a projective measurement on the same system.

2.4 Graphs and Orthogonal Representation

First, we recall some basic notions from graph theory. These will later be useful when we discuss orthogonal representations and binary colourings, which are also relevant in quantum foundations and quantum information [139, 236].

Definition 2.14. A graph $\mathcal{G}$ is defined by a set of vertices $\mathbf{V}(\mathcal{G}) = \{v_1, v_2, \ldots, v_n\}$ and edges $\mathbf{E}(\mathcal{G})$.

In a graph, an edge is denoted by a pair $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$. It connects the two vertices v_i and v_j, which are also called its endpoints. The number of vertices, $n = |\mathbf{V}(\mathcal{G})|$, is called the order of the graph. If the number of vertices in the graph is finite, then $\mathcal{G}$ is called a finite graph. We will only consider finite graphs. In general, a graph may also have infinite vertices. The degree of a vertex $v_i \in \mathbf{V}(\mathcal{G})$ is the number of edges having v_i as an endpoint. Vertices $v_i, v_j \in \mathbf{V}(\mathcal{G})$ are adjacent (or neighbours) if $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$. A graph may have multiple edges between the same pair of vertices. A *loop* is an edge with identical endpoints, *i.e.* $(v_i, v_i) \in \mathbf{E}(\mathcal{G})$ for some $i \in [n]$. Also, some of the edges $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ could be considered as an ordered tuple, as in the case of directed graphs. If all edges $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ are unordered tuples, then $\mathcal{G}$ is an undirected graph. Additionally, in weighted graphs, the edges are associated with a number. Here, we will consider only certain types of graphs called *simple graphs*.

Definition 2.15. A simple graph is an undirected and unweighted graph which does not have any loops or multiple edges between any two vertices.

We now recall a few specific types of graphs that will be used later.

Definition 2.16. A complete graph $\mathcal{G}$ is a simple graph in which every pair of distinct vertices is adjacent, *i.e.*, $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ for all distinct $v_i, v_j \in \mathbf{V}(\mathcal{G})$

Definition 2.17. The complement $\bar{\mathcal{G}}$ of a simple graph $\mathcal{G}$ is also a simple graph with the same vertex set $\mathbf{V}(\bar{\mathcal{G}}) = \mathbf{V}(\mathcal{G}) = \{v_1, v_2, \cdots, v_n\}$, and $(v_i, v_j) \in \mathbf{E}(\bar{\mathcal{G}})$ *iff* $v_i \neq v_j$ and $(v_i, v_j) \notin \mathbf{E}(\mathcal{G})$.

A graph is called self-complementary if it is isomorphic (see Definition 2.22 below) to its complement. A graph $\mathcal{G}$ is k-regular if each vertex in the graph has degree k. For simple graphs, the degree of a vertex is also the number of its neighbours.

Definition 2.18. An induced subgraph $\mathcal{H}$ of a graph $\mathcal{G}$ has vertex set $\mathbf{V}(\mathcal{H}) \subseteq \mathbf{V}(\mathcal{G})$ and $(v_i, v_j) \in \mathbf{E}(\mathcal{H})$ *iff* $v_i, v_j \in \mathbf{V}(\mathcal{H})$ and $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$.

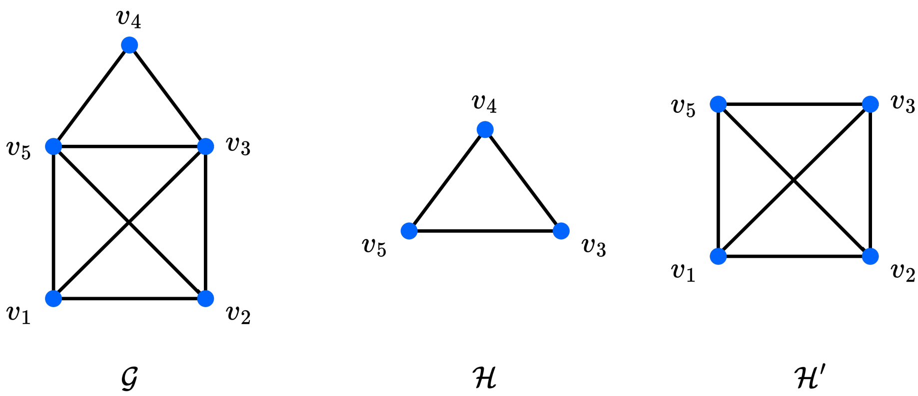

Figure 2.1: A graph $\mathcal{G}$ of order $n = 5$ along with two induced subgraphs $\mathcal{H}$ and $\mathcal{H}'$.

Definition 2.19. A clique of a graph $\mathcal{G}$ is an induced subgraph of $\mathcal{G}$ which is a complete graph. The size of a clique is the order of the induced subgraph.

Definition 2.20. A maximum clique of $\mathcal{G}$ is a clique of the largest size (order) in the graph. The size of a maximum clique is the clique number $\omega(\mathcal{G})$.

As an example, consider the graph $\mathcal{G}$ in Figure 2.1 with $\mathbf{V}(\mathcal{G}) = \{v_1, v_2, v_3, v_4, v_5\}$ and $\mathbf{E}(\mathcal{G}) = \{(v_i, v_j) : i, j \in \{1, 2, 3, 5\}, i < j\} \cup \{(v_i, v_j) : i, j \in \{3, 4, 5\}, i < j\}$. $\mathcal{H}$ and $\mathcal{H}'$ are induced subgraphs of $\mathcal{G}$. Both are cliques of $\mathcal{G}$, and $\mathcal{H}'$ is a maximum clique of $\mathcal{G}$. The clique number of $\mathcal{G}$ is $\omega(\mathcal{G}) = 4$.

An independent set of $\mathcal{G}$ is a subset of vertices in $\mathbf{V}(\mathcal{G})$ that are pairwise non-adjacent. Such subsets correspond to cliques of the complementary graph $\bar{\mathcal{G}}$. A simple graph $\mathcal{G}$ of order n can also be represented using an $n \times n$ matrix $A_{\mathcal{G}}$ called the *adjacency matrix* of $\mathcal{G}$. In this matrix, $(A_{\mathcal{G}})_{i,j} = 1$ if $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ and 0 otherwise for $i, j \in [n]$. Note that the adjacency matrix is symmetric, *i.e.* $A_{\mathcal{G}} = A_{\mathcal{G}}^T$, and its diagonal entries $(A_{\mathcal{G}})_{i,i} = 0$ for the simple graph. The eigenvalues of $A_{\mathcal{G}}$ are real and form the *spectrum of the graph*. Between two graphs $\mathcal{G}$ and $\mathcal{G}'$, a homomorphism is defined as follows:

Definition 2.21. A graph homomorphism from a simple graphs $\mathcal{G}$ to a graph $\mathcal{G}'$ is a function $f : \mathbf{V}(\mathcal{G}) \to \mathbf{V}(\mathcal{G}')$ such that if $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$, for $v_i, v_j \in \mathbf{V}(\mathcal{G})$, then $(f(v_i), f(v_j)) \in \mathbf{E}(\mathcal{G}')$.

If such a map exists, $\mathcal{G}$ is homomorphic to the graph $\mathcal{G}'$. If the inverse $f^{-1} : \mathbf{V}(\mathcal{G}') \to \mathbf{V}(\mathcal{G})$ is also a homomorphism from $\mathcal{G}'$ to $\mathcal{G}$, then f is called a graph isomorphism. In this case, $\mathcal{G}$ and $\mathcal{G}'$ are isomorphic.

Definition 2.22. A graph isomorphism $f : \mathbf{V}(\mathcal{G}) \to \mathbf{V}(\mathcal{G})$ from a simple graph $\mathcal{G}$ to itself is called a graph automorphism.

A graph automorphism of $\mathcal{G}$ is a permutation of its vertices that preserves adjacency. The set of all automorphisms is closed under composition and forms the automorphism group of $\mathcal{G}$. A graph $\mathcal{G}$ is vertex transitive if for each pair of vertices $v_i, v_j \in \mathbf{V}(\mathcal{G})$, there exists an automorphism $f : \mathbf{V}(\mathcal{G}) \to \mathbf{V}(\mathcal{G})$ such that $f(v_i) = v_j$. Examples of vertex-transitive and self-complementary graphs include Paley graphs. A Paley graph $\mathcal{G}_{Paley(q)}$, where $q = 4k + 1 = p^m$ for some prime p and positive $k, m \in \mathbb{N}$, has q vertices. Each vertex is associated with a distinct element of the finite field $\mathbb{F}_q$. Two vertices are adjacent if their difference is a quadratic residue (non-zero square) in the finite field $\mathbb{F}_q$. We next discuss orthogonal representations of a graph. An orthogonal representation maps the vertices of a graph to vectors in a d-dimensional inner product space $\mathbb{F}^d$, defined over a field $\mathbb{F}$, while satisfying some additional constraints [236, 237].

Definition 2.23. Given a graph $\mathcal{G}$, an orthogonal representation (OR) is a function $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{F}^d$, where $\mathbb{F}^d$ is a d-dimensional inner product space defined over the field $\mathbb{F}$, such that adjacency implies orthogonality, *i.e.* if $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ then $\langle \phi(v_i) | \phi(v_j) \rangle = 0$ where $\langle . | . \rangle$ is the inner product in $\mathbb{F}^d$.

An upper bound on the minimum dimension d in which $\mathcal{G}$ has *orthogonal representation* is $d \leq \chi(\mathcal{G})$ as each colour can be mapped to a distinct orthonormal vector to obtain an orthogonal representation. Here, $\chi(\mathcal{G})$ is the chromatic number of $\mathcal{G}$, which is the minimum number of colours needed to assign different colours to all adjacent vertices in $\mathcal{G}$. In some OR, non-adjacent vertices may also be assigned orthogonal vectors. Although an OR may assign the same vector to different vertices, we assume distinct assignments. This non-degeneracy condition appears in the definition of a general position orthogonal representation [238], where representing vectors must be linearly independent.

Definition 2.24. An orthogonal representation $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{F}^d$ is a faithful orthogonal representation (FOR) if (i) it is injective and (ii) for all $v_i, v_j \in \mathbf{V}(\mathcal{G})$, $\langle \phi(v_i) | \phi(v_j) \rangle = 0 \implies (v_i, v_j) \in \mathbf{E}(\mathcal{G})$.

If a FOR ϕ of a graph $\mathcal{G}$ satisfies $|\phi(v_i)| = 1$ for all $v_i \in \mathbf{V}(\mathcal{G})$, then it is called a faithful orthonormal representation. With mild abuse of language, we will use the term *faithful orthogonal representation* to refer to such orthonormal representations. Moreover, if $\mathcal{G}$ has a faithful orthogonal representation ϕ in $\mathbb{R}^d$, the same function ϕ is trivially a faithful orthogonal representation in $\mathbb{C}^d$. For a graph $\mathcal{G}$, the faithful orthogonal range is the minimum dimension d such that there exists a faithful orthogonal representation in $\mathbb{R}^d$ [237]. Let us denote it by $\xi(\mathcal{G})$. It satisfies the following relation:

$$\omega(\mathcal{G}) \leq \xi(\mathcal{G}) \leq |\mathbf{V}(\mathcal{G})| \tag{2.23}$$

This follows trivially since all vertices in a maximum clique must be assigned orthogonal vectors. In a real vector space of dimension $|\mathbf{V}(\mathcal{G})|$, there are enough orthogonal vectors to satisfy

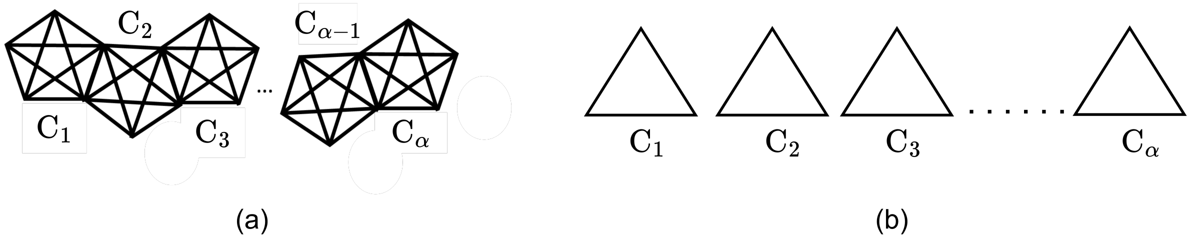

Figure 2.2: Two examples of graph $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$. In example (a), $\beta = 5$ and $\gamma = 2$. In example (b), $\beta = 3$ and $\gamma = 0$. The graph in example (b) will be denoted as $\mathcal{G}_{disc.}^{(\alpha,\beta=3)}$ since all the maximum cliques are disconnected in this graph.

orthogonality for all adjacent vertices [237]. In general, finding the faithful orthogonal range is hard. The graphs with maximum vertex degree $deg(\mathcal{G})$ have FOR in dimension 2 $deg(\mathcal{G})$ [238]. A graph $\mathcal{G}$ has FOR in dimension $\tilde{d}$ if its complementary graph $\bar{\mathcal{G}}$ is $(|\mathbf{V}(\mathcal{G})| - \tilde{d})$ connected [238]. A graph $\mathcal{G}$ is k-connected ($k < |\mathbf{V}(\mathcal{G})|$) if it remains connected even after removing fewer than k vertices. We now discuss a family of graphs and their faithful orthogonal range, which will be relevant to some of the results in Chapter 4.

Definition 2.25. Let $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$ be a graph where $\alpha, \beta, \gamma \in \mathbb{N}$, $\alpha, \beta \geq 2$, and $\gamma \in [0, \frac{\beta}{2})$. The graph consists of a chain of α maximum cliques, each of size β. Consecutive maximum cliques C_i and C_{i+1} share exactly γ vertices for $i \in \{1, 2, \cdots, \alpha - 1\}$, and the rest of the maximum cliques do not share any additional vertices and edges.

When $\gamma = 0$, all the maximum cliques of $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$ are disconnected. We denote this case by $\mathcal{G}_{disc.}^{(\alpha,\beta)}$. See Figure 2.2 for examples. We will now determine the faithful orthogonal range of the graph $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$.

Lemma 2.1. *The faithful orthogonal range of the graph* $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$ *is* $\xi(\mathcal{G}_{\gamma}^{(\alpha,\beta)}) = \beta$.

Proof. [2] For $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$, the faithful orthogonal range is at least β since the vertices in a maximum clique must be assigned mutually orthogonal vectors. As $\omega(\mathcal{G}_{\gamma}^{(\alpha,\beta)}) = \beta$, any faithful orthogonal representation requires a vector space dimension of at least β. We now show that $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$ has an FOR in $\mathbb{R}^{\beta}$ using the method of induction.

Case $\alpha = 2$:
Consider this simplest case. Let C_1 and C_2 denote the two maximum cliques, with vertex sets $\mathbf{V}(\mathrm{C}_1)$ and $\mathbf{V}(\mathrm{C}_2)$ respectively. From Definition 2.25, these cliques share $|\mathbf{V}(\mathrm{C}_1) \cap \mathbf{V}(\mathrm{C}_2)| = \gamma$ vertices. Also, the graph contains no other edges besides those within each clique. We begin

[2]Alternative proof using results from [238] was given in [184]. Here we prove it using elementary techniques.

by assigning an orthonormal basis of $\mathbb{R}^\beta$ to the vertices in C_1. In other words, we map each $v \in \mathbf{V}(C_1)$ to a unit vector $f(v) \in \mathbb{R}^\beta$. The shared vertex $v \in \mathbf{V}(C_1) \cap \mathbf{V}(C_2)$ is already assigned a unit vector in this process. Let $V_{1,2} = span\{f(v) : v \in \mathbf{V}(C_1) \cap \mathbf{V}(C_2)\} \subset \mathbb{R}^\beta$ be the subspace spanned by these shared vertices. It has dimension γ and its orthogonal complement $V_{1,2}^\perp$ is a subspace of dimension $\beta - \gamma$.

The $\beta - \gamma$ vertices in $\mathbf{V}(C_2) \backslash \mathbf{V}(C_1)$ can be assigned mutually orthogonal unit vectors in $V_{1,2}^\perp$. Notice that each of these vectors is by definition orthogonal to the vectors already assigned to the γ vertices in $\mathbf{V}(C_1) \cap \mathbf{V}(C_2)$. However, individually they must be non-orthogonal to the vectors assigned to vertices in $\mathbf{V}(C_1) \backslash \mathbf{V}(C_2)$ in order to satisfy the faithfulness condition. We will show that such an assignment exists using the argument we present next.

Step 1: We follow this step if $\beta - \gamma > 2$, otherwise we continue to **Step** 2. For the first $\beta - \gamma - 2$ vertices in $\mathbf{V}(C_2) \backslash \mathbf{V}(C_1) = \{v_1, v_2, \cdots, v_{\beta-\gamma}\}$ we proceed in the following manner. We define $V_{1,2}^{(0)} := V_{1,2}^\perp$, which is a subspace of dimension $\beta - \gamma$ (> 2). At the t^{th} iteration, suppose $(t-1)$ mutually orthogonal unit vectors have already been chosen from $V_{1,2}^{(0)}$ and assigned to the first $(t-1)$ vertices in $\mathbf{V}(C_2) \backslash \mathbf{V}(C_1)$. In other words, $\forall \ i \in [t-1]$, $v_i \in \mathbf{V}(C_2) \backslash \mathbf{V}(C_1)$ is assigned $f(v_i) \in V_{1,2}^{(0)}$ such that: (i) $\langle f(v_i)|f(v_j)\rangle = 0$ for $i \neq j \in [t-1]$ and (ii) $\langle f(v')|f(v_i)\rangle \neq 0$ for $v' \in \mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$. We now choose the unit vector that will be assigned to $v_t \in \mathbf{V}(C_2) \backslash \mathbf{V}(C_1)$. Let $V_{1,2}^{(t)} = \{x \in V_{1,2}^{(0)} : \langle x|f(v_i)\rangle = 0 \ \forall \ i \in [t-1]\}$ be the orthogonal complement of $span\{f(v_1), f(v_2), \cdots, f(v_{t-1})\}$ in $V_{1,2}^{(0)}$. The dimension of the subspace $V_{1,2}^{(t)}$ is $\beta - \gamma - (t-1)$. If $f(v_t) \in V_{1,2}^{(t)}$ then $\langle f(v_i)|f(v_t)\rangle = 0$ for all $i \in [t-1]$. For faithfulness, the following should be satisfied: $\langle f(v_t)|f(v')\rangle \neq 0$ for $v' \in \mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$. For each $v' \in \mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$, the set of vectors $x \in V_{1,2}^{(t)}$ with $\langle x|f(v')\rangle = 0$ defines a hyperplane (closed set) of lower dimension in $V_{1,2}^{(t)}$, and must be avoided as assignment to v_t for satisfying faithfulness. Since there are finite vertices in $\mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$, the number of such hyperplanes is also finite. The union of all these forbidden hyperplanes thus forms a closed proper subspace of $V_{1,2}^{(t)}$ and cannot cover the entire space $V_{1,2}^{(t)}$. As $V_{1,2}^{(t)}$ contains infinite feasible points outside this union, any vector in this set can be assigned as $f(v_t)$ to the vertex v_t.

Step 2: For the last two vertices in the set $\mathbf{V}(C_2) \backslash \mathbf{V}(C_1)$, let $V_{1,2}^{(\beta-\gamma-1)} := \{x \in V_{1,2}^\perp : \langle x|f(v_i)\rangle = 0 \ \forall \ i \in [\beta - \gamma - 2]\}$ be the orthogonal complement of all previously assigned vectors in $V_{1,2}^\perp$. It is a subspace of dimension 2. The vertices $v_{\beta-\gamma-1}$ and $v_{\beta-\gamma}$ can be assigned vectors $f(v_{\beta-\gamma-1}) = x\cos\delta + x^\perp \sin\delta$ and $f(v_{\beta-\gamma}) = -x\sin\delta + x^\perp\cos\delta$, respectively, where $x, x^\perp \in V_{1,2}^{(\beta-\gamma-1)}$ are mutually orthogonal unit vectors. Here, $\delta \in [0, 2\pi]$ such that $\langle f(v_{\beta-\gamma-1})|f(v')\rangle \neq 0$ and $\langle f(v_{\beta-\gamma})|f(v')\rangle \neq 0$ for all $v' \in \mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$. This is required for satisfying the condition of faithfulness. The union of the solutions of $\langle f(v_{\beta-\gamma-1})|f(v')\rangle = 0$ and $\langle f(v_{\beta-\gamma})|f(v')\rangle = 0$ for each $v' \in \mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}) \backslash \mathbf{V}(C_2)$ is finite. The set of these finite values of δ is forbidden. As the set $[0, 2\pi]$ is dense, removing these finite points from the continuous set $[0, 2\pi]$ still leaves infinite feasible values for assignment as δ. The function f is a FOR of this graph.

Assume that for some $\alpha \geq 2$, $\mathcal{G}_{\gamma}^{(\alpha,\beta)}$ has FOR $f : \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha,\beta)}) \rightarrow \mathbb{R}^{\beta}$.

Case $\alpha \rightarrow \alpha + 1$:

We will now construct a FOR for the case $\alpha + 1$, *i.e.*, the graph is $\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}$. Let C_{α} and $\mathrm{C}_{\alpha+1}$ be the last two maximum cliques in $\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}$. From Definition 2.25, $|\mathbf{V}(\mathrm{C}_{\alpha}) \cap \mathbf{V}(\mathrm{C}_{\alpha+1})| = \gamma$. The vertices in $\mathbf{V}(\mathrm{C}_{\alpha})$, and the shared vertices $\mathbf{V}(\mathrm{C}_{\alpha}) \cap \mathbf{V}(\mathrm{C}_{\alpha+1})$, are assigned mutually orthogonal unit vectors $\{f(v) \in \mathbb{R}^{\beta} : v \in \mathbf{V}(\mathrm{C}_{\alpha})\}$. We define $V_{\alpha,\alpha+1} := \mathit{span}\{f(v) \in \mathbb{R}^{\beta} : v \in \mathbf{V}(\mathrm{C}_{\alpha}) \cap \mathbf{V}(\mathrm{C}_{\alpha+1})\}$, which is a γ-dimensional subspace of $\mathbb{R}^{\beta}$. Let $V_{\alpha,\alpha+1}^{(0)} := V_{\alpha,\alpha+1}^{\perp} \subset \mathbb{R}^{\beta}$ be its orthogonal complement subspace of dimension $\beta - \gamma$.

The $\beta - \gamma$ vertices in $\mathbf{V}(\mathrm{C}_{\alpha+1}) \setminus \mathbf{V}(\mathrm{C}_{\alpha}) = \{v_1, v_2, \cdots, v_{\beta-\gamma}\}$ must be assigned orthonormal vectors in $V_{\alpha,\alpha+1}^{(0)}$ such that: (i) they are mutually orthogonal and (ii) they are non-orthogonal to all $f(v')$ where $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$. Similar to the **Step** 1 of case $\alpha = 2$, the first $\beta - \gamma - 2$ vertices in $\mathbf{V}(\mathrm{C}_{\alpha+1}) \setminus \mathbf{V}(\mathrm{C}_{\alpha})$ can be assigned unit vectors in the following manner if $\beta - \gamma > 2$. Otherwise, we skip this step. At the t^{th} iteration, for $i \in [t-1]$ the vertices v_i have been already assigned unit vectors in $f(v_i) \in V_{\alpha,\alpha+1}^{(0)}$ such that: (i) $\langle f(v_i)|f(v_j)\rangle = 0$ for $i \neq j \in [t-1]$ and (ii) $\langle f(v')|f(v_i)\rangle \neq 0$ for $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$. A unit vector will be assigned now to $v_t \in \mathbf{V}(\mathrm{C}_{\alpha+1}) \setminus \mathbf{V}(\mathrm{C}_{\alpha})$ from $V_{\alpha,\alpha+1}^{(t)} = \{x \in V_{\alpha,\alpha+1}^{(0)} : \langle x|f(v_i)\rangle = 0 \ \forall \ i \in [t-1]\}$. $V_{\alpha,\alpha+1}^{(t)}$ is $\beta - \gamma - (t-1)$ dimensional subspace which is the orthogonal complement of $\mathit{span}\{f(v_1), f(v_2), \cdots, f(v_{t-1})\}$ in $V_{\alpha,\alpha+1}^{(0)}$. The unit vector $f(v_t)$ must satisfy the following: $\langle f(v_t)|f(v')\rangle \neq 0$ for $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$ for faithfulness. $\mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$ has finite set of vertices. For each $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$ and vector $x \in V_{\alpha,\alpha+1}^{(t)}$, $\langle f(x)|f(v')\rangle = 0$ defines a hyperplane of forbidden vectors in the subspace $V_{\alpha,\alpha+1}^{(t)}$. As before, a finite union of such hyperplanes, each corresponding to a $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$, is a proper subset of $V_{\alpha,\alpha+1}^{(t)}$. Thus, even after removing the forbidden set, infinitely many valid choices for $f(v_t)$ remain that satisfy the faithfulness condition.

For assigning vectors to the last two vertices, we consider a 2-dimensional subspace. We pick an orthonormal basis $x, x^{\perp}$ of $V_{\alpha,\alpha+1}^{(\beta-\gamma-1)} := \{x' \in V_{\alpha,\alpha+1}^{\perp} : \langle x'|f(v_i)\rangle = 0 \ \forall \ i \in [\beta - \gamma - 2]\}$ and assign $f(v_{\beta-\gamma-1}) = x\cos\delta' + x^{\perp}\sin\delta', f(v_{\beta-\gamma}) = -x\sin\delta' + x^{\perp}\cos\delta'$. Here, $\delta \in [0, 2\pi]$ such that $\langle f(v_{\beta-\gamma-1})|f(v')\rangle \neq 0$ and $\langle f(v_{\beta-\gamma})|f(v')\rangle \neq 0$ for all $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$. Each $v' \in \mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})$ gives at most two forbidden values of δ'. In other words, these are solutions of the equations $\langle f(v_{\beta-\gamma-1})|f(v')\rangle = 0$ and $\langle f(v_{\beta-\gamma})|f(v')\rangle = 0$. Since $|\mathbf{V}(\mathcal{G}_{\gamma}^{(\alpha+1,\beta)}) \setminus \mathbf{V}(\mathrm{C}_{\alpha+1})|$ is finite, the forbidden set of angles is finite. It is possible to choose

$\delta' \in [0, 2\pi]$ outside this forbidden set to satisfy the faithfulness condition, as $[0, 2\pi]$ is dense.

Thus, by defining f on the vertices of $\mathrm{C}_{\alpha+1}$ using these vectors, and keeping the assignment unchanged on the vertices of $\mathcal{G}_\gamma^{(\alpha,\beta)}$, yields a FOR for $\mathcal{G}_\gamma^{(\alpha+1,\beta)}$ in $\mathbb{R}^\beta$. Using induction, every graph $\mathcal{G}_\gamma^{(\alpha,\beta)}$ with $0 \leq \gamma < \frac{\beta}{2}$ has a faithful orthogonal representation in $\mathbb{R}^\beta$. Combined with the lower bound on faithful orthogonal range, this proves the lemma. ■

Observation 2.1. *If graph $\mathcal{G}$ with clique number $\omega(\mathcal{G})$ also has FOR in dimension $\omega(\mathcal{G})$, then any two distinct maximum cliques C, C' in the graph can have at most $\omega(\mathcal{G}) - 2$ vertices in common. Indeed, the contrary would imply the vertices that are not in common between the two cliques must be assigned the same vector if the representation is forced to be in dimension $\omega(\mathcal{G})$, thus violating faithfulness. Also, every vertex $v \in \mathbf{V}(\mathrm{C}) \setminus \mathbf{V}(\mathrm{C}')$ is adjacent to at most $\omega(\mathcal{G}) - 2$ vertices in $\mathbf{V}(\mathrm{C}')$ for such a graph.*

Another topic in graph theory which is also studied in quantum foundations is the notion of assigning *colours* to the vertices of a graph. Consider two colours denoted by 0 and 1. A binary colouring assigns one of these colours to each vertex of a graph. Binary colouring of graphs is related to quantum contextuality [139, 239]. Later, we will introduce the notion of a clique label based on the binary colouring of the vertices in a maximum clique of a graph.

Definition 2.26. A binary colouring for graph $\mathcal{G}$, if it exists, is a function $\mathsf{f} : \mathbf{V}(\mathcal{G}) \to \{0, 1\}$ satisfying the following:

(i) adjacent vertices are not assigned 1 simultaneously, *i.e.* $(v_i, v_j) \in \mathbf{E}(\mathcal{G}) \implies \mathsf{f}(v_i)\mathsf{f}(v_j) = 0$.

(ii) Exactly one vertex in every maximum clique C of the graph is assigned colour 1, *i.e.* $\sum_{v \in \mathbf{V}(\mathrm{C})} \mathsf{f}(v) = 1$.

In the above definition, $\mathbf{V}(\mathrm{C}) \subseteq \mathbf{V}(\mathcal{G})$ denotes the vertices in the maximum clique C. Based on Definition 2.26, we can define binary colouring for each maximum clique of a graph $\mathcal{G}$. In this case, $\mathbf{V}(\mathcal{G}) = \{v_1, v_2, \cdots, v_n\}$, where $n \in \mathbb{N}$, and C is a maximum clique of the graph of size $\omega(\mathcal{G})$. Also, let $\mathbf{V}(\mathrm{C}) = \{v_{i_1}, v_{i_2}, \cdots v_{i_{\omega(\mathcal{G})}}\} \subseteq \mathbf{V}(\mathcal{G})$ denote the set of vertices of maximum clique C such that $i_1 < i_2 < \cdots < i_{\omega(\mathcal{G})}$ and $i_k \in [n]$ for $k \in [\omega(\mathcal{G})]$. In other words, we have arranged the vertices in the maximum clique C in increasing order. A binary colouring for the maximum clique C is a function $\mathsf{f} : \mathbf{V}(\mathrm{C}) \to \{0, 1\}$ that assigns value 1 to exactly one vertex, as in Definition 2.26. The set of all binary colourings for the clique can be denoted by F_C. Clearly, $|\mathsf{F}_\mathrm{C}| = \omega(\mathcal{G})$. We can define clique labelling in the following manner, based on the binary colourings of a maximum clique in the graph $\mathcal{G}$.

Definition 2.27. For a maximum clique C of graph $\mathcal{G}$, clique labelling is a map $g_\mathrm{C} : \mathsf{F}_\mathrm{C} \to \{0, 1, \cdots, \omega(\mathcal{G}) - 1\}$ such that $g_\mathrm{C}(\mathsf{f}) = k - 1$ for $\mathsf{f} \in \mathsf{F}_\mathrm{C}$ if $\mathsf{f}(v_{i_k}) = 1$ where $v_{i_k} \in \mathbf{V}(\mathrm{C}) =$

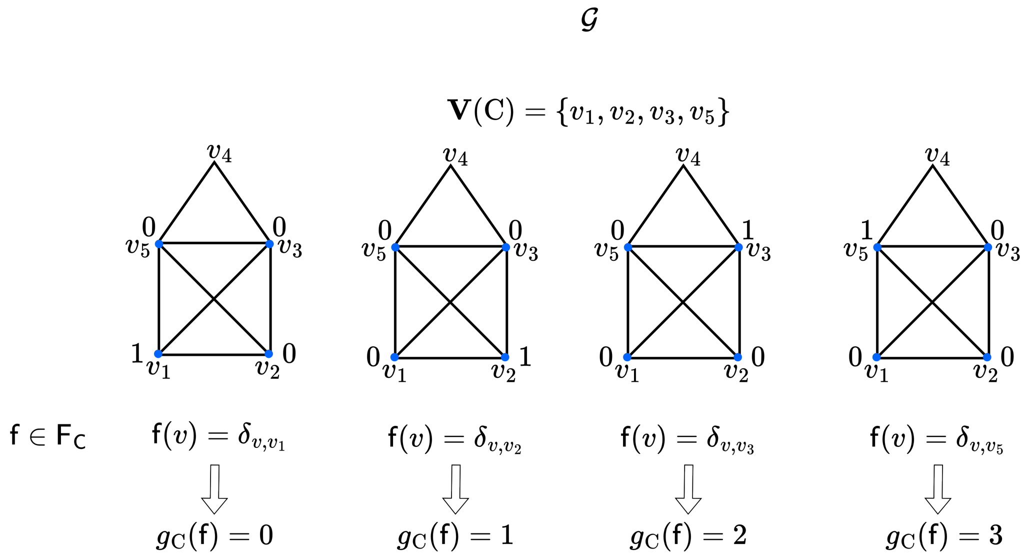

Figure 2.3: In graph $\mathcal{G}$, the maximum clique C of size $\omega(\mathcal{G}) = 4$ has vertices $\mathbf{V}(\mathrm{C}) = \{v_1, v_2, v_3, v_5\}$. The clique label $g_\mathrm{C}(\mathsf{f}) = 0, 1, 2, 3$, respectively, corresponds to binary colourings $\mathsf{f} \in \mathsf{F}_\mathrm{C}$ such that $\mathsf{f}(v) = \delta_{v,v_1}$, $\mathsf{f}(v) = \delta_{v,v_2}$, $\mathsf{f}(v) = \delta_{v,v_3}$ and $\mathsf{f}(v) = \delta_{v,v_5}$, where $v \in \mathbf{V}(\mathrm{C})$.

$\{v_{i_1}, v_{i_2}, \cdots v_{i_{\omega(G)}}\}$. A clique label corresponds to a binary colouring of the maximum clique through this function.

The lowest clique label 0 corresponds to the binary colouring that assigns binary colour 1 to the vertex with the lowest index in the maximum clique, *i.e.*, v_{i_1}. Similarly, the highest clique label $\omega(\mathcal{G}) - 1$ corresponds to the binary colouring that assigns binary colour 1 to the vertex with the highest index in the maximum clique, *i.e.* $v_{i_{\omega(\mathcal{G})}}$. For example, consider the maximum clique C of size $\omega(\mathcal{G}) = 4$ with vertices $\mathbf{V}(\mathrm{C}) = \{v_1, v_2, v_3, v_5\}$ as shown in Figure 2.3. Clique label $g_\mathrm{C}(\mathsf{f}) = 0, 1, 2, 3$, respectively, corresponds to binary colourings $\mathsf{f} \in \mathsf{F}_\mathrm{C}$ such that $\mathsf{f}(v) = \delta_{v,v_1}$, $\mathsf{f}(v) = \delta_{v,v_2}$, $\mathsf{f}(v) = \delta_{v,v_3}$ and $\mathsf{f}(v) = \delta_{v,v_5}$, where $v \in \mathbf{V}(\mathrm{C})$.

Although it is always possible to define binary colouring of maximum cliques for a graph $\mathcal{G}$ individually, a binary colouring for the entire graph may not exist. Such graphs are particularly relevant in demonstrating the impossibility of non-contextual hidden variable models for quantum theory [139]. Consider a classical hidden variable theory **TH** for describing quantum mechanics. In this theory, the state space for a d-dimensional quantum system is given by Λ. Each state $\lambda \in \Lambda$ completely determines the outcomes of all quantum measurements that could be performed on this system. In other words, for any projective measurement $\{E_i = |\psi_i\rangle\langle\psi_i| : |\psi_i\rangle \in \mathbb{C}^d\}_{i=1}^d$, where $\sum_{i=1}^d E_i = I_d$, a state λ deterministically determines its outcomes, *i.e.* $p(i|\lambda) \in \{0, 1\}$ $\forall i \in [d]$. Exactly one projector, say E_i, is assigned an outcome probability of 1. Now, if we consider two different projective measurements in $d \geq 3$, $\{E_i = |\psi_i\rangle\langle\psi_i| : |\psi_i\rangle \in \mathbb{C}^d\}_{i=1}^d$ and $\{E'_i = |\psi'_i\rangle\langle\psi'_i| : |\psi'_i\rangle \in \mathbb{C}^d\}_{i=1}^d$ that have some common projectors or effects. In this case, the hidden variable λ also assigns outcome probabilities $0, 1$ to the projectors in both measurements as before. For state λ, if the probability of the outcome for each projector is independent of the context, *i.e.*, the measurement in which it appears, then the hidden variable model is called non-contextual.

Now, consider a graph whose vertices correspond to some rank 1 orthogonal projectors (or equivalently vectors). For such a graph $\mathbf{V}(\mathcal{G}) = \{v_i \equiv E_i = |\psi_i\rangle\langle\psi_i| : |\psi_i\rangle \in \mathbb{C}^d\}_{i=1}^n$. The edges of the graph are given by orthogonality relations among the projectors, *i.e.* $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$ *iff* $\langle\psi_j|\psi_i\rangle = 0$. Assume the maximum cliques are of size $\omega(\mathcal{G}) = d$ for the graph. Thus, each maximum clique of $\mathcal{G}$, consisting of d orthogonal projectors, corresponds to a projective measurement on a d-dimensional quantum system. Every binary colouring of the graph, if it exists, corresponds to a non-contextual assignment of outcome probability to the projectors by a hidden variable λ. However, there exists some graph $\mathcal{G}$ which is not KS colourable[139], *i.e.*, it cannot have a binary colouring as in Definition 2.26. Thus, the outcome probabilities for each measurement, corresponding to a maximum clique in the graph, when performed on a quantum state, cannot be explained by any non-contextual hidden variable model. Such graphs exhibit quantum contextuality. Initially, the Kochen-Specker proof of quantum contextuality was presented using a graph with 117 vertices [139] and later on, a simpler proof was formulated while using graphs of order 18 [240]. Subsequently, this notion of non-contextuality was generalised in [140]. A comprehensive discussion on quantum contextuality with examples of KS non-colourable graphs can be found in [239].

2.5 Communication Based Information Processing Tasks

In this section, we discuss some one-way communication scenarios involving two parties, Alice and Bob, also known as prepare and measure scenarios (PM). Along with communication, the parties may sometimes have access to previously shared physical systems. Before describing the specific communication scenarios, let us first briefly discuss the communication channels available to Alice and Bob.

2.5.1 Classical communication

Suppose Alice has access to classical systems and wants to communicate with Bob using them. Such communications are modelled by a classical channel. If Alice sends a symbol $\tau \in T_A$, where T_A is a finite alphabet set, through a channel $\mathscr{T}$, then Bob would receive a symbol $\tau' \in T_B$ with probability $\mathscr{T}(\tau'|\tau)$. This classical channel $\mathscr{T} : T_A \to T_B$ is thus described by the conditional probability distribution $\{\mathscr{T}(\tau'|\tau)\}_{\tau,\tau'}$. Here, we assume that the channel is memoryless, *i.e.*, it satisfies the i.i.d. assumption over multiple uses. For simplicity, we take $T_A = T_B = T$. A channel $\mathscr{T} : T \to T$ is noiseless if $\{\mathscr{T}(\tau'|\tau) = \delta_{\tau,\tau'}\}_{\tau,\tau'}$. In the case of the noiseless channel, Alice transmits a classical system of operational dimension $d = |T|$ without any disturbance.

Assume there is a noiseless binary channel $\mathscr{T}$, *i.e.* $T = \{0, 1\}$ and $\{\mathscr{T}(\tau'|\tau) = \delta_{\tau,\tau'}\}_{\tau,\tau'}$. Suppose Alice wants to send a sequence of k i.i.d. messages $\{\mathcal{X}_1, \cdots, \mathcal{X}_k\}$ drawn from a finite set M_A, where each message $\mathsf{x} \in \mathsf{M}_A$ is chosen with probability $p(\mathsf{x})$. In other words, the source entropy is $H(\mathcal{X}) = -\sum_{\mathsf{x}\in\mathsf{M}_A} p(\mathsf{x}) \log_2 p(\mathsf{x})$. According to Shannon's source coding theorem, if the rate (average number of bits per message) is higher than $H(\mathcal{X})$ then there exist encoding and decoding schemes whose decoding error probability vanishes as $k \to \infty$. Conversely, if the rate is strictly less than $H(\mathcal{X})$, no coding scheme can achieve vanishing error [190].

Now consider the general case when Alice has access to a discrete memoryless channel $\mathscr{T} : T \to T$ described by transition probabilities $\{\mathscr{T}(\tau'|\tau)\}_{\tau,\tau'}$. The Shannon capacity of the channel $C(\mathscr{T})$ gives the maximum number of bit per use of the channel that Alice can send to Bob with negligible error in communication, while using some optimal encoding and decoding, in the limit of long message sequences. Let the input random variable $\mathcal{X}$ for Alice take values in T with distribution $\{p(\tau)\}_{\tau\in T}$, and let the output random variable for Bob $\mathcal{Y}$ be induced by $p_{\mathcal{Y}|\mathcal{X}}(\tau'|\tau) = \mathscr{T}(\tau'|\tau)$. The capacity $C(\mathscr{T}) = \max_{\{p(\tau)\}} I(\mathcal{X} : \mathcal{Y}) = \max_{\{p(\tau)\}}[H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X})]$ [190]. Here, $I(\mathcal{X} : \mathcal{Y})$ denotes the mutual information between $\mathcal{X}$ adn $\mathcal{Y}$. For our discussions, we will restrict ourselves to a single use of the channel. For a noiseless channel with input alphabet size $|T|$, the capacity is $\log_2 |T|$ bit.

2.5.2 Quantum communication

Suppose Alice has access to quantum systems to communicate with Bob. Such communication is modelled using a quantum channel. A quantum channel is described by a CPTP map $\mathbf{\Phi} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^{d'})$. Let $\{K_i | K_i : \mathbb{C}^d \to \mathbb{C}^{d'}\}_{i=1}^n$ be the Kraus operators of the map satisfying $\sum_{i=1}^n K_i^\dagger K_i = I_d$. If Alice sends a system in state $\rho \in D(\mathbb{C}^d)$ using the channel, then Bob receives the state $\mathbf{\Phi}(\rho) = \sum_{i=1}^n K_i \rho K_i^\dagger \in D(\mathbb{C}^{d'})$. Analogous to a noiseless classical channel, a noiseless quantum channel is essentially given by the identity map $\mathbf{I}_d : \mathbb{C}^d \to \mathbb{C}^d$. A quantum channel can be used to encode and transmit either quantum information, which is some unknown quantum state, or classical information to Bob.

We consider the communication of classical information using a quantum memoryless channel $\mathbf{\Phi} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^{d'})$. Say Alice wants to send a random message $\mathcal{X}$ chosen from the finite set M_A with probability $p(\mathsf{x})$. Alice encodes a message $\mathsf{x} \in \mathsf{M}_A$ using a state $\rho_{\mathsf{x}} \in \mathcal{D}(\mathbb{C}^d)$ and transmits it through $\mathbf{\Phi}$. Over n uses we assume *product-state encoding* by Alice, *i.e.*, $\rho_{\mathsf{x}_1} \otimes \cdots \otimes \rho_{\mathsf{x}_n}$ as input to $\mathbf{\Phi}^{\otimes n}$, while Bob is allowed to perform joint measurements on the n output systems for decoding. The Holevo-Schumacher-Westmoreland (HSW) theorem gives the product state classical capacity $C(\mathbf{\Phi})$ of the channel as the maximum achievable rate (bits per channel use) while using product inputs and joint decoding with vanishing decoding error probability in the asymptotic limit [190]. Formally, $C(\mathbf{\Phi}) = \max_{p(\mathsf{x}),\rho_\mathsf{x}} \{S(\mathbf{\Phi}(\sum_\mathsf{x} p(\mathsf{x})\ \rho_\mathsf{x})) - \sum_\mathsf{x} p(\mathsf{x}) S(\mathbf{\Phi}(\rho_\mathsf{x}))\}$ where $S(\rho) = -\operatorname{Tr}[\rho \log_2 \rho]$ is the von Neumann entropy. The optimisation is over the ensembles $\{p(\mathsf{x}), \rho_\mathsf{x}\}_\mathsf{x}$ of possible inputs to the channel and their probabilities. For the identity map $\mathbf{I}_d : \mathbb{C}^d \to \mathbb{C}^d$, the classical capacity is $\log_2 d$ bits per use. We will only consider a single use of such a communication channel.

An important no-go result in quantum information theory is Holevo's theorem [159] for accessible information. For a single use of a noiseless d-dimensional quantum channel and without preshared entanglement, it is impossible to communicate more than $\log_2 d$ bit per use by sending a d-dimensional system. Say Alice sends a random message $\mathcal{X}$ chosen from the finite set M_A with probability $p(\mathsf{x})$. Alice encodes message $\mathsf{x} \in \mathsf{M}_A$ using a state $\rho_\mathsf{x} \in \mathcal{D}(\mathbb{C}^d)$ and transmits it through the identity channel $\mathbf{I}_d : \mathbb{C}^d \to \mathbb{C}^d$. Bob performs a POVM $\{E_i \geq 0, \sum_i E_i = I_d\}$ on the received system and produces the outcome. Let $\mathcal{Y}$ denote the outcome random variable of Bob. Then the mutual information $I(\mathcal{X} : \mathcal{Y})$ is always less than or equal to $S(\sum_\mathsf{x} p(\mathsf{x})\rho_\mathsf{x}) - \sum_\mathsf{x} p(\mathsf{x}) S(\rho_\mathsf{x})$. Although this limits the information-carrying capacity of quantum channels, several works have established the advantage of quantum communication in PM scenarios.

2.5.3 Prepare and Measure Scenarios

We now discuss prepare-and-measure (PM) scenario, which will be useful to define certain communication-based information processing tasks presented in Chapters 3 and 5. A prepare-and-measure (PM) scenario involves two separate parties, Alice and Bob. In general, the parties may also have access to a two-way channel, which can be used multiple times. For our considerations, Alice has a one-way communication channel that she can use to send information to Bob. The one-way channel may be quantum or classical, depending on the resources allowed in this scenario. In addition to the communication channel, the parties may share correlations for assistance. Such shared resources can be classical (shared randomness), quantum or even post-quantum (general no-signalling correlations). In general, communication and shared resources may be considered costly or constrained in amount, depending on the operational setting.

In the PM scenario, Alice receives some random input $\mathsf{m_a} \in \mathsf{M}_A$. Bob receives a random input $\mathsf{m_b} \in \mathsf{M}_B$ and produces an output $\mathsf{n} \in \mathsf{N}$ (see Figure 2.4). For the tasks, we will define $\mathsf{M}_A, \mathsf{M}_B$ and N will be finite sets. In this scenario, various tasks can be defined based on the conditional probability distribution $\{p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b})\}_{\mathsf{m_a},\mathsf{m_b},\mathsf{n}}$ which is obtained. For comparing classical and quantum communication resources, we will use the notion of operational dimension as in

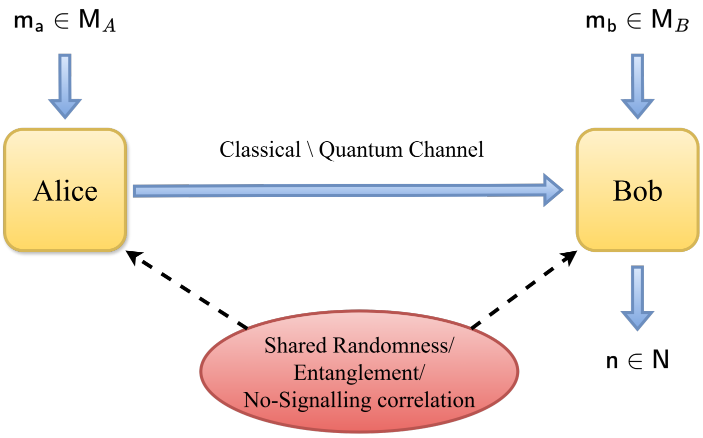

Figure 2.4: A schematic representation of a prepare-and-measure scenario where Alice and Bob receive input $m_a \in M_A$ and $m_b \in M_B$, respectively. Bob produces output $n \in N$. Based on the operational setting, Alice and Bob may have access to a one-way classical and/or quantum channel. Also, they may have (whenever allowed) access to different types of shared resources.

Definition 2.7.

In Chapter 3, we first consider a PM scenario where Alice and Bob have no shared correlations and may use either classical or quantum communication. Here, communication is regarded as costly. For the information processing task which will be defined, we compare the minimum amount of classical and quantum communication needed to accomplish it. We then consider the case where Alice's communication to Bob is limited. In this case, the shared resources are allowed but also considered costly. In this operational setting, we investigate the amount of shared resources required to complete the task with the limited communication.

In Chapter 5, we consider a PM scenario where Alice has access to a limited amount of classical communication. In this case, Alice and Bob may share classical randomness, quantum entanglement, or more general no-signalling correlations. We explore the advantage of non-local correlations over shared randomness in assisting the classical channel for different communication tasks.

Information processing tasks in PM scenarios with classical communication and classical shared resources have been widely studied such as in classical communication complexity [157]. In the general scenario, beyond our scope, the parties may have multiple uses of a two-way noiseless classical channel. These tasks are defined by a relation $\mathcal{R}$ which can be identified as a subset of $M_A \times M_B \times N$, *i.e.* $\mathcal{R} \subset M_A \times M_B \times N$. The objective is for Bob to output $n \in N$ such that given Alice's and Bob's random input, $m_a \in M_A$ and $m_b \in M_B$ respectively, $(m_a, m_b, n) \in \mathcal{R}$.

The classical communication complexity of relation $\mathcal{R}$ is the minimum amount of classical communication required to accomplish this for all valid inputs. Corresponding to this minimum communication, there is a protocol ensuring Bob's output is related to the input of both parties. If the computation has to be without any error, *i.e.*, Bob always outputs $\mathsf{n} \in \mathsf{N}$ with $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$, the complexity is called zero-error communication complexity. Sometimes a small amount of error ϵ is allowed in the computation, and the complexity is referred to as bounded error communication complexity if $\epsilon < \frac{1}{3}$. Note that input pairs of Alice and Bob can be sometimes restricted to a subset $\mathsf{D} \subseteq \mathsf{M}_A \times \mathsf{M}_B$. In such problems, there is a promise on the input for Alice and Bob, and $(\mathsf{m_a}', \mathsf{m_b}') \notin \mathsf{D}$ is an illegal input pair. The illegal inputs may not even be related to any $\mathsf{n} \in \mathsf{N}$ under the relation $\mathcal{R}$.

Suppose the parties also use resources in addition to classical communication. Private coins for Alice and Bob are locally accessible random variables, denoted $a \in \mathsf{A}$ and $b \in \mathsf{B}$, sampled from distributions $\{p(a)\}_{a\in\mathsf{A}}$ and $\{p(b)\}_{b\in\mathsf{B}}$, respectively. A public coin, in contrast, is a shared random variable, so that $a = b$ and $\mathsf{A} = \mathsf{B}$, drawn from a joint distribution $\{p(a)\}_a$. A classical public coin can be modelled as a shared hidden variable $\lambda \in \Lambda$, with each party applying fixed measurements, say x for Alice and y for Bob. We denote the classical communication complexity of a relation without shared resources, with private coins access, and with a public coin, respectively, by $\mathbf{D}(\mathcal{R})$, $\mathbf{R}_{priv}(\mathcal{R})$ and $\mathbf{R}_{pub}(\mathcal{R})$.

Consider the cases when the parties have multiple allowed uses of a two-way noiseless bit channel $\mathscr{T} : \{0,1\} \to \{0,1\}$, with no access to randomness. Any protocol that enables Bob to compute an output n related by $\mathcal{R}$ to the inputs $\mathsf{m_a}$ and $\mathsf{m_b}$ can be represented as a binary tree [157]. Each internal vertex of the binary tree is labelled with either Alice or Bob, indicating which party sends the next bit at that stage of the protocol. For each vertex v labelled by Alice or Bob, there is a function $f_v : \mathsf{M}_A \to \{0,1\}$ or $f_v : \mathsf{M}_B \to \{0,1\}$, respectively. The value at each vertex ($f_v(\mathsf{m_a})$ for Alice or $f_v(\mathsf{m_b})$ for Bob) determines whether the party sends bit 0 or 1. The protocol subsequently follows the left child of the vertex if the communicated bit is 0, else it follows the right child. Each leaf of the binary tree is labelled with a value in the set N, which represents Bob's output. The communication cost of the protocol is the height h of the binary tree. The communication complexity is the minimum possible height over all such trees. Consider a matrix $A^{\mathcal{R}}_{\mathsf{M}_A\times\mathsf{M}_B}$ where rows correspond to the inputs of Alice and columns to the inputs of Bob. Its entries are the list of valid outcomes of Bob corresponding to the input pair. A protocol with communication cost h bit partitions $A^{\mathcal{R}}_{\mathsf{M}_A\times\mathsf{M}_B}$ into almost 2^h monochromatic rectangles (see [157] for a detailed introduction).

When the parties have access to private or public coins, a similar description of a protocol can be given. The local function at each vertex now also depends on the randomly generated classical variable. The function defined for each vertex v associated with Alice or Bob is given by $f_v : \mathsf{M}_A \times \mathsf{A} \to \{0,1\}$ or $f_v : \mathsf{M}_B \times \mathsf{B} \to \{0,1\}$, respectively. Here, $a \in \mathsf{A}$ and $b \in \mathsf{B}$ are the random variables for Alice and Bob. In the case of a public coin, $a = b$ with $\mathsf{A} = \mathsf{B}$. As before, the height of the binary tree gives the cost of a protocol. The communication complexity with private or public coins is defined as the minimum tree height over all valid protocols using the respective resources. A detailed overview of classical communication complexity and its application to circuit depth can be found in [157].

The above description simplifies in the case of one-way classical communication complexity. Without any accessible randomness, $\mathbf{D}(\mathcal{R})$ for a relation $\mathcal{R}$ is the minimum over all valid deterministic protocols. A valid protocol enables Bob to compute $\mathsf{n} \in \mathsf{N}$ such that $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$. A valid deterministic protocol using $\log_2 |T|$ bit communication (or equivalently a noiseless channel $\mathscr{T} : T \to T$) is described by an encoding scheme $\mathbb{E} : \mathsf{M}_A \to T$ for Alice and a decoding scheme $\mathbb{D} : T \times \mathsf{M}_B \to \mathsf{N}$ for Bob. Alice uses $\mathbb{E}$ to decide the symbol she sends to Bob based on her input. Similarly, Bob uses $\mathbb{D}$ to decide his output based on the received symbol and his input. With private coins, $\mathbf{R}_{priv}(\mathcal{R})$ is defined by minimising over all valid protocols that use private randomness. Here, a protocol with $\log_2 |T|$ bits corresponds to a local distribution over deterministic encoding and decoding schemes. Alice chooses an encoding scheme $\mathbb{E}_a : \mathsf{M}_A \to T$, depending on her random variable $a \in \mathsf{A}$. Similarly, Bob chooses a decoding scheme $\mathbb{D}_b : T \times \mathsf{M}_B \to \mathsf{N}$, depending on his random variable $b \in \mathsf{B}$. When the parties have access to public coins (shared randomness), the random variables used by Alice and Bob are perfectly correlated, *i.e.*, $a = b$. A valid protocol using $\log_2 |T|$ bits is then a correlated choice of deterministic encoding $\mathbb{E}_a$ by Alice and decoding $\mathbb{D}_a$ by Bob, based on the shared random variable $a \in \mathsf{A}$. The value $\mathbf{R}_{pub}(\mathcal{R})$ is obtained by minimising $\log_2 |T|$ over all such protocols.

The communication complexity of relations has also been studied using quantum channels [241, 242]. From Holevo's theorem, a d-dimensional quantum system carries no more classical information than a d-dimensional classical system. Nonetheless, several results show non-trivial advantages of quantum communication in this setting [16, 17]. In the next chapter, we present an example of unbounded separation between quantum and classical communication in a communication complexity like problem without shared randomness.

Chapter 3

Unbounded Quantum Advantage in Communication

Chapter Note

This chapter is partially based on work which was done in collaboration with Dr. Nitica Sakharwade, Dr. Some Sankar Bhattacharya, Dr. Ravishankar Ramanathan, and Prof. Dr. hab Paweł Horodecki [184].

In this chapter, we show an arbitrary separation between quantum and classical communication for tasks defined in a one-way prepare-and-measure scenario without pre-shared resources. The quantum communication required for the task is constant, while the classical communication grows with the order of the underlying graph used to define the task. We also show an unbounded separation between quantum entanglement and shared randomness, necessary as assistance to bounded classical communication for a task defined using a specific family of graphs.

These results fit within the broader context of information theory. Shannon's communication model uses classical systems as a carrier of information [243]. Substituting them with quantum systems for communication opens new possibilities. This motivated the development of quantum Shannon theory [244] and inspired studies on the advantages of encoding classical information in quantum systems. The use of quantum communication for distributed computation was first proposed in [241, 242].

Suppose the parties have access to a one-way noiseless quantum channel in the PM scenario as discussed in Chapter 2 (Section 2.5.3). The *distributed computation task* is defined using a relation $\mathcal{R} \subset \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$. Given inputs $\mathsf{m_a} \in \mathsf{M}_A$ for Alice and $\mathsf{m_b} \in \mathsf{M}_B$ for Bob, the objective is for Bob to output $\mathsf{n} \in \mathsf{N}$ such that $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$, using minimum communication.

In a quantum protocol, Alice prepares a state $|\mathsf{m_a}\rangle$ based on her input $\mathsf{m_a}$. She applies a joint operation on this state and an ancilla initialised in $|0\rangle$, and sends the resulting system ρ' to Bob. Upon receiving the quantum system ρ', Bob performs a joint operation on this system, an ancilla, and his prepared state $|\mathsf{m_b}\rangle$, which depends on his input $\mathsf{m_b} \in \mathsf{M}_B$. He uses the outcome of the operation to determine his output n. The communication cost of the protocol is the logarithm of the dimension of the transmitted quantum system, evaluated in the worst case over inputs. The quantum communication complexity is the minimum cost over all protocols that ensure such computation. A general setting for such problems allows multiple uses of two-way quantum communication.

In [16], the authors studied both a total function and a partial function. The total function $DISJ$ they considered is defined as $DISJ(\mathsf{m_a}, \mathsf{m_b}) = \vee_{i\in[2^n]}(\mathsf{m_{a(\mathit{i})}} \wedge \mathsf{m_{b(\mathit{i})}})$, where $\mathsf{m_{a(\mathit{i})}}$ and $\mathsf{m_{b(\mathit{i})}}$ denote the i^{th} bit of Alice and Bob's input $\mathsf{m_a}, \mathsf{m_b} \in \{0,1\}^{2^n}$. The classical two-way bounded-error communication complexity of this function is lower bounded by $\Omega(2^n)$ bit [245]. Using a quantum algorithm similar to Grover's search [13], the distributed computation can be carried out with bounded error using less than $O(n2^{n/2})$ qubits. This yields nearly a quadratic quantum advantage. For the partial function EQ', the input pairs for the parties are restricted to the domain $\mathsf{D} = \{(\mathsf{m_a}, \mathsf{m_b}) : \Delta(\mathsf{m_a}, \mathsf{m_b}) \in \{0, \frac{n}{2}\}\} \subset \mathsf{M}_A \times \mathsf{M}_B = \{0,1\}^{2^n} \times \{0,1\}^{2^n}$. Here, $\Delta(\mathsf{m_a}, \mathsf{m_b})$ denotes the Hamming distance. $EQ'(\mathsf{m_a}, \mathsf{m_b}) = 1$ if $\Delta(\mathsf{m_a}, \mathsf{m_b}) = 0$ and otherwise $EQ'(\mathsf{m_a}, \mathsf{m_b}) = 0$. The zero-error two-way classical communication complexity of this problem was proven to be lower bounded by $\Omega(2^n)$ bit, while the zero-error quantum communication complexity was shown to be upper bounded by $O(n)$ qubit, leading to an exponential quantum advantage.

An exponential separation was shown in [17] between bounded-error quantum and bounded-error classical communication complexity in a two-way communication scenario. Alice's input is $\mathsf{m_a} = (\vec{v}, \mathbf{V}_0, \mathbf{V}_1)$, where vector $\vec{v} \in \mathbb{R}^n$ is a unit vector and $\mathbf{V}_0, \mathbf{V}_1$ are two orthogonal subspaces of $\mathbb{R}^n$, each of dimension $\frac{n}{2}$. Bob's input is an orthogonal matrix $\mathsf{m_b} = U : \mathbb{R}^n \to \mathbb{R}^n$. For $i = 0, 1$, Bob's output should be $\mathsf{n} = i$ if the distance between $U\vec{v}$ and $\mathbf{V}_i$ is less than or equal to a fixed constant $\alpha \in (0, \frac{1}{\sqrt{2}})$; otherwise either outcome is acceptable. It was shown that the bounded error quantum communication complexity is $\Theta(\log n)$ qubit. Particularly, bounded error classical communication complexity was proven to be at least $\Omega(\sqrt{n})$ bit, with an upper bound of $O(n^{\frac{3}{4}})$ bit. Although the input sets here are infinite, the separation persists under finite-bit approximations as well. In [47], an exponential gap was established between one-way quantum communication and bounded-error classical communication using the Hidden Matching problem. This problem is defined by a relation $\mathcal{R}$. For an even number n, Alice's input is $\mathsf{m_a} \in \mathsf{M}_A = \{0,1\}^n$. Bob's input $\mathsf{m_b} \in \mathsf{M}_B$ is a perfect matching of n vertices, *i.e.*, a partition $\{(i_1, j_1), (i_2, j_2), \cdots, (i_{n/2}, j_{n/2})\}$ of $[n]$ into $\frac{n}{2}$ disjoint pairs. Bob's output is

$\mathsf{n} = (\mathsf{n}_{(1)}, \mathsf{n}_{(2)}, \mathsf{n}_{(3)}) \in \mathsf{N} = [n] \times [n] \times \{0,1\}$. The triple $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$ if $(\mathsf{n}_{(1)}, \mathsf{n}_{(2)}) \in \mathsf{m_b}$ and $\mathsf{n}_{(3)} = \mathsf{m}_{\mathsf{a}(\mathsf{n}_{(1)})} \oplus_2 \mathsf{m}_{\mathsf{a}(\mathsf{n}_{(2)})}$. The one-way quantum communication complexity of the relation is $O(\log n)$ qubit, while the one-way bounded error classical communication complexity is $\Omega(\sqrt{n})$ bit. Subsequently, in [160], an exponential separation of the same form ($O(\log n)$ vs $\Omega(\sqrt{n})$) was shown between one-way quantum and classical bounded error communication complexity for a paticular partial Boolean function, which is a variant of the Hidden matching problem. In [18], an exponential separation was shown between zero-error one-way quantum communication complexity ($O(\log n)$ qubits) and bounded-error two-way classical communication complexity ($\Omega(\frac{n^{\frac{1}{8}}}{\sqrt{\log n}})$ bit). Various other works have demonstrated such exponential separations between quantum and classical communication complexity by considering different relations [19] or setups such as simultaneous message passing [20].

Besides distributed computation, communication-based information processing tasks can be defined as the simulation of some specific correlations. In [21–23], the classical communication cost of simulating the correlations from measurements performed on a shared Bell state, for example $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|01\rangle_{AB} - |10\rangle_{AB}\Big)$, was studied. In the PM scenario, Alice's input $\mathsf{m_a} \in \mathsf{M}_A$ and Bob's input $\mathsf{m_b} \in \mathsf{M}_B$ correspond to von Neumann measurements on qubits, and their outputs must reproduce the statistics of measuring the Bell state. In [22], it was demonstrated that the correlation can be simulated with an *average* of 1.48 bits of classical communication. However, the protocol may require arbitrarily large communication in some rounds. In [21], the authors gave a protocol that simulates the correlation with exactly 8 bits of bounded classical communication. However, their protocol requires an infinite amount of shared randomness [24]. In [24], it was further shown that $O(2^n)$ bit of average one-way communication, without shared randomness, suffice to simulate the statistics from n shared Bell pairs and arbitrary local POVMs. The same task can be performed with $O(n)$ qubit of quantum communication. In [25], it was shown that 1 bit of communication is enough to simulate correlations from a shared Bell state and projective measurements. More recently, 2 bit of bounded classical communication were shown to be necessary and sufficient to simulate correlations from a shared Bell state and arbitrary POVMs [26]. Some works have also studied the classical communication cost of simulating correlations that arise when quantum systems are communicated and then measured in a PM scenario. In [27], it was shown that 6.38 bit of *average communication* are needed for a teleportation task where Alice's input $\mathsf{m_a} = |\psi\rangle$ is the description of a qubit state, Bob's input $\mathsf{m_b}$ is a qubit POVM, and Bob's output must reproduce the statistics of the measurement on $|\psi\rangle$. If Bob's input is restricted to von Neumann measurements, the cost reduces to 2.19 bit on *average*. Later, improved protocols showed that only 2 bits are sufficient for such simulations when measurements are arbitrary POVM on a qubit [26].

Such works raise the question of how large the separation between quantum and classical communication can be in information processing tasks. Some examples of infinite separations were shown in [28, 29], but with *internal information cost* as the measure rather than communication cost. Interestingly, in [24], an instance of arbitrary separation was shown in a one-way PM scenario without shared randomness. They considered the communication cost of the NOT-EQUAL problem and showed that 1 qubit communication is sufficient, while $\log n$ bit of classical communication is necessary. In this task, the input set for Alice and Bob is $\mathsf{M}_A = \mathsf{M}_\mathsf{B} = \{0,1\}^n$. So, the input size of $2n$ bit grows exponentially relative to the classical communication cost of $\log n$ bit. This motivates the search for other tasks where the separation between classical and quantum communication grows arbitrarily large. A natural question is whether, in general, unbounded separation can be shown while the required classical communication exceeds $\log_2 n$ when the size of input scales as $O(n)$. Another question is regarding the existence of tasks that require a d-dimensional quantum system (beyond a qubit), while the corresponding classical communication cost grows arbitrarily large. In other words, if it is possible to demonstrate such unbounded separation for any d-dimensional quantum system, beyond a qubit. In this chapter, we explore these questions.

We consider tasks in the PM scenario where Alice has a one-way communication channel, and no shared correlations (shared randomness or entanglement) are available. The channel may be classical or quantum, and we compare them using the notion of operational dimension (see Definition 2.7). The tasks are defined using relations $\mathcal{R}(\mathcal{G})$ specified by the distributed *Clique Labelling Problem* (CLP) for an orthogonality graph $\mathcal{G}$. We first consider zero-error *distributed computation* of relation $\mathcal{R}(\mathcal{G})$. We show that there is no quantum advantage in this task for certain graphs. Both classical and quantum zero-error communication complexity of relation $\mathcal{R}(\mathcal{G})$ equal $\log_2 \omega(\mathcal{G})$ bit/qubit. We then consider a variant of the task for $\mathcal{R}(\mathcal{G})$, which we refer to as *relation reconstruction*. For this task, classical communication cost without shared and private randomness grows as $\log_2 |\mathbf{V}(\mathcal{G})|$ bit with the order of the graph $|\mathbf{V}(\mathcal{G})|$. The classical cost without shared randomness is at least $\log_2 K(\mathcal{G})$ bit, where $K(\mathcal{G})$ is the disjointness number of $\mathcal{G}$ (see Definition 3.4). This number satisfies $K(\mathcal{G}) \geq \max\{\omega(\mathcal{G}),\ \log_2 |\mathbf{V}(\mathcal{G})| + \frac{1}{2}\log_2\log_2 |\mathbf{V}(\mathcal{G})| + O(1)\}$. In contrast, $\log_2 \omega(\mathcal{G})$ qubits are sufficient whenever $\mathcal{G}$ has a faithful orthogonal representation in dimension $\omega(\mathcal{G})$. We identify families of graphs where the order $|\mathbf{V}(\mathcal{G})|$ grows unboundedly while the faithful orthogonal range remains fixed. For such families, each graph is characterised by two parameters: the clique number $\beta(\geq 2) \in \mathbb{N}$, which is a constant, and the number of maximum cliques $\alpha(\geq 2) \in \mathbb{N}$. Considering $\mathcal{R}(\mathcal{G})$ for this family of graphs, we show an unbounded quantum advantage in relation reconstruction. Specifically, quantum communication remains constant at $\log_2 \beta$ qubit, while the classical communication required is at least $\max\{\log_2 \beta,\ \log_2(\log_2 \alpha + \frac{1}{2}\log_2\log_2 \alpha + c)\}$ bit, where c is a constant. The input size of the task is $2\log_2 \alpha + c'$ bits, where c' is a constant.

However, the unbounded communication advantage disappears when shared resources are allowed between parties. We then examine the amount of shared randomness required to assist a bounded classical channel in relation reconstruction. For certain families of graphs, this necessary shared randomness grows as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit, where α is the number of maximum cliques. For a particular family of graphs, we show that one e-bit of entanglement (a two-qubit maximally entangled state), together with a one-bit channel, is sufficient for relation reconstruction. In contrast, the shared randomness assistance required with one-bit communication grows arbitrarily large with the number of maximum cliques.

We have organised the chapter in the following manner. In Section 3.1, we introduce the tasks and the Clique Labelling Problem. Section 3.2 shows that there is no quantum advantage in the distributed computation of a relation based on the CLP. We then present an instance of unbounded quantum advantage in relation reconstruction. In Section 3.3, we calculate the amount of correlation assistance required for bounded classical communication. We show an instance of arbitrary separation between entanglement assistance and shared randomness assistance for a task with one-bit communication. Finally, in Section 3.4, we discuss some applications.

3.1 Distributed Computation and Relation Reconstruction

We now present two tasks *distributed computation* and *relation reconstruction*, which we will discuss in this chapter. Consider a PM scenario with two separated parties, Alice and Bob, where communication is allowed from Alice to Bob. The communication is treated as costly and may be classical or quantum. The parties do not have access to any shared resources or public coins. They may have local randomness or private coins. We focus on two tasks defined by a relation $\mathcal{R} \subset \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$ known to both parties. In both the distributed computation and the relation reconstruction task, Alice and Bob receive input $\mathsf{m_a} \in \mathsf{M}_A$ and $\mathsf{m_b} \in \mathsf{M}_B$, respectively, which are randomly chosen following a uniform probability distribution. Bob produces an output $\mathsf{n} \in \mathsf{N}$. The resulting conditional probability distribution during the tasks is denoted by $\{p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b})\}_{\mathsf{m_a},\mathsf{m_b},\mathsf{n}}$. A schematic representation of the scenario is given in Figure 3.1. In our case, for each pair of inputs $\mathsf{m_a} \in \mathsf{M}_A, \mathsf{m_b} \in \mathsf{M}_B$, there is at least one valid output n such that $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$. Unlike functions, in some cases, more than one n may satisfy the relation.

The first task we consider is the distributed computation of relation $\mathcal{R}$, where Bob outputs n such that $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$. We specifically focus on the zero-error setting, where Bob never outputs n such that $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \notin \mathcal{R}$[1]. Thus,

[1]Also in Section III of [184]

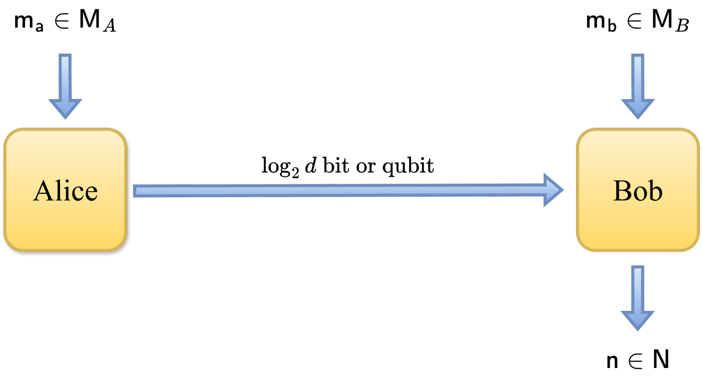

Figure 3.1: A schematic representation of a prepare-and-measure scenario where Alice has access to a one-way classical and/or quantum channel. Here, the parties do not have access to shared resources. Alice and Bob receive input $m_a \in M_A$ and $m_b \in M_B$, respectively. Bob produces output $n \in N$.

$$p(n|m_a, m_b) = 0 \quad \forall \quad (m_a, m_b, n) \notin \mathcal{R} \tag{3.1}$$

Alice and Bob can use communication and follow a protocol to achieve this. The cost of a protocol is the maximum communication required to complete the zero-error distributed computation for worst-case inputs $m_a \in M_A$ and $m_b \in M_B$.

Definition 3.1. The one-way zero-error communication complexity of a relation or CCR is the minimum communication from Alice to Bob required such that for all $m_a \in M_A$ and $m_b \in M_B$, Bob's output n always satisfies $(m_a, m_b, n) \in \mathcal{R}$.

The optimal protocol is the one with the minimum communication cost. We will refer to the communication complexity as classical CCR or quantum CCR, depending on whether the communication is classical or quantum. Trivially, classical/quantum CCR is at most $\log_2 |M_A|$ bit/qubit, since Alice could send her entire input to Bob. However, an efficient protocol may require less communication.

A function is a special case of a relation, where each input pair has a unique correct answer for Bob to guess. For general relations, some input pairs may have multiple correct answers. Consequently, we now define a new task for relations. This would be the second task we consider. It is a stronger version of zero-error distributed computation, called *relation reconstruction*[2]. In this task, Bob's outputs $n \in N$ across different rounds must collectively cover all valid answers

[2]Also in Section III of [184]

for every input pair $\mathsf{m_a} \in \mathsf{M}_A$ and $\mathsf{m_b} \in \mathsf{M}_B$. In other words, for relation reconstruction, the probability distribution $\{p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b})\}_{\mathsf{m_a},\mathsf{m_b},\mathsf{n}}$ must satisfy both Equation (3.1) and the condition

$$p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b}) > 0 \ \ \forall \ \ (\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R} \tag{3.2}$$

Clearly, relation reconstruction reduces to zero-error distributed computation when the relation is a function. We call the minimum communication cost of relation reconstruction the strong communication complexity of the relation.

Definition 3.2. For $\mathcal{R}$, the one-way zero-error strong communication complexity of relation[3] or S-CCR is the minimum communication required from Alice to Bob such that: (i) for all $\mathsf{m_a} \in \mathsf{M}_A$ and $\mathsf{m_b} \in \mathsf{M}_B$, Bob's output $\mathsf{n} \in \mathsf{N}$ always satisfies $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$, and (ii) for each input pair $(\mathsf{m_a}, \mathsf{m_b})$, the outputs across different rounds statistically span $\{\mathsf{n} \in \mathsf{N} : (\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}\}$ in the sense of Equation (3.2).

In each round, the input-output tuple $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n})$ is recorded. Over many rounds, these tuples define the observed statistics $\{p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b})\}_{\mathsf{m_a},\mathsf{m_b},\mathsf{n}}$. If Alice and Bob succeed, their input-output statistics allow a third party, called the *Reconstructor*, to infer the relation $\mathcal{R} \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$. The Reconstructor, who is previously unaware of the relation $\mathcal{R}$, learns about the input-output sets and the related tuples $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$ from the recorded list, which is shared after a sufficiently large number of rounds. Successful reconstruction of relation $\mathcal{R}$ requires $p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b}) > 0$ *iff* $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}$. We define the following payoff function (See Section III of [184]) for relation reconstruction:

$$\mathcal{S}_{\mathcal{R}} = \min_{(\mathsf{m_a},\mathsf{m_b},\mathsf{n}) \in \mathcal{R}} p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b}) \tag{3.3}$$

The payoff corresponds to the *least likely* valid outcome produced by Bob. Hence, for every input pair $(\mathsf{m_a}, \mathsf{m_b})$, each valid output n occurs with probability at least $\mathcal{S}_{\mathcal{R}}$. *A positive payoff is therefore necessary for the successful reconstruction of relation.* Optimising over all protocols gives the maximum achievable value $\mathcal{S}^*_{\mathcal{R}}$, called the algebraic maximum of payoff. This maximum is reached by a trivial protocol where Alice sends her entire input to Bob, and Bob samples all valid outputs uniformly. We will mostly be interested in the protocols with the least communication cost. Such protocols may not achieve $\mathcal{S}^*_{\mathcal{R}}$. The maximum payoff attainable with $\log_2 d$ bit of classical communication, denoted $\mathcal{S}^{Cl,d}_{\mathcal{R}}$, may differ from the maximum payoff $\mathcal{S}^{Q,d}_{\mathcal{R}}$ attainable with $\log_2 d$ qubit of quantum communication.

The payoff $\mathcal{S}_{\mathcal{R}}$ defined in Equation (3.3) is connected to the success probability $p(\mathsf{Success}^{(k)})$ of relation reconstruction by the Reconstructor after $k \in \mathbb{N}$ rounds are completed. Assume i.i.d. inputs $\mathsf{m_a} \in \mathsf{M}_A$ and $\mathsf{m_b} \in \mathsf{M}_B$ are chosen uniformly in each round. The parties follow

[3] Also in Section III of [184]

a protocol according to which for input-pair $\mathsf{m_a} \in \mathsf{M}_A$ and $\mathsf{m_b} \in \mathsf{M}_B$, Bob outputs $\mathsf{n} \in \mathsf{N}$ in each round with the conditional distribution $\{p(\mathsf{n}|\mathsf{m_a}, \mathsf{m_b})\}_{\mathsf{m_a},\mathsf{m_b},\mathsf{n}}$. For convenience, let us write the relation as $\mathcal{R} = \{r_1, r_2, \ldots, r_J\}$ as an ordered set of tuples in $\mathcal{R} \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$, where $r_i = (\mathsf{m_a}^{(i)}, \mathsf{m_b}^{(i)}, \mathsf{n}^{(i)})$. For successful reconstruction after k rounds, every $r_i \in \mathcal{R}$ must occur at least once. The probability that r_i occurs in a round, denoted by p_i is

$$p_i = p(\mathsf{m_a}^{(i)}, \mathsf{m_b}^{(i)}, \mathsf{n}^{(i)}) = p(\mathsf{n}^{(i)}|\mathsf{m_a}^{(i)}, \mathsf{m_b}^{(i)})\ p(\mathsf{m_a}^{(i)})\ p(\mathsf{m_b}^{(i)}) \tag{3.4}$$

$$= \frac{p(\mathsf{n}^{(i)}|\mathsf{m_a}^{(i)}, \mathsf{m_b}^{(i)})}{|\mathsf{M}_A||\mathsf{M}_B|} \tag{3.5}$$

The probability that r_i never occurs in k rounds corresponds to the probability of failure of reconstruction for this tuple after k rounds. It is given by

$$p(\mathsf{Fail}_i^{(k)}) = (1 - p_i)^k \tag{3.6}$$

Note $p_{min} = \min_{i' \in J} \{p_{i'}\}$ equals $\frac{\mathcal{S}_\mathcal{R}}{|\mathsf{M}_A||\mathsf{M}_B|}$. Since the tuple $r_i \in \mathcal{R}$ corresponding to p_{min} must occur at least once in the k rounds for successful reconstruction, we get

$$p(\mathsf{Success}^{(k)}) \le 1 - (1 - p_{min})^k \le 1 - e^{-kp_{min}} = 1 - e^{-k\frac{\mathcal{S}_\mathcal{R}}{|\mathsf{M}_A||\mathsf{M}_B|}} \tag{3.7}$$

Here we used the fact that $(1 - p_{min})^k \le e^{-kp_{min}}$ as $0 \le p_{min} \le 1$ and $k \in \mathbb{N}$. Thus, $\mathcal{S}_\mathcal{R} > 0$ or equivalently $p_{min} > 0$ is necessary for successful reconstruction. Moreover,

$$p(\mathsf{Success}^{(k)}) = 1 - p(\bigcup_{i=1}^{J} \mathsf{Fail}_i^{(k)}) \ge 1 - \sum_{i=1}^{J} (1 - p_i)^k \ge 1 - J(1 - p_{min})^k \ge 1 - Je^{-kp_{min}} \tag{3.8}$$

$$p(\mathsf{Success}^{(k)}) \ge 1 - Je^{-k\frac{\mathcal{S}_\mathcal{R}}{|\mathsf{M}_A||\mathsf{M}_B|}} \tag{3.9}$$

Combining Equations (3.7) and (3.9),

$$1 - Je^{-k\frac{\mathcal{S}_\mathcal{R}}{|\mathsf{M}_A||\mathsf{M}_B|}} \le p(\mathsf{Success}^{(k)}) \le 1 - e^{-k\frac{\mathcal{S}_\mathcal{R}}{|\mathsf{M}_A||\mathsf{M}_B|}} \tag{3.10}$$

From this equation, if $\mathcal{S}_\mathcal{R} > 0$ then $p(\mathsf{Success}^{(k)}) \to 1$ as $k \to \infty$. Thus, the payoff $\mathcal{S}_\mathcal{R}$ is a *faithful* quantifier of relation reconstruction.

3.1.1 Clique Labelling Problem

We now introduce the relation that we will use for distributed computation and relation reconstruction. The relation $\mathcal{R}(\mathcal{G})$ is defined by the distributed *clique labelling problem* (CLP).

The problem is specified using a simple graph $\mathcal{G}$ (see Definition 2.15) with vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$. Each vertex of $\mathcal{G}$ belongs to at least one maximum clique in the graph. If the graph has α maximum cliques and clique number $\omega(\mathcal{G})$, we denote the set of maximum cliques as $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$. The set of clique labels for each maximum clique in $\mathcal{G}$ is $\mathcal{O} = \{0, 1, \cdots, \omega(\mathcal{G})-1\}$. Each label corresponds to a binary colouring of the clique C_i according to Definition 2.27. We will refer to the graph $\mathcal{G}$, along with a faithful orthogonal representation (FOR) in the minimum dimension, as an orthogonality graph. We will consider the faithful orthogonal representations $\phi : V(\mathcal{G}) \to \mathbb{C}^d$ for a graph. Clearly, if a graph $\mathcal{G}$ has a faithful orthogonal representation ϕ in $\mathbb{R}^d$, the same function ϕ is also a faithful orthogonal representation in $\mathbb{C}^d$.

The distributed CLP is defined in a PM scenario where Alice and Bob know the orthogonality graph $\mathcal{G}$. The input and output sets for the problem are: $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}$, $\mathsf{M}_B = \mathcal{C}$ and $\mathsf{N} = \mathcal{O}$. The inputs are chosen randomly following a uniform distribution. Alice receives $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) \in \mathsf{M}_A$, *i.e.*, a maximum clique $\mathsf{m_{a(1)}} \in \mathcal{C}$ and a clique label $\mathsf{m_{a(2)}} \in \mathcal{O}$ associated with the maximum clique, both chosen randomly. Bob receives a maximum clique $\mathsf{m_b} \in \mathsf{M}_B = \mathcal{C}$. Bob's output $\mathsf{n} \in \mathsf{N} = \mathcal{O}$ is a clique label for his input maximum clique. The output must satisfy the constraints of *consistent labelling of pairwise cliques*[4] described below. Each clique label is mapped to a binary colouring of the maximum clique through Definition 2.27, and this correspondence will be useful to state the constraints in terms of binary colouring.

(i) If Alice's and Bob's input maximum cliques share some vertices, then Bob's binary colouring (either 0 or 1) of each shared vertex must be identical to Alice's colouring (from her input clique label) of these vertices.

(ii) If their input maximum cliques are distinct and a vertex from Alice's input maximum clique is adjacent to some vertex from Bob's input maximum clique, then both the vertices sharing the edge must not be simultaneously assigned binary colour 1 (by the input-output clique labels of Alice and Bob).

(iii) In the remaining cases, Bob can output any valid clique label for his input maximum clique, independent of Alice's input clique label.

If Alice and Bob receive the same maximum clique, *i.e.* $\mathsf{m_{a(1)}} = \mathsf{m_b}$, then by condition (i) the output clique label of Bob n must be the same as the input clique label of Alice $\mathsf{m_{a(2)}}$. If their input cliques are different and some vertex of Alice's input maximum clique is adjacent to a vertex of Bob's input maximum clique in the graph $\mathcal{G}$, then their labels must ensure that such adjacent vertices are not both assigned the colour 1 at the same time.

[4] Also in Section III of [184]

The relation $\mathcal{R}(\mathcal{G}) \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N} = (\mathcal{C} \times \mathcal{O}) \times \mathcal{C} \times \mathcal{O}$ can be directly defined from the distributed CLP for the graph $\mathcal{G}$. The tuple $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) = ((\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}), \mathsf{m_b}, \mathsf{n})$ belongs to $\mathcal{R}(\mathcal{G}) \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$ if Alice's and Bob's input maximum clique, Alice's input clique label and output clique label of Bob satisfy the constraints of consistent labelling of pairwise cliques described above. For convenience, we write the tuples as $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})$ in this chapter by dropping extra "()". In the distributed computation of $\mathcal{R}(\mathcal{G})$, given Alice's input maximum clique $\mathsf{m_{a(1)}}$ with clique label $\mathsf{m_{a(2)}}$, Bob must output clique label n for his input maximum clique $\mathsf{m_b}$ such that $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})$. In relation reconstruction, for Alice's input $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}})$, Bob's output clique label n for his maximum clique $\mathsf{m_b}$ in different rounds should additionally span all valid labels $\{\mathsf{n} \in \mathsf{N} : (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})\}$. For $\mathcal{R}(\mathcal{G})$, the payoff of relation reconstruction, as in Equation (3.3), is given below:

$$\mathcal{S}_{\mathcal{R}(\mathcal{G})} = \min_{(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})} p(\mathsf{n}|\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}) \tag{3.11}$$

Since the payoff in Equation (3.3) is a faithful measure of relation reconstruction, we have $\mathcal{S}_{\mathcal{R}(\mathcal{G})} > 0$ whenever the relation $\mathcal{R}(\mathcal{G})$ can be successfully reconstructed. As before, the algebraic maximum for this payoff is $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$. This maximum value can be expressed directly in terms of the elements of $\mathcal{R}(\mathcal{G})$[5]. For each input $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}})$ of Alice and $\mathsf{m_b}$ of Bob, we define $\eta(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}) := |\{\mathsf{n} \in \mathsf{N} : (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})\}|$. Then,

$$\eta = \max_{\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}} \eta(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}) \tag{3.12}$$

$$0 \leq \mathcal{S}_{\mathcal{R}(\mathcal{G})} \leq \frac{1}{\eta} = \mathcal{S}^*_{\mathcal{R}(\mathcal{G})} \tag{3.13}$$

Consider the graph $\mathcal{G}^{(\alpha,\beta)}_{disc.}$, as in Definition 2.25, where $\omega(\mathcal{G}) = \beta$ and all maximum cliques are disconnected. For $\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})$, $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})} = \frac{1}{\beta}$. The payoff is optimal when $\mathcal{S}_{\mathcal{R}(\mathcal{G})} = \frac{1}{\eta} = \mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$.

For the distributed computation or relation reconstruction, Alice and Bob follow some protocol. Following a protocol, suppose Bob outputs a clique label n for his input maximum clique $\mathsf{m_b}$ with conditional probability $p(\mathsf{n}|\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b})$ when $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}})$ is the input of Alice. The set $\{p(\mathsf{n}|\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b})\}_{\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}}$ can be arranged into a table $\mathbf{W}$. The rows are indexed by Alice's inputs $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}})$, and the columns by Bob's input-output pairs $(\mathsf{m_b}, \mathsf{n})$. In both cases, the index first runs over the clique labels $\mathcal{O} = \{0, 1, \ldots, \omega(\mathcal{G}) - 1\}$ in increasing order and then moves to the next maximum clique in $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$. We will use this table in some of the proofs later in this chapter. In terms of the elements $\mathbf{W}_{(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}), (\mathsf{m_b}, \mathsf{n})} := p(\mathsf{n}|\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b})$ of this table, the condition for a protocol to accomplish zero-error distributed computation of $\mathcal{R}(\mathcal{G})$ is stated as **(T0)**. The condition for relation reconstruction requires both **(T0)** and **(T1)**.

[5] Also in Section III of [184]

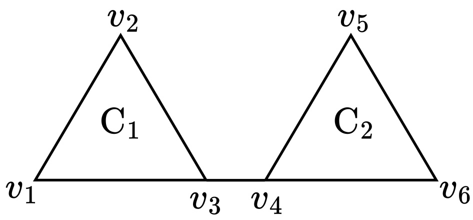

Figure 3.2: A graph $\mathcal{G}$ with $\omega(\mathcal{G}) = 3$ and two maximum cliques C_1 and C_2.

(T0) : $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},\mathsf{m}_{\mathsf{a}(2)}),(\mathsf{m}_\mathsf{b},\mathsf{n})} = 0$ if $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_\mathsf{b}, \mathsf{n}) \notin \mathcal{R}(\mathcal{G})$

(T1) : $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},\mathsf{m}_{\mathsf{a}(2)}),(\mathsf{m}_\mathsf{b},\mathsf{n})} > 0$ if $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_\mathsf{b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})$

Notice that any protocol that accomplishes relation reconstruction, *i.e.*, $\mathcal{S}_{\mathcal{R}(\mathcal{G})} > 0$, can also perform the distributed computation of the relation $\mathcal{R}(\mathcal{G})$. We denote an entire row of $\mathbf{W}$ indexed by $\mathsf{m}_\mathsf{a} = (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$ as $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},\mathsf{m}_{\mathsf{a}(2)}),(ALL)}$. Similarly, an entire column of $\mathbf{W}$ indexed by $(\mathsf{m}_\mathsf{b}, \mathsf{n})$ is denoted as $\mathbf{W}_{(ALL),(\mathsf{m}_\mathsf{b},\mathsf{n})}$. The table $\mathbf{W}$ contains $\alpha\omega(\mathcal{G})$ rows and $\alpha\omega(\mathcal{G})$ columns. For a fixed $\mathsf{m}_{\mathsf{a}(1)} = C_i$ and $\mathsf{m}_\mathsf{b} = C_j$, the block of $\mathbf{W}$ formed by $\{\mathbf{W}_{(C_i,k),(C_j,l)} : k, l \in \mathcal{O}\}$ is denoted as $\mathbf{W}_{(C_i,ALL),(C_j,ALL)}$. There are α^2 such blocks in $\mathbf{W}$. If $\mathbf{T(0)}$ is satisfied, then the diagonal blocks $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},ALL),(\mathsf{m}_\mathsf{b},ALL)}$ with $\mathsf{m}_{\mathsf{a}(1)} = \mathsf{m}_\mathsf{b}$ are equal to $I_{\omega(\mathcal{G})}$. $I_{\omega(\mathcal{G})}$ is $\omega(\mathcal{G}) \times \omega(\mathcal{G})$ identity matrix. This reflects the fact that if Alice and Bob receive the same clique, Bob must reproduce Alice's clique label exactly.

As an example, consider the graph $\mathcal{G}$ shown in Figure 3.2. For this graph, $\mathbf{V}(\mathcal{G}) = \{v_1, v_2, \cdots, v_6\}$, $\omega(\mathcal{G}) = 3$, the set of maximum cliques is $\mathcal{C} = \{C_1, C_2\}$, and the set of clique labels is $\mathcal{O} = \{0, 1, 2\}$. Also, $\mathbf{V}(C_1) = \{v_1, v_2, v_3\}$ and $\mathbf{V}(C_2) = \{v_4, v_5, v_6\}$. From Definition 2.27, for C_1, the binary colourings $\mathsf{f}(v) = \delta_{v,v_1}$,$\mathsf{f}(v) = \delta_{v,v_2}$ and $\mathsf{f}(v) = \delta_{v,v_3}$, where $v \in \mathbf{V}(C_1)$, are mapped to the clique labels $g_{C_1}(\mathsf{f}) = 0$, $g_{C_1}(\mathsf{f}) = 1$ and $g_{C_1}(\mathsf{f}) = 2$. Similarly for C_2, the binary colourings $\mathsf{f}(v) = \delta_{v,v_4}$,$\mathsf{f}(v) = \delta_{v,v_5}$ and $\mathsf{f}(v) = \delta_{v,v_6}$, where $v \in \mathbf{V}(C_2)$, are mapped to clique labels $g_{C_2}(\mathsf{f}) = 0$, $g_{C_2}(\mathsf{f}) = 1$ and $g_{C_2}(\mathsf{f}) = 2$. For this graph, the relation $\mathcal{R}(\mathcal{G})$ is given below:

$$\begin{aligned}\mathcal{R}(\mathcal{G}) = \{&(C_1,0,C_1,0),(C_1,1,C_1,1),(C_1,2,C_1,2),(C_1,0,C_2,0),(C_1,0,C_2,1),(C_1,0,C_2,2),\\&(C_1,1,C_2,0),(C_1,1,C_2,1),(C_1,1,C_2,2),(C_1,2,C_2,1),(C_1,2,C_2,2),(C_2,0,C_2,0),\\&(C_2,1,C_2,1),(C_2,2,C_2,2),(C_2,0,C_1,0),(C_2,0,C_1,1),(C_2,1,C_1,0),(C_2,1,C_1,1),\\&(C_2,1,C_1,2),(C_2,2,C_1,0),(C_2,2,C_1,1),(C_2,2,C_1,2)\}\end{aligned} \tag{3.14}$$

The Table 3.1 shows $\mathbf{W}$ for a protocol that accomplishes zero-error distributed computation or relation reconstruction. The entries marked $*$ can take values in $[0, 1]$ subject to normalisation. $\mathbf{W}$ then satisfies condition **(T0)**, and the protocol can be used to achieve zero-error distributed

$(\mathsf{m}_{\mathsf{a}(1)},\mathsf{m}_{\mathsf{a}(2)})/(\mathsf{m}_\mathsf{b},\mathsf{n})$	$(\mathrm{C}_1,0)$	$(\mathrm{C}_1,1)$	$(\mathrm{C}_1,2)$	$(\mathrm{C}_2,0)$	$(\mathrm{C}_2,1)$	$(\mathrm{C}_2,2)$
$(\mathrm{C}_1,0)$	1	0	0	$*$	$*$	$*$
$(\mathrm{C}_1,1)$	0	1	0	$*$	$*$	$*$
$(\mathrm{C}_1,2)$	0	0	1	0	$*$	$*$
$(\mathrm{C}_2,0)$	$*$	$*$	0	1	0	0
$(\mathrm{C}_2,1)$	$*$	$*$	$*$	0	1	0
$(\mathrm{C}_2,2)$	$*$	$*$	$*$	0	0	1

Table 3.1: An example of the table $\mathbf{W}$ of conditional probabilities for a protocol that accomplishes the distributed computation of $\mathcal{R}(\mathcal{G})$ and relation reconstruction of $\mathcal{R}(\mathcal{G})$. The graph $\mathcal{G}$ is shown in Figure 3.2. For distributed computation, the entries marked $*$ take any values in $[0,1]$ subject to normalisation. For relation reconstruction, these entries must instead take values in $(0,1)$ while still satisfying normalisation.

computation of the relation. If the entries marked $*$ are strictly in $(0,1)$, the condition **(T1)** is also satisfied, and the protocol can accomplish relation reconstruction. In this case, the algebraic maximum of the payoff is $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})} = \frac{1}{3}$.

3.2 Classical and Quantum CCR and S-CCR of $\mathcal{R}(\mathcal{G})$

We now study the classical and quantum communication costs required for zero-error distributed computation of $\mathcal{R}(\mathcal{G})$ and the reconstruction of the relation. We first calculate the minimum classical and quantum communication needed for the zero-error distributed computation of the relation $\mathcal{R}(\mathcal{G})$ when Alice has a one-way noiseless channel to Bob. In the PM scenario, the parties have no shared resources or public coins, such as shared randomness or entanglement. They can use local randomness, and communication is considered costly in this setting. We show that the classical CCR is the same as the quantum CCR for some graphs, *i.e.*, there is no quantum advantage in the distributed computation of $\mathcal{R}(\mathcal{G})$.

For certain graphs $\mathcal{G}$, we then consider the task of reconstruction of relation. We compute a lower bound on the classical S-CCR and give an upper bound on the quantum S-CCR. These results yield a sufficient condition for quantum advantage in this task. In particular, we identify families of graphs where the gap between classical and quantum S-CCR is arbitrarily large.

3.2.1 Communication Complexity of $\mathcal{R}(\mathcal{G})$

We now discuss the classical and quantum zero-error one-way communication complexity of the relation $\mathcal{R}(\mathcal{G})$ (see Section 3.1.1) in the absence of shared resources. We consider the relations $\mathcal{R}(\mathcal{G})$ defined for the following graphs here. Consider an orthogonality graph $\mathcal{G}$ with vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$. The graph has $\alpha(\geq 2)$ maximum cliques $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$ with clique number $\omega(\mathcal{G})(\geq 2)$. Every vertex lies in at least one maximum clique, and its chromatic number equals $\omega(\mathcal{G})$. The set of clique labels for each maximum clique is $\mathcal{O} = \{0, 1, \ldots, \omega(\mathcal{G})-1\}$. For the relation $\mathcal{R}(\mathcal{G}) \subset \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$, $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}$, $\mathsf{M}_B = \mathcal{C}$ and $\mathsf{N} = \mathcal{O}$.

Theorem 3.1. *For $\mathcal{R}(\mathcal{G})$, the classical and quantum CCR without shared resources are $\log_2 \omega(\mathcal{G})$ bit and $\log_2 \omega(\mathcal{G})$ qubit, respectively.*

Proof. [6] We first note that both classical and quantum CCR are lower bounded by $\log_2 \omega(\mathcal{G})$ bit or qubit. Equivalently, Alice must send at least an $\omega(\mathcal{G})$-dimensional classical or quantum system to Bob for the zero-error distributed computation of $\mathcal{R}(\mathcal{G})$. Here, Alice's and Bob's input set are $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}$, $\mathsf{M}_B = \mathcal{C}$, respectively, and Bob's output set is $\mathsf{N} = \mathcal{O}$. Consider the case when Alice's input maximum clique $\mathsf{m}_{\mathsf{a}(1)} = \mathrm{C}_i \in \mathcal{C}$ is the same as Bob's input maximum clique $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_i \in \mathcal{C}$. From condition (i) of consistent labelling of pairwise cliques for distributed CLP (see Section 3.1.1), Bob's output n must be equal to Alice's input clique label $\mathsf{m}_{\mathsf{a}(2)}$. Since Alice's input clique label $\mathsf{m}_{\mathsf{a}(2)} \in \mathcal{O}$ can take $\omega(\mathcal{G})$ distinct random values, she must encode the inputs $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) \in \{(\mathrm{C}_i, 0), (\mathrm{C}_i, 1), \cdots, (\mathrm{C}_i, \omega(\mathcal{G})-1)\}$ using mutually distinguishable states. This is necessary to ensure that Bob can recover the label exactly whenever $\mathsf{m}_{\mathsf{b}} = \mathsf{m}_{\mathsf{a}(1)} = \mathrm{C}_i$. Hence, Alice needs at least an $\omega(\mathcal{G})$-dimensional system, or equivalently $\log_2 \omega(\mathcal{G})$ bit or qubit for zero-error distributed computation of $\mathcal{R}(\mathcal{G})$. We will show that $\log_2 \omega(\mathcal{G})$ bit are sufficient for distributed computation. Any protocol that uses an $\omega(\mathcal{G})$-dimensional classical system for communication can be simulated by an $\omega(\mathcal{G})$-dimensional quantum system. Thus, sufficiency for the classical case immediately implies that the quantum CCR is also $\log_2 \omega(\mathcal{G})$ qubits.

Note that Alice and Bob know the graph $\mathcal{G}$ and their strategy is represented by a table $\mathbf{W}$ of conditional probability distribution $\{p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})\}_{\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}}$, where n is Bob's output given Alice's input $\mathsf{m}_{\mathsf{a}} = (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$ and Bob's input m_{b}. The rows of $\mathbf{W}$ are denoted by $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}),(ALL)}$, each corresponding to an input of Alice. Similarly, columns of $\mathbf{W}$ are denoted by $\mathbf{W}_{(ALL),(\mathsf{m}_{\mathsf{b}}, \mathsf{n})}$, each corresponding to an input-output tuple of Bob. $\mathbf{W}$ has $\alpha\omega(\mathcal{G})$ rows and $\alpha\omega(\mathcal{G})$ columns. The row and columns are arranged such that the index increasingly runs over all values of $\mathsf{m}_{\mathsf{a}(2)}$ and n, respectively, before updating to the next value of $\mathsf{m}_{\mathsf{a}(1)}$ or m_{b}. The block $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, ALL),(\mathsf{m}_{\mathsf{b}}, ALL)}$ consists of elements $\{\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}),(\mathsf{m}_{\mathsf{b}}, \mathsf{n})} : \mathsf{m}_{\mathsf{a}(2)}, \mathsf{n} \in \mathcal{O}\}$. The diagonal blocks $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}=\mathrm{C}_i, ALL),(\mathsf{m}_{\mathsf{b}}=\mathrm{C}_i, ALL)}$ are equal to $I_{\omega(\mathcal{G})}$ for satisfying **(T0)**. This condition ensures that Bob's output clique label matches Alice's input clique label whenever they receive the same maximum clique. Satisfying the condition **(T0)** is equivalent to accomplishing the distributed computation task.

If the $\alpha\omega(\mathcal{G})$ rows of $\mathbf{W}$ can be grouped into $\omega(\mathcal{G})$ partitions such that rows in the same partition are identical across all columns, either enforced by **(T0)** or by the freedom in assigning values to conditional probabilities, then a protocol using $\log_2 \omega(\mathcal{G})$ bits exists. Alice communicates the

[6] Also in Theorem 1 of [184]

partition containing her input, and Bob can then output a clique label for her input maximum clique $\mathsf{m_b}$ following the probability values from the corresponding column of the table $\mathbf{W}$, thus satisfying the consistency conditions. The converse is trivially true. Given a protocol using $\log_2 \omega(\mathcal{G})$ bits exists, for each $\mathrm{C}_i \in \mathcal{C}$, inputs $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) \in \{(\mathrm{C}_i, 0), (\mathrm{C}_i, 1), \cdots, (\mathrm{C}_i, \omega(\mathcal{G})-1)\}$ must be deterministically encoded with distinct symbols by Alice and sent to Bob and so that whenever $\mathsf{m_b} = \mathrm{C}_i$, Bob can perfectly guess Alice's input clique label. Thus, the rows corresponding to inputs that are encoded using the same symbol will have identical entries across all columns in table $\mathbf{W}$.

The $\alpha\omega(\mathcal{G})$ rows of $\mathbf{W}$ satisfying **(T0)** cannot be grouped into fewer than $\omega(\mathcal{G})$ number of partitions such that any two rows in the same partition are identical. This follows from the diagonal blocks $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}=\mathrm{C}_i, ALL),(\mathsf{m_b}=\mathrm{C}_i, ALL)}$ being equal to $I_{\omega(\mathcal{G})}$ so that $\mathbf{W}$ satisfies condition **(T0)**. Since each row of $I_{\omega(\mathcal{G})}$ is different, the rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$, where $k \in \mathcal{O}$, of $\mathbf{W}$ are also distinct. For a protocol using $\log_2 \omega(\mathcal{G})$ bit, let the disjoint partitions be $\mathcal{Z}_r$, where $r \in \{0, 1, \cdots, \omega(\mathcal{G}) - 1\}$. Each partition contains identical rows of $\mathbf{W}$. For every $\mathrm{C}_i \in \mathcal{C}$, the $\omega(\mathcal{G})$ rows $\mathbf{W}_{(\mathrm{C}_i,0),(ALL)}, \mathbf{W}_{(\mathrm{C}_i,1),(ALL)}, \cdots, \mathbf{W}_{(\mathrm{C}_i,\omega(\mathcal{G})-1),(ALL)}$ must each belong to a different partition $\mathcal{Z}_r$ as they are all distinct. Every partition $\mathcal{Z}_r$ contains exactly one row from this set. Before proceeding, we record some properties of $\mathbf{W}$ when the rows can be grouped into $\omega(\mathcal{G})$ disjoint partitions $\mathcal{Z}_r$ while respecting the above restrictions. Imposing this requirement introduces the following constraints on the structure of $\mathbf{W}$ that can be used to perform distributed computation:

(i) Consider an off-diagonal block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)}$ where $\mathrm{C}_i, \mathrm{C}_{j\neq i} \in \mathcal{C}$. For any $k \in \mathcal{O}$, exactly one element in block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)}$ among those common to the row $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ can be non-zero. If instead two entries are non-zero, say $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} \neq 0, \mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l')} \neq 0$ for $l, l' \in \mathcal{O}$. Then the row $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ cannot belong to any partition $\mathcal{Z}_r$ already containing a row $\mathbf{W}_{(\mathrm{C}_j,\tilde{l}),(ALL)}$, where $\tilde{l} \in \mathcal{O}$. For each $\tilde{l} \in \mathcal{O}$, there is a column $\mathbf{W}_{(ALL),(\mathrm{C}_j,l'')}$ where the rows $\mathbf{W}_{(\mathrm{C}_j,\tilde{l}),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ have different entries. The reason is that none of the rows in the block $\mathbf{W}_{(\mathrm{C}_j,ALL),(\mathrm{C}_j,ALL)} = I_{\omega(\mathcal{G})}$ has positive entries in two different columns. This would in turn force the number of partitions to increase to $\omega(\mathcal{G}) + 1$.

(ii) Consider an off-diagonal block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)}$ where $\mathrm{C}_i, \mathrm{C}_{j\neq i} \in \mathcal{C}$. For any $l' \in \mathcal{O}$, exactly one entry in block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)}$ among those common with the column $\mathbf{W}_{(ALL),(\mathrm{C}_j,l')}$ can be non-zero. Say this is not the case and $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l')} \neq 0, \mathbf{W}_{(\mathrm{C}_i,k'),(\mathrm{C}_j,l')} \neq 0$ for instance, where $k, k' \in \mathcal{O}$. The rows containing these two non-zero entries $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ as well as $\mathbf{W}_{(\mathrm{C}_i,k'),(ALL)}$ can only belong to the partition containing $\mathbf{W}_{(\mathrm{C}_j,l'),(ALL)}$. Since these rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_i,k'),(ALL)}$ are distinct, they must belong to different partitions, again forcing more than $\omega(\mathcal{G})$ partitions. Thus, if the number of partitions is limited to ω, then each off-diagonal block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)}$ must be some permutation matrix $\pi_{(\mathrm{C}_i,\mathrm{C}_j)}$.

(iii) The table should satisfy $\mathbf{W}^T = \mathbf{W}$. Otherwise there would be an element $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = 1 \neq \mathbf{W}_{(\mathrm{C}_j,l),(\mathrm{C}_i,k)}$, where $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$ and $k, l \in \mathcal{O}$. The row $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ must belong to the same partition as row $\mathbf{W}_{(\mathrm{C}_j,l),(ALL)}$ as $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = 1 = \mathbf{W}_{(\mathrm{C}_j,l)(\mathrm{C}_j,l)} = 1$ whereas $\mathbf{W}_{(\mathrm{C}_j,l')(\mathrm{C}_j,l)} = 0$ for all $l'(\neq l) \in \mathcal{O}$. However, the rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_j,l),(ALL)}$ have different entries in column $\mathbf{W}_{(ALL),(\mathrm{C}_i,k)}$ as $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_i,k)} = 1 \neq \mathbf{W}_{(\mathrm{C}_j,l),(\mathrm{C}_i,k)}$. Thus, the row $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ ultimately cannot belong to any partition that contains row $\mathbf{W}_{(\mathrm{C}_j,l''),(ALL)}$ where $l'' \in \mathcal{O}$, forcing more than $\omega(\mathcal{G})$ partitions. Moreover, once the permutation matrices in the first row of blocks are fixed, the identity in block diagonals require that the block permutations must satisfy $\pi_{(\mathrm{C}_i,\mathrm{C}_j)}{=}\pi^{-1}_{(\mathrm{C}_1,\mathrm{C}_i)}\pi_{(\mathrm{C}_1,\mathrm{C}_j)}$. Thus, the entire family of permutations is determined by the choice of permutations with respect to a reference clique, say C_1.

For example, Alice and Bob consider a set $\{(v_{t,1}, v_{t,2}, \cdots, v_{t,\alpha}) : v_{t,j} \in \mathbf{V}(\mathrm{C}_j), \quad t \in \{1, 2, \cdots, \omega(\mathcal{G})\}\}$ for $\mathcal{G}$ such that for all t, j, j' $v_{t,j} = v_{t,j'}$ whenever $v_{t,j} \in \mathbf{V}(\mathrm{C}_{j'})$ else $v_{t,j}$ and $v_{t,j'}$ are non-adjacent. Also $v_{t,j} \neq v_{t',j'}$ whenever $t \neq t'$ and $j \neq j'$. Say clique label $l, k \in \mathcal{O}$ correspond to binary colouring of C_i and C_j that assign binary colour 1 to vertex $v_{t,i}$ and $v_{t',j}$, respectively. Then $\mathbf{W}_{(\mathrm{C}_j,l),(\mathrm{C}_i,k)} = 1$ *iff* $v_{t,i}$ and $v_{t',j}$ are in the same tuple, *i.e*, $t = t'$.

Now, we will construct a specific kind $\omega(\mathcal{G})$ disjoint partitions $\mathcal{Z}_r$ where $r \in \{0, 1, \cdots, \omega(\mathcal{G}-1)\}$ of the inputs of Alice based on a table that has the form as stated above.

Step 1: For all $k \in \{0, 1, \cdots, \omega(\mathcal{G}) - 1\}$, $(\mathsf{m}_{\mathsf{a}(1)} = \mathrm{C}_1, \mathsf{m}_{\mathsf{a}(2)} = k) \in \mathcal{Z}_k$.

Step 2: $\forall j \in \{2, \cdots, \alpha\}$, if $\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_j,ALL)}$ is the permutation matrix $\pi_{\mathrm{C}_1,\mathrm{C}_j}$ and k' is the k^{th} element of $\pi_{\mathrm{C}_1,\mathrm{C}_j} \times (0\ 1\ \cdots\ \omega(\mathcal{G}) - 1)^T$, then the input $(\mathsf{m}_{\mathsf{a}(1)} = \mathrm{C}_j, \mathsf{m}_{\mathsf{a}(2)} = k') \in \mathcal{Z}_k$.

Given an input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$, Alice communicates the partition corresponding to the input. Bob can thus pick a valid label n for his input maximum clique $\mathsf{m_b}$ using the information of the row. It is essential that each row $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},0),(ALL)}, \mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},1),(ALL)}, \cdots \mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},\omega(\mathcal{G})-1),(ALL)}$ belongs to a different partition otherwise Bob will not be able to satisfy consistency condition.

As an example, consider the graph $\mathcal{G}$ in Figure 3.3. For this graph, $\mathbf{V}(\mathcal{G}) = \{v_1, v_2, v_3, v_4, v_5\}$, $\omega(\mathcal{G}) = 3$, the set of maximum cliques is $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2\}$, and the set of clique labels is $\mathcal{O} = \{0, 1, 2\}$. Here, we have $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}, \mathsf{M}_B = \mathcal{C}$ and $\mathsf{N} = \mathcal{O}$. Table 3.2 shows $\mathbf{W}$ for a protocol fixed by Alice and Bob that achieves zero-error distributed computation or relation reconstruction. The rows are partitioned as follows: $\mathcal{Z}_0 = \{\mathbf{W}_{(\mathrm{C}_1,0),(ALL)}, \mathbf{W}_{(\mathrm{C}_2,2),(ALL)}\}$, $\mathcal{Z}_1 = \{\mathbf{W}_{(\mathrm{C}_1,1),(ALL)}, \mathbf{W}_{(\mathrm{C}_2,1),(ALL)}\}$ and $\mathcal{Z}_2 = \{\mathbf{W}_{(\mathrm{C}_1,2),(ALL)}, \mathbf{W}_{(\mathrm{C}_2,0),(ALL)}\}$. On receiving input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$, Alice sends symbol $\tau = r$ if $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)},\mathsf{m}_{\mathsf{a}(2)}),(ALL)}$ belongs to partition $\mathcal{Z}_r$.

$(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})/(\mathsf{m}_\mathsf{b}, \mathsf{n})$	$(\mathrm{C}_1, 0)$	$(\mathrm{C}_1, 1)$	$(\mathrm{C}_1, 2)$	$(\mathrm{C}_2, 0)$	$(\mathrm{C}_2, 1)$	$(\mathrm{C}_2, 2)$
$(\mathrm{C}_1, 0)$	1	0	0	0	0	1
$(\mathrm{C}_1, 1)$	0	1	0	0	1	0
$(\mathrm{C}_1, 2)$	0	0	1	1	0	0
$(\mathrm{C}_2, 0)$	0	0	1	1	0	0
$(\mathrm{C}_2, 1)$	0	1	0	0	1	0
$(\mathrm{C}_2, 2)$	1	0	0	0	0	1

Table 3.2: An example of the table **W** of conditional probabilities for a protocol that performs the distributed computation of $\mathcal{R}(\mathcal{G})$, based on the graph in Figure 3.3.

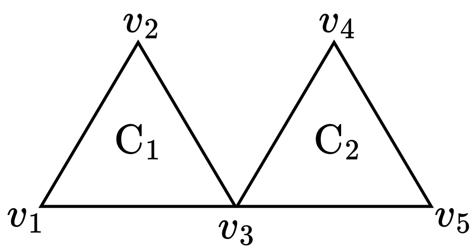

Figure 3.3: A graph $\mathcal{G}$ with $\omega(\mathcal{G}) = 3$ and two maximum cliques C_1 and C_2.

After receiving τ, Bob identifies the corresponding row and uses it to output the clique label n for his input clique m_b. This guarantees that the consistency condition is always satisfied. Hence, $\log_2 3$ bit of communication are sufficient for the distributed computation of $\mathcal{R}(\mathcal{G})$.

■

Notice that given $\log_2 \omega(\mathcal{G})$ bits of communication without shared resources, Alice and Bob must use deterministic encoding and decoding for achieving zero-error distributed computation of $\mathcal{R}(\mathcal{G})$. In particular, whenever input maximum cliques $\mathsf{m}_{\mathsf{a}(1)} = \mathsf{m}_\mathsf{b}$, Bob's output clique label n must be equal to Alice's input clique label $\mathsf{m}_{\mathsf{a}(2)}$. Thus, Alice must encode the inputs $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) \in \{(\mathrm{C}_i, 0), (\mathrm{C}_i, 1), \cdots, (\mathrm{C}_i, \omega(\mathcal{G}) - 1)\}$ in different symbols. If Alice randomises her encoding locally, Bob cannot determine Alice's input clique label and guess it perfectly when they receive the same maximum clique. Likewise, if Bob locally randomises his decoding of the received symbol, he will fail to reproduce Alice's clique label whenever their input cliques match.

Theorem 3.1 shows that there is no quantum advantage in the distributed computation of $\mathcal{R}(\mathcal{G})$ for the graphs we considered. We now show that a quantum advantage exists in the relation reconstruction task. The FOR of the graph plays no role in distributed computation but becomes relevant for relation reconstruction, where quantum communication enables unbounded advantage.

3.2.2 Strong Communication Complexity of $\mathcal{R}(\mathcal{G})$

In this section, we study the classical and quantum one-way strong communication complexity of the relation $\mathcal{R}(\mathcal{G})$ when the parties do not have shared resources. We focus on specific classes of graphs $\mathcal{G}$ for the relation reconstruction task. The orthogonality graph $\mathcal{G}$ which we would consider has vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$. It contains $\alpha(\geq 2)$ maximum cliques $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$, and the set of clique labels for each maximum clique is $\mathcal{O} = \{0, 1, \ldots, \omega(\mathcal{G}) - 1\}$, where $\omega(\mathcal{G}) \geq 2$. The graph has at least two distinct maximum cliques, say $\mathrm{C}_i, \mathrm{C}_j$, such that some vertex in C_i is non-adjacent to at least two vertices in C_j. This orthogonality graph $\mathcal{G}$ is assumed to satisfy the following conditions:

(G0) : Every vertex of $\mathcal{G}$ belongs to at least one maximum clique in the graph.

(G1) : For each maximum clique C_i and each vertex $v_k \in \mathbf{V}(\mathrm{C}_i)$, take any vertex $v_{k'}$ in another maximum clique C_j but not in C_i. Then there exists a vertex $v_l \in \mathbf{V}(\mathcal{G})$ which is adjacent only to v_k.

For graph $\mathcal{G}$ satisfying **(G1)**, symmetry of the condition implies that there exists a vertex $v_{l'} \in \mathbf{V}(\mathcal{G})$ which is adjacent only to $v_{k'}$. For such a graph, the relation $\mathcal{R}(\mathcal{G}) \subset \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$, is defined with $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}$, $\mathsf{M}_B = \mathcal{C}$ and $\mathsf{N} = \mathcal{O}$. In the relation reconstruction, given Alice's input $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) \in \mathsf{M}_A$ and Bob's input $\mathsf{m_b} \in \mathsf{M}_B$, Bob must never output $\mathsf{n} \in \mathsf{N}$ if $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \notin \mathcal{R}(\mathcal{G})$. Moreover, for each such input pair, Bob's outputs across different rounds must span $\{\mathsf{n} : (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})\}$. In other words, $p(\mathsf{n}|\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}) > 0$ for all $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})$. Before turning to the classical communication cost of relation reconstruction, we introduce a notion for a graph that will be useful in determining the classical S-CCR without shared resources.

Definition 3.3. Let $\mathcal{G}$ be a graph with vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$. A *disjointness representation* of $\mathcal{G}$ is an assignment of distinct non-empty subsets $T_v \subseteq T$ of a finite set T, which we refer to as ground set, to the vertices $v \in \mathbf{V}(\mathcal{G})$ such that, for any two distinct vertices $v_i, v_j \in \mathbf{V}(\mathcal{G})$: (i) $T_{v_i} \nsubseteq T_{v_j}$, $T_{v_j} \nsubseteq T_{v_i}$, *i.e.* $\{T_v\}$ is an *anti-chain* and (ii) $T_{v_j} \cap T_{v_i} = \varnothing$ if $(v_i, v_j) \in \mathbf{E}(\mathcal{G})$.

Definition 3.4. The *disjointness number* $K(\mathcal{G})$ is the size of the smallest set T such that $\mathcal{G}$ admits a disjointness representation over some ground set T.

Disjointness representation is a set-theoretic analogue of orthogonal representation with an additional anti-chain property. The disjointness number is thus almost set-theoretically equivalent of the orthogonal range of a graph.

Observation 3.1. *For any graph $\mathcal{G}$, $K(\mathcal{G}) \geq \omega(\mathcal{G})$, since the vertices of a maximum clique must be assigned pairwise disjoint non-empty sets. This requires at least $\omega(\mathcal{G})$ distinct elements in T.*

Now consider a disjointness representation $\{T_v : v \in \mathbf{V}(\mathcal{G})\}$ of graph $\mathcal{G}$ defined over the ground set T. By Definition, the family $\{T_v\}_{v\in\mathbf{V}(\mathcal{G})}$ forms an anti-chain. For any distinct $v, v' \in \mathbf{V}(\mathcal{G})$ we have $T_v \not\subseteq T_{v'}$ and $T_{v'} \not\subseteq T_v$. By Sperner's theorem, the maximum size of an anti-chain on a k-element set T is $\binom{k}{\lfloor k/2\rfloor}$ [246]. Hence,

$$|\mathbf{V}(\mathcal{G})| \;\leq\; \binom{k}{\lfloor k/2\rfloor} \;\leq\; 2^k \tag{3.15}$$

We define

$$k^* = \min\Big\{k :\; |\mathbf{V}(\mathcal{G})| \leq \binom{k}{\lfloor k/2\rfloor}\Big\} \tag{3.16}$$

Since $K(\mathcal{G})$ is the minimum possible size of such a ground set T, then $K(\mathcal{G}) \geq k^*$. This is because $k \geq k^*$ is necessary just to satisfy the condition (i) in Definition 3.3. However, it does not guarantee that condition (ii) is satisfied for such k. Taking the logarithm on both sides in Equation (3.15), it follows that

$$K(\mathcal{G}) \geq k^* \geq \log_2 |\mathbf{V}(\mathcal{G})| \tag{3.17}$$

Using refinement of Stirling's formula (sometimes called Robbins bounds), for $k \geq 1$ we have $\sqrt{2\pi k}(\frac{k}{e})^k e^{\frac{1}{12k+1}} < k! < \sqrt{2\pi k}(\frac{k}{e})^k e^{\frac{1}{12k}}$. This implies an upper bound on the central binomial coefficient, $\binom{k}{\lfloor k/2\rfloor} \leq c\frac{2^k}{\sqrt{k}}$ for some constant $c > 0$ [247]. Now take k such that $\binom{k}{\lfloor k/2\rfloor} \geq |\mathbf{V}(\mathcal{G})|$. For k satisfying Equation (3.15),

$$|\mathbf{V}(\mathcal{G})| \leq c\frac{2^k}{\sqrt{k}} \implies \log_2 |\mathbf{V}(\mathcal{G})| \leq k - \frac{\log_2 k}{2} + \log_2 c \tag{3.18}$$

$$\implies k \geq \log_2 |\mathbf{V}(\mathcal{G})| + \frac{\log_2 k}{2} - \log_2 c \tag{3.19}$$

As $k \geq 1$, it follows that $k \geq \log_2 |\mathbf{V}(\mathcal{G})| - \log_2 c$. Substituting this in Equation (3.19),

$$k \geq \log_2 |\mathbf{V}(\mathcal{G})| + \frac{\log_2 k}{2} - \log_2 c \geq \log_2 |\mathbf{V}(\mathcal{G})| + \frac{\log_2(\log_2 |\mathbf{V}(\mathcal{G})| - \log_2 c)}{2} - \log_2 c \tag{3.20}$$

For large $|\mathbf{V}(\mathcal{G})|$, $\log_2(\log_2 |\mathbf{V}(\mathcal{G})| - \log_2 c) = \log_2 \log_2 |\mathbf{V}(\mathcal{G})| + O(1)$. So,

$$k \geq \log_2 |\mathbf{V}(\mathcal{G})| + \frac{\log_2 \log_2 |\mathbf{V}(\mathcal{G})|}{2} + O(1) \tag{3.21}$$

Since the inequality (3.21) also holds for $k = k^*$, and $K(\mathcal{G}) \geq k^*$, we obtain the asymptotic lower bound

$$K(\mathcal{G}) \geq \log_2 |\mathbf{V}(\mathcal{G})| + \frac{\log_2 \log_2 |\mathbf{V}(\mathcal{G})|}{2} + O(1) \tag{3.22}$$

Now we show that the minimum classical cost of relation reconstruction, *i.e.* S-CCR, without shared resources grows at least logarithmically with the disjointness number $K(\mathcal{G})$ of the graph.

Theorem 3.2. *For $\mathcal{R}(\mathcal{G})$, where $\mathcal{G}$ satisfies conditions **(G0)** and **(G1)**, the classical S-CCR without shared resources and without private randomness for Alice is $\log_2 |\mathbf{V}(\mathcal{G})|$ bits*[7]*. More generally, the classical S-CCR without shared resources is at least $\log_2 K(\mathcal{G})$ bit.*

Proof. Before the task begins, Alice and Bob may fix a protocol. Such a protocol can be represented by a table $\mathbf{W}$ with $\alpha\omega(\mathcal{G})$ rows and $\alpha\omega(\mathcal{G})$ columns. Its entries are conditional probabilities $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$. If $\mathbf{W}$ satisfies conditions **(T0)** and **(T1)** (see Section 3.1.1), then it accomplishes relation reconstruction. Given input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$, Alice communicates the associated row, and Bob with probability $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$ outputs clique label n based on his input m_{b}. The deterministic strategy used in the proof of Theorem 3.1 cannot be applied here. In that case, $\mathbf{W}$ assigns probability 1 to a single outcome n for each input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$ of Alice and m_{b} of Bob. For relation reconstruction, however, we require that $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}) > 0$ for all $n \in \mathsf{N}$ when $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \in \mathcal{R}(\mathcal{G})$.

Given a table $\mathbf{W}$, if two rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_{i'},k'),(ALL)}$, where $\mathrm{C}_i, \mathrm{C}_{i'} \in \mathcal{C}$ and $k, k' \in \mathcal{O}$, are identical then Alice can encode them using the same symbol for communication. Bob can then decode the row from the received symbol and output his clique label based on his input. In the graph $\mathcal{G}$, a vertex v may belong to two distinct maximum cliques, say C_i and $\mathrm{C}_{i'}$. Then there exist labels $k \in \mathcal{O}$ for C_i and $k' \in \mathcal{O}$ for $\mathrm{C}_{i'}$ corresponding to binary colourings for respective cliques that assign colour 1 to v. Then rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_{i'},k'),(ALL)}$ can be given identical entries while satisfying $\mathbf{T(0)}$ and $\mathbf{T(1)}$. This can be seen as follows. Corresponding v, the input-output tuples $(\mathrm{C}_i, k, \mathrm{C}_i, k), (\mathrm{C}_{i'}, k', \mathrm{C}_{i'}, k'), (\mathrm{C}_i, k, \mathrm{C}_{i'}, k'), (\mathrm{C}_{i'}, k', \mathrm{C}_i, k) \in \mathcal{R}(\mathcal{G})$. For satisfying condition **(T0)**, the associated entries $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_i,k)} = \mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_{i'},k')} = \mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_{i'},k')} = \mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_i,k)} = 1$. Moreover, since adjacency of v with vertices in other cliques of $\mathcal{G}$ is fixed, we have for every $\mathrm{C}_{j(\neq i,i')} \in \mathcal{C}$ and $l \in \mathcal{O}$: $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \mathcal{R}(\mathcal{G}) \iff (\mathrm{C}_{i'}, k', \mathrm{C}_j, l) \notin \mathcal{R}(\mathcal{G})$ and $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \mathcal{R}(\mathcal{G}) \iff (\mathrm{C}_{i'}, k', \mathrm{C}_j, l) \in \mathcal{R}(\mathcal{G})$. Thus, to satisfy **(T0)** set $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = \mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_j,l)} = 0$ for all $\mathrm{C}_j \in \mathcal{C}, l \in \mathcal{O}$ if $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \mathcal{R}(\mathcal{G})$. For satisfying **(T1)**, assign same positive values $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = \mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_j,l)} > 0$ for all $\mathrm{C}_j \in \mathcal{C}, l \in \mathcal{O}$ when $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \mathcal{R}(\mathcal{G})$. In particular, the non-zero entries of the two rows can always be chosen identically while satisfying **(T1)**, and the zero entries coincide for both rows, satisfying **(T0)**.

[7] This result is in Lemma 2 of [184]

Consequently, these two rows can be encoded using the same symbol by Alice.

By repeating this process, Alice and Bob can eliminate redundant rows corresponding to the same vertex. Each row of $\mathbf{W}$ that is associated with a vertex $v \in \mathbf{V}(\mathcal{G})$ can be given identical assignments while satisfying **(T0)** and **(T1)**. The resulting table $\mathbf{W}$ therefore has $|\mathbf{V}(\mathcal{G})|$ distinct rows, each representing a unique vertex of $\mathcal{G}$. Alice encodes identical rows using the same symbol, and Bob decodes the received symbol to output a clique label with the appropriate probability based on his input. Hence, $\log_2 |\mathbf{V}(\mathcal{G})|$ bit of communication are sufficient.

We now prove that for $\mathcal{R}(\mathcal{G})$, $\log_2 |\mathbf{V}(\mathcal{G})|$ bit of communication are necessary for relation reconstruction when Alice does not have private randomness and the parties do not have shared resources. Since $\mathcal{G}$ satisfies **(G0)**, each vertex v belongs to some maximum clique $\mathrm{C}_i \in \mathcal{C}$. Thus, corresponding to each vertex $v \in \mathbf{V}(\mathcal{G})$, there is some input of Alice $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$, where the clique label $k \in \mathcal{O}$ corresponds to binary colouring of C_i that assigns colour 1 to v. Hence, the row $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ is associated with vertex v. For any maximum clique C_i, each of Alice's input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$, with $k \in \mathcal{O}$ must be encoded with a distinct symbol. Otherwise, when $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_i$, Bob could not output Alice's clique label $\mathsf{n} = k$ with certainty. Thus, Alice's inputs corresponding to the vertices that belong to the same maximum clique must be encoded using distinct symbols. Since $\mathcal{G}$ also satisfies **(G1)**, consider a vertex $v_t \in \mathbf{V}(\mathrm{C}_i)$ and a vertex $v'_t \in \mathbf{V}(\mathrm{C}_{i'})$ with $v'_t \notin \mathbf{V}(\mathrm{C}_i)$. Then there is a vertex v in some maximum clique C_j which is adjacent to v_t and non-adjacent to $v_{t'}$. By symmetry, there is also a vertex v' in some maximum clique $\mathrm{C}_{j'}$ adjacent to $v_{t'}$ but not to v_t. Let Alice's input $\mathsf{m}_{\mathsf{a}} = (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$ and $\mathsf{m}_{\mathsf{a}} = (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_{i'}, k')$ correspond to the vertex v_t and $v_{t'}$. Here, clique label k and k' assign binary colour 1 to the vertex v_t and $v_{t'}$ in maximum clique C_i and $\mathrm{C}_{i'}$, respectively. Let Bob's input is $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_j$ and clique label l assigns value 1 to the vertex v in this maximum clique C_j. Then for satisfying **(T0)**, $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = p(l|\mathrm{C}_i, k, \mathrm{C}_j) = 0$ and $\mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_j,l)} = p(l|\mathrm{C}_{i'}, k', \mathrm{C}_j) > 0$ for satisfying **(T1)**. Since $\mathbf{W}_{(\mathrm{C}_i,k),(\mathrm{C}_j,l)} = p(l|\mathrm{C}_i, k, \mathrm{C}_j) = 0$ and $\mathbf{W}_{(\mathrm{C}_{i'},k'),(\mathrm{C}_j,l)} = p(l|\mathrm{C}_{i'}, k', \mathrm{C}_j) > 0$, thus the rows $\mathbf{W}_{(\mathrm{C}_i,k),(ALL)}$ and $\mathbf{W}_{(\mathrm{C}_{i'},k'),(ALL)}$ must be distinct. Therefore, Alice cannot encode inputs (C_i, k) and $(\mathrm{C}_{i'}, k')$ with the same symbol: doing so would either force Bob to output $\mathsf{n} = l$ with nonzero probability, violating **(T0)**, or never output it, preventing reconstruction. Combining both cases, inputs corresponding to distinct vertices in $\mathbf{V}(\mathcal{G})$ must be encoded with distinct symbols. Thus, the classical system used for communication must have dimension at least $|\mathbf{V}(\mathcal{G})|$, which implies that $\log_2 |\mathbf{V}(\mathcal{G})|$ bits are necessary.

We now argue that local randomisation over deterministic encodings and decodings cannot reduce the communication cost below $\log_2 K(\mathcal{G})$ bits for relation reconstruction. In other words, using fewer than $\log_2 K(\mathcal{G})$ bits together with private randomness cannot satisfy both **(T0)** and **(T1)**. If it were possible, then $\mathcal{G}$ would admit a disjointness representation over a ground set smaller than $K(\mathcal{G})$. However, it is impossible from Definition 3.4.

Suppose Alice employs a convex mixture of deterministic encoding protocols, each using less than $\log_2 K(\mathcal{G})$ bits. Let T denote the set of alphabets available to her, with $|T| < K(\mathcal{G})$. The conditional distribution can be written as

$$p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}) = \sum_{\tau \in T} p_e(\tau|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) p_d(\mathsf{n}|\tau, \mathsf{m}_{\mathsf{b}}) \tag{3.23}$$

where $p_e(\tau|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$ is the probability with which Alice sends symbol τ given her input and $p_d(\mathsf{n}|\tau, \mathsf{m}_{\mathsf{b}})$ is Bob's decoding probability. For a symbol $\tau \in T$, we define the set $\mathsf{M}_{A,\tau} = \{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k) \in \mathsf{M}_A : p_e(\tau|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) > 0\}$, *i.e.* all inputs that Alice may encode as τ. To satisfy condition **(T0)**, Bob's decoding of τ must ensure $p_d(\mathsf{n}|\tau, \mathsf{m}_{\mathsf{b}}) = 0$ whenever $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \notin \mathcal{R}(\mathcal{G})$ for some $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) \in \mathsf{M}_{A,\tau}$. Since $\mathcal{G}$ satisfies **(G0)**, each vertex v belongs to some maximum clique $\mathrm{C}_i \in \mathcal{C}$. Thus, corresponding to each vertex v there is some input of Alice $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$, where $k \in \mathcal{O}$ is the clique label assigning colour 1 to v in C_i. For each input, we define the support $T_{(\mathrm{C}_i,k)} = \{\tau \in T : p_e(\tau|\mathrm{C}_i, k) > 0\}$.

Consider any two $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$ and $(\mathrm{C}_{i'}, k')$ of Alice which correspond to two adjacent vertices $v, v' \in \mathbf{V}(\mathcal{G})$. Here, clique label k and k' assign colour 1 to vertex $v \in \mathbf{V}(\mathrm{C}_i)$ and $v' \in \mathbf{V}(\mathrm{C}_{i'})$, respectively. If Bob's input $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_i$, then the consistency conditions require $p(k|\mathrm{C}_i, k, \mathrm{C}_i) = 1 \implies p_d(k|\tau, \mathrm{C}_i) = 1 \ \forall \ \tau \in T_{(\mathrm{C}_i,k)}$. On the other hand, for the Alice's input $(\mathrm{C}_{i'}, k')$, consistency requires $p(k|\mathrm{C}_{i'}, k', \mathrm{C}_i) = 0 \implies p_d(k|\tau, \mathrm{C}_i) = 0 \ \forall \ \tau \in T_{(\mathrm{C}_{i'},k')}$. Hence,

(i) $T_{(\mathrm{C}_{i'},k')} \cap T_{(\mathrm{C}_i,k)} = \varnothing$, whenever inputs $(\mathrm{C}_{i'}, k')$ and (C_i, k) correspond to adjacent vertices $v', v \in \mathbf{V}(\mathcal{G})$ respectively, in order to satisfy condition **(T0)**.

Take any two non-adjacent vertices $\tilde{v}, \tilde{v}' \in \mathbf{V}(\mathcal{G})$ and for each of them consider an input of Alice, say $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}) = (\mathrm{C}_i, k)$ and $(\mathrm{C}_{i'}, k')$. Since the vertices are non-adjacent, $\tilde{v} \in \mathbf{V}(\mathrm{C}_i) \setminus \mathbf{V}(\mathrm{C}_{i'})$ and $\tilde{v}' \in \mathbf{V}(\mathrm{C}_{i'}) \setminus \mathbf{V}(\mathrm{C}_i)$. Here, clique label k and k' assign colour 1 to vertex $\tilde{v} \in \mathbf{V}(\mathrm{C}_i)$ and $\tilde{v}' \in \mathbf{V}(\mathrm{C}_{i'})$, respectively. Using **(G1)**, there exists vertex v in maximum clique C_j and vertex v' in maximum clique $\mathrm{C}_{j'}$ such that v is adjacent to $\tilde{v}$ and non-adjacent to $\tilde{v}'$, while v' is adjacent to $\tilde{v}'$ but not to $\tilde{v}$. Say, clique labels $l, l' \in \mathcal{O}$ assign binary colour 1, respectively, to v in maximum clique C_j and v' in maximum clique $\mathrm{C}_{j'}$. Suppose Bob's input $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_j$. Then to satisfy **(T0)**, $p(l|\mathrm{C}_i, k, \mathrm{C}_j) = 0 \implies p_d(l|\tau, \mathrm{C}_j) = 0 \ \forall \ \tau \in T_{(\mathrm{C}_i,k)}$. However, for successful relation reconstruction $p(l|\mathrm{C}_{i'}, k', \mathrm{C}_j) > 0 \implies p_d(l|\tau, \mathrm{C}_j) > 0$ for some $\tau \in T_{(\mathrm{C}_{i'},k')}$. Thus, for successful relation reconstruction $T_{(\mathrm{C}_{i'},k')} \setminus T_{(\mathrm{C}_i,k)} \neq \varnothing$. Now say Bob's input is $\mathsf{m}_{\mathsf{b}} = \mathrm{C}_{j'}$. Then to satisfy **(T0)**, $p(l'|\mathrm{C}_{i'}, k', \mathrm{C}_{j'}) = 0 \implies p_d(l'|\tau, \mathrm{C}_{j'}) = 0 \ \forall \ \tau \in T_{(\mathrm{C}_{i'},k')}$. However, for successful relation reconstruction and $p(l'|\mathrm{C}_i, k, \mathrm{C}_{j'}) > 0 \implies p_d(l'|\tau, \mathrm{C}_{j'}) > 0$ for some $\tau \in T_{(\mathrm{C}_i,k)}$. Thus, for successful relation reconstruction $T_{(\mathrm{C}_i,k)} \setminus T_{(\mathrm{C}_{i'},k')} \neq \varnothing$.

(ii) Thus, $T_{(\mathrm{C}_{i'},k')} \not\subseteq T_{(\mathrm{C}_i,k)}$ and $T_{(\mathrm{C}_i,k)} \not\subseteq T_{(\mathrm{C}_{i'},k')}$ for inputs $(\mathrm{C}_{i'}, k')$ and (C_i, k) corresponding to two non-adjacent vertices $\tilde{v}, \tilde{v}' \in \mathbf{V}(\mathcal{G})$, respectively, in order to satisfy **(T1)**.

Suppose relation reconstruction is possible using set of alphabets in T, where $|T| < K(\mathcal{G})$, *i.e.* using less than $\log_2 K(\mathcal{G})$ bits. Then conditions (i) and (ii) must hold for all inputs. For each vertex $v \in \mathbf{V}(\mathcal{G})$ consider an input (C_i, k) corresponding to it and associate the set $T_v \equiv T_{(\mathrm{C}_i,k)} \subset T$. From condition (i), if v and v' are adjacent, then $T_v \cap T_{v'} = \varnothing$. From condition (ii), if v and v' are non-adjacent, then $T_v \not\subseteq T_{v'}$ and $T_{v'} \not\subseteq T_v$. Hence, the family $\{T_v : v \in \mathbf{V}(\mathcal{G})\}$ forms a disjointness representation of $\mathcal{G}$ over T. But this contradicts the definition of $K(\mathcal{G})$. It is the minimum size of a ground set required for having such a representation for the graph $\mathcal{G}$. Since $|T| < K(\mathcal{G})$, such a representation is impossible using T. Thus, relation reconstruction cannot be achieved with fewer than $\log_2 K(\mathcal{G})$ bits when no shared resources are available.

Consequently, for the relation $\mathcal{R}(\mathcal{G})$, S-CCR without shared randomness and private randomness (for Alice) is $\log_2 |\mathbf{V}(\mathcal{G})|$ bit. More generally, S-CCR without shared randomness is $\geq \log_2 K(\mathcal{G})$ bit. ■

Thus, the zero-error classical strong communication complexity of $\mathcal{R}(\mathcal{G})$ scales logarithmically with the disjointness number $K(\mathcal{G})$ of the graph. We now turn to the reconstruction of relations using quantum communication protocols. We consider $\mathcal{R}(\mathcal{G})$ for graph $\mathcal{G}$ of the type described earlier, which satisfy conditions **(G0)** and **(G1)**.

Lemma 3.1. *Let $\mathcal{G}$ be a graph satisfying **(G0)** and **(G1)** with faithful orthogonal range $\omega(\mathcal{G})$, then for $\mathcal{R}(\mathcal{G})$ the quantum S-CCR without shared resources (and private randomness) is $\log_2 \omega(\mathcal{G})$ qubit.*

Proof. Alice and Bob know the graph $\mathcal{G}$ and a faithful orthogonal representation (FOR) $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{C}^{\omega(\mathcal{G})}$ of $\mathcal{G}$. The FOR ϕ assigns a unit vector $\phi(v) \in \mathbb{C}^{\omega(\mathcal{G})}$ to each vertex $v \in \mathbf{V}(\mathcal{G})$ such that $\langle \phi(v)|\phi(v')\rangle = 0$ if and only if v and v' are adjacent in $\mathcal{G}$. Say, Alice is given input $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) = (\mathrm{C}_i, k)$, where $k \in \mathcal{O}$ and $\mathrm{C}_i \in \mathcal{C}$. The clique label k corresponds to the binary colouring that assigns 1 to vertex v of the maximum clique $\mathrm{C}_i \in \mathcal{C}$. Alice encodes it using the quantum state $\rho_{(\mathrm{C}_i,k)} = |\phi(v)\rangle\langle\phi(v)| \in D(\mathbb{C}^{\omega(\mathcal{G})})$ and sends it to Bob. If Bob receives input $\mathsf{m_b} = \mathrm{C}_j$, then he performs the measurement $\{E = |\phi(v')\rangle\langle\phi(v')| : v' \in \mathbf{V}(\mathrm{C}_j)\}$. On obtaining outcome corresponding to the projector $|\phi(v')\rangle\langle\phi(v')|$, Bob outputs $\mathsf{n} = l$, where the clique label $l \in \mathcal{O}$ for C_j assigns colour 1 to the vertex v'. This quantum strategy always satisfies the consistency condition. This is so because $\mathrm{Tr}[(|\phi(v)\rangle\langle\phi(v)|)(|\phi(v')\rangle\langle\phi(v')|)]$ is 1 if $v = v'$ and 0 if v is adjacent to v'. In all other cases, it is positive. Thus, the protocol accomplishes relation reconstruction. Also, $\log_2 \omega(\mathcal{G})$ qubit are necessary. This is because all the inputs $(\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) \in \{(\mathrm{C}_i, 0), (\mathrm{C}_i, 1), \cdots, (\mathrm{C}_i, \omega(\mathcal{G}) - 1)\}$ must be encoded using mutually orthogonal states. ■

Using the results we discussed so far, we obtain a sufficient condition for quantum advantage in the relation reconstruction task.

Theorem 3.3. *Let $\mathcal{G}$ be a graph satisfying conditions **(G0)** and **(G1)** with faithful orthogonal range $\omega(\mathcal{G})$. Then quantum advantage in relation reconstruction of $\mathcal{R}(\mathcal{G})$ without shared randomness exists whenever $K(\mathcal{G}) > \omega(\mathcal{G})$. If Alice has no private randomness, quantum advantage already exists whenever $|\mathbf{V}(\mathcal{G})| > \omega(\mathcal{G})$.*

Proof. The proof of the claim follows from Theorem 3.2 and Lemma 3.1. For $\mathcal{R}(\mathcal{G})$, the classical S-CCR without shared randomness is at least $\log_2 K(\mathcal{G})$ bit, and without private randomness for Alice it is $\log_2 |\mathbf{V}(\mathcal{G})|$ bit. In contrast, the quantum S-CCR is $\log_2 \omega(\mathcal{G})$ qubit. ■

From Equation 3.22 and Observation 3.1, the disjointness number $K(\mathcal{G})$ is lower-bounded by $\max\{\omega(\mathcal{G}),\ \log_2 |\mathbf{V}(\mathcal{G})| + \frac{1}{2}\log_2 \log_2 |\mathbf{V}(\mathcal{G})| + O(1)\}$. On the other hand, finding the minimum dimension in which a graph admits a faithful orthogonal representation (FOR) is a hard problem in general. Nevertheless, there are results that provide upper bounds on the faithful orthogonal range [238, 248]. In [238], it was shown that $\mathcal{G}$ has an FOR in dimension d if its complement $\bar{\mathcal{G}}$ is $(n-d)$-connected. Hence, all graphs where the parameter $d = \omega(\mathcal{G})$, has a FOR in dimension $\omega(\mathcal{G})$.

3.2.3 Unbounded Quantum Advantage in Relation Reconstruction

Having a sufficient condition for quantum advantage in the reconstruction of $\mathcal{R}(\mathcal{G})$, we now ask how large the separation can be between classical and quantum resources in this task. For this purpose consider the family of graph $\mathcal{G}_\gamma^{(\alpha,\beta)}$, as in Definition 2.25, where $\alpha, \beta, \gamma \in \mathbb{N}$, $\alpha \geq 2, \beta \geq 2$, and $\gamma \in [0, \frac{\beta}{2})$. $\mathcal{G}_\gamma^{(\alpha,\beta)}$ consists of a chain of α maximum cliques, each of size β. The consecutive maximum cliques C_i and C_{i+1} share exactly γ vertices for $i \in \{1, 2, \cdots, \alpha - 1\}$, and the rest of the maximum cliques do not share any additional vertices and edges. When $\gamma = 0$, all the maximum cliques of $\mathcal{G}_\gamma^{(\alpha,\beta)}$ are disconnected. We refer to such graphs as $\mathcal{G}_{disc.}^{(\alpha,\beta)}$.

Theorem 3.4. *There exists a class of graphs $\mathcal{G}$ satisfying **(G0)** and **(G1)** for which the separation between classical and quantum one-way S-CCR for $\mathcal{R}(\mathcal{G})$ is unbounded in both cases: (i) when the parties do not share any resources, and (ii) when Alice additionally does not have private randomness.*

Proof. Consider the family of graphs $\mathcal{G}_\gamma^{(\alpha,\beta)}$ with parameter $\alpha \in \mathbb{N}$ and fixed constants $\beta(\geq 2) \in \mathbb{N}$ and $\gamma \in [0, \frac{\beta}{2})$. Each graph in the family has clique number $\omega(\mathcal{G}_\gamma^{(\alpha,\beta)}) = \beta$ and its order is $|\mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)})| = \alpha(\beta - \gamma) + \gamma$. From Lemma 2.1, the graph $\mathcal{G}_\gamma^{(\alpha,\beta)}$ has FOR in dimension β. Thus, the faithful orthogonal range of $\mathcal{G}_\gamma^{(\alpha,\beta)}$ is $\omega(\mathcal{G}_\gamma^{(\alpha,\beta)}) = \beta$. From Observation 3.1 and

Equation (3.22),

$$K(\mathcal{G}_\gamma^{(\alpha,\beta)}) \geq \max\{\omega(\mathcal{G}_\gamma^{(\alpha,\beta)}), \log_2|\mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)})| + \frac{1}{2}\log_2\log_2|\mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)}| + O(1)\} \tag{3.24}$$

$$\implies K(\mathcal{G}_\gamma^{(\alpha,\beta)}) \geq \max\{\beta, \log_2\{\alpha(\beta-\gamma)+\gamma\} + \frac{1}{2}\log_2(\log_2\{\alpha(\beta-\gamma)+\gamma\}) + O(1)\} \tag{3.25}$$

As $\beta \in \mathbb{N} \setminus \{0,1\}$ and $\gamma \in (0, \beta/2)$ are constants, $\log_2\{\alpha(\beta-\gamma)+\gamma\} \geq \log_2\alpha + \log_2(\beta-\gamma)$. Also, $\log_2(\log_2\{\alpha(\beta-\gamma)+\gamma\}) \geq \log_2\log_2\alpha$. Substituting them in Equation (3.25),

$$K(\mathcal{G}_\gamma^{(\alpha,\beta)}) \geq \max\{\beta, \log_2\alpha + \log_2(\beta-\gamma) + \frac{1}{2}\log_2\log_2\alpha + O(1)\} \tag{3.26}$$

From Theorem 3.2, the classical S-CCR for $\mathcal{R}(\mathcal{G}_\gamma^{(\alpha,\beta)})$ is $\log_2|\mathbf{V}(\mathcal{G}_\gamma^{(\alpha,\beta)})| = \log_2[\alpha(\beta-\gamma)+\gamma]$ bit when Alice has no private randomness and the parties share no resources. This quantity grows with α. More generally, without shared resources, the classical S-CCR for $\mathcal{R}(\mathcal{G}_\gamma^{(\alpha,\beta)})$ is at least $\log_2 K(\mathcal{G}_\gamma^{(\alpha,\beta)})$ bit, which also increases with α (see Equation (3.26)). From Lemma 3.1, the quantum S-CCR for $\mathcal{R}(\mathcal{G}_\gamma^{(\alpha,\beta)})$ is $\log_2\beta$ qubit, which is independent of α. Hence, the gap between classical and quantum S-CCR grows unboundedly with α, establishing unbounded separation. ■

In the above proof, the input set size is $\log_2(\alpha^2\beta)$, where β is fixed. The separation between classical and quantum communication is at least

$$\max\{\log_2\beta, \log_2[\log_2\alpha + \log_2(\beta-\gamma) + \frac{1}{2}\log_2\log_2\alpha] + O(1)\} - \log_2\beta \tag{3.27}$$

when no shared resources are available. If Alice also lacks private randomness, then the separation is $\log_2[\alpha(\beta-\gamma)+\gamma] - \log_2\beta$.

We now consider the payoff for relation reconstruction as defined in Equation (3.11). Let $\mathcal{G}$ be a graph with clique number and faithful orthogonal range both equal to $\omega(\mathcal{G})$, and satisfying conditions **(G0)** and **(G1)**. For $\mathcal{R}(\mathcal{G})$, the maximum payoff using $\log_2\omega(\mathcal{G})$ qubits depends on the choice of an optimal FOR in dimension $\omega(\mathcal{G})$. Given a quantum strategy as in Lemma 3.1, based on a FOR $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{C}^{\omega(\mathcal{G})}$, the payoff is determined by the minimum overlap between vectors assigned to non-adjacent vertices of $\mathcal{G}$. Let v be the vertex in maximum clique $\mathrm{C}_i \in \mathcal{C}$ corresponding to Alice's input $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) = (\mathrm{C}_i, k)$, *i.e.*, clique label k corresponds to the binary colouring that assigns 1 to vertex v in C_i. Similarly, for Bob's input $\mathsf{m_b} = \mathrm{C}_j$ and output $\mathsf{n} = l$, the clique label $l \in \mathcal{O}$ for C_j assigns colour 1 to the vertex v'. Then

$$\mathcal{S}_{\mathcal{R}(\mathcal{G})} = \min_{(\mathsf{m_{a(1)}},\mathsf{m_{a(2)}},\mathsf{m_b},\mathsf{n})\in\mathcal{R}(\mathcal{G})} p(\mathsf{n}|\mathsf{m_{a(1)}},\mathsf{m_{a(2)}},\mathsf{m_b})$$

$$= \min_{(\mathrm{C}_i,k,\mathrm{C}_i,l)\in\mathcal{R}(\mathcal{G})} \mathrm{Tr}\left[|\phi(v)\rangle\langle\phi(v)|\phi(v')\rangle\langle\phi(v')|\right] \tag{3.28}$$

$$= \min_{(v,v')\notin\mathbf{E}(\mathcal{G})} |\langle\phi(v)|\phi(v')\rangle|^2 \tag{3.29}$$

For $\mathcal{R}(\mathcal{G})$, the maximum payoff with $\log_2 \omega(\mathcal{G})$ qubits is achieved by choosing a faithful orthogonal representation $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{C}^{\omega(\mathcal{G})}$ that maximises the minimum overlap between vectors corresponding to non-adjacent vertices of $\mathcal{G}$. The optimisation problem for the maximum payoff, denoted $\mathcal{S}_{\mathcal{R}(\mathcal{G})}^{\mathbb{C}^{\omega(\mathcal{G})},max}$, can be stated as follows:

$$\mathcal{S}_{\mathcal{R}(\mathcal{G})}^{\mathbf{C}^{\omega(\mathcal{G})},max} = \max_{\phi:\mathbf{V}(\mathcal{G})\to\mathbb{C}^{\omega(\mathcal{G})}} \left\{ \min_{(v,v')\notin\mathbf{E}(\mathcal{G})} |\langle\phi(v)|\phi(v')\rangle|^2 \right\} \tag{3.30}$$

In general, finding faithful orthogonal representations for a graph is a hard problem. This optimisation depends on finding FOR that maximises the minimum overlap between two non-adjacent vertices and is non-trivial to solve for a general graph.

3.3 Relation Reconstruction with Shared Resources

So far, we have considered relation reconstruction using classical or quantum communication without shared resources. We now turn to the PM scenario described in Section 3.1, where the parties may use shared resources or public coins. These resources may be classical (shared randomness) or quantum (shared entanglement). In most communication problems with public coins, the focus is on the minimum communication required to perform a task, while allowing unrestricted access to the shared resource. Here, however, we treat both communication and public coins as costly resources. In particular, we consider the PM scenario where the amount of communication (classical or quantum) is limited, and investigate how much public coin (shared randomness or entanglement) is required to accomplish relation reconstruction.

3.3.1 Shared Randomness Assisted Classical Communication

We first discuss the amount of shared randomness, or classical public-coin assistance, required for relation reconstruction with classical communication. For $\mathcal{R}(\mathcal{G})$, when the amount of classical communication is limited to $\log_2 \omega(\mathcal{G})$ bits and $K(\mathcal{G}) > \omega(\mathcal{G})$, we give a lower bound on the shared randomness necessary for relation reconstruction. The necessary shared randomness increases as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bits in the number of maximum cliques α in the graph. We consider the relation $\mathcal{R}(\mathcal{G})$, which is defined for some specific graphs in this section. Let graph $\mathcal{G}$ has vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$, and there are $\alpha(\geq 2)$ maximum cliques $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$ in the graph. The set of clique labels is $\mathcal{O} = \{0, 1, \cdots, \omega(\mathcal{G}) - 1\}$. The clique number, chromatic number and faithful orthogonal range are equal to $\omega(\mathcal{G})$. The graph satisfies conditions **(G0)** and **(G1)**. And $K(\mathcal{G}) > \omega(\mathcal{G})$ for the graph. From Theorem 3.1, the zero-error classical communication complexity of $\mathcal{R}(\mathcal{G})$ without shared randomness is $\log_2 \omega(\mathcal{G})$ bit. Equivalently, $\log_2 \omega(\mathcal{G})$ bit is minimum communication required for satisfying **(T0)**. From Theorem 3.2, the zero-error classical strong communication complexity without shared randomness is at least $\log_2 K(\mathcal{G})$ bit. In other words, it is the minimum communication

$(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})/(\mathsf{m}_{\mathsf{b}}, \mathsf{n})$	$(C_1, 0)$	$(C_1, 1)$	$(C_1, 2)$	$(C_2, 0)$	$(C_2, 1)$	$(C_2, 2)$
$(C_1, 0)$	1	0	0	0	$*$	$*$
$(C_1, 1)$	0	1	0	0	$*$	$*$
$(C_1, 2)$	0	0	1	1	0	0
$(C_2, 0)$	0	0	1	1	0	0
$(C_2, 1)$	$*$	$*$	0	0	1	0
$(C_2, 2)$	$*$	$*$	0	0	0	1

Table 3.3: The table $\mathbf{W}$ corresponding to a protocol that performs the relation reconstruction of $\mathcal{R}(\mathcal{G})$ based on the graph $\mathcal{G}$ is shown in Figure 3.3. The entries marked $*$ have values in $(0, 1)$. For the protocol that achieves algebraic maximum for the payoff, all entries marked $*$ have value 0.5.

required to simultaneously satisfy **(T0)** and **(T1)**. We now show that with $\log_2 \omega(\mathcal{G})$ bits and arbitrary shared randomness, both **(T0)** and **(T1)** can be satisfied (see Observation 3.2). This resource can also achieve the algebraic maximum for the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$. We then calculate the minimum amount of shared randomness assistance necessary to simultaneously satisfy **(T0)** and **(T1)**, or equivalently, accomplish relation reconstruction.

Observation 3.2. *For $\mathcal{R}(\mathcal{G})$, a strategy with $\log_2 \omega(\mathcal{G})$ bit can satisfy **(T0)**. Alice and Bob may adopt a deterministic strategy, described by an $\alpha\omega(\mathcal{G}) \times \alpha\omega(\mathcal{G})$ table $\mathbf{W}$ of conditional probabilities $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$. With shared randomness, they can pre-select a set of all deterministic strategies (or equivalently, tables) that satisfy the consistency condition or **(T0)**. Using the shared randomness, they can switch between these strategies in a correlated way across different rounds. Over multiple rounds, they can satisfy **(T1)**. Quite trivially, the maximum amount of shared randomness required is the logarithm of the number of all such deterministic strategies that individually satisfy the consistent colouring of pairwise clique conditions or equivalently **(T0)**.*

As an example, consider the graph $\mathcal{G}$ in Figure 3.3. Any strategy that satisfies **(T0)** and **(T1)** has the form shown in Table 3.3, where the entries marked by $*$ are strictly in $(0, 1)$ subject to normalisation. A table for a classical deterministic strategy is given in Table 3.2, and another in Table 3.4. With 1 bit shared randomness, Alice and Bob can switch between these two strategies and satisfy **(T0)** and **(T1)** simultaneously. If they use both the strategies with equal probability, then they also achieve the payoff $S^*_{\mathcal{R}(\mathcal{G})} = 0.5$. Now we provide a lower bound on the amount of shared randomness assistance required for reconstruction of relation $\mathcal{R}(\mathcal{G})$ when $\log_2 \omega(\mathcal{G})$ bit of communication is allowed.

Theorem 3.5. *Let graph $\mathcal{G}$ has α maximum cliques with clique number $\omega(\mathcal{G}) = \beta'$ and $K(\mathcal{G}) > \beta'$. Suppose $\mathcal{G}$ has faithful orthogonal range $\omega(\mathcal{G}) = \beta'$ and satisfies conditions **(G0)** and **(G1)**. Then the necessary (though not always sufficient) amount of shared randomness required to achieve relation reconstruction of $\mathcal{R}(\mathcal{G})$ with $\log_2 \omega(\mathcal{G})$ bit of communication is equal*

$(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})/(\mathsf{m}_{\mathsf{b}}, \mathsf{n})$	$(C_1, 0)$	$(C_1, 1)$	$(C_1, 2)$	$(C_2, 0)$	$(C_2, 1)$	$(C_2, 2)$
$(C_1, 0)$	1	0	0	0	1	0
$(C_1, 1)$	0	1	0	0	0	1
$(C_1, 2)$	0	0	1	1	0	0
$(C_2, 0)$	0	0	1	1	0	0
$(C_2, 1)$	1	0	0	0	1	0
$(C_2, 2)$	0	1	0	0	0	1

Table 3.4: Another example of the table $\mathbf{W}$ corresponding to a classical deterministic protocol that performs the distributed computation of $\mathcal{R}(\mathcal{G})$ based on the graph $\mathcal{G}$ is shown in Figure 3.3.

to the minimum shared randomness required for relation reconstruction of $\mathcal{R}(\mathcal{G}_{\text{disc.}}^{(\alpha,\beta=2)})$ *with* 1 *bit of communication.*

Proof. For $\mathcal{R}(\mathcal{G})$, the maximum amount of shared randomness required for relation reconstruction with $\log_2 \omega(\mathcal{G})$ bit of communication is given by the number of deterministic strategies, or equivalently the number of tables $\mathbf{W}$, that Alice and Bob can use while satisfying the constraints from the proof of Theorem 3.1, especially (i),(ii) and (iii). We observe the following based on these constraints on the table corresponding to a deterministic protocol which uses $\log_2 \omega(\mathcal{G})$ bit. Among all graphs $\mathcal{G}$ with the same clique number and number of maximum cliques, the reconstruction of relation $\mathcal{R}(\mathcal{G})$ based on a graph in which all maximum cliques are disconnected from each other requires the largest amount of shared randomness. This is because the corresponding $\mathbf{W}$ satisfying both **(T0)** and **(T1)** has the maximum number of entries marked $*$, which must be assigned values in $(0, 1)$. Conversely, when each maximum clique shares the maximum number of vertices with all other maximum cliques, the required amount of shared randomness is minimal for the reconstruction of relation $\mathcal{R}(\mathcal{G})$ based on this graph. This is because the number of entries marked $*$, which take values from $(0, 1)$, in $\mathbf{W}$ that satisfies both **(T0)** and **(T1)** is the least. All maximum cliques in such a graph can share $\beta' - 2$ vertices in common (from Observation 2.1). If the graph has both clique number and faithful orthogonal representation equal to β', then by Observation 2.1, any two maximum cliques can share up to $\beta' - 2$ vertices. See Figure 3.4 for an example of such a graph.

Thus, a lower bound on the amount of shared randomness required for the reconstruction of relation $\mathcal{R}(\mathcal{G})$ based on any graph $\mathcal{G}$ with α maximum clique and clique number β' can be obtained by considering the graph where all the maximum cliques share $\beta' - 2$ vertices in common. For this graph $\mathcal{G}$, $|\mathbf{V}(\mathcal{G})| = \omega(\mathcal{G}) + 2(\alpha - 1) = \beta' - 2(\alpha - 1)$. For the reconstruction of relation $\mathcal{R}(\mathcal{G})$ defined on this graph $\mathcal{G}$, let $\mathbf{W}$ represent the form of table satisfying both **(T0)** and **(T1)**. The number and structure of the free entries $*$ in $\mathbf{W}$ is identical to that in a table $\mathbf{W}'$ that satisfies both **(T0)** and **(T1)** for reconstruction of relation $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$. Hence, the set of classical deterministic strategies, and therefore the shared randomness required for

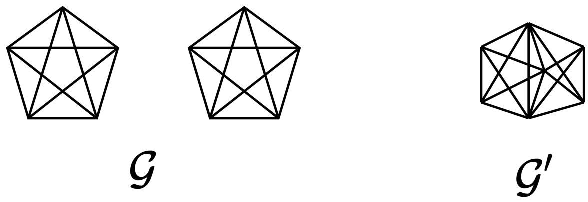

Figure 3.4: Graph $\mathcal{G}$ and $\mathcal{G}'$ both have two maximum cliques and clique number 5. The graph $\mathcal{G}$ has two maximum cliques that are disconnected. The two maximum cliques in $\mathcal{G}'$ share 3 vertices.

$(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})/(\mathsf{m}_{\mathsf{b}}, \mathsf{n})$	$(C_1, 0)$	$(C_1, 1)$	$(C_2, 0)$	$(C_2, 1)$
$(C_1, 0)$	1	0	$*$	$*$
$(C_1, 1)$	0	1	$*$	$*$
$(C_2, 0)$	$*$	$*$	1	0
$(C_2, 1)$	$*$	$*$	0	1

Table 3.5: The table **W** corresponding to a protocol that performs the relation reconstruction of $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$ is shown in Figure 3.5. The entries marked $*$ have values in $(0,1)$. For the protocol that achieves algebraic maximum for the payoff, all entries marked $*$ have value 0.5.

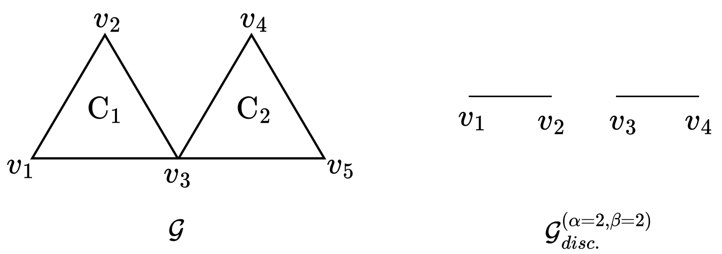

Figure 3.5: Graph $\mathcal{G}$ with clique number $\omega(\mathcal{G}) = 3$ contains two maximum cliques that share $\omega(\mathcal{G}) - 2 = 1$ vertices. For $\mathcal{R}(\mathcal{G})$, the conditional probability table **W** satisfying both **(T0)** and **(T1)** is equivalent, with respect to the free entries marked $*$, to the corresponding table for $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$.

reconstruction, is the same for both cases. For example, in the graphs shown in Figure 3.5, the tables corresponding to relation reconstruction, *i.e.* satisfying both **(T0)** and **(T1)**, are shown in Table 3.3 and Table 3.5. These two tables are equivalent with respect to the entries marked by $*$, which take values in $(0,1)$. Thus, a lower bound on shared randomness for reconstruction of the relation $\mathcal{R}(\mathcal{G})$ while using $\log_2 \beta'$ bit communication can be obtained by considering the amount of shared randomness needed for reconstruction of relation $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$. ■

We can now provide the lower bound on the amount of shared randomness assistance required for reconstruction of relation $\mathcal{R}(\mathcal{G})$ when $\log_2 \omega(\mathcal{G})$ bit of communication is allowed, expressed in terms of the number of maximum cliques in the graph.

Corollary 3.1. *Let $\mathcal{G}$ be a graph with α maximum cliques, clique number $\omega(\mathcal{G}) = \beta'$, faithful orthogonal range β', and $K(\mathcal{G}) > \omega(\mathcal{G})$, satisfying conditions* ***(G0)*** *and* ***(G1)****. Then at least* $\log_2(\lceil \log_2 \alpha \rceil + 1)$ *bit*[8] *of shared randomness are necessary (though not always sufficient) to accomplish relation reconstruction of $\mathcal{R}(\mathcal{G})$ with $\log_2 \beta'$ bit of communication.*

Proof. To lower bound the shared randomness required for reconstructing $\mathcal{R}(\mathcal{G})$ with $\log_2 \beta'$ bits of communication, we use Theorem 3.5. Specifically, we calculate the amount of shared randomness needed for reconstruction of relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ with 1 bit communication.

The reconstruction of relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ using one bit of communication requires a convex mixture of deterministic protocols, each individually using one bit. The final conditional probability table $\tilde{\mathbf{W}}$ must have only positive entries in the off-diagonal blocks $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$, and each diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_i,ALL)}$ must be equal to I_2, where $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$, in order to satisfy **(T0)** and **(T1)**. Thus, for every off-diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$, there must exist deterministic strategies with the following property: in one table $p(l|\mathrm{C}_i, k, \mathrm{C}_j) = 1$ and in another $p(l|\mathrm{C}_i, k, \mathrm{C}_j) = 0$, where $k, l \in \{0, 1\}$. Now, consider the table $\mathbf{W}$ corresponding to a deterministic protocol with 1 bit of communication that satisfies **(T0)**. For each such strategy, the off-diagonal block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$ is either I_2 or the matrix σ_1 (Pauli-x). For such $\mathbf{W}$, $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)} = \mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_i,ALL)} \oplus_2 \mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_j,ALL)}$ where we use the map $I_2 \to 0$ and $\sigma_1 \to 1$. Since communication is limited to one bit, the constraints in the proof of Theorem 3.1 (*i.e.* conditions (i), (ii), (iii)) imply that the entire table $\mathbf{W}$ is fixed once the blocks in its first row are chosen. To ensure that every off-diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_1,ALL),(\mathrm{C}_{j\neq 1},ALL)}$ in the final table has positive entries, the convex mixture must include at least one deterministic table with $W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{j\neq 1},ALL)} = I_2$ and another deterministic table with $W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{j\neq 1},ALL)} = \sigma_1$. Similar requirement for any off-diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_{i\neq 1},ALL),(\mathrm{C}_{j\neq i},ALL)}$ implies that, the convex mixture must include at least one deterministic table with $\mathbf{W}_{(\mathrm{C}_{i\neq 1},ALL),(\mathrm{C}_{j\neq i},ALL)} = W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{i\neq 1},ALL)} \oplus_2 W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{j\neq i},ALL)} = I_2$ and another deterministic table such that $\mathbf{W}_{(\mathrm{C}_{i\neq 1},ALL),(\mathrm{C}_{j\neq i},ALL)} = W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{i\neq 1},ALL)} \oplus_2 W_{(\mathrm{C}_1,ALL),(\mathrm{C}_{j\neq i},ALL)} = \sigma_1$.

Mapping $I_2 \mapsto 0$ and $\sigma_1 \mapsto 1$, any deterministic table $\mathbf{W}$ corresponds to a binary string of length α: $\{W_{(\mathrm{C}_1,ALL),(\mathrm{C}_1,ALL)}, W_{(\mathrm{C}_1,ALL),(\mathrm{C}_2,ALL)}, \cdots, W_{(\mathrm{C}_1,ALL),(\mathrm{C}_\alpha,ALL)}\}$ where the first entry is always 0. To achieve relation reconstruction with minimum shared randomness, we seek a binary matrix $\mathbf{N}$ with α columns and as few rows as possible, where each row represents a deterministic table $\mathbf{W}$. The matrix must satisfy the following conditions: every pair $j, j' \in [\alpha]$, there exist rows i, i' such that $\mathbf{N}_{i,j} \oplus_2 \mathbf{N}_{i,j'} = 1$ and $\mathbf{N}_{i',j} \oplus_2 \mathbf{N}_{i',j'} = 0$. If there are r rows in

[8] In [184], there was an overstatement in the scaling of the amount of shared randomness in terms of the number of the maximum clique which has been corrected here.

the matrix, then we can consider the columns as α number of bit strings each of length r. The first column is $(0, 0, \cdots, 0)^T$. Since each string $z \in \{0,1\}^r$ has a complement $\bar{z}$, there are 2^{r-1} disjoint pairs. We can select at most one string from each pair as a column of $\mathbf{N}$. Hence, at most 2^{r-1} distinct assignments as columns are possible. Consequently, $\alpha \leq 2^{r-1} \implies r \geq \lceil \log_2 \alpha \rceil + 1$. Therefore, at least $\lceil \log_2 \alpha \rceil + 1$ deterministic tables are required, and the amount of shared randomness necessary is $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit. ■

We now discuss the amount of shared randomness required for reconstructing relation $\mathcal{R}(\mathcal{G})$ and achieving the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})}$ when $\log_2 \omega(\mathcal{G}^{(\alpha,\beta=2)}_{disc.}) = 1$ bit of communication is allowed. This lower bound is linked to the existence of certain orthogonal arrays. Let us first briefly discuss orthogonal arrays (OA).

Definition 3.5. Let $\mathbf{A}$ be an $N \times k$ array with entries from a set S and $s = |S|$. $\mathbf{A}$ is an *orthogonal array* with s levels, strength $t \in \{0, 1, \ldots, k\}$ and index $L \in \mathbb{N}$ if every $N \times t$ sub-array of $\mathbf{A}$ contains each possible t-tuple formed using elements of S exactly L times. It is denoted by $OA(N, k, s, t)$.

The notion of orthogonal arrays has been widely studied in quantum information theory, particularly in entanglement theory [249–251] and quantum error correction [252]. We focus on $OA(N, k, s, t)$ with $s = 2$, $S = \{0, 1\}$, and strength $t = 2$. Let N^*_k denote the minimum N for which an orthogonal array $OA(N^*_k, k, 2, 2)$ exists. In such an array, every $N^*_k \times 2$ sub-array contains each of the tuples $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$ appearing with the same frequency. N^*_k is connected to the necessary amount of shared randomness required for reconstructing relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ and achieving the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})}$ with $\log_2 \omega(\mathcal{G}^{(\alpha,\beta=2)}_{disc.}) = 1$ bit of communication.

Corollary 3.2. *For $\mathcal{G}^{(\alpha,\beta=2)}_{disc.}$ at least one bit of shared randomness is necessary when $\alpha = 2$, and at least $\log_2 N^*_k$ bit of shared randomness are necessary when $\alpha > 2$, to achieve relation reconstruction of $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ and obtain the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})} = 0.5$ with one bit of communication.*

Proof. The reconstruction of relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ with one bit of communication requires a convex mixture of deterministic protocols, each using one bit. To satisfy **(T0)** and **(T1)** and obtain payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})} = 0.5$, the final table of conditional probabilities $\tilde{\mathbf{W}}$ should be of the following form. All entries in the off-diagonal blocks $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$ must be 0.5, and each diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_i,ALL)}$ equals I_2, where $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$. This is possible if, for every off-diagonal block $\tilde{\mathbf{W}}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$, we have uniform mixture of deterministic tables $\mathbf{W}$ where half of them have $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)} = I_2$ and rest have $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)} = \sigma_1$. Each entry in the block then has a weight 0.5. Now, consider the table $\mathbf{W}$ corresponding to a deterministic protocol with 1 bit of communication. For such $\mathbf{W}$, the off-diagonal block $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_{j\neq i},ALL)}$ is either I_2 or the matrix σ_1. Also, $\mathbf{W}_{(\mathrm{C}_i,ALL),(\mathrm{C}_j,ALL)} = \mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_i,ALL)} \oplus_2 \mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_j,ALL)}$ where $I_2 \to 0$ and $\sigma_1 \to 1$. Since communication is restricted to 1 bit, using the constraints established in the, the constraints in the proof of Theorem 3.1 (*i.e.* conditions (i), (ii), (iii)), the full table $\mathbf{W}$ is fixed once the blocks in its first row are assigned.

For $\alpha = 2$, the uniform convex mixture of two deterministic tables $\mathbf{W}$, one with $\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_2,ALL)} = I_2$ and another with $\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_2,ALL)} = \sigma_1$, achieves payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha=2,\beta=2)}_{disc.})} = 0.5$. For $\alpha = 3$, the uniform convex mixture of four deterministic table $\mathbf{W}$, each assigning $\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_2,ALL)} = I_2$ or σ_1 and $\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_3,ALL)} = I_2$ or σ_1, lead to payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha=3,\beta=2)}_{disc.})} = 0.5$. For $\alpha \geq 2$, we need minimal collection of deterministic tables $\mathbf{W}$ such that for every pair $\mathrm{C}_{j\neq 1}, \mathrm{C}_{j'\neq 1} \in \mathcal{C}$, we have the following. There are equal number of tables $\mathbf{W}$ where $(\mathbf{W}_{(\mathrm{C}_1,ALL),(\mathrm{C}_j,ALL)}, \mathbf{W}_{(\mathrm{C}_1,ALL)(\mathrm{C}_{j'},ALL)}) \in \{(I_2, I_2), (I_2, \sigma_1), (\sigma_1, I_2), (\sigma_1, \sigma_1)\}$. Mapping $I_2 \to 0$ and $\sigma_1 \to 1$, this reduces to the orthogonal array problem discussed earlier. Specifically, we want an array with $\alpha - 1$ columns and the minimum number of rows $N^*_{\alpha-1}$ such that each pair of columns ($t = 2$) contains all four tuples $\{(0,0), (0,1), (1,0), (1,1)\}$ in their rows occurring with equal frequency. Thus for $\alpha > 2$, the reconstruction of relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ using one bit communication requires $\log_2 N^*_{\alpha-1}$ bit of shared randomness to obtain the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})}$.
■

For graphs $\mathcal{G}$ with faithful orthogonal range $\omega(\mathcal{G})$, where $K(\mathcal{G}) > \omega(\mathcal{G})$ and conditions **(G0)** and **(G1)** hold, assistance from shared resources becomes necessary when communication is limited to $\log_2 \omega(\mathcal{G})$ bits for reconstructing relation $\mathcal{R}(\mathcal{G})$. Interestingly, there are graphs where bounded quantum communication also requires shared randomness to reach the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$. In such cases, even with one-way $\log_2 \omega(\mathcal{G})$ bit or qubit, shared resources are needed to accomplish relation reconstruction. In fact, for some graphs with $\omega(\mathcal{G}) = 2$, one bit of communication assisted by shared randomness achieves a higher payoff than any protocol restricted to only single-qubit communication without any shared resource.

Theorem 3.6. *There exist graphs $\mathcal{G}$ with clique number and faithful orthogonal range $\omega(\mathcal{G})$, satisfying conditions **(G0)** and **(G1)**, for which shared randomness is necessary to achieve relation reconstruction of $\mathcal{R}(\mathcal{G})$ and attain the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$ when $\log_2 \omega(\mathcal{G})$ bits or qubit of communication are allowed.*

Proof. We will prove this by considering the graph $\mathcal{G}^{(\alpha,\beta)}_{disc.}$ with $\alpha = \beta + 2$. The graph has clique number β and faithful orthogonal range β (using Lemma 2.1). For $\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})$, the algebraic maximum payoff is $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})} = \frac{1}{\beta}$. When the allowed communication is $\log_2 \beta$ qubits without shared randomness, every protocol as in proof of Lemma 3.1 yields a payoff $\mathcal{S}_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})} < \frac{1}{\beta}$. The quantum protocol in this case relies on some FOR for encoding and decoding. However, in $\mathbb{C}^\beta$ there are at most $\beta + 1$ mutually unbiased bases (MUBs), which are insufficient to implement unbiased encoding and decoding for $\alpha = \beta + 2$. Hence the assignment $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}) = \frac{1}{\beta}$ wherever $\mathsf{m}_{\mathsf{a}(1)} \neq \mathsf{m}_{\mathsf{b}}$ and $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \in \mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta)})$ cannot be achieved with $\log_2 \beta$ qubit communication. Moreover, $\log_2 \beta$ bit communication without shared randomness also fails to accomplish relation reconstruction (Theorem 3.2). In contrast, if shared randomness is allowed, $\log_2 \beta$ bit communication suffices to achieve relation reconstruction and obtain the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta)})} = \frac{1}{\beta}$. ■

As an example, consider the graph $\mathcal{G}_{disc.}^{(\alpha=4,\beta=2)}$. From Corollary 3.2, 1 bit communication and $\log_2 N_3^*$ bit shared randomness can accomplish reconstruction of the relation $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha=4,\beta=2)})$ and obtain payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha=4,\beta=2)})} = 0.5$. However, without shared randomness 1 qubit communication can accomplish reconstruction of the relation $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha=4,\beta=2)})$. However, the payoff obtained is always less than 0.5 as only 3 MUBs exist in $\mathbb{C}^2$.

3.3.2 Arbitrary Separation Between Entanglement and Shared Randomness

For $\mathcal{R}(\mathcal{G})$, we saw that shared randomness, or classical public coins, can assist classical communication in relation reconstruction when the communication is limited to $\log_2 \omega(\mathcal{G})$ bits. A natural question is how much assistance is required from quantum entanglement, or quantum public coins, for the same task. We show that a fixed amount of entanglement can advantageously assist limited classical communication in a family of tasks that would otherwise need an unbounded amount of shared randomness.

Theorem 3.7. *There exists a family of graphs $\mathcal{G}$ with clique number and faithful orthogonal range $\omega(\mathcal{G}) = \beta$, satisfying conditions **(G0)** and **(G1)**, for which the gap between entanglement assistance and shared randomness required to achieve relation reconstruction of $\mathcal{R}(\mathcal{G})$ with $\log_2 \beta$ bit of communication is unbounded. In particular, for relation reconstruction of $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$ with one bit communication the gap between entanglement and shared randomness assistance scales as $\log_2(\lceil \log_2 \alpha \rceil + 1)$ with the number of maximum cliques α.*

Proof. We will prove this by considering the family of graphs $\mathcal{G}_{disc.}^{(\alpha,\beta=2)}$ where $\alpha \geq 2$. These graphs have clique number $\beta = 2$ and faithful orthogonal range 2 (using Lemma 2.1). For $\mathcal{R}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$, at least $\log_2(\lceil \log_2 \alpha \rceil + 1)$ bit shared randomness is necessary for relation reconstruction when only 1 bit communication is allowed. This follows from Corollary 3.1. In contrast, the relation reconstruction is possible with 1 bit communication when Alice and Bob share maximally entangled state $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$, *i.e.* 1 e-bit entanglement. Each vertex $v \in \mathbf{V}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$ is adjacent to exactly a unique vertex $v' \in \mathbf{V}(\mathcal{G}_{disc.}^{(\alpha,\beta=2)})$. For each such pair of vertices, Alice and Bob assign them distinct orthogonal pairs of states from an equatorial circle of the Bloch sphere. Say $|\psi_v\rangle = c_{(v,v')}|0\rangle + (1 - c_{(v,v')})|1\rangle$ and $|\psi_{v'}\rangle = (1 - c_{(v,v')})|0\rangle - c_{(v,v')}|1\rangle$ where $c_{(v,v')} \in (0, 1)$.

Suppose, Alice is given input $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) = (\mathrm{C}_i, k)$, where $k \in \mathcal{O}$ and $\mathrm{C}_i \in \mathcal{C}$. The clique label k corresponds to the binary colouring that assigns 1 to vertex v of the maximum clique $\mathrm{C}_i \in \mathcal{C}$. Then Alice performs the measurement $\{|\psi_v\rangle\langle\psi_v|, |\psi_{v'}\rangle\langle\psi_{v'}|\}$ on her part of the entangled state $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|\psi_v\psi_{v'}\rangle_{AB} - |\psi_{v'}\psi_v\rangle_{AB})$. If she gets an outcome corresponding to $|\psi_v\rangle\langle\psi_v|$ then she sends 0, otherwise she sends symbol 1. Upon receiving Alice's sent symbol, Bob applies σ_1 to his part of the entangled system only if the received symbol is 1. Next, if Bob receives input $\mathsf{m_b} = \mathrm{C}_j$, then he performs the measurement $\{|\psi(\tilde{v})\rangle\langle\psi(\tilde{v})|, 1 - |\psi(\tilde{v})\rangle\langle\psi(\tilde{v})|\}$, where $\tilde{v} \in \mathbf{V}(\mathrm{C}_j)$, on his part of shared entangled state. On obtaining outcome corresponding to the projector $|\psi(\tilde{v})\rangle\langle\psi(\tilde{v})|$, Bob outputs $\mathsf{n} = l$, where the label $l \in \mathcal{O}$ for C_j assigns colour 1 to the vertex $\tilde{v}$. Otherwise, he outputs the label corresponding to the neighbour of $\tilde{v}$. This quantum strategy always satisfies the consistency condition and accomplishes relation reconstruction. The protocol is essentially remote state preparation on a great circle of the Bloch sphere [253, 254].

When the allowed communication is only 1 bit, the amount of shared randomness necessary increases with the number of maximum cliques α in the graph. On the other hand, 1 e-bit shared entanglement is sufficient for the task. With increasing α, this leads to an unbounded separation between the two shared resources. ■

3.4 Some Applications

We now discuss some applications of the tasks presented so far. The first application is the operational detection of mutually unbiased bases (MUBs) through relation reconstruction. We then discuss the role of the tasks as dimension witnesses and in the semi-device-independent detection of non-classical communication, as well as shared resources.

3.4.1 Operational Detection of MUBs

A set of orthonormal bases in $\mathbb{C}^d$ is called mutually unbiased bases (MUBs) if, for any two bases in the set, the squared absolute value of the inner product between a vector from one basis and a vector from the other is always $\frac{1}{d}$. Equivalently, two projective measurements on a quantum system with Hilbert space $\mathbb{C}^d$ are mutually unbiased if their projectors correspond to two MUBs. MUBs have been widely used in various information processing tasks as well as quantum cryptography [255–260]. We discuss the operational detection of MUBs through input-output statistics observed in a communication task defined in the PM scenario. For some specific graph $\mathcal{G}$, we consider relation reconstruction for $\mathcal{R}(\mathcal{G})$ while using $\log_2 \omega(\mathcal{G})$ qubit of communication and no shared resources. Achieving the payoff $\mathcal{S}^*_{\mathcal{R}(\mathcal{G})}$ in this setting implies the use of MUBs in $\mathbb{C}^{\omega(\mathcal{G})}$.

Observation 3.3. *The graph $\mathcal{G}^{(\alpha,\beta)}_{disc.}$ has both clique number and faithful orthogonal range in $\omega(\mathcal{G}^{(\alpha,\beta)}_{disc.}) = \beta$. Given $\log_2 \beta$ qubit communication without any shared resources, if a quantum* *protocol achieves payoff* $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})} = \frac{1}{\beta}$ *in relation reconstruction for* $\mathcal{R}(\mathcal{G}^{(\alpha,\beta)}_{disc.})$*, then Bob's measurement must correspond to MUBs in* $\mathbb{C}^\beta$*.*

For example, consider relation reconstruction for $\mathcal{R}(\mathcal{G}^{(\alpha=3,\beta=2)}_{disc.})$ when only 1 bit communication from Alice to Bob is allowed and no shared resources are available. Given a maximum clique and its label, Alice chooses the MUB associated with her input clique. She prepares the eigenstate of that MUB corresponding to the vertex assigned binary colour 1 by the clique label and sends it to Bob. Bob then performs the projective measurement corresponding to the MUB of his input clique and outputs the clique label based on the outcome. The achievable payoff in this case is $\mathcal{S}^*_{\mathcal{R}(\mathcal{G}^{(\alpha=3,\beta=2)}_{disc.})} = \frac{1}{2}$. Conversely, to obtain this payoff, every probability $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$ that should tak positive value specifically in $(0, 1)$ for relation reconstruction, must be assigned 0.5. These input-output tuples correspond to non-adjacent vertices of the graph. Hence, the preparations for such non-adjacent vertices must have an overlap of $\frac{1}{2}$.

3.4.2 Dimension Witness and Semi-device-Independent Detection of Non-classical Resources

The information-processing tasks defined in the PM scenario can also be used for identifying some features of the resources employed in the task. In such tasks, one natural question is regarding the operational dimension of the communicated system. Another is whether the resource used is classical or non-classical. For a quantum system, the operational dimension coincides with the dimension of the associated Hilbert space. The problem of dimension witnessing has been widely studied [153, 261–263]. Some approaches make no assumptions about the internal workings of the devices and infer the dimension or non-classicality directly from input-output correlations. These are called device-independent schemes. Others allow for certain reasonable assumptions about the devices, leading to semi-device-independent schemes. We now show how the tasks discussed earlier can serve as semi-device-independent witnesses for both dimension and non-classicality.

Consider a graph $\mathcal{G}$ that satisfies **(G0)** and **(G1)** and has faithful orthogonal range $\omega(\mathcal{G})$. If the communication is strictly less than $\log_2 K(\mathcal{G})$ and no shared resources are allowed, accomplishing relation reconstruction for $\mathcal{R}(\mathcal{G})$ implies that the communicated system must be non-classical (see Theorem 3.2). Next, consider the reconstruction of relation $\mathcal{R}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.})$ when 1 bit of communication and limited shared resources, say $\log_2 d$ bits or e-bit, are available. A positive payoff $\mathcal{SR}(\mathcal{G}^{(\alpha,\beta=2)}_{disc.}) > 0$ implies that the shared resource must be non-classical whenever $\log_2(\lceil\log_2 \alpha\rceil + 1) > d$ (see the proof of Theorem 3.7). For a dimension witness, consider zero-error distributed computation of $\mathcal{R}(\mathcal{G})$, where the graph satisfies the constraints in Section 3.2.1. If some protocol achieves zero-error distributed computation of $\mathcal{R}(\mathcal{G})$, then the communicated system must be at least $\omega(\mathcal{G})$-dimensional.

Summary: We explored distributed computation and relation reconstruction in the PM scenario for relations defined through the Clique Labelling Problem on graphs. We first considered the case without shared resources. While no quantum advantage was found for zero-error distributed computation, we showed an unbounded separation between classical and quantum communication resources in relation reconstruction for certain graph families. This advantage disappears when shared resources are allowed. However, we showed that for relations defined using specific graphs, the amount of shared resources required grows with the number of maximum cliques when classical communication is bounded. We then demonstrated an unbounded separation between classical and quantum shared resources. In particular, for a family of graphs, a constant amount of shared entanglement can assist one bit of classical communication to accomplish relation reconstruction, whereas achieving the same with shared randomness requires an arbitrary amount. Finally, we discussed some applications of these tasks in detecting MUBs and as semi-device-independent witnesses of dimension and non-classicality.

Chapter 4

Detection of Non-Projective-Simulable Measurements

Chapter Note

This chapter is based on work which was done in collaboration with Dr. Some Sankar Bhattacharya and Prof. Dr. hab Paweł Horodecki [185].

The set of all measurements which can be performed on a system is specified by the physical theory. In a classical theory, a measurement on a d-dimensional system can have at most d irreducible outcomes. Measurement with more than d outcomes in this theory is simply a post-processing of some d-outcome extremal measurement performed on the same system. Similarly, there exists POVMs with $N(> d)$ outcomes on a d-dimensional system that can be simulated by using post-processing and a d-outcome projective measurement (PVM) on the d-dimensional system. We refer to such POVMs as *projective-simulable* (PS) measurements (see Definition 2.1). However, quantum theory also allows for measurements on a d-dimensional system with more than d irreducible outcomes. There exist POVMs with $N > d$ outcomes that cannot be realised by post-processing the outcomes of any d-outcome PVM performed on the same system. We call such measurements *non-projective-simulable* (nPS) (see Definition 2.2 and Section 2.1). In this chapter, we will discuss some operational detection schemes for such measurements using information processing tasks defined in a Bell scenario.

General POVMs beyond projective measurements serve as useful resources in many information processing tasks. Their applications include quantum state discrimination [166–169], entanglement detection [170], quantum tomography [171–175], cryptographic protocols [176], port-based teleportation [177–179], quantum metrology [180–182], and randomness certification [183]. They are also resourceful in areas such as quantum computing [264–266] and quantum information theory [151, 267]. The broad range of applications makes the problem of detecting such measurements quite relevant. A trivial approach is to prepare random states and perform tomography [268] of the measurement device. However, this method requires a lot of resources. This motivates the search for more efficient detection schemes.

In the last Chapter 3, we saw that information processing tasks can be used to reveal non-classicality of the resources used during the task. Similar methods have been applied to the detection of general quantum POVMs. In a one-way PM scenario with limited quantum communication, a robust scheme for semi-device-independent detection of three and four-outcome qubit POVMs was proposed in [269]. A similar detection scheme for a seven-outcome ququad non-projective measurement in the PM scenario was presented in [270]. These works require, in addition, some known measurements to characterise the unknown device. Beyond PM scenarios, detection has also been studied in the Bell scenario, where space-like separated parties perform local measurements based on their inputs on a shared quantum state [271, 272]. This approach typically assumes bounds on the local dimension of the shared system. Additionally, the inputs are chosen randomly in both the PM scenario and the Bell scenario. Thus, these approaches require a source of randomness.

Any classical device can only generate pseudo-random strings. A genuine source of randomness, in contrast, requires a non-classical resource and therefore can be regarded as costly. Even a small amount of such randomness can be useful to detect non-classical resources [273, 274]. In Bell scenarios, which are often used to witness non-classicality, this reliance on randomness amounts to using one non-classical resource to certify another. This circularity has motivated alternate approaches that avoid input randomness. For instance, [275] proposed a network scenario with multiple separated parties and independent preparation sources, where the parties perform fixed local measurements. More recently, [276, 277] studied a Bell-type setup without any inputs. Only assuming an upper bound on the local dimension of the shared system, this setting was used to certify the non-classicality of both the preparation and the measurement.

In this chapter, following [276, 277], we adopt a Bell scenario without inputs to provide *randomness-free* detection schemes for non-projective-simulable measurements under the assumption of an upper bound on the local operational dimension of the shared system. Multiple copies of the measurement device are distributed between spatially separated parties. Each party receives a subsystem and performs the given measurement on their subsystem. This setup does not require seed randomness or pre-characterised measurements.

We begin by introducing the operational setup and the information processing task in Section 4.1. In this setting, we show that any correlation generated using a classical system of local operational dimension d can also be achieved by performing a projective-simulable measurement on a shared quantum system of local dimension d, and vice versa. In Section 4.2 we propose detection schemes in a bipartite scenario for three and four-outcome qubit nPS, assuming the local dimension of the shared system is 2. Some of these schemes are robust against arbitrary depolarising noise. In Section 4.3 we consider five-outcome qutrit nPS and discuss their detection in both bipartite and tripartite scenarios, assuming the local dimension of the shared system is 3. We first present a detection scheme while considering the case where both parties perform the identical measurement in a bipartite scenario. We provide numerical evidence for the robustness of this scheme against arbitrary depolarising noise. We then relax this assumption and propose detection schemes for five-outcome qutrit nPS in both bipartite and tripartite settings. For these tasks, we numerically obtain bounds on the payoff achievable with projective-simulable measurements, and we show violations of these bounds using a qutrit shared state and a qutrit POVM. Finally, in Section 4.4 we extend the notion of nPS to GPTs. We then use an operational task to show that the square-bit model or box world theory is unphysical.

4.1 Information Processing Task Without Input Randomness

In a standard Bell scenario, parties receive inputs and return outputs. Such scenarios usually assume freedom of choice, which requires the inputs to be chosen at random. We consider a modification in which the parties perform fixed measurements, thereby eliminating the requirement for seed randomness. In this setting, the parties are still free to apply local operations. We use such scenarios to define information processing tasks, under the additional assumption that the local operational dimension of the shared system is bounded.

4.1.1 The Setup

We consider an operational scenario with n spatially separated parties $A_1, A_2, \ldots, A_n$ sharing an n-partite system prepared by a device $\mathcal{P}_{A_1,A_2,\ldots,A_n}$. Each subsystem has a local operational dimension d (see Figure 4.1). The parties have local sources of randomness but no shared correlations. Each party A_i, where $i \in [n]$, has an uncharacterised measurement device $\mathcal{M}_{A_i}$ that can produce an outcome in $\{0, 1, \ldots, k-1\}$, where $k > d$. After measuring their subsystem, party A_i outputs an outcome $a_i \in \{0, 1, \ldots, k-1\}$. The resulting joint distribution is denoted by $\{P(a_1, a_2, \ldots, a_n)\}_{a_1,a_2,\ldots,a_n}$. Let us denote this setup by $[n, d, k]$. In this chapter, we restrict attention to bipartite and tripartite cases, with parties labelled A, B, and C, and their outcomes a, b, and c, respectively. In some instances, we also assume that the measurement devices are identical for all parties.

The shared system and measurement device could be described by classical, quantum or post-quantum theory. However, we mainly consider shared resources and measurements to be either

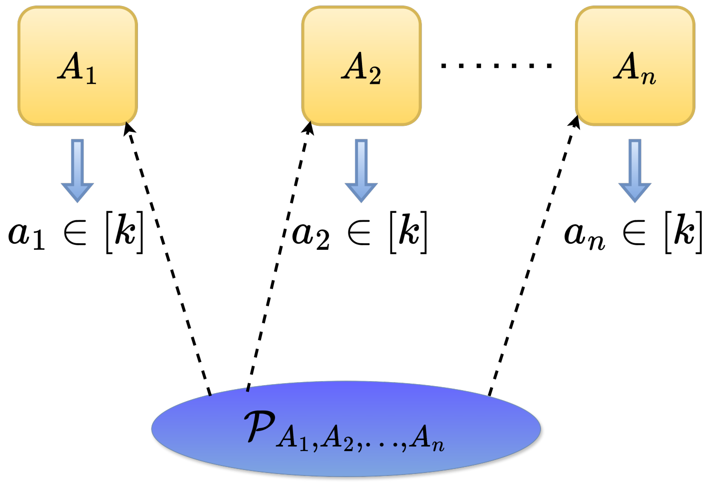

Figure 4.1: A schematic representation of the task $\mathbb{G}[n,d,k]$. The n spatially separated parties $A_1, A_2, \ldots, A_n$ share an n-partite system prepared using $\mathcal{P}_{A_1,A_2,\ldots,A_n}$, where each subsystem has operational dimension d. Each party A_i uses an uncharacterised k-outcome measurement device $\mathcal{M}_{A_i}$ with $k > d$ and outputs $a_i \in \{0, 1, \ldots, k-1\}$. The joint distribution of outcomes is denoted by $\{P(a_1, a_2, \cdots, a_n)\}_{a_1,a_2,\cdots,a_n}$.

classical or quantum. We define

$$\mathfrak{C}(k,n) := \{\{P(a_1, a_2, \cdots, a_n)\}_{a_1,a_2,\cdots,a_n}\} \tag{4.1}$$

as the set of all valid k-outcome joint probability distributions. Let $\mathfrak{C}_d^{Cl}(k,n) \subset \mathfrak{C}(k,n)$ denote the subset achievable using n-partite classical systems with local operational dimension d. Similarly, $\mathfrak{C}_d^{Q}(k,n)$ and $\mathfrak{C}_d^{PQ}(k,n)$ denote the set of $\{P(a_1, a_2, \cdots, a_n)\}_{a_1,a_2,\cdots,a_n}$ achievable using, respectively, k-outcome quantum measurements and k-outcome projective-simulable measurements on n-partite quantum systems with local dimension d. Then $\mathfrak{C}_d^{nPS}(k,n) = \mathfrak{C}_d^{Q}(k,n) \setminus \mathfrak{C}_d^{PQ}(k,n)$ is the set of correlations that can only be generated with k-outcome non-projective-simulable measurements on such systems. We now show that the correlations generated by k-outcome projective-simulable measurements on n-partite quantum systems with local dimension d are the same as those achievable using n-partite classical systems of the same local dimension. Thus, projective-simulable measurements are effectively classical in this scenario $[n,d,k]$. Consequently, in a task defined in setup $[n,d,k]$, any advantage while sharing quantum systems with local dimension d must stem from nPS measurements.

Theorem 4.1. *In the setup $[n,d,k]$ with $k > d$, $\mathfrak{C}_d^{Cl}(k,n) = \mathfrak{C}_d^{PQ}(k,n)$.*

Proof. We present the proof for $n = 2$. The case of general $n > 2$ follows by a straightforward extension of this argument. Consider a correlation $\{P(a,b)\}_{a,b=0}^{k-1} \in \mathfrak{C}_d^{PQ}(k,2)$. Such a correlation can be realised using a shared quantum state $\rho_{AB} \in D(\mathbb{C}^d \otimes \mathbb{C}^d)$ and local projective-simulable measurements $\mathcal{M}_A := \{E_a = \sum_{t=0}^{d-1} p(a|t)|\psi_t\rangle\langle\psi_t|\}_{a=0}^{k-1}$ for A and $\mathcal{M}_B := \{F_b = \sum_{t'=0}^{d-1} p'(b|t')|\phi_{t'}\rangle\langle\phi_{t'}|\}_{b=0}^{k-1}$ for B. Here, $\{p(a|t)\}_{a=0}^{k-1}$ and $\{p'(b|t')\}_{b=0}^{k-1}$ are valid conditional probability distributions for all $t, t' \in \{0, 1, \cdots, d-1\}$. The sets $\{|\psi_t\rangle\}_{t=0}^{d-1}$ and $\{|\phi_{t'}\rangle\}_{t'=0}^{d-1}$ form orthonormal bases of $\mathbb{C}^d$. We now show that $\{P(a,b)\}_{a,b=0}^{k-1}$ can also be obtained by sharing a classical state $\omega_{AB}^{Cl} = (p_{0,0}, p_{0,1}, \ldots, p_{d-1,d-1})$, where $\sum_{i=0}^{d-1}\sum_{j=0}^{d-1} p_{i,j} = 1$, with the local dimension d, and performing measurement $W_A^{d\to k}$ by A and $W_B^{d\to k}$ by B. In classical theory, the probability rule is given by $(W_A^{d\to k} \otimes W_B^{d\to k})(\omega_{AB}^{Cl})^T$ (see Section 2.2.1).

The probability $P(a,b)$ in quantum theory is given by

$$P(a,b) = \mathrm{Tr}\Big[\rho_{AB}\; E_a \otimes F_b\Big] = \mathrm{Tr}\Big[\rho_{AB} \sum_t p(a|t)|\psi_t\rangle\langle\psi_t| \otimes \sum_{t'} p'(b|t')|\phi_{t'}\rangle\langle\phi_{t'}|\Big] \tag{4.2}$$

$$= \sum_{t,t'} p(a|t)p'(b|t')\,\mathrm{Tr}\Big[\rho_{AB}\; |\psi_t\rangle\langle\psi_t| \otimes |\phi_{t'}\rangle\langle\phi_{t'}|\Big] \tag{4.3}$$

We write $\rho_{AB} = \sum_{ijkl=0}^{d-1} \alpha_{ijkl}\; (|\psi_i\phi_j\rangle\langle\psi_k\phi_l|)$ and substitute it in the Equation (4.3).

$$P(a,b) = \sum_{t,t'} p(a|t)p'(b|t')\,\mathrm{Tr}\Big[\rho_{AB}|\psi_t\phi_{t'}\rangle\langle\psi_t\phi_{t'}|\Big]$$

$$= \sum_{t,t'} p(a|t)p'(b|t') \sum_{ijkl=0}^{d-1} \alpha_{ijkl}\;\; \mathrm{Tr}\Big[(|\psi_i\phi_j\rangle\langle\psi_k\phi_l|)|\psi_t\phi_{t'}\rangle\langle\psi_t\phi_{t'}|\Big] \tag{4.4}$$

$$= \sum_{t,t'} p(a|t)p'(b|t') \sum_{ijkl=0}^{d-1} \alpha_{ijkl}\;\; \delta_{t,k}\delta_{l,t'}\delta_{i,t}\delta_{j,t'} \tag{4.5}$$

$$= \sum_{t,t'} p(a|t)p'(b|t')\alpha_{tt'tt'} \tag{4.6}$$

$$= \sum_{t,t'} p(a|t)p'(b|t') \sum_{ij=0}^{d-1} \alpha_{ijij}\;\; \mathrm{Tr}\Big[|\psi_i\phi_j\rangle\langle\psi_i\phi_j| \cdot |\psi_t\phi_{t'}\rangle\langle\psi_t\phi_{t'}|\Big] \tag{4.7}$$

$$= \sum_{t,t'} p(a|t)p'(b|t') \sum_{ij=0}^{d-1} \alpha_{ijij}\;\; \mathrm{Tr}\Big[|ij\rangle\langle ij| \cdot |tt'\rangle\langle tt'|\Big] \tag{4.8}$$

$$= \sum_{t,t'} p(a|t)p'(b|t')\,\mathrm{Tr}\Big[\tilde{\rho}_{AB}\; |mn\rangle\langle mn|\Big] \quad \Big(\tilde{\rho}_{AB} = \sum_{ij=0}^{d-1} \alpha_{ijij}\; |ij\rangle\langle ij|\Big) \tag{4.9}$$

$$= \sum_{t,t'} (W_A^{d\to k})_{a,t}(W_B^{d\to k})_{b,t'}\;\; p_{t,t'} \tag{4.10}$$

In Equation (4.3), we used the classical state $\omega_{AB}^{Cl} = (p_{0,0}, p_{0,1}, \ldots, p_{d-1,d-1})$ where $p_{i,j} = \alpha_{ijij}$ for $i, j \in \{0, 1, \cdots, d-1\}$. The measurement for A is a local stochastic map $W_A^{d\to k}$ where $(W_A^{d\to k})_{i,j} = p(a = i|t = j)$ for $i \in \{0, \cdots, k-1\}$ and $j \in \{0, \cdots, d-1\}$. Similarly, the measurement for B is $W_B^{d\to k}$ with entries $(W_B^{d\to k})_{i,j} = p'(b = i|t' = j)$ for $i \in \{0, \cdots, k-1\}$ and $j \in \{0, \cdots, d-1\}$. The probability $P(a, b)$ is also the probability of outcomes a and b when $W_A^{d\to k}$ and $W_B^{d\to k}$ are applied to ω_{AB}^{Cl}. Hence, $\mathfrak{C}_d^{PQ}(k, 2) \subseteq \mathfrak{C}_d^{Cl}(k, 2)$.

The converse $\mathfrak{C}_d^{Cl}(k, 2) \subseteq \mathfrak{C}_d^{PQ}(k, 2)$ is easy to show. Consider a correlation $\{P(a, b)\}_{a,b=0}^{k-1} \in \mathfrak{C}_d^{Cl}(k, 2)$ is obtained from a classical state $\omega_{AB}^{Cl} = (p_{0,0}, p_{0,1}, \ldots, p_{d-1,d-1})$ with local operational dimension d and local measurements $W_A^{d\to k}$ and $W_B^{d\to k}$ by A and B. Consider the diagonal quantum state $\tilde{\rho}_{AB} = \sum_{ij=0}^{d-1} p_{i,j}\ |ij\rangle\langle ij|$ in the computational basis, along with projective measurements in that basis followed by local post-processing via $W_A^{d\to k}$ for A and $W_B^{d\to k}$ for B. This reproduces the same correlation $\{P(a, b)\}_{a,b=0}^{k-1}$. The argument extends to arbitrary n by assuming local projective-simulable measurements for each party acting on a shared multipartite state. The proof proceeds exactly as discussed above. ■

From Theorem 4.1, the correlations generated by local projective-simulable measurements on an n-partite quantum system of local dimension d are identical to those achievable with an n-partite classical system of the same dimension. In this sense, k-outcome local projective-simulable measurements on d-dimensional quantum systems are *effectively classical* resources. Hence, any quantum advantage observed in a task defined by the scenario $[n, d, k]$ certifies the use of non-projective-simulable measurements on a d-dimensional quantum system. Suppose that the measurement devices are identical for all parties in the setup $[n, d, k]$. Here, the set of all correlations obtained from k-outcome projective-simulable measurements on a shared system with local dimension d is the same as the set of classically simulable correlations obtained by applying identical post-processing to a shared classical state with local dimension d. When identical projective-simulable measurements are used on a quantum system with d-dimensional subsystems, this is equivalent to requiring that the post-processing of associated projective measurement outcomes is identical for all parties. From this observation and the proof of Theorem 4.1, it follows that the correlations generated by identical projective-simulable measurements on a shared quantum system with local dimension d are the same as those generated by a shared classical state with local dimension d and identical post-processing of outcomes.

4.1.2 The Task $\mathbb{G}[n, d, k]$

An operational scenario $[n, d, k]$ consists of n parties sharing a system with local operational dimension d and each party produces outcomes in $\{0, 1, \cdots, k-1\}$, where $k > d$. Such a scenario can be used to define a task $\mathbb{G}[n, d, k]$. The objective of such a task can be one of the following: (i) it could be to generate correlations in a chosen subset $Cor[n, d, k] \subseteq \mathfrak{C}(k, n)$, or alternatively, (ii) there could be a payoff function $\mathcal{S}_{(n,d,k)}$ defined over the probability distribution $P = \{P(a_1, a_2, \cdots, a_n)\}_{a_1,a_2,\cdots,a_n}$ for the task. In the second case, the goal is to maximise the payoff using the available shared resources and measurements. If the payoff has an upper bound

$\mathcal{S}^{PQ,max}_{(n,d,k)}$ for correlation in the set in $\mathfrak{C}^{PQ}_d(k,n)$, then any violation of this bound by correlations in $\mathfrak{C}^{Q}_d(k,n)$ implies detection of an nPS measurement on a d-dimensional quantum system. For each task, we will specify the objective explicitly at the beginning. When the measurement devices are identical for all parties, we denote the task by $\mathbb{G}^{sym}[n,d,k]$.

From Theorem 4.1, $\mathfrak{C}^{Cl}_d(k,n) = \mathfrak{C}^{PQ}_d(k,n)$. Given a payoff function $\mathcal{S}_{(n,d,k)}$,

$$\begin{aligned} \mathcal{S}^{PQ,max}_{(n,d,k)} &= \max_{P=\{P(a_1,a_2,\cdots,a_n)\}_{a_1,a_2,\cdots,a_n}\in\mathfrak{C}^{PQ}_d(k,n)} \mathcal{S}_{(n,d,k)}(P) \\ &= \max_{P=\{P(a_1,a_2,\cdots,a_n)\}_{a_1,a_2,\cdots,a_n}\in\mathfrak{C}^{Cl}_d(k,n)} \mathcal{S}_{(n,d,k)}(P) = \mathcal{S}^{Cl,max}_{(n,d,k)} \end{aligned} \tag{4.11}$$

We will optimise over $\mathfrak{C}^{Cl}_d(k,n)$ to get the projective-simulable bound $\mathcal{S}^{PQ,max}_{(n,d,k)}$ on the payoff for a given task. Recall from Section 2.2.1 that the state space of an n-partite classical system with local operational dimension d is $(\Omega_{d^n})_{A_1A_2\cdots A_n} = \{\omega^{Cl}_{A_1A_2\cdots A_n} = (p_{i_1,i_2,\cdots,i_n})_{i_1,i_2,\cdots,i_n\in\{0,1,\cdots,d-1\}} \in [0,1]^{d^n} : \sum_{i_1,i_2,\cdots,i_n\in\{0,1,\cdots,d-1\}} p_{i_1,i_2,\cdots,i_n} = 1\}$. In the scenario $[n,d,k]$, the local measurement $(\mathcal{M}^{d\to k}_{A_i})_{Cl}$ of party A_i is described by a column stochastic map $W^{d\to k}_{A_i}$ with entries $(W^{d\to k}_{A_i})_{r,s} = w^{A_i}_{r,s}$, where $\sum_{r=0}^{k-1} w^{A_i}_{r,s} = 1 \ \forall\ s \in \{0,1,\cdots,d-1\}$ and $w^{A_i}_{r,s} \geq 0\ \forall r \in \{0,1,\cdots,k-1\}, s \in \{0,1,\cdots,d-1\}$. The outcome distribution $P = \{P(a_1,a_2,\cdots,a_n)\}_{a_1,a_2,\cdots,a_n} \in \mathfrak{C}^{Cl}_d(k,n)$ in this case is given by $P = (\otimes_{i=1}^n W^{d\to k}_{A_i})(\omega^{Cl}_{A_1A_2\cdots A_n})^T$. Thus, $\mathcal{S}^{Cl,max}_{(n,d,k)}$ is obtained by optimising over all states in $(\Omega_{d^n})_{A_1A_2\cdots A_n}$ and local measurements $\{W^{d\to k}_{A_i}\}_{i\in[n]}$.

Corollary 4.1. *Consider task $\mathbb{G}[n,d,k]$ with some payoff function $\mathcal{S}_{(n,d,k)}$. If for some $P = \{P(a_1,a_2,\cdots,a_n)\}_{a_1,a_2,\cdots,a_n} \in \mathfrak{C}^{Q}_d(k,n)$ we have $\mathcal{S}_{(n,d,k)}(P) > \mathcal{S}^{Cl,max}_{(n,d,k)}$, then the local measurements used by the parties must be nPS.*

Proof. From Theorem 4.1, $\mathfrak{C}^{Cl}_d(k,n) = \mathfrak{C}^{PQ}_d(k,n)$. From Equation (4.11), $\mathcal{S}^{PQ,max}_{(n,d,k)} = \mathcal{S}^{Cl,max}_{(n,d,k)}$. Thus, any violation of the classical bound $\mathcal{S}_{(n,d,k)}(P) > \mathcal{S}^{Cl,max}_{(n,d,k)}$ certifies the use of k-outcome nPS measurements, given that the shared system has local operational dimension d. ∎

For the scenario $[n=2,d=2,k]$, we use the following known results to characterise classically achievable correlations and the payoff functions associated with tasks defined in this setup. We will use them in later sections.

Lemma 4.1. *(Guha et al. [276]) All the bipartite classical states $\omega^{Cl}_{AB} \in (\Omega_{d^2})_{AB}$ with local operational dimension 2 can be generated from some state $(p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p)$ with $p \in [0,1]$ by applying local stochastic maps.*

To compute the classical bound $\mathcal{S}^{PQ,max}_{(n=2,d,k)}$ for the setup $[n=2,d=2,k]$, it is enough to optimise over classical states of the form $(p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p)$ where $p \in [0,1]$ and local column stochastic maps for the parties.

Lemma 4.2. *(Guha et al. [276]) In the setup $[n = 2, d = 2, k]$, any correlations $P = \{P(a,b)\}_{a,b} \in \mathfrak{C}(k, n = 2)$ with $P(a,b) = 0$ for $a = b$ and $P(a,b) > 0$ for $a \neq b$ cannot be obtained by sharing a classical state of the form $(p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1 - p)$ where $p \in [0,1]$. Such correlations do not belong to $\mathfrak{C}_d^{Cl}(k, n = 2)$.*

Theorem 4.2. *(Guha et al. [276]) In the setup $[n = 2, d = 2, k = 3]$, consider a payoff function $\mathcal{S}_{(2,2,3)}(P) = \min_{\substack{a,b \\ a\neq b}} P(a,b)$ defined over correlations $P = \{P(a,b)\}_{a,b} \in \mathfrak{C}(k = 3, n = 2)$. The classical bound is $\mathcal{S}_{(2,2,3)}^{Cl,max} = \frac{1}{8}$.*

Next, we discuss detection schemes for three and four-outcome qubit nPS measurements.

4.2 Detecting Qubit Non-Projective-Simulable Measurements

We will now discuss various detection schemes for three and four-outcome qubit nPS measurements in the scenario $[n = 2, d = 2, k = 3]$ and $[n = 2, d = 2, k = 4]$, respectively. We begin by discussing a task for detecting three-outcome qubit nPS in the bipartite case $[n = 2, d = 2, k = 3]$, with parties A and B. A source $\mathcal{P}_{A,B}$ prepares a state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ which is shared between them. Each party performs a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^2) : E_a \geq 0\}_{a=0}^2$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^2) : F_b \geq 0\}_{b=0}^2$ for B. These measurements generate the joint distribution $\{P(a,b)\}_{a,b=0}^2$, where $P(a,b) = \mathrm{Tr}[\rho_{AB}(E_a \otimes F_b)]$. For the task $\mathbb{G}[2,2,3]$, we define the following payoff function over the distribution $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}(3,2)$.

$$\mathcal{S}_{(2,2,3)}(P) = \min_{\substack{a,b \\ a\neq b}} P(a,b) \tag{4.12}$$

Here, the maximum achievable payoff is achievable $\frac{1}{6}$, obtained when $P(a,b) = \frac{1}{6}$ for $a \neq b$ and $P(a,b) = 0$ otherwise, *i.e.* the anti-correlated outcomes are equiprobable. Any other $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}(3,2)$ yields a lower payoff. This task can serve as a detection scheme for three-outcome qubit nPS measurements.

Theorem 4.3. *Consider the payoff function $\mathcal{S}_{(2,2,3)}$ defined in Equation (4.12) for the scenario $[n = 2, d = 2, k = 3]$. Then, $\mathcal{S}_{(2,2,3)}^{PQ,max} \leq \frac{1}{8}$, while there exists $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}_2^Q(3,2)$ such that $\mathcal{S}_{(2,2,3)}(P) = \frac{1}{6} > \mathcal{S}_{(2,2,3)}^{PQ,max}$.*

Proof. [1] Using Theorem 4.1 and Equation (4.11), we have $\mathcal{S}_{(2,2,3)}^{PQ,max} = \mathcal{S}_{(2,2,3)}^{Cl,max}$. By using Theorem 4.2, $\mathcal{S}_{(2,2,3)}^{PQ,max} \leq \frac{1}{8}$. The bound is violated by using shared state $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$ together with the *trine* POVMs, *i.e.*, $\mathcal{M}_A = \{E_a = \frac{2}{3}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \cos[\frac{2\pi}{3}a]|0\rangle + \sin[\frac{2\pi}{3}a]|1\rangle, a \in \{0,1,2\}\}$ and $\mathcal{M}_B = \{F_b = \frac{2}{3}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle =$

[1] Also given in Section 3.2 of [185]

$\cos[\frac{2\pi}{3}b]\,|0\rangle + \sin[\frac{2\pi}{3}b]\,|1\rangle, b \in \{0,1,2\}\}$. This yields $P = \{P(a,b)\}_{a,b=0}^{2}$ with $P(a,b) = \frac{1}{6}$ for $a \neq b$ achieving payoff $\mathcal{S}_{(2,2,3)}(P) = \frac{1}{6}$. ■

In the above proof, if trine POVMs are used, the payoff remains higher even with the noisy shared state $p\,|\Psi^-\rangle\langle\Psi^-| + (1-p)\frac{\mathbb{I}}{2} \otimes \frac{\mathbb{I}}{2}$, provided $p > \frac{1}{4}$. Consequently, the task defined by the payoff function in Equation (4.12) can therefore serve as a detection scheme for three-outcome qubit nPS measurements, whenever the projective-simulable bound $\mathcal{S}_{(2,2,3)}^{PQ,max}$ is violated. By considering a similar payoff function as in [276] for the scenario $[n=2, d=2, k=4]$, one can also detect four-outcome qubit nPS measurements.

4.2.1 Detecting Qubit nPS Measurements Without Noise

We now define two tasks where the goal is to simulate any correlation from a specific set. We will use the task defined in the scenario $[n=2, d=2, k=3]$ to detect 3-outcome qubit nPS measurements and the task in scenario $[n=2, d=2, k=4]$ to detect 4-outcome qubit nPS measurements. A source $\mathcal{P}_{A,B}$ prepares a state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ which is shared between A and B. Each party performs a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^2) : E_a \geq 0\}_{a=0}^{k-1}$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^2) : F_b \geq 0\}_{b=0}^{k-1}$ for B, where $k \in \{3,4\}$. In the scenarios $[n=2, d=2, k=3]$ and $[n=2, d=2, k=4]$, we consider the sets of target correlations $Cor_1[2,2,3] \subset \mathfrak{C}(3,2)$ and $Cor_1[2,2,4] \subset \mathfrak{C}(4,2)$, respectively. For $k \in \{3,4\}$, here we define the set

$$Cor_1[2,2,k] := \Big\{\{P(a,b)\}_{a,b=0}^{k-1} \in \mathfrak{C}(k,2) : P(a,b) = 0 \text{ if } a = b, P(a,b) > 0 \text{ if } a \neq b\Big\} \tag{4.13}$$

In the correlation simulation task $\mathbb{G}[2,2,k]$ defined by $Cor_1[2,2,k]$, the objective is to produce some correlation in this set. k-outcome qubit projective-simulable measurements cannot generate correlations in this set when the shared state is any $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$. We then show that some correlations in $Cor_1[2,2,k]$ can be realised using k-outcome qubit POVMs on a shared state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Thus, these correlation simulation tasks can serve as detection schemes for $k=3$ and $k=4$-outcome qubit nPS measurements.

Theorem 4.4. *For $k \in \{3,4\}$, in the task $\mathbb{G}[2,2,k]$ defined using the set $Cor_1[2,2,k]$,*

$$\mathfrak{C}_2^{PQ}(k,2) \bigcap Cor_1[2,2,k] = \varnothing$$

Proof. [2] From Theorem 4.1, we get $\mathfrak{C}_d^{Cl}(k,2) = \mathfrak{C}_d^{PQ}(k,2)$. Now, using Lemma 4.1, we can assume *wlog* that the shared classical state is $\omega_{AB}^{Cl} = (p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p)$, where $p \in [0,1]$. Lemma 4.2 shows that any correlation $P = \{P(a,b)\}_{a,b} \in \mathfrak{C}(k, n=2)$ with $P(a,b) = 0$ when $a = b$ and $P(a,b) > 0$ when $a \neq b$ cannot be obtained from ω_{AB}^{Cl} using any local stochastic maps by A and B. Thus, $\mathfrak{C}_2^{PQ}(k,2) \bigcap Cor_1[2,2,k] = \varnothing$. ■

[2] Also discussed in Section 3.2 – 3.3 of [185]

Theorem 4.5. *For $k \in \{3,4\}$, in the task $\mathbb{G}[2,2,k]$ defined using the set $Cor_1[2,2,k]$,*

$$\mathfrak{C}_2^Q(k,2)\bigcap Cor_1[2,2,k] \neq \varnothing$$

Proof. [3] Let $k \in \{3,4\}$. Consider the shared state $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$. Party A performs the k-outcome extremal qubit POVM $\mathcal{M}_A = \{E_a = p_a(I_2 + \vec{n}_a \cdot \vec{\sigma})\}_{a=0}^{k-1}$, and B performs $\mathcal{M}_B = \{F_b = p_b(I_2 + \vec{n}_b \cdot \vec{\sigma})\}_{b=0}^{k-1}$. Here, $p_i > 0$, $\sum_i p_i = 1$, $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, $\vec{n}_i \in \mathbb{R}^3$, $|\vec{n}_i|^2 = 1$, $\sum_i p_i \vec{n}_i = 0$ and $\vec{n}_i \neq \vec{n}_j$ $\forall i,j \in \{0,1,\dots,k-1\}$. The resulting correlation $\{P(a,b)\}_{a,b=0}^{k-1}$ satisfies $P(a,b) = 0$ when $a = b$ and $P(a,b) > 0$ when $a \neq b$. Thus, $\{P(a,b)\}_{a,b}^{k-1} \in Cor_1[2,2,k]$. ■

For $k = 3,4$, using Theorem 4.4 and 4.5, the correlation simulation task $\mathbb{G}[2,2,k]$ defined using the set $Cor_1[2,2,k]$ can be used for the detection of k-outcome qubit nPS measurements in the absence of noise.

4.2.2 Robust Detection Scheme for Qubit nPS Measurements

We now present detection schemes for 3 and 4-outcome qubit nPS measurements that are robust against arbitrary depolarising noise (except for the completely depolarising noise). As in Section 4.2.1, the aim is to simulate a correlation from a prescribed set of target correlations. The task in scenario $[n=2, d=2, k=3]$ is used to detect 3-outcome qubit nPS measurements, and the task in scenario $[n=2, d=2, k=4]$ is used to detect 4-outcome qubit nPS measurements. A source $\mathcal{P}_{A,B}$ prepares a state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ which is shared between A and B. Each party performs a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^2) : E_a \geq 0\}_{a=0}^{k-1}$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^2) : F_b \geq 0\}_{b=0}^{k-1}$ for B, where $k \in \{3,4\}$. In the scenarios $[n=2, d=2, k=3]$ and $[n=2, d=2, k=4]$, we define the target sets $Cor_2[2,2,3] \subset \mathfrak{C}(3,2)$ and $Cor_2[2,2,4] \subset \mathfrak{C}(4,2)$, respectively. For $k \in \{3,4\}$, here we define the set $Cor_2[2,2,k]$ consisting of correlations $\{P(a,b)\}_{a,b=0}^{k-1} \in \mathfrak{C}(k,2)$ such that

$$P(a=i, b=i) = P(a=i', b=i') \neq \frac{1}{k^2} \tag{4.14}$$

$$\text{and } P(a=i, b=j) = P(a=i', b=j') \neq \frac{1}{k^2} \;\forall\; i, i', j(\neq i), j'(\neq i') \in \{0,1,\cdots,k-1\} \tag{4.15}$$

In other words, $Cor_2[2,2,k]$ is the set of correlations where all correlated outcomes occur with the same probability, and all anti-correlated outcomes occur with the same probability. In the correlation simulation task $\mathbb{G}[2,2,k]$ based on $Cor_2[2,2,k]$, the goal is to obtain any correlation from the set $Cor_2[2,2,k]$. We show that k-outcome qubit projective-simulable measurements cannot generate such correlations when the shared state is any $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$. In contrast, certain correlations in $Cor_2[2,2,k]$ can be realised with k-outcome qubit POVMs on a shared state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Thus, these tasks provide a method to detect $k=3$ and $k=4$ outcome

[3] Also discussed in Section 3.2 – 3.3 of [185]

qubit nPS measurements. We further show that the detection remains robust under arbitrary depolarising noise.

4.2.2.1 Detecting 3-outcome Qubit nPS Mesurements with $Cor_2[2,2,3]$

Let us consider the task $\mathbb{G}[2,2,3]$ based on the target correlation set $Cor_2[2,2,3]$ and discuss the detection scheme for 3-outcome qubit nPS measurement.

Theorem 4.6. *In the task $\mathbb{G}[2,2,3]$ defined by the set $Cor_2[2,2,3]$,*

$$\mathfrak{C}_2^{PQ}(3,2) \bigcap Cor_2[2,2,3] = \varnothing$$

Proof. [4] From Theorem 4.1, we have $\mathfrak{C}_d^{PQ}(3,2) = \mathfrak{C}_d^{Cl}(3,2)$. Using Lemma 4.1, we assume *wlog* that the shared classical state is $\omega_{AB}^{Cl} = (p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p)$, where $p \in [0,1]$. Let the measurements of A and B be described by column stochastic map $W_A^{2\to3}$ and $W_B^{2\to3}$, respectively. Let the entries of the column stochastic maps be $(W_A^{2\to3})_{r,s} = w_{r,s}$ and $(W_B^{2\to3})_{r,s} = w'_{r,s}$, respectively, where $r \in \{0,1,2\}, s \in \{0,1\}$. These entries satisfy the usual constraints for a stochastic map:

$$\sum_{r=0}^{2} w_{r,s} = 1 \ \forall\ s \in \{0,1\} \text{ and } w_{r,s} \geq 0\ \forall\ r \in \{0,1,2\}, s \in \{0,1\} \tag{4.16}$$

$$\sum_{r=0}^{2} w'_{r,s} = 1 \ \forall\ s \in \{0,1\} \text{ and } w'_{r,s} \geq 0\ \forall\ r \in \{0,1,2\}, s \in \{0,1\} \tag{4.17}$$

The outcome distribution $P = \{P(a,b)\}_{a,b=0}^{2} \in \mathfrak{C}_d^{Cl}(3,2)$ obtained from the shared state and measurement is given by

$$P = (W_A^{2\to3} \otimes W_B^{2\to3})(\omega_{AB}^{Cl})^T \tag{4.18}$$

For the correlation $P = \{P(a,b)\}_{a,b=0}^{2}$ in Equation (4.18) to lie in set $Cor_2[2,2,3]$, the following conditions must be satisfied:

$$P(0,0) = P(1,1) = P(2,2) \neq \frac{1}{9} \tag{4.19}$$

$$P(0,1) = P(0,2) = P(1,0) = P(1,2) = P(2,0) = P(2,1) \neq \frac{1}{9} \tag{4.20}$$

From Equations (4.19) and (4.20), the marginals for both A and B must satisfy

$$P(a) = \frac{1}{3}\ \forall\ a \in \{0,1,2\} \text{ and } P(b) = \frac{1}{3}\ \forall\ b \in \{0,1,2\} \tag{4.21}$$

[4] Also discussed in Section 3.2 of [185]

Using Equation (4.21) together with constraints from Equations (4.16) and (4.17), we find that $0 < p < 1$ must hold. For $p = 0$ or $p = 1$, the state ω^{Cl}_{AB} cannot satisfy the constraints in the Equations (4.19) and (4.20). Intuitively, the state ω^{Cl}_{AB} is a probability distribution with zero mutual information when $p = 0$ or $p = 1$, while the correlations in $Cor_2[2, 2, 3]$ have non-zero mutual information. Since mutual information cannot increase under local operations [244], such correlations cannot be obtained. Substituting the condition from Equation (4.21) into the expression in Equation (4.18), and imposing the constraints from Equations (4.16) and (4.17), we obtain

$$w_{0,1} = \frac{1}{(1-p)}(\frac{1}{3} - p\ w_{0,0}) \text{ and } w_{1,1} = \frac{1}{(1-p)}(\frac{1}{3} - p\ w_{1,0}) \tag{4.22}$$

$$w'_{0,1} = \frac{1}{(1-p)}(\frac{1}{3} - p\ w'_{0,0}) \text{ and } w'_{1,1} = \frac{1}{(1-p)}(\frac{1}{3} - p\ w'_{1,0}) \tag{4.23}$$

We substitute Equations (4.22) and (4.23) into Equations (4.19) and (4.20), and solve them using a symbolic solver (Mathematica) under the constraints in Equations (4.16) and (4.17). We consider two cases: $P(0, 0) > P(0, 1)$ and $P(0, 0) < P(0, 1)$. In both cases, the system of equations has no solution. ■

Theorem 4.7. *In the task $\mathbb{G}[2, 2, 3]$ defined by the set $Cor_2[2, 2, 3]$,*

$$\varnothing \neq \mathfrak{C}^Q_2(3, 2) \bigcap Cor_2[2, 2, 3] \subsetneq Cor_2[2, 2, 3]$$

Proof. [5] First, we show that some correlations in $Cor_2[2, 2, 3]$ do not belong to $\mathfrak{C}^Q_2(3, 2)$. In other words, they cannot be obtained by sharing a bipartite quantum state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ and performing POVMs $\mathcal{M}_A$ by A and $\mathcal{M}_B$ by B. For a correlation $P = \{P(a, b)\}^2_{a,b=0} \in Cor_2[2, 2, 3]$, let $P(0, 0) = \alpha$ and $P(0, 1) = \beta$. Using the definition of $Cor_2[2, 2, 3]$ and normalisation condition, we obtain,

$$3\alpha + 6\beta = 1 \implies \alpha = \frac{1}{3} - 2\beta \tag{4.24}$$

For $P = \{P(a, b)\}^2_{a,b=0} \in Cor_2[2, 2, 3]$, the mutual information is $2 \log_2 3 + 3\alpha \log_2 \alpha + 6\beta \log_2 \beta$. When the parties share a state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$, they can only apply local operations. Since mutual information is non-increasing under local completely positive trace-preserving maps, its value cannot exceed $\log_2 d$, where d is the subsystem dimension [244]. For qubits, this bound is 1. Hence, any correlation in $Cor_2[2, 2, 3]$ with mutual information greater than 1 cannot be obtained using a shared two-qubit state. In particular, correlations $P = P(a, b)^2_{a,b=0} \in Cor_2[2, 2, 3]$ satisfying

$$2 \log_2 3 + 3\alpha \log_2 \alpha + 6\beta \log_2 \beta > 1 \tag{4.25}$$

do not belong to $\mathfrak{C}^Q_2(3, 2)$, as they cannot be generated from any two-qubit state and local POVMs (see region $R4$ in Figure 4.2). Now, we will show that some correlations in

[5] Also discussed in Section 3.2 of [185]

$a\backslash b$	0	1	2
0	$\frac{1}{9}(1-p)$	$\frac{1}{18}(2+p)$	$\frac{1}{18}(2+p)$
1	$\frac{1}{18}(2+p)$	$\frac{1}{9}(1-p)$	$\frac{1}{18}(2+p)$
2	$\frac{1}{18}(2+p)$	$\frac{1}{18}(2+p)$	$\frac{1}{9}(1-p)$

Table 4.1: Correlations $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$ generated using shared state $(\rho_p)_{AB}$ and qubit trine POVM by A and B when $p \in (0,1]$.

$a\backslash b$	0	1	2
0	$\frac{1}{9}(1+p)$	$\frac{1}{18}(2-p)$	$\frac{1}{18}(2-p)$
1	$\frac{1}{18}(2-p)$	$\frac{1}{9}(1+p)$	$\frac{1}{18}(2-p)$
2	$\frac{1}{18}(2-p)$	$\frac{1}{18}(2-p)$	$\frac{1}{9}(1+p)$

Table 4.2: Correlations $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$ generated using shared state $(\tilde{\rho}_p)_{AB}$ and qubit trine POVM by A and B when $p \in (0,1]$.

$P = \{P(a,b)\}_{a,b=0}^{2} \in \mathfrak{C}_2^Q(3,2)$ also belong to $Cor_2[2,2,3]$.

Let the shared state be $(\rho_p)_{AB} = p\,|\Psi^-\rangle_{AB}\langle\Psi^-| + (1-p)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$, where $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$ is the two-qubit singlet state and $p \in (0,1]$. Both parties perform the qubit trine POVM: $\mathcal{M}_A = \{E_a = \frac{2}{3}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \cos[\frac{2\pi}{3}a]\,|0\rangle + \sin[\frac{2\pi}{3}a]\,|1\rangle, a \in \{0,1,2\}\}$, and $\mathcal{M}_B = \{F_b = \frac{2}{3}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = \cos[\frac{2\pi}{3}b]\,|0\rangle + \sin[\frac{2\pi}{3}b]\,|1\rangle, b \in \{0,1,2\}\}$. The resulting correlation is shown in Table 4.1. For $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. Using the above state and measurements, all correlations with $\alpha \in [0, \frac{1}{9})$ can be realised. The limiting value $\alpha = 0$ is reached for $p = 1$ (see region $R1$ in the Figure 4.2).

Now, let the shared state be $(\tilde{\rho}_p)_{AB} := p\,|\Psi^+\rangle_{AB}\langle\Psi^+| + (1-p)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$, where $|\Psi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB})$ and $p \in (0,1]$. Both parties perform the qubit trine POVM: $\mathcal{M}_A = \{E_a = \frac{2}{3}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \cos[\frac{2\pi}{3}a]\left|\frac{0+1}{\sqrt{2}}\right\rangle + \sin[\frac{2\pi}{3}a]\left|\frac{0-1}{\sqrt{2}}\right\rangle, a \in \{0,1,2\}\}$, and $\mathcal{M}_B = \{F_b = \frac{2}{3}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = \cos[\frac{2\pi}{3}b]\left|\frac{0+1}{\sqrt{2}}\right\rangle + \sin[\frac{2\pi}{3}b]\left|\frac{0-1}{\sqrt{2}}\right\rangle, b \in \{0,1,2\}\}$. The resulting correlation is given in Table 4.2. For $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. Using the above state and measurements, all correlations with $\alpha \in (\frac{1}{9}, \frac{2}{9}]$ can be simulated using this shared state and measurements. The upper limit $\alpha = \frac{2}{9}$ is reached for $p = 1$. (see region $R2$ in the Figure 4.2).

■

From Theorem 4.6 and 4.7, correlations in $\mathfrak{C}_2^Q(3,2) \bigcap Cor_2[2,2,3]$, such as those shown in Table 4.1 and 4.2, are generated using 3-outcome qubit nPS measurements on shared two-qubit state. Beyond these cases, it is unknown whether all correlations in $Cor_2[2,2,3]$ with mutual information at most 1 belong to $\mathfrak{C}_2^Q(3,2)$. In particular, it remains open whether correlations

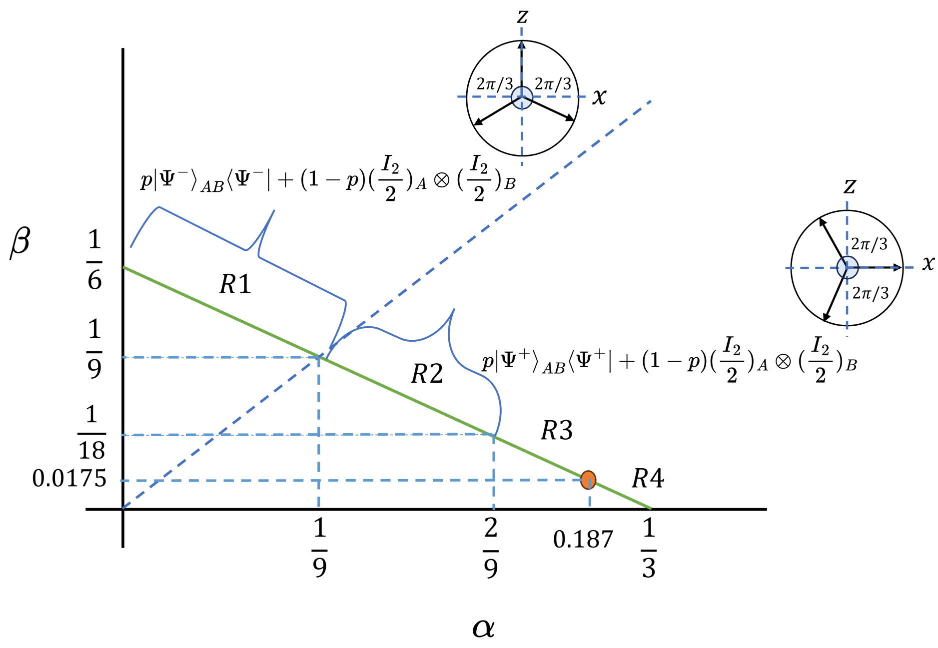

Figure 4.2: Correlations in $Cor_2[2,2,3]$ that can be simulated using 3-outcome qubit nPS measurements. For $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. Correlations with α in region $R1$ can be simulated using a qubit trine POVM on the shared state $(\rho_p)_{AB}$, while correlations with α in $R2$ can be simulated using a rotated qubit trine POVM on $(\tilde{\rho}_p)_{AB}$. Correlations with α in region $R4$ cannot be obtained using local qubit POVMs on any shared two-qubit state. Whether correlations with α in region $R3$ belong to $\mathfrak{C}_2^Q(3,2)$ is unknown.

with $\alpha > \frac{2}{9}$, shown as region $R3$ in Figure 4.2, can be simulated using 3-outcome qubit POVMs.

Robustness against arbitrary depolarising noise: From the discussion above, correlations in $Cor_2[2,2,3]$ are not qubit projective-simulable when the parties share a two-qubit state. For example, suppose the goal is to generate correlations $\{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$ with $P(0,0) = 0$. The parties can share the two-qubit singlet $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$ and perform a qubit trine POVMs: $\mathcal{M}_A = \{E_a = \frac{2}{3}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \cos[\frac{2\pi}{3}a]|0\rangle + \sin[\frac{2\pi}{3}a]|1\rangle, a \in \{0,1,2\}\}$, and $\mathcal{M}_B = \{F_b = \frac{2}{3}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = \cos[\frac{2\pi}{3}b]|0\rangle + \sin[\frac{2\pi}{3}b]|1\rangle, b \in \{0,1,2\}\}$. In practice, both the state and the measurements may be noisy. A standard model is depolarising noise under which the shared state becomes $(\rho_{\epsilon_s})_{AB} = \epsilon_s|\Psi^-\rangle_{AB}\langle\Psi^-| + (1-\epsilon_s)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$ where $\epsilon_s \in (0,1]$. Similarly, depolarising noise affects the measurements and the noisy measurement is given by $\mathcal{M}_A = \{E_a = (\Pi_a)_A : a \in \{0,1,2\}\}$ and $\mathcal{M}_B = \{F_b = (\Pi_b)_B : b \in \{0,1,2\}\}$, where $(\Pi_i)_X$ for $X = A, B$ and $i \in \{0,1,2\}$ is given below with noise parameters $\epsilon_A, \epsilon_B \in (0,1]$:

$a \backslash b$	0	1	2
0	$\frac{1}{9}(1-\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$
1	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{9}(1-\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$
2	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{18}(2+\epsilon_s\epsilon_A\epsilon_B)$	$\frac{1}{9}(1-\epsilon_s\epsilon_A\epsilon_B)$

Table 4.3: Correlations $P = \{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$ generated using the noisy shared state $(\rho_{\epsilon_s})_{AB}$ and noisy qubit trine POVMs. The parameters $\epsilon_s, \epsilon_A, \epsilon_B$ denote the noise in the shared state and in A's and B's measurements, respectively.

$$(\Pi_0)_X = \begin{pmatrix} \frac{1+\epsilon_X}{3} & 0 \\ 0 & \frac{1-\epsilon_X}{3} \end{pmatrix}, (\Pi_1)_X = \begin{pmatrix} \frac{2-\epsilon_X}{6} & \frac{-\epsilon_X}{2\sqrt{3}} \\ \frac{-\epsilon_X}{2\sqrt{3}} & \frac{2+\epsilon_X}{6} \end{pmatrix}, (\Pi_2)_X = \begin{pmatrix} \frac{2-\epsilon_X}{6} & \frac{\epsilon_X}{2\sqrt{3}} \\ \frac{\epsilon_X}{2\sqrt{3}} & \frac{2+\epsilon_X}{6} \end{pmatrix} \tag{4.26}$$

The noisy state and measurements generate the correlations shown in Table 4.3. For any $\epsilon_s, \epsilon_A, \epsilon_B \in (0,1]$, the resulting correlations lie in $Cor_2[2,2,3]$. Hence, this detection scheme is robust against arbitrary depolarising noise.

If the parties want to achieve correlations $\{P(a,b)\}_{a,b=0}^{2} \in Cor_2[2,2,3]$ with $P(0,0) = \frac{2}{9}$, they can share the two-qubit state $|\Psi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} + |10\rangle_{AB})$. Both parties perform qubit trine POVM: $\mathcal{M}_A = \{E_a = \frac{2}{3}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \cos[\frac{2\pi}{3}a]\left|\frac{0+1}{\sqrt{2}}\right\rangle + \sin[\frac{2\pi}{3}a]\left|\frac{0-1}{\sqrt{2}}\right\rangle, a \in \{0,1,2\}\}$, and $\mathcal{M}_B = \{F_b = \frac{2}{3}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = \cos[\frac{2\pi}{3}b]\left|\frac{0+1}{\sqrt{2}}\right\rangle + \sin[\frac{2\pi}{3}b]\left|\frac{0-1}{\sqrt{2}}\right\rangle, b \in \{0,1,2\}\}$. Even with this state and measurement, the scheme remains robust against arbitrary depolarising noise, following the process discussed above. More generally, let ρ_{AB} be any two-qubit shared state, and let $\mathcal{M}_A = \{E_a\}_{a=0}^{2}$ and $\mathcal{M}_B = \{F_b\}_{b=0}^{2}$ be qubit POVMs, where $\sum_i E_i = \sum_i F_i = I_2$, $E_i > 0$ and $F_i > 0$. The resulting correlations are $P(a,b) = \mathrm{Tr}[\rho_{AB}(E_a \otimes F_b)]$. A fixed shared state can be used to detect 3-outcome qubit nPS measurements with this scheme whenever the resulting correlations satisfy the constraints in Equations (4.14) and (4.15).

4.2.2.2 Detecting 4-outcome Qubit nPS Measurements with $Cor_2[2,2,4]$

Next, we consider the task $\mathbb{G}[2,2,4]$, based on the target set $Cor_2[2,2,4]$ defined in Equations (4.14) and (4.15), to present a detection scheme for 4-outcome qubit nPS measurements. This scheme is also robust against arbitrary depolarising noise.

Theorem 4.8. *In the task $\mathbb{G}[2,2,4]$ defined by $Cor_2[2,2,4]$,*

$$\mathfrak{C}_2^{PQ}(4,2) \bigcap Cor_2[2,2,4] = \varnothing$$

Proof. [6] From Theorem 4.1, we have $\mathfrak{C}_d^{PQ}(4,2) = \mathfrak{C}_d^{Cl}(4,2)$. Using Lemma 4.1, we can assume *wlog* that the shared classical state is $\omega_{AB}^{Cl} = (p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p)$, where

[6] Also discussed in Section 3.3 of [185]

$p \in [0,1]$. Let A and B perform the column stochastic map $W_A^{2\to4}$ and $W_B^{2\to4}$, respectively. Let the entries of the column stochastic maps be $(W_A^{2\to4})_{r,s} = w_{r,s}$ and $(W_B^{2\to4})_{r,s} = w'_{r,s}$, respectively, where $r \in \{0,1,2,3\}, s \in \{0,1\}$. These entries satisfy:

$$\sum_{r=0}^{3} w_{r,s} = 1 \;\forall\; s \in \{0,1\} \text{ and } w_{r,s} \geq 0 \;\forall\; r \in \{0,1,2,3\}, s \in \{0,1\} \tag{4.27}$$

$$\sum_{r=0}^{3} w'_{r,s} = 1 \;\forall\; s \in \{0,1\} \text{ and } w'_{r,s} \geq 0 \;\forall\; r \in \{0,1,2,3\}, s \in \{0,1\} \tag{4.28}$$

The resulting outcome distribution $P = \{P(a,b)\}_{a,b=0}^{3} \in \mathfrak{C}_d^{Cl}(4,2)$ is given by

$$P = (W_A^{2\to4} \otimes W_B^{2\to4})(\omega_{AB}^{Cl})^T \tag{4.29}$$

For the correlation $P = \{P(a,b)\}_{a,b=0}^{3}$ in Equation (4.29) to belong to $Cor_2[2,2,4]$, the following conditions must be satisfied:

$$P(0,0) = P(1,1) = P(2,2) = P(3,3) \neq \frac{1}{16} \tag{4.30}$$

$$\begin{aligned} &P(0,1) = P(0,2) = P(0,3) = P(1,0) = P(1,2) = P(1,3) = P(2,0) = P(2,1) = \\ &P(2,3) = P(3,0) = P(3,1) = P(3,2) \neq \frac{1}{16} \end{aligned} \tag{4.31}$$

From Equations (4.30) and (4.31), the marginals for both A and B must be uniform.

$$P(a) = \frac{1}{4} \;\forall\; a \in \{0,1,2,3\} \text{ and } P(b) = \frac{1}{4} \;\forall\; b \in \{0,1,2,3\} \tag{4.32}$$

Using Equation (4.32) together with constraints from Equations (4.27) and (4.28), gives $0 < p < 1$. For $p = 0$ or $p = 1$, the state ω_{AB}^{Cl} cannot satisfy the Equations (4.30) and (4.31). Intuitively, the state ω_{AB}^{Cl} is a probability distribution with zero mutual information when $p = 0$ or $p = 1$. The correlations in $Cor_2[2,2,4]$ have non-zero mutual information. Since mutual information cannot increase under local operations [244], such correlations cannot be obtained when $p = 0$ or $p = 1$. Now, substituting the Equation (4.32) into the expression in Equation (4.29), and applying the constraints in Equations (4.27) and (4.28), we obtain

$$w_{r,1} = \frac{1}{(1-p)}\left(\frac{1}{4} - p\, w_{r,0}\right) \text{ and } w'_{r,1} = \frac{1}{(1-p)}\left(\frac{1}{4} - p\, w'_{r,0}\right) \;\forall\; r \in \{0,1,2,3\} \tag{4.33}$$

We substitute Equations (4.33) into Equations (4.30) and (4.31), and solve them using a symbolic solver (Mathematica) under the constraints in Equations (4.27) and (4.28). We consider two cases: $P(0,0) > P(0,1)$ and $P(0,0) < P(0,1)$. In both cases, the system of equations has no solution. ■

Theorem 4.9. *In the task $\mathbb{G}[2,2,4]$ defined by $Cor_2[2,2,4]$,*

$$\varnothing \neq \mathfrak{C}_2^{Q}(4,2) \bigcap Cor_2[2,2,4] \subsetneq Cor_2[2,2,4]$$

Proof. [7] We first show that some correlations in $Cor_2[2,2,4]$ do not lie in $\mathfrak{C}_2^Q(4,2)$, *i.e.*, they cannot be realised by sharing a two-qubit state $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$ and locally performing qubit POVMs $\mathcal{M}_A$ and $\mathcal{M}_B$. For $P = \{P(a,b)\}_{a,b=0}^3 \in Cor_2[2,2,4]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. We get the following from the definition of $Cor_2[2,2,4]$ and using normalisation condition

$$4\alpha + 12\beta = 1 \implies \alpha = \frac{1}{4} - 3\beta \tag{4.34}$$

For $P = \{P(a,b)\}_{a,b=0}^3 \in Cor_2[2,2,4]$, the mutual information is $4 + 4\alpha\log_2\alpha + 12\beta\log_2\beta$. Sharing $\rho_{AB} \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$, the parties can only apply local operations. Since mutual information is non-increasing under local completely positive trace-preserving maps, its value cannot exceed $\log_2 d$, where d is the subsystem dimension [244]. For qubits, this bound is 1. Hence, any correlation in $Cor_2[2,2,4]$ with mutual information greater than 1 cannot be obtained using a shared two-qubit state. In particular, correlations $P = P(a,b)_{a,b=0}^2 \in Cor_2[2,2,4]$ satisfying $4 + 4\alpha\log_2\alpha + 12\beta\log_2\beta > 1$ do not belong to $\mathfrak{C}_2^Q(4,2)$ (see region $R4$ in Figure 4.3). Now, we will show that some correlations in $\mathfrak{C}_2^Q(4,2)$ also lie in $Cor_2[2,2,3]$.

Let the shared state be $(\rho_p)_{AB} = p\,|\Psi^-\rangle_{AB}\langle\Psi^-| + (1-p)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$, where $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$ is the two-qubit singlet state and $p \in (0,1]$. Both parties perform the qubit SIC-POVM: $\mathcal{M}_A = \{E_a = \frac{1}{2}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = |0\rangle \text{ for } a = 0, |\psi_a\rangle = \sqrt{\frac{1}{3}}\,|0\rangle + \sqrt{\frac{2}{3}}\,e^{\frac{2\pi i}{3}(a-1)}\,|1\rangle \text{ for } a \in \{1,2,3\}\}$, and $\mathcal{M}_B = \{F_b = \frac{1}{2}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = |0\rangle \text{ for } b = 0, |\psi_b\rangle = \sqrt{\frac{1}{3}}\,|0\rangle + \sqrt{\frac{2}{3}}\,e^{\frac{2\pi i}{3}(b-1)}\,|1\rangle \text{ for } b \in \{1,2,3\}\}$. The resulting correlation is shown in Table 4.4. For $P = \{P(a,b)\}_{a,b=0}^3 \in Cor_2[2,2,4]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. Using the above-mentioned state and measurements, all correlations with $\alpha \in [0, \frac{1}{16})$ can be realised. The limiting value $\alpha = 0$ is reached for $p = 1$ (see region $R1$ in the Figure 4.3).

Now, consider the shared state $(\tilde{\rho}_p)_{AB} := p\,|\Phi^+\rangle_{AB}\langle\Phi^+| + (1-p)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$, where $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$ and $p \in (0,1]$. The parties perform the qubit SIC-POVM given below:

$$\begin{aligned}\mathcal{M}_A = &\{E_a = \frac{1}{2}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \text{ for } a = 0,\\ &|\psi_a\rangle = \frac{(1+\sqrt{2}e^{\frac{2\pi i}{3}(a-1)})\,|0\rangle + (1-\sqrt{2}e^{\frac{2\pi i}{3}(a-1)})\,|1\rangle}{\sqrt{6}} \text{ for } a \in \{1,2,3\}\}\end{aligned} \tag{4.35}$$

$$\begin{aligned}\mathcal{M}_B = &\{F_b = \frac{1}{2}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \text{ for } b = 0,\\ &|\psi_b\rangle = \frac{(1+\sqrt{2}e^{\frac{2\pi i}{3}(1-b)})\,|0\rangle + (1-\sqrt{2}e^{\frac{2\pi i}{3}(1-b)})\,|1\rangle}{\sqrt{6}} \text{ for } b \in \{1,2,3\}\}\end{aligned} \tag{4.36}$$

The resulting correlation is shown in Table 4.5. For $P = \{P(a,b)\}_{a,b=0}^3 \in Cor_2[2,2,4]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. With this state and measurement, all correlations with $\alpha \in (\frac{1}{16}, \frac{1}{8}]$

[7] Also discussed in Section 3.3 of [185]

$a\backslash b$	0	1	2	3
0	$\frac{1}{16}(1-p)$	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$
1	$\frac{1}{48}(3+p)$	$\frac{1}{16}(1-p)$	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$
2	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$	$\frac{1}{16}(1-p)$	$\frac{1}{48}(3+p)$
3	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$	$\frac{1}{48}(3+p)$	$\frac{1}{16}(1-p)$

Table 4.4: Correlations $P = \{P(a,b)\}_{a,b=0}^{3} \in Cor_2[2,2,4]$ generated using shared state $(\rho_p)_{AB}$ and qubit SIC-POVM by A and B when $p \in (0,1]$.

$a\backslash b$	0	1	2	3
0	$\frac{1}{16}(1+p)$	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$
1	$\frac{1}{48}(3-p)$	$\frac{1}{16}(1+p)$	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$
2	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$	$\frac{1}{16}(1+p)$	$\frac{1}{48}(3-p)$
3	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$	$\frac{1}{48}(3-p)$	$\frac{1}{16}(1+p)$

Table 4.5: Correlations $P = \{P(a,b)\}_{a,b=0}^{3} \in Cor_2[2,2,4]$ generated using shared state $(\tilde{\rho}_p)_{AB}$ and rotated qubit SIC-POVM by A and B when $p \in (0,1]$.

can be simulated. The upper limit $\alpha = \frac{1}{8}$ is reached for $p = 1$. (see region $R2$ in the Figure 4.3).

■

From Theorem 4.8 and 4.9, correlations in $\mathfrak{C}_2^Q(4,2) \bigcap Cor_2[2,2,4]$, such as those shown in Table 4.4 and 4.5, are generated using 4-outcome qubit nPS measurement on a shared two-qubit state. Beyond these examples, it is not known whether all correlations in $Cor_2[2,2,4]$ with mutual information at most 1 belong to $\mathfrak{C}_2^Q(4,2)$. In particular, it remains open whether correlations with $\alpha > \frac{1}{8}$, shown as region $R3$ in Figure 4.3, can be simulated using 4-outcome qubit POVMs.

Robustness against arbitrary depolarising noise: As discussed above, correlations in $Cor_2[2,2,4]$ are not qubit projective-simulable when the shared system is a two-qubit system, while some can be obtained with 4-outcome qubit nPS measurements. For example, to generate correlations $\{P(a,b)\}_{a,b=0}^{3} \in Cor_2[2,2,4]$ with $P(0,0) = 0$, the parties may share $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB})$ and perform the qubit SIC-POVM: $\mathcal{M}_A = \{E_a = \frac{1}{2}|\psi_a\rangle\langle\psi_a| : |\psi_a\rangle = |0\rangle$ for $a = 0, |\psi_a\rangle = \sqrt{\frac{1}{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{\frac{2\pi i}{3}(a-1)}|1\rangle$ for $a \in \{1,2,3\}\}$, and $\mathcal{M}_B = \{F_b = \frac{1}{2}|\psi_b\rangle\langle\psi_b| : |\psi_b\rangle = |0\rangle$ for $b = 0, |\psi_b\rangle = \sqrt{\frac{1}{3}}|0\rangle + \sqrt{\frac{2}{3}}e^{\frac{2\pi i}{3}(b-1)}|1\rangle$ for $b \in \{1,2,3\}\}$. In practice, both the state and the measurements may be noisy. Under depolarising noise, the shared state is $(\rho_{\epsilon_s})_{AB} = \epsilon_s|\Psi^-\rangle_{AB}\langle\Psi^-| + (1-\epsilon_s)(\frac{I_2}{2})_A \otimes (\frac{I_2}{2})_B$ where $\epsilon_s \in (0,1]$. Similarly, under depolarising noise the measurements are $\mathcal{M}_A = \{E_a = (\Pi_a)_A : a \in \{0,1,2,3\}\}$ and $\mathcal{M}_B = \{F_b = (\Pi_b)_B : b \in \{0,1,2,3\}\}$. $(\Pi_i)_X$ for $X = A, B$ and $i \in \{0,1,2,3\}$ is given below with noise parameters $\epsilon_A, \epsilon_B \in (0,1]$:

$a\backslash b$	0	1	2	3
0	$\frac{1}{16}(1-\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$
1	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{16}(1-\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$
2	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{16}(1-\epsilon)$	$\frac{1}{48}(3+\epsilon)$
3	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{48}(3+\epsilon)$	$\frac{1}{16}(1-\epsilon)$

Table 4.6: Correlations $P = \{P(a,b)\}_{a,b=0}^{4} \in Cor_2[2,2,4]$ generated using noisy shared state $(\rho_{\epsilon_s})_{AB}$ and noisy qubit SIC-POVMs. The parameters $\epsilon_s, \epsilon_A, \epsilon_B$ denote the noise in the shared state and in A's and B's measurements, respectively, and $\epsilon = \epsilon_s\epsilon_A\epsilon_B$.

$$(\Pi_0)_X = \begin{pmatrix} \frac{1+\epsilon_X}{4} & 0 \\ 0 & \frac{1-\epsilon_X}{4} \end{pmatrix}, (\Pi_1)_X = \begin{pmatrix} \frac{3-\epsilon_X}{12} & \frac{\epsilon_X}{3\sqrt{2}} \\ \frac{\epsilon_X}{3\sqrt{2}} & \frac{3+\epsilon_X}{12} \end{pmatrix},$$
$$(\Pi_2)_X = \begin{pmatrix} \frac{3-\epsilon_X}{12} & \frac{i^{\frac{1}{3}}\epsilon_X}{3\sqrt{2}} \\ \frac{i(i+\sqrt{3})\epsilon_X}{6\sqrt{2}} & \frac{3+\epsilon_X}{12} \end{pmatrix}, (\Pi_3)_X = \begin{pmatrix} \frac{3-\epsilon_X}{12} & \frac{i(i+\sqrt{3})\epsilon_X}{6\sqrt{2}} \\ -\frac{i^{\frac{1}{3}}\epsilon_X}{3\sqrt{2}} & \frac{3+\epsilon_X}{12} \end{pmatrix} \tag{4.37}$$

These noisy resources produce the correlations in Table 4.6. For any $\epsilon_s, \epsilon_A, \epsilon_B \in (0,1]$, the resulting correlations lie in $Cor_2[2,2,4]$. Thus, the detection scheme is robust against arbitrary depolarising noise. If the goal is instead to obtain correlation in $Cor_2[2,2,4]$ with $P(0,0) = \frac{1}{8}$, the parties may share $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$ and perform the qubit SIC-POVMs defined in Equation (4.35) and (4.36). In this case as well, the detection scheme remains robust under arbitrary depolarising noise.

We saw that correlations in $Cor_2[2,2,3]$ cannot be generated using local 3-outcome qubit projective-simulable measurements on a two-qubit shared state. From Theorem 4.1, $\mathfrak{C}_d^{PQ}(3,2) = \mathfrak{C}_d^{Cl}(3,2)$. This raises the question of whether some correlations in $Cor_2[2,2,3]$ can be simulated when the parties share two two-level bipartite classical states, where both states are uncorrelated with each other. We further assume that the measurement devices are identical copies. Each party, A and B, measures each subsystem independently and then jointly post-processes the outcomes. As the target correlation $Corr_3[2,2,3]$, consider $P(0,0) = P(1,1) = P(2,2) = 0$ and $P(0,1) = P(0,2) = P(1,0) = P(1,2) = P(2,0) = P(2,1) = \frac{1}{6}$. We now show that under these conditions, *i.e.* identical measurement devices and shared pairs of classical two-level systems, target correlation in $Corr_3[2,2,3]$ cannot be simulated.

Corollary 4.2. *The correlation in* $Cor_2[2,2,3]$ *cannot be simulated by sharing bipartite classical states* $\omega_{AB}^{Cl} \otimes \tilde{\omega}_{AB}^{Cl}$*, where each subsystem has local dimension* 2*, when the parties use identical classical measurements. The same correlation can, however, be realised by performing identical* 3*-outcome qubit nPS measurements on a shared two-qubit state.*

Proof. The target correlation $P(0,0) = P(1,1) = P(2,2) = 0,\ P(0,1) = P(0,2) = P(1,0) = P(1,2) = P(2,0) = P(2,1) = \frac{1}{6}$ can be realised using shared state $|\Psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} -$

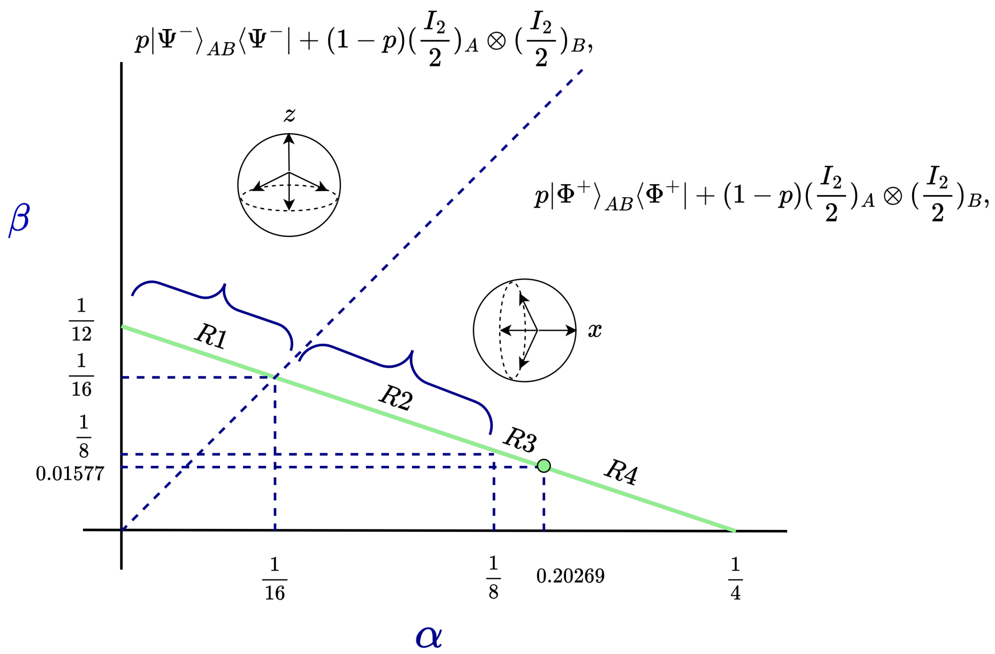

Figure 4.3: Correlations in $Cor_2[2,2,4]$ that can be simulated using 4-outcome qubit nPS measurements. For $P = \{P(a,b)\}_{a,b=0}^4 \in Cor_2[2,2,4]$, let $P(0,0) = \alpha$ and $P(0,1) = \beta$. Correlations with α in region $R1$ can be simulated using a qubit SIC-POVM on the shared state $(\rho_p)_{AB}$, while correlations with α in $R2$ can be simulated using a rotated qubit SIC-POVM on $(\tilde{\rho}_p)_{AB}$. Correlations with α in region $R4$ cannot be obtained using local qubit POVMs on any shared two-qubit state. Whether correlations with α in region $R3$ belong to $\mathfrak{C}_2^Q(4,2)$ is unknown.

$|10\rangle_{AB}$) when A and B perform 3-outcome qubit trine measurements (see proof of Theorem 4.7).

Using Lemma 4.1, we may assume *wlog* that the pair of two-level classical states is $\omega_{AB}^{Cl} \otimes \tilde{\omega}_{AB}^{Cl} = (p_{0,0} = p, p_{0,1} = 0, p_{1,0} = 0, p_{1,1} = 1-p) \otimes (p'_{0,0} = p', p'_{0,1} = 0, p'_{1,0} = 0, p'_{1,1} = 1-p') = (pp', 0, 0, 0, 0, p(1-p'), 0, 0, 0, 0, (1-p)p', 0, 0, 0, 0, (1-p)(1-p'))$ where $p, p' \in [0,1]$. If $p = 0$ or 1, the state ω_{AB}^{Cl} can be locally prepared. From Lemma 4.2, classical state $\tilde{\omega}_{AB}^{Cl}$ cannot simulate the target correlation. Hence, using $\omega_{AB}^{Cl} \otimes \tilde{\omega}_{AB}^{Cl}$ with $p \in \{0,1\}$ and $p' \in [0,1]$ cannot reproduce it either. By symmetry, the same holds if $p' \in \{0,1\}$. Thus, both parameters must satisfy $p, p' \in (0,1)$. Since the measurement devices are identical, both can be described by a column-stochastic map $W_A^{4\to3} = W_B^{4\to3} = W^{4\to3}$ with entries $(W^{4\to3})_{r,s} = w_{r,s}$ where $r \in \{0,1,2\}$, $s \in \{0,1,2,3\}$ and they satisfy

$$\sum_{r=0}^{2} w_{r,s} = 1 \;\forall\; s \in \{0,1,2,3\} \text{ and } w_{r,s} \geq 0 \;\forall\; r \in \{0,1,2\}, s \in \{0,1,2,3\} \tag{4.38}$$

The correlation $P = \{P(a,b)\}_{a,b=0}^{2}$ generated is

$$P = (W^{4\to3} \otimes W^{4\to3})(\omega_{AB}^{Cl} \otimes \tilde{\omega}_{AB}^{Cl})^T \tag{4.39}$$

Taking $p, p' \in (0,1)$ and solving for $P(0,0) = 0$ using expression from Equation (4.39), we obtain $w_{0,0} = w_{0,1} = w_{0,2} = w_{0,3} = 0$. Similarly, solving for $P(1,1) = P(2,2) = 0$ yields $w_{r,s} = 0$ for all r, s. This contradicts the normalisation condition in Equation (4.38). Hence, the target correlation cannot be simulated in this setting. ■

4.3 Detecting Qutrit Non-Projective-Simulable Measurements

We now discuss detection schemes for 5-outcome qutrit nPS measurements. We begin with the bipartite scenario $[n = 2, d = 3, k = 5]$, assuming that the measurement devices of both parties are identical. In this case, we present a correlation simulation task that serves as a detection scheme for 5-outcome qutrit nPS measurements and provide numerical evidence showing that the scheme is robust against arbitrary depolarising noise. Next, we consider both bipartite $[n = 2, d = 3, k = 5]$ and tripartite $[n = 3, d = 3, k = 5]$ scenarios, without assuming that the measurement devices are identical. Here, we define tasks using payoff functions. We numerically compute the projective-simulable bounds on these payoff functions and show that they can be violated using qutrit POVMs on shared qutrit states. In general, as the local dimension d increases, the number of parameters associated with correlations generated by arbitrary projective-simulable measurements grows. As a result, determining whether a given target correlation can be realised by projective-simulable measurements becomes computationally demanding. The same difficulty also exists while optimising for the projective-simulable bound of some payoff function used to define a task.

4.3.1 Detecting 5-outcome Qutrit nPS Measurements with $\mathbb{G}^{sym}[2,3,5]$

Consider the scenario $[n = 2, d = 3, k = 5]$ with two spatially separated parties, A and B. A source $\mathcal{P}_{A,B}$ prepares a two-qutrit state $\rho_{AB} \in D(\mathbb{C}^3 \otimes \mathbb{C}^3)$ which is shared between them. Each party has access to a measurement device described by a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^3) : E_a \geq 0\}_{a=0}^{4}$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^3) : F_b \geq 0\}_{b=0}^{4}$ for B. We additionally assume that the devices of both parties are identical. For this task $\mathbb{G}^{sym}[2,3,5]$, the target correlation set is $Cor_1[2,3,5]$, which contains a single correlation $\{P(a,b)\}_{a,b=0}^{4}$ defined as

$a\backslash b$	0	1	2	3	4
0	$\frac{1-\epsilon^2\nu}{9}$	$\frac{2+\epsilon^2\nu}{36}$	$\frac{2+\epsilon^2\nu}{36}$	$\frac{2+\epsilon\nu+3\epsilon^2\nu}{36}$	$\frac{2-\epsilon\nu-\epsilon^2\nu}{36}$
1	$\frac{2+\epsilon^2\nu}{36}$	$\frac{1-\epsilon^2\nu}{36}$	$\frac{2+\epsilon^2\nu}{72}$	$\frac{4+2\epsilon\nu-3\epsilon^2\nu}{144}$	$\frac{4-2\epsilon\nu+\epsilon^2\nu}{144}$
2	$\frac{2+\epsilon^2\nu}{36}$	$\frac{2+\epsilon^2\nu}{72}$	$\frac{1-\epsilon^2\nu}{36}$	$\frac{4+2\epsilon\nu-3\epsilon^2\nu}{144}$	$\frac{4-2\epsilon\nu+\epsilon^2\nu}{144}$
3	$\frac{2+\epsilon\nu+3\epsilon^2\nu}{36}$	$\frac{4+2\epsilon\nu-3\epsilon^2\nu}{144}$	$\frac{4+2\epsilon\nu-3\epsilon^2\nu}{144}$	$\frac{1+\epsilon\nu-2\epsilon^2\nu}{36}$	$\frac{2+\epsilon^2\nu}{72}$
4	$\frac{2-\epsilon\nu-\epsilon^2\nu}{36}$	$\frac{4-2\epsilon\nu+\epsilon^2\nu}{144}$	$\frac{4-2\epsilon\nu+\epsilon^2\nu}{144}$	$\frac{2+\epsilon^2\nu}{72}$	$\frac{1-\epsilon\nu}{36}$

Table 4.7: Correlations $\{P(a,b)\}_{a,b=0}^4 \in Cor_2[2,3,5]$, where $\nu, \epsilon \in (0,1]$.

$$
\begin{aligned}
&P(0,0) = P(1,1) = P(2,2) = P(3,3) = P(4,4) = P(0,4) = P(4,0) = 0\\
&P(0,1) = P(0,2) = P(1,0) = P(2,0) = \frac{1}{12}, \qquad P(0,3) = P(3,0) = \frac{1}{6}\\
&P(1,2) = P(2,1) = P(3,4) = P(4,3) = \frac{1}{24}\\
&P(1,3) = P(1,4) = P(2,3) = P(2,4) = P(3,1) = P(3,2) = P(4,1) = P(4,2) = \frac{1}{48}
\end{aligned}
\tag{4.40}
$$

The set of correlations $Cor_2[2,3,5]$ is shown in Table 4.7, where $\nu, \epsilon \in (0,1]$. The correlation $\{P(a,b)\}_{a,b=0}^4 \in Cor_1[2,3,5]$ is a special case of this set, obtained by substituting $\nu = \epsilon = 1$. We will first show that the target correlation $Cor_1[2,3,5]$ cannot be generated using two identical 5-outcome qutrit projective-simulable measurements on a shared two-qutrit state.

Theorem 4.10. *In the task $\mathbb{G}^{sym}[2,3,5]$ defined by the set $Cor_1[2,3,5]$, the target correlation in $Cor_1[2,3,5]$ cannot be generated using identical 5-outcome qutrit projective-simulable measurements on a shared two-qutrit state.*

Proof. [8] As a consequence of Theorem 4.1, any correlation generated by 5-outcome qutrit projective-simulable measurements on a two-qutrit shared state can also be obtained from a shared three-level classical system. We represent the shared classical state as $\omega_{A,B}^{Cl} = (p_{i,j})_{i,j\in\{0,1,2\}}$ with $\sum_{i,j=0}^2 p_{i,j} = 1$ and $p_{i,j} \geq 0$ for all $i,j \in \{0,1,2\}$. Since the measurements of A and B are identical, both are described by a column-stochastic map $W_A^{3\to5} = W_B^{3\to5} = W^{3\to5}$ with entries $(W^{3\to5})_{r,s} = w_{r,s}$ where $r \in \{0,1,2,3,4\}, s \in \{0,1,2\}$. They satisfy

$$
\sum_{r=0}^{4} w_{r,s} = 1 \ \forall\ s \in \{0,1,2\} \text{ and } w_{r,s} \geq 0 \ \forall\ r \in \{0,1,2,3,4\}, s \in \{0,1,2\}
\tag{4.41}
$$

The resulting correlation $P = \{P(a,b)\}_{a,b=0}^4$ is given by

$$
P = (W^{3\to5} \otimes W^{3\to5})(\omega_{A,B}^{Cl})^T
\tag{4.42}
$$

[8] Also discussed in Section 4.2 of [185]

Substituting the values from Equation (4.40) into (4.42), the expressions for the correlated outcomes are:

$$
\begin{aligned}
P(0,0) &= p_{0,0}w_{0,0}^2 + p_{0,1}w_{0,1}w_{0,0} + p_{0,2}w_{0,2}w_{0,0} + p_{1,0}w_{0,1}w_{0,0} + p_{2,0}w_{0,2}w_{0,0} + \\
&\qquad p_{1,1}w_{0,1}^2 + p_{1,2}w_{0,1}w_{0,2} + p_{2,1}w_{0,1}w_{0,2} + p_{2,2}w_{0,2}^2 = 0 \\
P(1,1) &= p_{0,0}w_{1,0}^2 + p_{0,1}w_{1,1}w_{1,0} + p_{0,2}w_{1,2}w_{1,0} + p_{1,0}w_{1,1}w_{1,0} + p_{2,0}w_{1,2}w_{1,0} + \\
&\qquad p_{1,1}w_{1,1}^2 + p_{1,2}w_{1,1}w_{1,2} + p_{2,1}w_{1,1}w_{1,2} + p_{2,2}w_{1,2}^2 = 0 \\
P(2,2) &= p_{0,0}w_{2,0}^2 + p_{0,1}w_{2,1}w_{2,0} + p_{0,2}w_{2,2}w_{2,0} + p_{1,0}w_{2,1}w_{2,0} + p_{2,0}w_{2,2}w_{2,0} + \\
&\qquad p_{1,1}w_{2,1}^2 + p_{1,2}w_{2,1}w_{2,2} + p_{2,1}w_{2,1}w_{2,2} + p_{2,2}w_{2,2}^2 = 0 \\
P(3,3) &= p_{0,0}w_{3,0}^2 + p_{0,1}w_{3,1}w_{3,0} + p_{0,2}w_{3,2}w_{3,0} + p_{1,0}w_{3,1}w_{3,0} + p_{2,0}w_{3,2}w_{3,0} + \\
&\qquad p_{1,1}w_{3,1}^2 + p_{1,2}w_{3,1}w_{3,2} + p_{2,1}w_{3,1}w_{3,2} + p_{2,2}w_{3,2}^2 = 0 \\
P(4,4) &= p_{0,0}w_{4,0}^2 + p_{0,1}w_{4,1}w_{4,0} + p_{0,2}w_{4,2}w_{4,0} + p_{1,0}w_{4,1}w_{4,0} + p_{2,0}w_{4,2}w_{4,0} + \\
&\qquad p_{1,1}w_{4,1}^2 + p_{1,2}w_{4,1}w_{4,2} + p_{2,1}w_{4,1}w_{4,2} + p_{2,2}w_{4,2}^2 = 0
\end{aligned}
\tag{4.43}
$$

The terms in Equations (4.43) are non-negative individually. Thus, every term must be zero for the correlated outcomes to be zero. This implies $p_{i,i} = 0$ for $i \in \{0,1,2\}$. If instead $p_{i,i} \neq 0$ for some $i \in \{0,1,2\}$, then $w_{j,i} = 0$ for all $j \in \{0,1,2,3,4\}$, contradicting the normalisation $\sum_j w_{j,i} = 1$ in Equation (4.41). Thus, $p_{i,i} = 0$ for $i \in \{0,1,2\}$. Substituting $p_{i,i} = 0$ into (4.42) and equating the expressions with the values in (4.40) using a symbolic solver (Mathematica), we find that there is no solution for $\omega_{A,B}^{Cl}$ and $W^{3\to5}$ which also satisfies the constraints in Equation (4.41). ■

Theorem 4.11. *The correlation in $Cor_1[2,3,5]$ can be simulated using two identical qutrit POVMs on a shared two-qutrit state.*

Proof. [9] Let the shared two-qutrit state be $|\psi\rangle_{AB} = \frac{1}{\sqrt{6}}(|01\rangle_{AB} + |02\rangle_{AB} + |10\rangle_{AB} + |12\rangle_{AB} + |20\rangle_{AB} - |21\rangle_{AB})$. The parties perform the following POVM: $\mathcal{M}_A = \{E_a = |\psi_a\rangle\langle\psi_a|\}_{a=0}^4$ for A and $\mathcal{M}_B = \{F_b = |\psi_b\rangle\langle\psi_b|\}_{b=0}^4$, where $|\psi_0\rangle = |0\rangle, |\psi_1\rangle = \frac{1}{\sqrt{2}}|1\rangle, |\psi_2\rangle = \frac{1}{\sqrt{2}}|2\rangle, |\psi_3\rangle = \frac{1}{2}(|1\rangle + |2\rangle), |\psi_4\rangle = \frac{1}{2}(|1\rangle - |2\rangle)$. This yields the correlation in Equation (4.40). ■

Evidence of robustness against noise: We now present numerical evidence that correlations in $Cor_2[2,3,5]$, defined in Table 4.7 with $\nu, \epsilon \in (0,1]$, are not obtainable using identical 5-outcome qutrit projective-simulable measurements. For several discrete values of $\nu, \epsilon \in \{0.01, 0.02, \ldots, 0.99, 1\}$, we checked whether the corresponding correlations $P = \{P(a,b)\}_{a,b=0}^4 \in Cor_2[2,3,5]$ could be generated using identical 5-outcome qutrit projective-simulable measurements. Following the approach in Theorem 4.10, we assumed a shared classical three-level state $\omega_{A,B}^{Cl} = (p_{i,j})_{i,j\in\{0,1,2\}}$ with $\sum_{i,j=0}^2 p_{i,j} = 1$ and $p_{i,j} \geq 0$. Since the measurements

[9] Also discussed in Section 4.2 of [185]

of A and B are identical, they are described by the same column-stochastic map $W^{3\to5}$ with entries $(W^{3\to5})_{r,s} = w_{r,s}$ where $r \in \{0,1,2,3,4\}, s \in \{0,1,2\}$. The entries satisfy $\sum_{r=0}^{4} w_{r,s} = 1 \ \forall \ s \in \{0,1,2\}$ and $w_{r,s} \geq 0$. The resulting correlation $P = P(a,b)_{a,b=0}^{4}$ is given by $P = (W^{3\to5} \otimes W^{3\to5})(\omega_{A,B}^{Cl})^{T}$. Equating this expression with the corresponding values in Table 4.7, we solved the system using a symbolic solver (Mathematica) for each choice of ν and ϵ in the set $\{0.01, 0.02, 0.03, \cdots, 0.99, 1\}$. For each chosen pair (ν, ϵ), we obtain that the system of equations does not have a solution. This provides numerical evidence that no two-qutrit state and identical 5-outcome qutrit projective-simulable measurements can generate correlations in $Cor_2[2,3,5]$.

On the other hand, correlations in $Cor_2[2,3,5]$ can be realised using identical 5-outcome qutrit POVMs. Let the noisy shared state be $(\rho_p) = p\,|\psi\rangle_{AB}\langle\psi| + (1-p)(\frac{I_3}{3})_A \otimes (\frac{I_3}{3})_B$ where $|\psi\rangle_{AB} = \frac{1}{\sqrt{6}}(|01\rangle_{AB} + |02\rangle_{AB} + |10\rangle_{AB} + |12\rangle_{AB} + |20\rangle_{AB} - |21\rangle_{AB})$ and $p \in (0,1]$. The parties perform POVM: $\mathcal{M}_A = \{E_a = \Pi_a^{\epsilon}\}_{a=0}^{4}$ and $\mathcal{M}_B = \{F_b = \Pi_b^{\epsilon}\}_{b=0}^{4}$, where $\epsilon \in (0,1]$ and

$$\Pi_0^{\epsilon} = \begin{pmatrix} \frac{1+2\epsilon}{3} & 0 & 0 \\ 0 & \frac{1-\epsilon}{3} & 0 \\ 0 & 0 & \frac{1-\epsilon}{3} \end{pmatrix}, \Pi_1^{\epsilon} = \begin{pmatrix} \frac{1-\epsilon}{6} & 0 & 0 \\ 0 & \frac{1+2\epsilon}{6} & 0 \\ 0 & 0 & \frac{1-\epsilon}{6} \end{pmatrix}, \Pi_2^{\epsilon} = \begin{pmatrix} \frac{1-\epsilon}{6} & 0 & 0 \\ 0 & \frac{1-\epsilon}{6} & 0 \\ 0 & 0 & \frac{1+2\epsilon}{6} \end{pmatrix} \tag{4.44}$$

$$\Pi_3^{\epsilon} = \begin{pmatrix} \frac{1-\epsilon}{6} & 0 & 0 \\ 0 & \frac{1+\frac{\epsilon}{2}}{6} & \frac{\epsilon}{4} \\ 0 & \frac{\epsilon}{4} & \frac{1+\frac{\epsilon}{2}}{6} \end{pmatrix}, \Pi_4^{\epsilon} = \begin{pmatrix} \frac{1-\epsilon}{6} & 0 & 0 \\ 0 & \frac{1+\frac{\epsilon}{2}}{6} & \frac{-\epsilon}{4} \\ 0 & \frac{-\epsilon}{4} & \frac{1+\frac{\epsilon}{2}}{6} \end{pmatrix} \tag{4.45}$$

The resulting correlations lie in $Cor_2[2,3,5]$. In essence, these operators can be written as $\Pi_i^{\epsilon} = \kappa_i(I_3 + \epsilon \sum_{j=1}^{8} s_j G_j)$ where $\Pi_i = \kappa_i(I_3 + \sum_{j=1}^{8} s_j G_j)$ are the effects from the proof of Theorem 4.11 and $\{G_j\}_{j=1}^{8}$ are the eight 3×3 Gell-Mann matrices.

4.3.2 Detecting 5-outcome Qutrit nPS Measurements with $\mathbb{G}[2,3,5]$

Consider the scenario $[n=2, d=3, k=5]$ with two spatially separated parties A and B. A source $\mathcal{P}_{A,B}$ prepares a two-qutrit state $\rho_{AB} \in D(\mathbb{C}^3 \otimes \mathbb{C}^3)$ which is shared between them. Each party has access to a measurement device described by a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^3) : E_a \geq 0\}_{a=0}^{4}$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^3) : F_b \geq 0\}_{b=0}^{4}$ for B. We define the task $\mathbb{G}[2,3,5)]$ using the following payoff function defined for correlation $P = \{P(a,b)\}_{a,b=0}^{4} \in \mathfrak{C}(5,2)$

$$\mathcal{S}_{(2,3,5)}(P) = [\min_{\substack{a\neq b \\ (a,b)\notin\{(0,4),(4,0)\}}} P(a,b)] - \sum_{i=0}^{4} P(i,i) \tag{4.46}$$

Projective-simulable Bound for the payoff: From Theorem 4.1, $\mathfrak{C}_d^{Cl}(5,2) = \mathfrak{C}_d^{PQ}(5,2)$ and from Equation (4.11), $\mathcal{S}_{(2,3,5)}^{Cl,max} = \mathcal{S}_{(2,3,5)}^{PQ,max}$. To obtain projective-simulable bound, we optimise

$\mathcal{S}_{(2,3,5)}(P)$ over correlations in $\mathfrak{C}_d^{Cl}(5,2)$. Take a general shared classical state with local operational dimension 3: $\omega_{A,B}^{Cl} = (p_{i,j})_{i,j\in\{0,1,2\}}$ with $\sum_{i,j=0}^{2} p_{i,j} = 1$ and $p_{i,j} \geq 0$ for all i,j. The local stochastic maps are $W_A^{3\to 5}$ for A and $W_B^{3\to 5}$ for B with entries $(W_A^{3\to 5})_{r,s} = w_{r,s}$ and $(W_B^{3\to 5})_{r,s} = w'_{r,s}$, where $r \in \{0,1,2,3,4\}$, $s \in \{0,1,2\}$. The entries satisfy

$$\text{(4.47)} \qquad \sum_{r=0}^{4} w_{r,s} = 1,\ \sum_{r=0}^{4} w'_{r,s} = 1\ \forall\ s \in \{0,1,2\} \text{ and } w_{r,s} \geq 0, w'_{r,s} \geq 0\ \forall\ r,s$$

The resulting correlation $P = \{P(a,b)\}_{a,b=0}^{4}$ is given by $P = (W_A^{3\to 5} \otimes W_B^{3\to 5})(\omega_{A,B}^{Cl})^T$. The optimisation is

$$\text{(4.48)} \qquad \mathcal{S}_{(2,3,5)}^{Cl,max} = \max_{P=\{P(a,b)\}_{a,b=0}^{4}\in\mathfrak{C}_d^{Cl}(5,2)} \Big[\big[\min_{\substack{a\neq b\\(a,b)\notin\{(0,4),(4,0)\}}} P(a,b)\big] - \sum_{i=0}^{4} P(i,i)\Big]$$

A numerical optimisation under these constraints yields

$$\text{(4.49)} \qquad \mathcal{S}_{(2,3,5)}^{Cl,max} = 1.84749 \times 10^{-10} \approx 0$$

Quantum Violation of $\mathcal{S}_{(2,3,5)}^{Cl,max}$*:* A quantum strategy exceeds this bound. Consider the two-qutrit shared state $|\psi\rangle_{AB} = \frac{1}{\sqrt{6}}(|01\rangle_{AB} + |02\rangle_{AB} + |10\rangle_{AB} + |12\rangle_{AB} + |20\rangle_{AB} - |21\rangle_{AB})$. Both parties perform the POVM $\mathcal{M}_A = \{E_a = |\psi_a\rangle\langle\psi_a|\}_{a=0}^{4}$ for A and $\mathcal{M}_B = \{F_b = |\psi_b\rangle\langle\psi_b|\}_{b=0}^{4}$. Here $|\psi_0\rangle = |0\rangle, |\psi_1\rangle = \frac{1}{\sqrt{2}}|1\rangle, |\psi_2\rangle = \frac{1}{\sqrt{2}}|2\rangle, |\psi_3\rangle = \frac{1}{2}(|1\rangle + |2\rangle), |\psi_4\rangle = \frac{1}{2}(|1\rangle - |2\rangle)$. The resulting correlation $P = \{P(a,b)\}_{a,b=0}^{4}$ is given by: $P(a,b) = \mathrm{Tr}\left[(E_a \otimes F_b)|\psi\rangle_{AB}\langle\psi|\right]$. The payoff using this strategy is

$$\text{(4.50)} \qquad \mathcal{S}_{(2,3,5)}(P) = \frac{1}{48} = 0.0208 > \mathcal{S}_{(2,3,5)}^{PQ,max}.$$

Since this strategy violates $\mathcal{S}_{(2,3,5)}^{PQ,max}$, therefore, the POVMs used are qutrit nPS measurements.

4.3.3 Detecting 5-outcome Qutrit nPS Measurements with $\mathbb{G}[3,3,5]$

Consider the scenario $[n = 3, d = 3, k = 5]$ with three spatially separated parties A, B and C. A source $\mathcal{P}_{A,B,C}$ prepares a three-qutrit state $\rho_{ABC} \in D(\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3)$ which is shared between them. Each party has access to a measurement device described by a POVM: $\mathcal{M}_A = \{E_a \in \mathcal{B}(\mathbb{C}^3) : E_a \geq 0\}_{a=0}^{4}$ for A and $\mathcal{M}_B = \{F_b \in \mathcal{B}(\mathbb{C}^3) : F_b \geq 0\}_{b=0}^{4}$ for B and $\mathcal{M}_C = \{G_c \in \mathcal{B}(\mathbb{C}^3) : G_c \geq 0\}_{c=0}^{4}$. We define the task $\mathbb{G}[3,3,5)]$ using the following payoff function defined for correlation $P = \{P(a,b,c)\}_{a,b,c=0}^{4} \in \mathfrak{C}(5,3)$

$$\text{(4.51)} \qquad \mathcal{S}_{(3,3,5)}(P) = \min_{(a,b,c)\in\mathcal{Z}} P(a,b,c)$$

where $\mathcal{Z} = \{(a,b,c) : a \neq b, c = 0\} \bigcup \{(a,b,c) : a \neq c, b = 0\} \bigcup \{(a,b,c) : b \neq c, a = 0\}$

Projective-simulable Bound for the payoff: From Theorem 4.1, $\mathfrak{C}_d^{Cl}(5,3) = \mathfrak{C}_d^{PQ}(5,3)$ and from Equation (4.11), $\mathcal{S}_{(3,3,5)}^{Cl,max} = \mathcal{S}_{(3,3,5)}^{PQ,max}$. To obtain $\mathfrak{C}_d^{PQ}(5,3)$, we optimise $\mathcal{S}_{(3,3,5)}(P)$ over correlations in $\mathfrak{C}_d^{Cl}(5,3)$. Take a general tripartite classical state with local operational dimension 3: $\omega_{A,B,C}^{Cl} = (p_{i,j,l})_{i,j,l\in\{0,1,2\}}$ with $\sum_{i,j,l=0}^{2} p_{i,j,l} = 1$ and $p_{i,j,l} \geq 0$ for all i,j,l. The local stochastic maps are $W_A^{3\to 5}$ for A, $W_B^{3\to 5}$ for B and $W_C^{3\to 5}$ for C with entries $(W_A^{3\to 5})_{r,s} = w_{r,s}$, $(W_B^{3\to 5})_{r,s} = w'_{r,s}$ and $(W_C^{3\to 5})_{r,s} = w''_{r,s}$, where $r \in \{0,1,2,3,4\}$, $s \in \{0,1,2\}$. The entries satisfy

(4.52)
$$\sum_{r=0}^{4} w_{r,s} = 1, \sum_{r=0}^{4} w'_{r,s} = 1, \sum_{r=0}^{4} w''_{r,s} = 1 \ \forall\ s \in \{0,1,2\} \text{ and } w_{r,s} \geq 0, w'_{r,s} \geq 0, w''_{r,s} \geq 0 \ \forall\ r,s$$

The resulting correlation $P = \{P(a,b,c)\}_{a,b,c=0}^{4}$ is given by $P = (W_{3\to 5}^{A} \otimes W_{3\to 5}^{B} \otimes W_{3\to 5}^{C})(\omega_{A,B,C}^{Cl})^T$. The optimisation is thus

$$\mathcal{S}_{(3,3,5)}^{Cl,max} = \max_{P=\{P(a,b,c)\}_{a,b,c=0}^{4} \in \mathfrak{C}_d^{Cl}(5,3)} \left[\min_{(a,b,c)\in\mathcal{Z}} P(a,b,c) \right] \tag{4.53}$$

A numerical optimisation under these constraints yields

$$\mathcal{S}_{(3,3,5)}^{Cl,max} = 0.015888 \tag{4.54}$$

Quantum Violation of $\mathcal{S}_{(3,3,5)}^{Cl,max}$*:* A quantum strategy exceeds this bound. Consider the two-qutrit shared state $|\psi\rangle_{ABC} = \frac{1}{\sqrt{6}}(|012\rangle_{ABC} + |120\rangle_{ABC} + |201\rangle_{ABC} - |021\rangle_{ABC} - |102\rangle_{ABC} - |210\rangle_{ABC})$. Both parties perform the POVM $\mathcal{M}_A = \{E_a = \frac{1}{2}\Pi_a\}_{a=0}^{4}$ for A, $\mathcal{M}_B = \{F_b = \frac{1}{2}\Pi_b\}_{b=0}^{4}$ for B and $\mathcal{M}_C = \{G_c = \frac{1}{2}\Pi_c\}_{c=0}^{4}$ for C, where $\Pi_0 = |0\rangle\langle 0|, \Pi_1 = |1\rangle\langle 1|, \Pi_2 = |2\rangle\langle 2|, \Pi_3 = (|1\rangle + |2\rangle)(\langle 1| + \langle 2|), \Pi_4 = (|1\rangle - |2\rangle)(\langle 1| - \langle 2|)$. The resulting correlation $P = \{P(a,b,c)\}_{a,b,c=0}^{4}$ is given by: $P(a,b,c) = \mathrm{Tr}\left[(E_a \otimes F_b \otimes G_c)|\psi\rangle_{ABC}\langle\psi|\right]$. The payoff using this strategy is

$$\mathcal{S}_{(3,3,5)}(P) = 0.020833 > \mathcal{S}_{(3,3,5)}^{PQ,max} \tag{4.55}$$

Since this strategy violates $\mathcal{S}_{(3,3,5)}^{PQ,max}$, therefore, the POVMs are qutrit nPS measurements.

4.4 Analogue of Non-Projective-Simulable Measurements for GPTs

Similar to the notion of projective-simulable measurements in quantum theory, we introduce a class of measurements in GPTs that we call *sharp simulable measurements*. Consider a theory **TH** with a system $Sys_A(\Omega_A, \mathcal{E}_A)$ of type A having operational dimension $\mathbf{d}_A$.

Definition 4.1. A $\mathbf{d}_A$-outcome measurement with effects $\{e_i \in \mathcal{E}_A\}_{i=1}^{\mathbf{d}_A}$, satisfying $\sum_{i=1}^{\mathbf{d}_A} e_i = u_A$, is called sharp if all the effects are extremal points of the effect space $\mathcal{E}_A$.

Sharp measurements generalise the role of projective measurements in quantum theory. Note that in other works, "sharp" is sometimes used to denote repeatable or minimally disturbing measurements as well. When the operational dimension $\mathbf{d}_A$ of a system $Sys_A(\Omega_A, \mathcal{E}_A)$ is known, we can define measurements with $N(> \mathbf{d}_A)$ outcomes obtained by post-processing the outcome probabilities of a sharp $\mathbf{d}_A$-outcome measurement.

Definition 4.2. Given a system $Sys_A(\Omega_A, \mathcal{E}_A)$ with operational dimension $\mathbf{d}_A$, an N-outcome measurement with $N > \mathbf{d}_A$ and effects $\{e_i \in \mathcal{E}_A\}_{i=1}^N$, satisfying $\sum_{i=1}^N e_i = u_A$, is called sharp simulable if there exists a sharp measurement $\{f_j \in \mathcal{E}_A\}_{j=1}^{\mathbf{d}_A}$, with $\sum_{j=1}^{\mathbf{d}_A} f_j = u_A$ and probability distribution $\{P(i|j)\}_{i=1}^N$ for $j \in [\mathbf{d}_A]$ such that $e_i = \sum_{j=1}^{\mathbf{d}_A} P(i|j) f_j$.

In quantum theory, for a system described by Hilbert space $\mathbb{C}^d$, sharp simulable measurements correspond to projective-simulable measurements (Definition 2.1). Any N-outcome measurement with $N > \mathbf{d}_A$ for the system $Sys_A(\Omega_A, \mathcal{E}_A)$ that cannot be expressed in the form of Definition 4.2 is called *not-sharp simulable.*

Now consider the box-world theory $\mathbf{TH} = \mathfrak{B}$ for instance, which is described in Section 2.2.3. We show that a task defined in the setup of Section 4.1.1 can be used to detect not-sharp simulable measurements in this theory. In particular, we use the setup $[n = 2, d = 2, k = 3]$, which involves two parties, A and B, sharing states with local operational dimension 2. Each party has a measurement device producing outcomes $a \in \{0, 1, 2\}$ for A and $b \in \{0, 1, 2\}$ for B. Recall that in box-world $\mathfrak{B}$, the state space Ω_A of an elementary system $Sys_A(\Omega_A, \mathcal{E}_A)$ is defined as $\Omega_A = Conv\{\omega_1, \omega_2, \omega_3, \omega_4\} \subset \mathbb{R}^3$ (see Equation (2.4)) and effect space $\mathcal{E}_A = Conv\{\mathbb{O}, u_A, e_1, e_2, e_3, e_4\} \subset \mathbb{R}^3$ (see Equations (2.5) and (2.6)). An elementary box-world system has operational dimension 2 and two sharp measurements: $\{e_1, e_3\}$ and $\{e_2, e_4\}$. A composite system $Sys_{AB}(\Omega_{AB}, \mathcal{E}_{AB})$ consists of two elementary systems $Sys_A(\Omega_A, \mathcal{E}_A)$ and $Sys_B(\Omega_B, \mathcal{E}_B)$. The state space is $\Omega_{AB} = Conv\{(\omega_i)_{AB} : i \in [24]\}$ where product extremal states $\{(\omega_i)_{AB}\}_{i=1}^{16}$ are given in Equation (2.7) and entangled states $\{(\omega_i)_{AB}\}_{i=17}^{24}$ are given in Equation (2.8). The effect space $\mathcal{E}_{AB} = \{(e)_{AB} = (e)_A \otimes (e)_B^T : (e)_A \in \mathcal{E}_A, (e)_B \in \mathcal{E}_B\}$.

In the scenario $[n = 2, d = 2, k = 3]$, analogous to the sets $\mathfrak{C}_2^Q(3, 2)$ and $\mathfrak{C}_2^{Cl}(3, 2)$ we define $\mathfrak{C}_2^{PR}(3, 2)$ as follows. It is the set of all joint probability distributions $P = \{P(a, b)\}_{a,b=0}^2$ achievable in the PR-model with local operational dimension 2:

Definition 4.3. $\mathfrak{C}_2^{PR}(3, 2)$ is the set of correlations $P = \{P(a, b)\}_{a,b=0}^2$ of the form

$$P(a, b) = \mathrm{Tr}\Big[(f_a)_A^T \otimes (g_b)_B \cdot \omega_{AB}\Big], a, b \in \{0, 1, 2\} \tag{4.56}$$

where $\omega_{AB} \in \Omega_{AB}$ is shared state, and the measurements of the parties are $\{(f_a)_A \in \mathcal{E}_A\}_{a=0}^2$, $\{(g_b)_B \in \mathcal{E}_B\}_{b=0}^2$, where $\sum_{a\in\{0,1,2\}} (f_a)_A = u_A$, $\sum_{b\in\{0,1,2\}} (g_b)_B = u_B$.

The set of sharp-simulable correlations in the PR-model is defined as follows

Definition 4.4. $\mathfrak{C}_2^{SPR}(3,2)$ is the set of correlations $P = \{P(a,b)\}_{a,b=0}^2$ of the form

$$P(a,b) = \mathrm{Tr}\Big[(f_a)_A^T \otimes (g_b)_B \cdot \omega_{AB}\Big], a, b \in \{0,1,2\} \tag{4.57}$$

where $\omega_{AB} \in \Omega_{AB}$ is the shared state, and the measurements of the parties are $\{(f_a)_A = \sum_{i\in\{0,1\}} p(a|i)\ (h_i)_A\}_{a=0}^2$, $\{(g_b)_B = \sum_{j\in\{0,1\}} p'(b|j)\ (h'_j)_B\}_{b=0}^2$. $\{p(a|i)\}_{a=0}^2$ and $\{p'(b|j)\}_{b=0}^2$ are a probability distribution for $i, j \in \{0,1\}$ and $\{(h_0)_A, (h_1)_A\}$ is a sharp measurement for system A and $\{(h'_0)_B, (h'_1)_B\}$ is a sharp measurement for system B. Here, $\sum_a (f_a)_A = u_A$, $\sum_b (g_b)_B = u_B$.

For A, there are two sharp measurements: $\{(e_1)_A, (e_3)_A\}$ and $\{(e_2)_A, (e_4)_A\}$. Similarly, B has two sharp measurements $\{(e_1)_B, (e_3)_B\}$ and $\{(e_2)_B, (e_4)_B\}$. We now show that there exist correlations in $\mathfrak{C}_2^{PR}(3,2)$ which cannot be obtained using any 3-outcome sharp simulable measurement on a shared bipartite state in box-world with local dimension 2.

Theorem 4.12. $\mathfrak{C}_2^{SPR}(3,2) \subsetneq \mathfrak{C}_2^{PR}(3,2)$.

Proof. We first show that $\mathfrak{C}_2^{Cl}(3,2) = \mathfrak{C}_2^{SPR}(3,2)$ and then prove that $\mathfrak{C}_2^{Cl}(3,2) \subsetneq \mathfrak{C}_2^{PR}(3,2)$. A classical system of operational dimension 2, *i.e.* a bit, can be embedded in a box-world elementary system with state space $\tilde{\Omega} := Conv\{\omega_1, \omega_3\}$ and the measurements being the post-processings of $\{e_1, e_3\}$. To obtain a correlation in $\mathfrak{C}_2^{Cl}(3,2)$, suppose the shared classical state is $\omega_{A,B}^{Cl} = (p_{i,j})_{i,j\in\{0,1\}}$ with $\sum_{i,j=0}^1 p_{i,j} = 1$ and $p_{i,j} \geq 0$ for all i, j. And the local stochastic maps are $W_A^{2\to 3}$ for A and $W_B^{2\to 3}$ for B with entries $(W_A^{2\to 3})_{r,s} = w_{r,s}$ and $(W_B^{2\to 3})_{r,s} = w'_{r,s}$, where $r \in \{0,1,2\}, s \in \{0,1\}$. Here $\sum_{r=0}^2 w_{r,s} = 1, \sum_{r=0}^2 w'_{r,s} = 1\ \forall\ s \in \{0,1\}$ and $w_{r,s} \geq 0, w'_{r,s} \geq 0$. The same correlation can be reproduced in box-world using the bipartite state $\omega_{AB} = p_{0,0}(\omega_1)_A \otimes (\omega_2)_B^T + p_{0,1}(\omega_1)_A \otimes (\omega_4)_B^T + p_{1,0}(\omega_3)_A \otimes (\omega_2)_B^T + p_{1,1}(\omega_3)_A \otimes (\omega_4)_B^T$ with local sharp measurements $\{(h_0)_A = e_1, (h_1)_A = e_3\}$ by A and $\{(h'_0)_B = e_2, (h'_1)_B = e_4\}$ by B followed by classical post-processing of outcomes: $W_A^{2\to 3}$ for A and $W_B^{2\to 3}$ for B. Thus, $\mathfrak{C}_2^{Cl}(3,2) \subseteq \mathfrak{C}_2^{SPR}(3,2)$.

Consider any joint probability distribution $\{P(a,b)\}_{a,b=0}^2$ obtained from a bipartite box-world state ω_{AB} and sharp-simulable measurements, say $\{(f_a)_A = \sum_{i\in\{0,1\}} p(a|i)\ (h_i)_A\}_{a=0}^2$ for A and $\{(g_b)_B = \sum_{j\in\{0,1\}} p'(b|j)\ (h'_j)_B\}_{b=0}^2$ for B. Here, $\{(h_0)_A, (h_1)_A\}$ be the sharp measurement for A and $\{(h'_0)_B, (h'_1)_B\}$ be the sharp measurement for B. Then the correlation $\{P(a,b)\}_{a,b=0}^2$ can be expressed as

$$P(a,b) = \mathrm{Tr}\Big[(f_a)_A^T \otimes (g_b)_B \cdot \omega_{AB}\Big] \tag{4.58}$$

$$= \sum_{i,j\in\{0,1\}} p(a|i)\ p'(b|j)\ \mathrm{Tr}\Big[(h_i)_A^T \otimes (h'_j)_B \cdot \omega_{AB}\Big] \tag{4.59}$$

Take bipartite classical state $\omega_{AB}^{Cl} = (p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1})$ with $p_{i,j} = \mathrm{Tr}\Big[(h_i)_A^T \otimes (h'_j)_B \cdot \omega_{AB}\Big]$ for $i, j \in \{0,1\}$ and local column stochastic maps $W_A^{2\to 3}$ for A and $W_B^{2\to 3}$ for B with entries

$(W_A^{2\to 3})_{r,s} = w_{r,s} = p(r|s)$ and $(W_B^{2\to 3})_{r,s} = w'_{r,s} = p'(r|s)$, where $r \in \{0,1,2\}, s \in \{0,1\}$. This classical state and measurement gives the same probability distribution as in the Equation (4.59). Thus, $\mathfrak{C}_2^{SPR}(3,2) \subseteq \mathfrak{C}_2^{Cl}(3,2)$.

We will show that $\mathfrak{C}_2^{Cl}(3,2) \subset \mathfrak{C}_2^{PR}(3,2)$. Consider the bipartite shared state $\omega_{AB} = \frac{p_1}{2}\left((\omega_6)_{AB} + (\omega_{16})_{AB}\right) + \frac{1-p_1}{2}\left((\omega_{22})_{AB} + (\omega_{23})_{AB}\right)$ and local measurements $\{(f_0)_A = \frac{1}{6}(e_1)_A + \frac{1}{2}(e_4)_A, (f_1)_A = \frac{1}{3}(e_1)_A + \frac{1}{3}(e_2)_A, (f_2)_A = \frac{1}{6}(e_2)_A + \frac{1}{2}(e_3)_A\}$ for A and $\{(g_0)_B = \frac{1}{2}(e_2)_B + \frac{1}{6}(e_3)_B, (g_1)_B = \frac{1}{3}(e_3)_B + \frac{1}{3}(e_4)_B, (g_2)_B = \frac{1}{2}(e_1)_B + \frac{1}{6}(e_4)_B$ for B. The correlation $P = \{P(a,b)\}_{a,b=0}^2$ obtained using the state and measurements is given by $P(a,b) = \mathrm{Tr}\Big[(f_a)_A^T \otimes (g_b)_B \cdot \omega_{AB}\Big]$.

Consider the payoff function $\mathcal{S}_{(2,2,3)} = \min_{\substack{a,b\in\{0,1,2\}\\ a\neq b}} P(a,b)$, defined over the observed joint probability distribution $\{P(a,b)\}_{a,b=0}^2$. Substituting the value $p_1 = 0.07$ in the shared state ω_{AB} defined above, leads to a correlation $\{P(a,b)\}_{a,b=0}^2$ for which $\mathcal{S}_{(2,2,3)}(P) = 0.14$. This exceeds the classical bound on the payoff $\mathcal{S}_{(2,2,3)}^{Cl,max} = \mathcal{S}_{(2,2,3)}^{PQ,max} = \frac{1}{8}$ (using Theorem 4.3 and Equation (4.11)). Thus, $\mathfrak{C}_2^{Cl}(3,2) \subsetneq \mathfrak{C}_2^{PR}(3,2)$.

■

Using an operational task, we show that the box-world theory is unphysical. Consider the setup $[n=2, d=2, k=3]$ with two parties, A and B, sharing states of local operational dimension 2. Each party has a measurement device producing outcomes $a \in \{0,1,2\}$ for A and $b \in \{0,1,2\}$ for B. For this task, we use the payoff function defined in Equation (4.12): $\mathcal{S}_{(2,2,3)}(P) = \min_{\substack{a,b\\ a\neq b}} P(a,b)$ for correlations $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}(3,2)$

Theorem 4.13. *There exist a correlation $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}_2^Q(3,2)$ such that $P \notin \mathfrak{C}_2^{PR}(3,2)$.*

Proof. From Theorem 4.3, there is a correlation $P = \{P(a,b)\}_{a,b=0}^2 \in \mathfrak{C}_2^Q(3,2)$ achieving $\mathcal{S}_{(2,2,3)}(P) = \frac{1}{6}$. We now optimise the payoff over correlations in $\mathfrak{C}_2^{PR}(3,2)$. This is done by considering all states in the convex state space $\omega_{AB} \in \Omega_{AB}$ and measurements $\{(f_a)_A \in \mathcal{E}_A\}_{a=0}^2$, $\{(g_b)_B \in \mathcal{E}_B\}_{b=0}^2$ with $\sum_{a\in\{0,1,2\}}(f_a)_A = u_A$, $\sum_{b\in\{0,1,2\}}(g_b)_B = u_B$. Numerical maximisation gives $\mathcal{S}_{(2,2,3)}^{PR,max} \approx 0.156 < \frac{1}{6}$. ■

We present a class of operational tasks based on simulating specific target correlations or achieving a payoff function. These tasks require no seed randomness and only assume a bound on the local dimension of the shared system. We first establish the equivalence between correlations generated by qudit projective-simulable measurements on a shared system of local dimension d and those obtained from a pre-shared classical d-level system. In the bipartite setting, we then provide detection schemes for three- and four-outcome qubit non-projective-simulable (nPS) measurements, which remain robust against arbitrary noise. We extend the discussion to five-outcome qutrit nPS measurements, both in bipartite and tripartite scenarios. For the case of uncharacterised devices, we numerically derive projective-simulable bounds on payoff functions and demonstrate their violation using qutrit POVMs on shared qutrit states. When the measurement devices are assumed to be identical, we show the impossibility of simulating certain target correlations with two-qutrit projective-simulable measurements. These correlations can then serve as detection schemes for qutrit nPS measurements. Finally, we show that some of these tasks can also rule out hypothetical theories without input randomness. For instance, the square-bit (box-world) theory, which generates PR correlations in Bell scenarios, is rendered unphysical if a payoff beyond a certain threshold is observed in the corresponding task.

Chapter 5

Correlation Assisted Classical Communication

Chapter Note

This chapter is partially based on work which was done in collaboration with Dr. Anubhav Chaturvedi, Dr. Some Sankar Bhattacharya and Prof. Dr. hab Paweł Horodecki [2]. A part of the chapter is based on unpublished results.

In Chapter 3, we analysed the advantage of quantum communication over classical communication in a one-way PM scenario. Previous works have also shown the advantage of quantum shared resources in assisting quantum communication. For example, in the PM scenario with one-way communication and shared entanglement, two-bit information can be transmitted by sending only one qubit through superdense coding [7]. Shared entanglement can also increase the capacity of quantum channels [30] and offer an advantage in other quantum communication tasks [31, 36].

Interestingly, although shared entanglement cannot be used for direct communication, it can lower the cost of classical communication in information-processing tasks or increase the probability of success. Initial demonstrations of such advantages appeared in communication complexity problems, both in bipartite [32] as well as multipartite scenarios [32–34]. These works showed that classical communication assisted by shared entanglement can outperform shared randomness-assisted classical communication. Subsequent works further explored such advantages of entanglement-assisted classical communication, often connecting them to violations of some Bell inequalities [39–44]. In a prepare-and-measure (PM) scenario, [28] showed an example of unbounded separation between one-way classical communication complexity assisted by shared entanglement and that assisted by shared randomness: while the former remained constant, the latter scaled linearly as $\Omega(n)$. Several other results have also demonstrated the advantages of shared entanglement to classical communication in the one-way PM scenario [35–38, 278–280].

Within PM scenarios involving one-way communication, several works have focused on the minimal setting where Bob has no input, which we refer to as *minimal scenario*. Here, Alice receives an input $\mathsf{m_a} \in \mathsf{M}_A$ and sends a message to Bob, who produces an output $\mathsf{n} \in \mathsf{N}$. When only shared randomness is allowed as assistance, sending a d-dimensional quantum system is equivalent to sending a d-dimensional classical system [281]. Any distribution $\{p(\mathsf{n}|\mathsf{m_a})\}_{\mathsf{n},\mathsf{m_a}}$ obtained using $\log_2 d$ qubit communication can also be realised using $\log_2 d$ bit communication. Thus, quantum advantage in this minimal PM setting must rely on non-classical resources such as entanglement. In this minimal PM scenario, works have shown the advantage of shared entanglement in assisting classical communication. For instance, [45, 46] considered correlation-assisted channel simulation. The task is to simulate a target channel, say $\bar{\mathscr{T}} : \bar{T} \to \bar{T}$, using some available resource classical channel $\mathscr{T} : T \to T$, described by the transition probabilities $\{\mathscr{T}(\tau'|\tau)\}_{\tau,\tau'}$, and shared correlations. When the target is some noiseless channel and the available classical channel is a noisy one, then the problem of zero-error simulation involves finding the maximum alphabet size (zero-error capacity) that can be communicated perfectly with a single use of the noisy channel $\mathscr{T} : T \to T$. It was shown that entanglement assistance can increase the zero-error capacity of such channels beyond what is possible with shared randomness alone. Moreover, access to post-quantum resources, such as no-signalling correlations, can lead to further improvements. These benefits also extend to asymptotic settings, where multiple independent uses of the channel are allowed [45, 46, 162]. The reverse zero-error channel simulation, where the target is a noisy channel and the available resource channel is noiseless, was also explored in this work. The goal is to find the noiseless channel with minimal capacity that suffices for such a simulation with a single use. Here, entanglement reduces the required capacity compared with shared randomness. Post-quantum no-signalling correlations require a resource channel of even lower capacity. However, these advantages vanish asymptotically, highlighting a fundamental contrast between the single-shot and asymptotic regimes. Similar to the correlation-assisted zero-error channel simulation problem, a task in the minimal one-way PM scenario was introduced in [282]. The advantage of entanglement-assisted classical communication for this task is related to quantum contextuality. This highlights a deeper foundational underpinning of quantum advantages in minimal communication scenarios. Very recently [283], it was shown that quantum and extremal no-signalling correlations can enhance the zero-error capacity of a classical channel that otherwise has zero capacity under shared randomness-assisted communication.

Recently, using a communication task defined in this minimal PM scenario [163] showed that shared randomness-assisted one-bit communication is suboptimal while correlations violating the CHSH inequality provide advantageous assistance to the one-bit classical channel. Subsequently, [165] showed that Hardy-type non-local correlations with dichotomic inputs and outputs provide an advantage over shared randomness in assisting a one-bit classical communication channel in a task. These works suggest a strong link between non-locality and the advantages achievable in correlation-assisted classical communication tasks, which we explore in this chapter.

We explore the interplay between non-locality and advantages in correlation-assisted classical communication by considering a one-way PM scenario with bounded classical communication where Bob has no input. First, we use a technique called *wire-cutting* to construct a Bell inequality tailored to a correlation-assisted bounded classical communication task. For tasks defined using a linear payoff function, we show that the advantage of entanglement assistance corresponds exactly to a violation of the associated Bell inequality, and vice versa. This connection allows us to study the advantages of non-local assistance in the communication task through the Bell inequality violations. Then we explore whether a non-local correlation can yield an advantage in some classical communication tasks. A key feature of classical channels is that the message can be read during transmission without disturbing the communication. This allows us to interpret the communicated letter as an observable and refer to this as *wire-reading*. We show that wire-reading reveals advantageous use of non-local correlations in scenarios where no such advantage exists otherwise. Specifically, including the classical message for the task uncovers correlations beyond those classically achievable, even when the set of all correlations in the initial task without wire-reading could be reproduced using shared randomness assistance to the classical channel.

Using wire-reading, we construct families of bounded classical communication tasks in the Bob-without-input PM scenario. In these tasks, non-local correlations provide an advantage, while shared randomness leads to strictly suboptimal payoffs. For each task, we derive the maximum payoff achievable with shared randomness assistance. In the first family, we prove that correlations from any non-local facet of the no-signalling polytope optimally assist the bounded classical channel by achieving the maximum payoff. However, for higher values of parameters defining the task, this advantage decreases rapidly, making it harder to demonstrate the advantage of shared entanglement. In particular, some correlations on the isotropic line connected to a non-local facet, though non-local, may fail to offer an advantage. We then propose a second family of tasks which can demonstrate the advantage of sharing some of these correlations. We introduce a class of tasks tailored to a non-local facet in the no-signalling polytope. Here, correlations in that facet are optimal and maximise the payoff. Considering tasks defined for non-local extremal or facet correlations with n inputs and dichotomic outputs, we show that all correlations on the isotropic line joining such an extremal/facet point and white noise remain resourceful as long as the noise fraction is below 0.5. Finally, we introduce a third family of tasks that highlights the role of Hardy non-local correlations with dichotomic

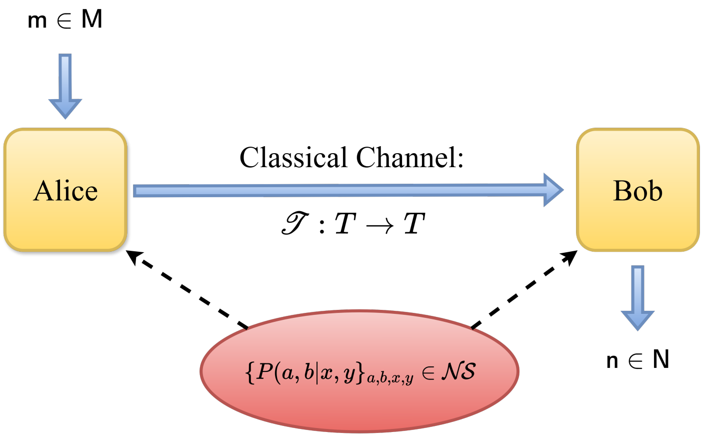

Figure 5.1: One-way prepare-and-measure (PM) scenario with two parties, Alice and Bob. They communicate through a classical channel $\mathscr{T} : T \rightarrow T$ and may share no-signalling correlations $\{P(a,b|x,y)\}_{a,b,x,y} \in \mathcal{NS}$, which can be local or non-local. Alice receives a random input $\mathsf{m} \in \mathsf{M}$, while Bob has no input and produces an output $\mathsf{n} \in \mathsf{N}$.

inputs and $d \in \mathbb{N}$ outputs in assisting a one-bit channel, when $d \geq 2$. In this case, the maximum payoff with shared randomness is zero, while Hardy correlations achieve a payoff equal to the Hardy violation. The advantage of using quantum-realisable Hardy-type correlations increases with the task parameter. We explicitly present several instances of quantum advantage for some of the tasks. For a particular task, we find that two-qutrit entangled states achieve the maximum payoff, whereas assistance from two-qubit entangled states yields a strictly lower payoff. This task can be used for the detection of the dimension of the shared entangled state.

The chapter is organised as follows. In Section 5.1, we describe the setup of correlation-assisted communication tasks and construct a Bell inequality tailored to such tasks using wire-cutting. We also introduce the notion of wire-reading. In Section 5.2, we present two families of communication tasks. The first family demonstrates the advantage of assistance from correlations belonging to any non-local facet of the no-signalling polytope. In the second family, each task is defined for a non-local facet of the polytope. In Section 5.3, we introduce a family of tasks showing the advantage of Hardy non-local correlations with dichotomic inputs in assisting a one-bit channel. Finally, in Section 5.4, we demonstrate the advantage of shared entanglement in some of these tasks.

5.1 Bell Inequalities for Classical Communication Tasks

We consider a one-way prepare-and-measure (PM) scenario with two parties, Alice and Bob. They communicate through a single use of a discrete classical channel $\mathscr{T} : T \rightarrow T$. The channel is described by $\mathscr{T}(\tau'|\tau)$, which represents the probability with which Bob receives symbol $\tau' \in T$ when Alice sends $\tau \in T$. The channel may be noisy in general. If it is noiseless, as in the case of the tasks we present later, the channel is characterised by $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$. In addition to the channel, the parties may share correlations as a resource. These can be classical (shared randomness), quantum, or even post-quantum (no-signalling). Sharing non-local correlations is regarded as costly, and we aim to show the advantage of such resources. In this scenario, Alice receives a random input $\mathsf{m} \in \mathsf{M}$, while Bob has no input and produces an output $\mathsf{n} \in \mathsf{N}$ (see Figure 5.1). We assume both M and N are finite sets. Let $\{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{m},\mathsf{n}}$ denote the conditional probability distribution of Bob's output given Alice's input. Such a distribution may be generated by local pre-processing and post-processing using the shared correlations together with communication. Different tasks can be defined based on this probability distribution. We consider tasks characterised by a success metric $\mathcal{S}$, which is a linear function of the conditional probability distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{m},\mathsf{n}}$, expressed as

$$\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n}} w_{\mathsf{n}}^{\mathsf{m}}\, p(\mathsf{n}|\mathsf{m}) \qquad \text{where} \quad w_{\mathsf{n}}^{\mathsf{m}} \in \mathbb{R}\ \forall\ \mathsf{n} \in \mathsf{N}, \mathsf{m} \in \mathsf{M} \tag{5.1}$$

We denote the communication task in this PM scenario, characterised by the payoff function in Equation (5.1), as $\mathbf{CS}_{M,N\mathscr{T}}$. Suppose the parties share a no-signalling correlation $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},y\in\mathsf{Y}} \in \mathcal{NS}$. Here, $x \in \mathsf{X}$, $y \in \mathsf{Y}$ denote inputs to the shared correlation for Alice and Bob, respectively, and $a \in \mathsf{A}, b \in \mathsf{B}$ denote the outputs of the no-signalling correlation for the respective parties (see Section 2.3). This correlation satisfies the no-signalling constraints given in Equations (2.13) and (2.14). A distribution $\{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ can be realised using the classical channel $\mathscr{T} : T \to T$ with no-signalling correlation assistance if there exists $P \in \mathcal{NS}$ and suitable strategies such that

$$p(\mathsf{n}|\mathsf{m}) = \sum_{a,b,x,y,\tau,\tau'} \Big[p(x|\mathsf{m})\, p_e(\tau|a,\mathsf{m})\, P(a,b|x,y) \mathscr{T}(\tau'|\tau)\, p(y|\tau')\, p_d(\mathsf{n}|\tau',b) \Big] \tag{5.2}$$

Here, $\{p(x|\mathsf{m})\}_{x\in\mathsf{X},\mathsf{m}\in\mathsf{M}}$ is the probability with which Alice selects input x to the no-signalling correlation upon receiving m. Similarly, $\{p(y|\tau')\}_{\tau'\in T, y\in\mathsf{Y}}$ is the probability with which Bob selects input y to the correlation after receiving symbol τ'. The set $\{p_e(\tau|a,\mathsf{m})\}_{\tau\in T, a\in\mathsf{A},\mathsf{m}\in\mathsf{M}}$ gives Alice's probability of encoding τ for the channel, based on her input m and output a from the correlation. The set $\{p_d(\mathsf{n}|\tau',b)\}_{\tau'\in T,\mathsf{n}\in\mathsf{N},b\in\mathsf{B}}$ gives Bob's probability of decoding n from his received symbol τ' and correlation output b.

Remark 5.1. *Consider a Bell scenario $\mathscr{B}_{\tilde{\mathsf{A}},\tilde{\mathsf{B}}}^{\tilde{\mathsf{X}},\tilde{\mathsf{Y}}}$ with Alice and Bob, where $\tilde{\mathsf{X}} = \mathsf{M}, \tilde{\mathsf{Y}} = T, \tilde{\mathsf{A}} = T$ and $\tilde{\mathsf{B}} = \mathsf{N}$. Thus, Alice and Bob receive inputs $m \in \mathsf{M}$ and $\tau' \in T$, and produce outputs $\tau \in T$ and $\mathsf{n} \in \mathsf{N}$, respectively. In this scenario $\mathscr{B}_{\tilde{\mathsf{A}},\tilde{\mathsf{B}}}^{\tilde{\mathsf{X}},\tilde{\mathsf{Y}}}$ suppose the parties share a no-signalling correlation $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},y\in\mathsf{Y}} \in \mathcal{NS}$ and apply the following local pre- and post-processing strategy [234, 235]. Alice queries $x \in \mathsf{X}$ to this correlation according to $\{p(x|\mathsf{m})\}_{x\in\mathsf{X},\mathsf{m}\in\mathsf{M}}$, based on her input $\mathsf{m} \in \mathsf{M}$. Similarly, Bob samples $y \in \mathsf{Y}$ according to* *distribution* $\{p(y|\tau')\}_{\tau'\in T, y\in \mathsf{Y}}$*, based on his input variables* $\tau' \in T$ *and uses it as his input to the correlation. Next, Alice post-processes her outcomes* $a \in \mathsf{A}$ *together with* m *to output* $\tau \in T$*, following distribution* $\{p_e(\tau|a,\mathsf{m})\}_{\tau\in T, a\in\mathsf{A},\mathsf{m}\in\mathsf{M}}$*. Bob post-processes his outcome* $b \in \mathsf{B}$ *together with* τ' *to output* $\mathsf{n} \in \mathsf{N}$*, following* $\{p_d(\mathsf{n}|\tau', b)\}_{\tau'\in T,\mathsf{n}\in\mathsf{N},b\in\mathsf{B}}$*, respectively. The resulting no-signalling correlation in* $\mathscr{B}^{\tilde{\mathsf{X}},\tilde{\mathsf{Y}}}_{\tilde{\mathsf{A}},\tilde{\mathsf{B}}}$ *is* $\{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\tau,\mathsf{n},\mathsf{m},\tau'}$*, which can be written as*

$$P(\tau,\mathsf{n}|\mathsf{m},\tau') = \sum_{a,b,x,y} [p(x|\mathsf{m})p_e(\tau|a,\mathsf{m})P(a,b|x,y)p(y|\tau')p_d(\mathsf{n}|\tau',b)] \tag{5.3}$$

We can rewrite Equation (5.2) using the transformed no-signalling correlation $\{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\tau,\mathsf{n},\mathsf{m},\tau'}$ given in Equation (5.3).

$$p(\mathsf{n}|\mathsf{m}) = \sum_{\tau,\tau'} \mathscr{T}(\tau'|\tau)\Big(\sum_{a,b,x,y}\Big[p(x|\mathsf{m})\,p_e(\tau|a,\mathsf{m})\,P(a,b|x,y)\,p(y|\tau')\,p_d(\mathsf{n}|\tau',b)\Big]\Big) \tag{5.4}$$

$$= \sum_{\tau,\tau'} \mathscr{T}(\tau'|\tau)P(\tau,\mathsf{n}|\mathsf{m},\tau') \tag{5.5}$$

Thus, the local pre- and post-processing maps transform the correlation $\{P(a,b|x,y)\}_{a,b,x,y}$ into the no-signalling correlation $\{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\tau,\mathsf{n},\mathsf{m},\tau'}$ (remark 5.1). The channel $\mathscr{T}$ couples Alice's post-processed outcome τ with Bob's preprocessing input τ' via the transition probability $\mathscr{T}(\tau'|\tau)$, thereby implementing a *wiring* of the effective correlation. Operationally, Alice receives input $\mathsf{m} \in \mathsf{M}$ and uses it as input to the transformed correlation $\{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\tau,\mathsf{n},\mathsf{m},\tau'}$. She sends her output $\tau \in T$ through the classical channel $\mathscr{T}$. Bob, upon receiving $\tau' \in T$, uses it as input to the correlation and outputs $\mathsf{n} \in \mathsf{N}$.

Note that Equation (5.2) is not applicable if the correlation $P = \{P(a,b|x,y)\}_{a,b,x,y}$ allows signalling from Bob to Alice, as the resulting quantity on the left may not be a valid conditional probability distribution. In such cases, the distribution $\{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ defined by the equation may fail to be a valid conditional probability distribution. To see this, consider a correlation P that is signalling from Bob to Alice. Then there exists $x^* \in \mathsf{X}$ such that for some $a \in \mathsf{A}$ there are $y', y'' \in \mathsf{Y}$ for which Alice's marginal distribution $P_A(a|x^*,y') \neq P_A(a|x^*,y'')$. From normalisation $\sum_{a\in\mathsf{A}} P_A(a|x^*,y) = 1\ \forall\ y \in \mathsf{Y}$. Define a set $\{y_a\}_{a\in\mathsf{A}}$ such that for each a, $y_a \in \mathsf{Y}$ maximises $P_A(a|x^*,y)$, *i.e.* $P_A(a|x^*,y_a) = \max_{y\in\mathsf{Y}} P_A(a|x^*,y)$. For the signalling correlation P, it is impossible that all y_a are equal to some $\tilde{y} \in \mathsf{Y}$, since that would imply $\sum_{a\in\mathsf{A}} P_A(a|x^*,y) < 1$ for some $y \in \mathsf{Y}$ where $y \neq \tilde{y}$. Moreover, $\sum_{a\in\mathsf{A}} P_A(a|x^*,y_a) > \sum_{a\in\mathsf{A}} P_A(a|x^*,y_{\tilde{a}}) = 1$, for some $\tilde{a} \in \mathsf{A}$. Now consider the PM scenario with $\mathsf{M} = \mathsf{X}$, $\mathsf{N} = \mathsf{B}$, and $\mathscr{T} : T \to T$ be a noiseless channel, *i.e.* $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$, with $T = \mathsf{Y}$. Let Alice and Bob choose their received inputs to P using $\{p(x|\mathsf{m}) = \delta_{x,\mathsf{m}}\}_{x,\mathsf{m}}$ and $\{p(y|\tau') = \delta_{y,\tau'}\}_{\tau',y}$ based on their recieved input/ message. Let Alice encode using $\{p_e(\tau|a,\mathsf{m}) = \delta_{\tau,y_a}\}_{\tau,a,\mathsf{m}}$ based on her input m and outcome a, and let Bob decode using $\{p_d(\mathsf{n}|\tau',b) = \delta_{\mathsf{n},b}\}_{\tau',\mathsf{n},b}$ based on his received message and output from P.

Substituting these into Equation (5.2) for $\mathsf{m} = x^*$ gives

$$p(\mathsf{n}|\mathsf{m} = x^*) = \sum_{\substack{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},\\ y\in\mathsf{Y},\tau,\tau'\in T}} [\delta_{x,\mathsf{m}}\, \delta_{\tau,y_a}\, P(a,b|x,y)\delta_{\tau',\tau}\, \delta_{y,\tau'}\, \delta_{\mathsf{n},b}] = \sum_{a\in\mathsf{A}} P(a,\mathsf{n}|x^*,y_a) \tag{5.6}$$

Now, we show that $\sum_{\mathsf{n}\in\mathsf{B}} p(\mathsf{n}|\mathsf{m} = x^*)$ is not normalised.

$$\sum_{\mathsf{n}\in\mathsf{B}} p(\mathsf{n}|\mathsf{m} = x^*) = \sum_{\mathsf{n}\in\mathsf{B}}\sum_{a\in\mathsf{A}} P(a,\mathsf{n}|x^*,y_a) = \sum_{a\in\mathsf{A}} P_A(a|x^*,y_a) > 1 \tag{5.7}$$

Thus, $p(\mathsf{n}|\mathsf{m} = x^*)$ is not a valid conditional probability distribution.

Suppose Alice and Bob share a local correlation $P = \{P(a,b|x,y)\}_{a,b,x,y} \in \mathcal{L}$ or, equivalently shared randomness, in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$, where $\mathcal{L}$ is the local polytope. Shared randomness allows them to coordinate their encodings and decodings by randomising over deterministic strategies. Their choices are determined by an ontic state $\lambda \in \Lambda$, distributed according to $\{p(\lambda)\}_{\lambda\in\Lambda}$. A conditional probability distribution $\mathbf{p}_\Lambda = \{p_\Lambda(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ is realisable through the classical channel $\mathscr{T} : T \to T$, described by $\mathscr{T}(\tau'|\tau)$, if there exists a distribution $\{p(\lambda)\}_{\lambda\in\Lambda}$, an encoding scheme $\{p_e(\tau|\mathsf{m},\lambda)\}_{\tau\in T,\lambda\in\Lambda,\mathsf{m}\in\mathsf{M}}$ for Alice and a decoding scheme $\{p_d(\mathsf{n}|\tau',\lambda)\}_{\mathsf{n}\in\mathsf{N},\tau'\in T,\lambda\in\Lambda}$ for Bob, such that

$$p_\Lambda(\mathsf{n}|\mathsf{m}) = \sum_{\lambda,\tau,\tau'} [p(\lambda)p_e(\tau|\mathsf{m},\lambda)\mathscr{T}(\tau'|\tau)p_d(\mathsf{n}|\tau',\lambda)] \tag{5.8}$$

We will denote a conditional probability that can be expressed as in Equation (5.8) by $\mathbf{p}_\Lambda = \{p_\Lambda(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ later on. The deterministic encoding and decoding strategies correspond to functions $\mathbb{E} : \mathsf{M} \to T$ and $\mathbb{D} : T \to \mathsf{N}$. The maximum payoff, or *local bound*, achievable in $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with shared randomness assistance to the channel is then

$$s_\Lambda = \max_{\mathbf{p}_\Lambda} \mathcal{S}(\mathbf{p}_\Lambda) = \max_{\{p_e(\tau|\lambda,\mathsf{m})\},\{p_d(\mathsf{n}|\tau',\lambda)\},\{p(\lambda)\}} \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p_\Lambda(\mathsf{n}|\mathsf{m}) \tag{5.9}$$

The payoff in Equation (5.1) is linear in the conditional probabilities $p(\mathsf{n}|\mathsf{m})$. When Alice and Bob have access to shared randomness, the local bound can be attained using an optimal deterministic encoding and decoding strategy. If the conditional probability distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ cannot be written in the form of Equation (5.8), then it is not realisable with arbitrary shared randomness assistance to the channel $\mathscr{T} : T \to T$. On the other hand, the same distribution can be realised with non-local correlation assistance if there exists $P \in \mathcal{NS}\setminus\mathcal{L}$ and suitable encoding and decoding strategies such that it can be written as in Equation (5.2). If the distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ realisable in this way achieves a payoff $\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p(\mathsf{n}|\mathsf{m}) > s_\Lambda$ in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$, then it demonstrates the advantage of non-classical correlation in assisting the channel. Next, we present a Bell inequality corresponding to the communication task using the *cutting the classical wire* procedure.

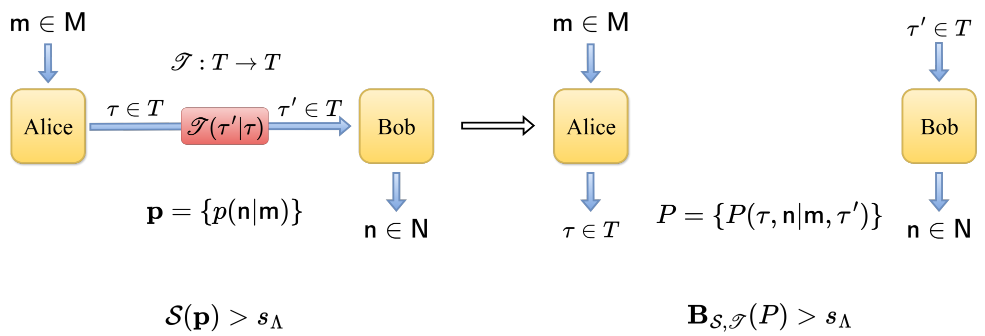

Figure 5.2: A schematic representation of the proof technique in Theorem 5.1, showing how *wire-cutting* leads to the associated Bell inequality. The task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff function $\mathcal{S}$ corresponds to the Bell functional $\mathbf{B}_{\mathcal{S},\mathscr{T}}$.

5.1.1 Cutting the classical wire

For the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff $\mathcal{S}$ defined in Equation (5.1), let the local bound s_Λ be given by Equation (5.9). To connect this task to a Bell inequality, consider a Bell scenario $\mathscr{B}^{\mathsf{M},T}_{T,\mathsf{N}}$. In this setting, Alice receives input $\mathsf{m} \in \mathsf{M}$ and outputs $\tau \in T$, while Bob receives $\tau' \in T$ and outputs $\mathsf{n} \in \mathsf{N}$. The resulting no-signalling correlation is $P = \{P(\tau, \mathsf{n}|\mathsf{m}, \tau')\}_{\tau\in T,\mathsf{n}\in\mathsf{N},\mathsf{m}\in\mathsf{M},\tau'\in T}$ as illustrated in Figure 5.2. In $\mathscr{B}^{\mathsf{M},T}_{T,\mathsf{N}}$, we the define following the Bell functional $\mathbf{B}_{\mathcal{S},\mathscr{T}}$ tailored to the communication task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$:

$$\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) = \sum_{\mathsf{m},\tau,\tau',\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} \mathscr{T}(\tau'|\tau)\; P(\tau, \mathsf{n}|\mathsf{m}, \tau') \tag{5.10}$$

The inequality $\mathcal{S}(\mathbf{p}_\Lambda) \leq s_\Lambda$ in the PM scenario then implies a Bell inequality based on Equation (5.10), where $\mathbf{p}_\Lambda = \{p_\Lambda(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ can be expressed as in Equation (5.8).

Lemma 5.1. *For the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff function $\mathcal{S}$ defined in Equation* (5.1) *and local bound s_Λ, the following Bell inequality holds for all no-signalling correlations $P = \{P(\tau, \mathsf{n}|\mathsf{m}, \tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau\in T,\tau'\in T}$ with a local hidden variable explanation:*

$$\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) \leq s_\Lambda, \tag{5.11}$$

where the Bell functional $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P)$ is given in Equation (5.10).

Proof. [1] All conditional distributions $=\{p_\Lambda(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ of the form in Equation (5.8) satisfy the inequality $\mathcal{S}(\mathbf{p}_\Lambda) \leq s_\Lambda$ for the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$. Now, any correlations $P = \{P(\tau, \mathsf{n}|\mathsf{m}, \tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau,\tau'\in T} \in \mathcal{L}$, with a *local-hidden variable* model can be written as

$$P(\tau, \mathsf{n}|\mathsf{m}, \tau') = \sum_\lambda p(\lambda) P(\tau|\mathsf{m}, \lambda) P(\mathsf{n}|\tau', \lambda) \tag{5.12}$$

[1]Also discussed as Lemma 1 in [2]

for some probability distributions $\{p(\lambda)\}_{\lambda\in\Lambda}$, $\{P(\tau|\mathsf{m},\lambda)\}_{\mathsf{m}\in\mathsf{M},\tau\in T,\lambda\in\Lambda}$, and $\{p(\mathsf{n}|\tau',\lambda)\}_{\mathsf{n}\in\mathsf{N},\tau'\in T,\lambda\in\Lambda}$. Suppose, for contradiction, that there exists a no-signalling correlation P admitting a local hidden-variable model of the form given in Equation (5.12) such that $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda$. Assume this correlation P is shared in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ defined in the PM scenario. In that case, using the channel $\mathscr{T}$, the encoding $\{p_e(\tau|\mathsf{m},\lambda) = P(\tau|\mathsf{m},\lambda)\}_{\tau,\mathsf{m},\lambda}$ for Alice, the decoding $\{p_d(\mathsf{n}|\tau',\lambda) = P(\mathsf{n}|\tau',\lambda)\}_{\mathsf{n},\tau',\lambda}$ for Bob, and shared randomness λ distributed according to $p(\lambda)$, the parties can generate a conditional distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$. It is of the same form as in Equation (5.8). Substituting this distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ in the payoff function $\mathcal{S}$, $\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p(\mathsf{n}|\mathsf{m}) = \mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda$. This contradicts the fact that $\mathcal{S}(\mathbf{p}_\Lambda) \leq s_\Lambda$ for all $\mathbf{p}_\Lambda$ of the form given in Equation (5.8). ■

Geometrically, Lemma 5.1 shows that the hyperplane $\mathcal{S}(\mathbf{p}) \leq s_\Lambda$, defined over conditional distributions $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ implies a corresponding hyperplane $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) \leq s_\Lambda$ in the no-signalling polytope $\mathcal{NS}$ for the Bell scenario $\mathscr{B}^{\mathsf{M},T}_{T,\mathsf{N}}$. We now show the connection between a violation of the inequality $\mathcal{S}(\mathbf{p}) \leq s_\Lambda$ using no-signalling correlation assistance in task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ and violation of the corresponding Bell inequality $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) \leq s_\Lambda$.

Theorem 5.1. *Let $P = \{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau\in T,\tau'\in T} \in \mathcal{NS}$. If P violates the Bell inequality defined in Equation (5.11), i.e. $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda$, then sharing P in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff function defined in Equation (5.1), yields a payoff strictly larger than the local bound s_Λ. Conversely, corresponding to a violation of the local bound s_Λ for the payoff function $\mathcal{S}$ defined for $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with no-signalling correlation assistance, i.e. $\mathcal{S}(\mathbf{p}) > s_\Lambda$, there exists a correlation $P = \{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau\in T,\tau'\in T} \in \mathcal{NS}$, obtained via cutting the classical wire, which violates the associated Bell inequality given in Equation (5.11), i.e. $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda$.*

Proof. [2] Consider a correlation $P = \{P(\tau,\mathsf{n}|\mathsf{m},\tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau\in T,\tau'\in T} \in \mathcal{NS}$ that violates the Bell inequality in Equation (5.11), *i.e.* $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda$. If P is shared in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$, the parties can use a strategy as follows. Alice uses her task input $\mathsf{m} \in \mathsf{M}$ directly as the input to the correlation P. Upon receiving output $\tau \in T$ from P, she sends it through the channel $\mathscr{T}$. Bob, upon receiving symbol $\tau' \in T$, uses it as his input to P. In the task, he produces outcome $\mathsf{n} \in \mathsf{N}$ from the correlation P as his final output. Thus, the resulting distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ during the task is given as $p(\mathsf{n}|\mathsf{m}) = \sum_{\tau,\tau'} \mathscr{T}(\tau'|\tau) P(\tau,\mathsf{n}|\mathsf{m},\tau')$. Using the expression for Bell functional $\mathbf{B}_{\mathcal{S},\mathscr{T}}$ as defined in Equation (5.10) and comparing it with the payoff achieved in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ using the above strategy, we get

$$\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p(\mathsf{n}|\mathsf{m}) = \mathbf{B}_{\mathcal{S},\mathscr{T}}(P) > s_\Lambda \tag{5.13}$$

Now suppose there is a conditional probability distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ for which $\mathcal{S}(\mathbf{p}) > s_\Lambda$ in the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$, that $\mathbf{p}$ can be realised using a no-signalling correlation $\tilde{P}$ together with the

[2] Also discussed as Theorem 1 in [2]

channel $\mathscr{T}: T \to T$, characterised by $\mathscr{T}(\tau'|\tau)$. In this case, $p(\mathsf{n}|\mathsf{m})$ admits a representation of the form in Equation (5.2), or equivalently in the form of Equation (5.5). The *wire-cutting* procedure (see Fig. 5.2) ensures the existence of a correlation $P = \{P(\tau, \mathsf{n}|\mathsf{m}, \tau')\}_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N},\tau\in T,\tau'\in T} \in \mathcal{NS}$ such that, when used with $\mathscr{T}$, it realises the distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ that achieves the payoff $\mathcal{S}(\mathbf{p}) > s_\Lambda$.

This can be seen as follows. From Equation (5.2), there exists an encoding and decoding strategy for Alice and Bob, using $\tilde{P}$ and $\mathscr{T}$, that realises $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$. Writing this expression in the form of Equation (5.5) gives $p(\mathsf{n}|\mathsf{m}) = \sum_{\tau,\tau'} \mathscr{T}(\tau'|\tau)P(\tau, \mathsf{n}|\mathsf{m}, \tau')$ where $\{P(\tau, \mathsf{n}|\mathsf{m}, \tau')\}_{\tau,\mathsf{n},\mathsf{m},\tau'}$ is another no-signalling correlation. The correlation $\tilde{P}$ is related to P by local pre- and post-processing (see Remark 5.1). Using $S(\mathbf{p}) > s_\Lambda$, we obtain

$$\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) = \sum_{\mathsf{m},\tau',\tau,\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}}\, \mathscr{T}(\tau'|\tau)\, P(\tau, \mathsf{n}|\mathsf{m}, \tau') = \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p(\mathsf{n}|\mathsf{m}) = \mathcal{S}(\mathbf{p}) > s_\Lambda \tag{5.14}$$

■

From the proof of Theorem 5.1 we conclude the following. For a communication task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff function given in Equation (5.1), the maximum payoff achievable with quantum correlation assistance equals the maximum quantum violation of the associated Bell inequality given in Equation (5.11) (see Equations (5.13),(5.14)). Let the maximum payoff attainable with quantum correlations in $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ be s_Q, achieved by sharing some quantum correlation $P \in \mathcal{NS}$. From Equation (5.14), $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P) = s_Q$. Suppose the maximum Bell violation for the Bell inequality in Equation (5.11) achievable with quantum correlations is $\mathbf{B}_{\mathcal{S},\mathscr{T}}(P') = \tilde{s}_Q \geq s_Q$, and is obtained using $P' \in \mathcal{NS}$. Then, by Equation (5.13), assistance from P' in $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ yields payoff $\tilde{s}_Q$. Hence, $\tilde{s}_Q = s_Q$. This equivalence facilitates the investigation of the maximal quantum advantage in $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ through the well-developed methods for bounding the maximum quantum violation of Bell inequalities, in particular non-commutative polynomial optimisation (NPO) [1, 37, 284, 285]. Next, we discuss *wire-reading*, which will be useful in constructing some of the correlation-assisted communication tasks introduced later.

5.1.2 Wire reading

In a classical channel $\mathscr{T}: T \to T$, the transmitted message can be observed without disturbing the communication. Thus, the communicated message can be recorded [286–288]. We refer to this as *wire-reading*, which is shown in Figure 5.3. Since the classical message can be accessed, we define a linear payoff for a task in the one-way PM scenario that depends both on the transmitted letter and on Bob's output. In the PM scenario as before, M and N denote Alice's inputs and Bob's outputs, respectively, and the parties share a one-way classical channel $\mathscr{T}: T \to T$. Here, $\mathscr{T}(\tau'|\tau)$ gives the probability with which Bob receives $\tau' \in T$ when Alice sends symbol $\tau \in T$. The payoff function for a task with wire-reading is defined on the

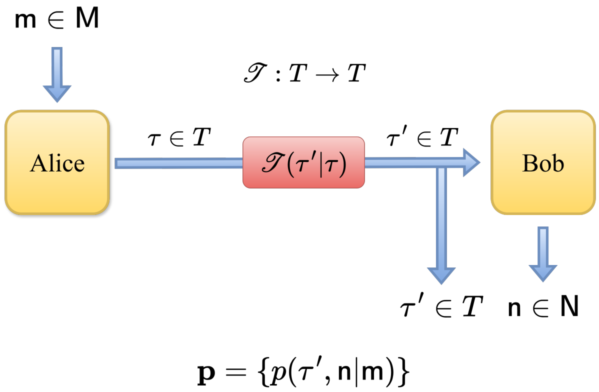

Figure 5.3: A task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with wire-reading in PM scenario considering distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m},\tau'}$.

distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\mathsf{n}\in\mathsf{N},\mathsf{m}\in\mathsf{M},\tau'\in T}$, where $p(\tau', \mathsf{n}|\mathsf{m})$ is the probability that Bob receives message τ' and outputs n when Alice receives input m. We denote the payoff function as $\mathcal{S}_W$ and it is given as

$$(5.15) \qquad \mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m},\tau',\mathsf{n}} w^{\mathsf{m}}_{\tau',\mathsf{n}}\, p(\tau', \mathsf{n}|\mathsf{m}) \quad \text{where} \quad w^{\mathsf{m}}_{\tau',\mathsf{n}} \in \mathbb{R} \;\; \forall \;\; \mathsf{n} \in \mathsf{N}, \mathsf{m} \in \mathsf{M}, \tau' \in T$$

This task with wire-reading is denoted as $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$. For a noiseless classical channel $\mathscr{T}$, it can be equivalently viewed as another communication task $\mathbf{CS}_{\tilde{\mathsf{M}},\tilde{\mathsf{N}},\mathscr{T}}$ in the PM scenario whose payoff function $\mathcal{S}$ is independent of the transmitted message.

Proposition 5.1. *For every task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with a noiseless channel $\mathscr{T}$, there exists a corresponding task $\mathbf{CS}_{\tilde{\mathsf{M}},\tilde{\mathsf{N}},\mathscr{T}}$ without wire-reading such that, given assistance from some $P \in \mathcal{NS}$, the optimal payoff s^{opt} obtained in the former equals to the optimal payoff s^{opt}_W in the latter task when the same correlation is used.*

Proof. In task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with input-output sets M and N and a noiseless classical channel $\mathscr{T} : T \to T$, *i.e.*, $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$, the payoff function is given as

$$(5.16) \quad \mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m}\in\mathsf{M},\tau'\in T,\mathsf{n}\in\mathsf{N}} w^{\mathsf{m}}_{\tau',\mathsf{n}}\, p(\tau', \mathsf{n}|\mathsf{m}) \quad \text{where} \quad w^{\mathsf{m}}_{\tau',\mathsf{n}} \in \mathbb{R} \;\; \forall \;\; \mathsf{n} \in \mathsf{N}, \mathsf{m} \in \mathsf{M}, \tau' \in T$$

Here, Bob's received message τ' is identical to Alice's sent message τ. Suppose the parties share a no-signalling correlation $P = \{P(a, b|x, y)\}_{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},y\in\mathsf{Y}} \in \mathcal{NS}$. Let the optimal payoff with assistance from P be s^{opt}_W. Since $\mathcal{S}_W(\mathbf{p})$ is linear in $p(\tau', \mathsf{n}|\mathsf{m})$, s^{opt}_W can be achieved while using some optimal deterministic strategy of the following form. Alice's input $x = \mathbb{FX}(\mathsf{m})$ to the correlation P is determined by a $\mathbb{FX} : \mathsf{M} \to \mathsf{X}$ and her input $\mathsf{m} \in \mathsf{M}$. Similarly, the function $\mathbb{FT} : \mathsf{M} \times \mathsf{A} \to T$ determines Alice's encoded message $\tau = \mathbb{FT}(\mathsf{m}, a)$ using m and the output $a \in \mathsf{A}$ from the correlation P. Bob uses a function $\mathbb{FY} : T \to \mathsf{Y}$ to determine his input $y = \mathbb{FY}(\tau')$ to the correlation P after receiving message $\tau' \in T$. Using the output $b \in \mathsf{B}$ from P based on the input, Bob outputs $\mathsf{n} = \mathbb{FN}(\tau', b)$ according to function $\mathbb{FN} : T \times \mathsf{B} \to \mathsf{N}$. The resulting conditional probability distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\mathsf{n}\in\mathsf{N},\mathsf{m}\in\mathsf{M},\tau'\in T}$ is

$$p(\tau', \mathsf{n}|\mathsf{m}) = \sum_{\substack{a,b,x,y,\\ \tau}} \left[\delta_{x,\mathbb{FX}(\mathsf{m})}\ \delta_{\tau,\mathbb{FT}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{\tau',\tau}\delta_{y,\mathbb{FY}(\tau')}\ \delta_{\mathsf{n},\mathbb{FN}(\tau',b)}\right] \tag{5.17}$$

$$= \sum_{a,b,x,y} \left[\delta_{x,\mathbb{FX}(\mathsf{m})}\ \delta_{\tau',\mathbb{FT}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\mathbb{FY}(\tau')}\ \delta_{\mathsf{n},\mathbb{FN}(\tau',b)}\right] \tag{5.18}$$

Now, $s_W^{opt} = \mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m}\in\mathsf{M},\tau'\in T,\mathsf{n}\in\mathsf{N}} w^{\mathsf{m}}_{\tau',\mathsf{n}}$.

The task corresponding to $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ in the PM scenario is denoted $\mathbf{CS}_{\tilde{\mathsf{M}},\tilde{\mathsf{N}},\mathscr{T}}$, where $\tilde{\mathsf{M}} = \mathsf{M} \cup T$ and $\tilde{\mathsf{N}} = T \times \mathsf{N}$ are the input and output sets for Alice and Bob. $\mathscr{T} : T \to T$ is the noiseless channel with $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$. If the conditional probability distribution of output $\mathsf{n} = (\mathsf{n}_{(0)}, \mathsf{n}_{(1)}) \in T \times \mathsf{N}$ given input $\mathsf{m} \in \mathsf{M} \cup T$ is $\bar{\mathbf{p}} = \{\bar{p}(\mathsf{n}|\mathsf{m}) = \bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m})\}_{\mathsf{n}\in\tilde{\mathsf{N}},\mathsf{m}\in\tilde{\mathsf{M}}}$, then the payoff function $\mathcal{S}$ for this task is defined as

$$\mathcal{S}(\bar{\mathbf{p}}) = \sum_{\substack{\mathsf{m}\in\tilde{\mathsf{M}},\\ (\mathsf{n}_{(0)},\mathsf{n}_{(1)})\in\tilde{\mathsf{N}}}} \tilde{w}^{\mathsf{m}}_{\mathsf{n}_{(0)},\mathsf{n}_{(1)}} \bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}), \quad \tilde{w}^{\mathsf{m}}_{\mathsf{n}_{(0)},\mathsf{n}_{(1)}} = \begin{cases} w^{\mathsf{m}}_{\mathsf{n}_{(0)},\mathsf{n}_{(1)}} & \text{for } \mathsf{m} \in \mathsf{M} \\ -\Theta(1 - \delta_{\mathsf{n}_{(0)},\mathsf{m}}) & \text{for } \mathsf{m} \in T \end{cases} \tag{5.19}$$

$-\Theta$ is a large penalty, which can be regarded as $-\infty$[3]. This enforces zero-error guessing of Alice's input by Bob, *i.e.* $\mathsf{n}_{(0)} = \mathsf{m}$ when $\mathsf{m} \in T$, in order to avoid the penalty. As before, consider that assistance from the shared correlation P is given, and Alice and Bob use the following strategy. Alice's input $x \in \mathsf{X}$ to P is determined by a function $\tilde{\mathbb{FX}} : \tilde{\mathsf{M}} \to \mathsf{X}$, where $\tilde{\mathbb{FX}}(\mathsf{m}) = \mathbb{FX}(\mathsf{m})$ for $\mathsf{m} \in \mathsf{M}$ and $\tilde{\mathbb{FX}}(\mathsf{m}) = x_1$ otherwise, for some fixed $x_1 \in \mathsf{X}$. Her encoding is given by $\tilde{\mathbb{FT}} : \tilde{\mathsf{M}} \times \mathsf{A} \to T$, which maps input m and correlation output a to a message $\tau \in T$. Specifically, $\tilde{\mathbb{FT}}(\mathsf{m}, a) = \mathbb{FT}(\mathsf{m}, a)$ when $\mathsf{m} \in \mathsf{M}$ and $\tilde{\mathbb{FT}}(\mathsf{m}, a) = \mathsf{m}$ when $\mathsf{m} \in T$. Bob determines his input using $\tilde{\mathbb{FY}} = \mathbb{FY} : T \to \mathsf{Y}$ based on the received message τ'. Given τ' and correlation output b, Bob outputs $\mathsf{n} = (\mathsf{n}_{(0)}, \mathsf{n}_{(1)})$ according to $\tilde{\mathbb{FN}} : T \times \mathsf{B} \to T \times \mathsf{N}$, defined as $\tilde{\mathbb{FN}} := \mathrm{id}_T \times \mathbb{FN}$, *i.e.*, $\tilde{\mathbb{FN}}(\tau', b) = (\tau', \mathbb{FN}(\tau', b))$. The resulting distribution $\bar{\mathbf{p}} = \{\bar{p}(\mathsf{n}|\mathsf{m}) = \bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m})\}_{\mathsf{n}\in\tilde{\mathsf{N}},\mathsf{m}\in\tilde{\mathsf{M}}}$ is

$$\bar{p}\ (\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{a,b,x,y,\tau,\tau'} \left[\delta_{x,\tilde{\mathbb{FX}}(\mathsf{m})}\ \delta_{\tau,\tilde{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{\tau',\tau}\delta_{y,\mathbb{FY}(\tau')}\ \delta_{(\mathsf{n}_{(0)},\mathsf{n}_{(1)}),\tilde{\mathbb{FN}}(\tau',b)}\right] \tag{5.20}$$

$$= \sum_{a,b,x,y,\tau'} \left[\delta_{x,\tilde{\mathbb{FX}}(\mathsf{m})}\ \delta_{\tau',\tilde{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\mathbb{FY}(\tau')}\ \delta_{\mathsf{n}_{(0)},\tau'}\ \delta_{\mathsf{n}_{(1)},\mathbb{FN}(\tau',b)}\right] \tag{5.21}$$

$$= \sum_{a,b,x,y} \left[\delta_{x,\tilde{\mathbb{FX}}(\mathsf{m})}\ \delta_{\mathsf{n}_{(0)},\tilde{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\mathbb{FY}(\mathsf{n}_{(0)})}\ \delta_{\mathsf{n}_{(1)},\mathbb{FN}(\mathsf{n}_{(0)},b)}\right] \tag{5.22}$$

[3]A similar penalty was used in [165] to enforce that Bob avoids certain outputs conditioned on Alice's input.

For $\mathsf{m} \in \mathsf{M}$, comparing Equation (5.22) with Equation (5.18), we see that $\bar{p}\ (\mathsf{n}_{(0)} = t_1, \mathsf{n}_{(1)} = t_2|\mathsf{m}) = p\ (\tau' = t_1, \mathsf{n} = t_2|\mathsf{m})$. Also, $\sum_{\mathsf{n}_{(1)} \in \mathsf{N}} \bar{p}\ (\mathsf{n}_{(0)} = \mathsf{m}, \mathsf{n}_{(1)}|\mathsf{m}) = 1$ for all $\mathsf{m} \in T$. Substituting these relations into Equation (5.19), we obtain

(5.23)
$$\mathcal{S}(\bar{\mathbf{p}}) = \sum_{\mathsf{m}\in\tilde{\mathsf{M}},(\mathsf{n}_{(0)},\mathsf{n}_{(1)})\in\tilde{\mathsf{N}}} \tilde{w}^{\mathsf{m}}_{\mathsf{n}_{(0)},\mathsf{n}_{(1)}}\ \bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{\mathsf{m}\in\mathsf{M},(\mathsf{n}_{(0)},\mathsf{n}_{(1)})\in\tilde{\mathsf{N}}} \tilde{w}^{\mathsf{m}}_{\mathsf{n}_{(0)},\mathsf{n}_{(1)}} \bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = s_W^{opt}$$

Thus, the above strategy with shared correlation P yields payoff s_W^{opt} in the task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$. If the optimal payoff in $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with assistance from P is denoted s^{opt}, then clearly $s_W^{opt} \leq s^{\mathrm{opt}}$. In the next step, we show that $s^{opt} \leq s_W^{opt}$, which implies $s^{opt} = s_W^{opt}$. Since payoff function $\mathcal{S}$ in Equation (5.19) is linear in $\bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m})$, the payoff s^{opt} can be achieved using a deterministic strategy with shared correlation P. Alice determines her input $x \in \mathsf{X}$ to P via a function $\bar{\mathbb{FX}} : \tilde{\mathsf{M}} \to \mathsf{X}$, depending on her input m. Her encoding $\tau \in T$ is chosen according to $\bar{\mathbb{FT}} : \tilde{\mathsf{M}} \times \mathsf{A} \to T$, based on m and the correlation output a. Bob uses $\bar{\mathbb{FY}} : T \to \mathsf{Y}$ to select his input $y = \bar{\mathbb{FY}}(\tau')$ to P after receiving $\tau' \in T$. Finally, his output $(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}) = \bar{\mathbb{FN}}(\tau', b) \in \tilde{\mathsf{N}}$ is determined by $\bar{\mathbb{FN}} : T \times \mathsf{B} \to T \times \mathsf{N}$, applied to (τ', b), where b is his outcome from P. The resulting conditional probability distribution $\bar{\mathbf{p}} = \{\bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m})\}$ is

$$\bar{p}\ (\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{a,b,x,y,\tau'} \left[\delta_{x,\bar{\mathbb{FX}}(\mathsf{m})}\ \delta_{\tau',\bar{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\bar{\mathbb{FY}}(\tau')}\ \delta_{(\mathsf{n}_{(0)},\mathsf{n}_{(1)}),\bar{\mathbb{FN}}(\tau',b)}\right] \tag{5.24}$$

When Alice's input $\mathsf{m} \in T$, Bob must output $\mathsf{n}_{(0)} = \mathsf{m}$ to avoid the penalty defined in Equation (5.19). To satisfy $\sum_{\mathsf{n}_{(1)}} \bar{p}(\mathsf{n}_{(0)} = \mathsf{m}, \mathsf{n}_{(1)}|\mathsf{m}) = 1\ \forall\ \mathsf{m} \in T$, we require $\tau = \tau' = \bar{\mathbb{FT}}(\mathsf{m}) = \zeta(\mathsf{m})$ for all $\mathsf{m} \in T$, where $\zeta : T \to T$ is a permutation. Bob's output must take the form $(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}) = \bar{\mathbb{FN}}(\tau', b) = (\zeta^{-1}(\tau'), \bar{\mathbb{FN}}_1(\tau', b))$ where $\bar{\mathbb{FN}}_1 : T \times \mathsf{B} \to \mathsf{N}$ determines his second output component $\mathsf{n}_{(1)}$. When $\mathsf{m} \in \mathsf{M}$, substituting it in Equation (5.24), we get

(5.25)
$$\bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{a,b,x,y,\tau'} \left[\delta_{x,\bar{\mathbb{FX}}(\mathsf{m})}\ \delta_{\tau',\bar{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\bar{\mathbb{FY}}(\tau')}\ \delta_{(\mathsf{n}_{(0)},\mathsf{n}_{(1)}),(\zeta^{-1}(\tau'),\bar{\mathbb{FN}}_1(\tau',b))}\right]$$
$$= \sum_{a,b,x,y,\tau'} \left[\delta_{x,\bar{\mathbb{FX}}(\mathsf{m})}\ \delta_{\tau',\bar{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\bar{\mathbb{FY}}(\tau')}\ \delta_{\mathsf{n}_{(0)},\zeta^{-1}(\tau')}\delta_{\mathsf{n}_{(1)},\bar{\mathbb{FN}}_1(\tau',b)}\right] \tag{5.26}$$
$$= \sum_{a,b,x,y} \left[\delta_{x,\bar{\mathbb{FX}}(\mathsf{m})}\ \delta_{\zeta(\mathsf{n}_{(0)}),\bar{\mathbb{FT}}(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\bar{\mathbb{FY}}(\zeta(\mathsf{n}_{(0)}))}\ \delta_{\mathsf{n}_{(1)},\bar{\mathbb{FN}}_1(\zeta(\mathsf{n}_{(0)}),b)}\right] \tag{5.27}$$
$$= \sum_{a,b,x,y} \left[\delta_{x,\bar{\mathbb{FX}}(\mathsf{m})}\ \delta_{\mathsf{n}_{(0)},\bar{\mathbb{FT}}'(\mathsf{m},a)}\ P(a, b|x, y)\ \delta_{y,\bar{\mathbb{FY}}'(\mathsf{n}_{(0)})}\ \delta_{\mathsf{n}_{(1)},\bar{\mathbb{FN}}_1'(\mathsf{n}_{(0)},b)}\right] \tag{5.28}$$

where $\bar{\mathbb{FY}}' = \bar{\mathbb{FY}} \circ \zeta$, and $\bar{\mathbb{FN}}_1'(\tau, b) = \bar{\mathbb{FN}}_1(\zeta(\tau), b)$. Similarly, $\bar{\mathbb{FT}}' = \zeta^{-1} \circ \bar{\mathbb{FT}}$. The resulting distribution $\bar{\mathbf{p}} = \{\bar{p}(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m})\}$ achieves payoff s^{opt}.

In task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$, the parties adopt the following strategy while sharing P. Alice uses a function $\mathbb{FX} : \mathsf{M} \to [X]$ for choosing the input to P, where $\mathbb{FX}(\mathsf{m}) = \bar{\mathbb{FX}}(\mathsf{m})$. Her encoding is given by

$\mathsf{m}\backslash(\tau', \mathsf{n})$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
1	$\frac{1}{4}$	0	0	$\frac{1}{4}$
2	$\frac{1}{4}$	0	$\frac{1}{4}$	0
3	0	$\frac{1}{4}$	0	$\frac{1}{4}$
4	0	$\frac{1}{4}$	$\frac{1}{4}$	0

Table 5.1: The coefficients $\{w^{\mathsf{m}}_{\tau',\mathsf{n}}\}_{\tau',\mathsf{m},\mathsf{n}}$ in payoff function $\mathcal{S}_W$ for task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ where $\mathsf{m} \in \mathsf{M} = \{1, 2, 3, 4\}$ and $\mathsf{n} \in \mathsf{N} = \{0, 1\}$ and $\tau' \in T = \{0, 1\}$.

$\mathbb{FT} : \mathsf{M} \times \mathsf{A} \to T$, with $\mathbb{FT}(\mathsf{m}, a) = \bar{\mathbb{FT}}'(\mathsf{m}, a)$ for $\mathsf{m} \in \mathsf{M}$ and $a \in \mathsf{A}$. Bob selects his input to P using $\mathbb{FY} : T \to \mathsf{Y}$, where $\mathbb{FY}(\tau') = \bar{\mathbb{FY}}(\tau')$ based on the received message τ'. His output n is determined by $\bar{\mathbb{FN}}_1' : T \times \mathsf{B} \to \mathsf{N}$, which depends on τ' and the outcome b from P. The resulting probability distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\mathsf{m},\tau',\mathsf{n}}$ has the same form as Equation (5.28) for $\mathsf{m} \in \mathsf{M}$. Since $\tilde{w}^{\mathsf{m}}_{t_1,t_2} = w^{\mathsf{m}}_{t_1,t_2}$ for $\mathsf{m} \in \mathsf{M}$ the payoff is again s^{opt}. Hence, for the correlation P, the optimal payoff in both tasks are equal, *i.e.* $s^{opt}_W = s^{opt}$. ■

From Proposition 5.1, the bounds on the payoff in task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with shared randomness, entanglement, or no-signalling assistance are equal to the corresponding bounds for task $\mathbf{CS}_{\tilde{M},\tilde{N},\mathscr{T}}$.

In communication tasks defined in the PM scenario, adding wire-reading can reveal an advantage from entanglement assistance, which may not exist otherwise. To illustrate, consider the task $\mathbf{CS}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ where Alice has a noiseless one-bit channel $\mathscr{T} : \{0, 1\} \to \{0, 1\}$, *i.e.*, $T = \{0, 1\}$ and $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$. Alice's input set is $\mathsf{M} = \{1, 2, 3, 4\}$, and Bob's output set is $\mathsf{N} = \{0, 1\}$. The payoff function $\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n}} w^{\mathsf{m}}_{\mathsf{n}} p(\mathsf{n}|\mathsf{m})$ is defined over probability distribution $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$. The maximum payoff achievable by optimising over all correlations and communication resources is $\mathcal{S}^{alg} = \sum_{\mathsf{m}\in\mathsf{M}}[\max_{\mathsf{n}\in\mathsf{N}}\{w^{\mathsf{m}}_{\mathsf{n}}\}]$ which we call the *algebraic maximum* for the payoff. In this task, $\mathcal{S}^{alg}$ can already be reached using shared randomness and the noiseless one-bit channel. Given input m, Alice sends $\tau = \tilde{\mathsf{n}}$ such that $w^{\mathsf{m}}_{\tilde{\mathsf{n}}} = \max_{\mathsf{n}\in\mathsf{N}}\{w^{\mathsf{m}}_{\mathsf{n}}\}$. Bob outputs $\mathsf{n} = \tau' = \tilde{\mathsf{n}}$ to achieve $\mathcal{S}^{alg}$. Hence, this task cannot witness any advantage from non-classical correlation assistance.

In the same PM scenario with the given input and output sets, consider the task $\mathbf{CW}_{\mathsf{M},\mathsf{N},\mathscr{T}}$ with payoff function $\mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n},\tau'} w^{\mathsf{m}}_{\tau',\mathsf{n}}\, p(\tau', \mathsf{n}|\mathsf{m})$ defined over distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$. The coefficients $w^{\mathsf{m}}_{\tau',\mathsf{n}}$ are given in Table 5.1. The algebraic maximum for payoff $\mathcal{S}^{alg}_W = 1$. Using shared randomness assistance to the channel, the maximum achievable payoff (local bound) is $\frac{3}{4}$. Consider the deterministic decoding strategy $\mathbb{D} : T \to \mathsf{N}$ by Bob where $\mathsf{n} = \mathbb{D}(\tau') = \tau'$. An optimal deterministic encoding for Alice then is $\mathbb{E} : \mathsf{M} \to T$ where $\mathbb{E}(1) = \mathbb{E}(4) = 0$ (or equivalently 1), $\mathbb{E}(2) = 0$ and $\mathbb{E}(3) = 1$. This strategy achieves payoff $\frac{3}{4}$. Similarly, the maximum payoff using any other deterministic encoding-decoding pair is at most $\frac{3}{4}$. Since $\mathcal{S}_W$ is linear in the conditional probabilities, the maximum achievable payoff with shared randomness is also $\frac{3}{4}$. The advantage of non-local correlations in this task follows from Theorem 5.3 and Corollary 5.1 in Section 5.2.1, which will be discussed later, by taking the case $d = 2, k = 2$.

5.2 Classical Communication Tasks Witnessing Non-locality

We now introduce two families of one-way classical communication tasks based on wire-reading. In both cases, the local bounds on the payoff achievable with shared randomness are strictly suboptimal. The first family shows that sharing any correlation from a non-local facet of the no-signalling polytope leads to an advantage. Each task in the second family is tailored to specific non-local facets of the no-signalling polytope.

5.2.1 Task: $\mathbf{CW}[d,k]$

Task Outline: In the task $\mathbf{CW}[d,k]$, Alice receives as input a random function $f:[d]\to[k]$. She sends a single symbol $z_1\in[d]$ to Bob through a noiseless classical channel $\mathscr{T}:[d]\to[d]$. Bob must then output a value $z_2\in[k]$ such that $z_2\neq f(z_1)$. Equivalently, Bob's task is to avoid returning the image of the element transmitted by Alice under f.

In task $\mathbf{CW}[d,k]$[4], the input and output sets are $\mathsf{M}=\{1,\cdots,k\}^d$ and $\mathsf{N}=[k]$, with $d,k\geq 2$ and $d,k\in\mathbb{N}$. The classical communication channel is noiseless, $\mathscr{T}:[d]\to[d]$, with $T=[d]$ and $\mathscr{T}(\tau'|\tau)=\delta_{\tau',\tau}$. Alice receives an input $\mathsf{m}=(\mathsf{m}_{(1)},\cdots,\mathsf{m}_{(d)})\in M$. Bob produces an output $\mathsf{n}\in\mathsf{N}$. Let $\mathbf{p}=\{p(\tau',\mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$ denote the conditional distribution where Bob receives message $\tau'\in[d]$ and outputs $\mathsf{n}\in\mathsf{N}$ given Alice's input $\mathsf{m}\in\mathsf{M}$. The payoff function for the task is

$$\mathcal{S}_W(\mathbf{p})=\sum_{\mathsf{m},\mathsf{n},\tau'} w^{\mathsf{m}}_{\tau',\mathsf{n}}\; p(\tau',\mathsf{n}|\mathsf{m}) \quad \text{where } w^{\mathsf{m}}_{\tau',\mathsf{n}}=\frac{1}{k^d}(1-\delta_{\mathsf{n},\mathsf{m}_{(\tau')}}) \tag{5.29}$$

For input $\mathsf{m}=(\mathsf{m}_{(1)},\cdots,\mathsf{m}_{(d)})$, the parties receive payoff $\frac{1}{k^d}$ if Bob's output $\mathsf{n}\neq\mathsf{m}_{(\tau')}$ (τ'-th component of m). Here $\tau'(=\tau)$ is the message transmitted and received through the noiseless channel $\mathscr{T}$. The case $d=2,k=2$ is inspired by the task in [163], where Alice instead has 6 possible inputs. The algebraic maximum for the payoff

$$\mathcal{S}^{alg}_W=\sum_{\mathsf{m}}\max_{\mathsf{n}\in\mathsf{N},\tau'\in[d]}\{w^{\mathsf{m}}_{\tau',\mathsf{n}}\}=\sum_{\mathsf{m}\in\mathsf{M}}\frac{1}{k^d}=1 \tag{5.30}$$

We now derive the local bound on the payoff when the parties use shared randomness in addition to the channel.

Theorem 5.2. *The maximum payoff with shared randomness assistance to classical channel $\mathscr{T}:[d]\to[d]$ in $\mathbf{CW}[d,k]$ is $s_\Lambda=(1-\frac{1}{k^d})$.*

Proof. Suppose Bob uses a deterministic decoding $\mathbb{D}:[d]\to\mathsf{N}$, so that on receiving $\tau'\in[d]$ he outputs $\mathsf{n}=\mathbb{D}(\tau')$. Consider an input $\mathsf{m}=(\mathsf{m}_{(1)},\ldots,\mathsf{m}_{(d)})\in\mathsf{M}$ of Alice such that $\mathsf{m}_{(i)}=\mathbb{D}(i)$ for each $i\in\{1,\cdots,d\}$. Let Alice use a deterministic encoding $\mathbb{E}:\mathsf{M}\to[d]$, sending $\tau=\mathbb{E}(\mathsf{m})$. For the above input m, Bob's output after getting message $\tau'=\tau$ is $\mathsf{n}=\mathbb{D}(\tau')=\mathsf{m}_{(\tau')}$.

[4] Also discussed in Section 3 of [2]

In this case, the conditional probability $p(\tau' = \mathbb{E}(\mathsf{m}), \mathsf{n} = \mathbb{D}(\tau')|\mathsf{m}) = p(\tau' = \mathbb{E}(\mathsf{m}), \mathsf{n} = \mathsf{m}_{(\tau')}|\mathsf{m}) = 1 \implies p(\tau', \mathsf{n} \neq \mathsf{m}_{(\tau')}|\mathsf{m}) = 0$. Thus, the payoff for this input m is 0 under the encoding $\mathbb{E}$ and decoding $\mathbb{D}$. Note that Alice's choice of $\mathbb{E}$ is independent of Bob's choice of $\mathbb{D}$. Moreover, for every deterministic decoding $\mathbb{D}$, there exists such an input m for which the payoff is 0, regardless of Alice's deterministic encoding. Hence, the payoff is upper bounded by $(\mathcal{S}_W^{alg} - \frac{1}{k^d}) = (1 - \frac{1}{k^d})$ for any deterministic encoding-decoding pair. Since the payoff function is linear in Equation (5.29), the maximum achievable payoff with shared randomness is attained by some deterministic strategy. Hence, the local bound for the task $\mathbf{CW}[d,k]$ satisfies $s_\Lambda \leq (1 - \frac{1}{k^d})$.

We now show that this upper bound can be saturated with shared randomness assistance. Let Bob's deterministic decoding be $\mathbb{D} : [d] \to \mathsf{N}$, defined by $\mathbb{D}(\tau') = 1 \ \forall \ \tau' \in [d]$. Alice's sends $\tau = \mathbb{E}(\mathsf{m})$ determined by $\mathbb{E} : \mathsf{M} \to [d]$, where $\mathbb{E}(\mathsf{m}) = 1$ if $\mathsf{m} = (1, 1, \cdots, 1)$, else $\mathbb{E}(\mathsf{m}) = \min\{i \in [d] : \mathsf{m}_{(i)} > 1\}$ for input $\mathsf{m} = (\mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \in \mathsf{M}$. For input $\mathsf{m} = (\mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \neq (1, 1, \cdots, 1)$, Alice sends $\tau = \mathbb{E}(\mathsf{m})$ and Bob outputs $\mathsf{n} = \mathbb{D}(\tau) = 1$. Since $\mathsf{m}_{(\tau)} > 1$, we have $\mathsf{n} = 1 \neq \mathsf{m}_{(\tau)}$ for $\mathsf{m} \neq (1, 1, \cdots, 1)$. The payoff using this strategy is $\sum_{\mathsf{m} \in \mathsf{M}, \mathsf{m} \neq (1,1,\cdots,1)} (\frac{1}{k^d}) = (1 - \frac{1}{k^d})$. ■

Before showing the advantage of using non-local correlation in this task, we prove the following property of non-local correlations.

Lemma 5.2. *Let $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$ be a function. For any correlation $P = \{P(a,b|x,y)\}_{a \in \mathsf{A}, b \in \mathsf{B}, x \in \mathsf{X}, y \in \mathsf{Y}}$ lying in a non-local facet of the no-signalling polytope $\mathcal{NS}$, there exists an input $x \in \mathsf{X}$ such that for all $a \in \mathsf{A}$ there exists a non-empty set $\Phi^{x,a}_{\mathbb{LB}}(P) = \{y \in \mathsf{Y} : P(a, b = \mathbb{LB}(y)|x,y) = 0\}$.*

Proof. Let $\{P_L^{(i)}\}$ and $\{P_{NL}^{(i)}\}$ denote the sets of local and non-local extremal points of the no-signalling polytope $\mathcal{NS}$. Each local extremal point $P_L^{(i)}$ is specified by a pair of functions $\mathbb{LA} : \mathsf{X} \to \mathsf{A}$ and $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$. Consider correlation $P = \{P(a,b|x,y)\}_{a \in \mathsf{A}, b \in \mathsf{B}, x \in \mathsf{X}, y \in \mathsf{Y}}$ lying on a non-local facet of $\mathcal{NS}$. Either $P \in \{P_{NL}^{(i)}\}$ or it can *only* be expressed as a convex mixture of such extremal points. Assume, for showing contradiction, that the lemma fails for some function $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$ and some P on a non-local facet. Thus, for every $x \in \mathsf{X}$ there exists at least one $a_x \in \mathsf{A}$ such that $\Phi^{x,a_x}_{\mathbb{LB}}(P) = \varnothing$. As a result, $\forall \ y \in \mathsf{Y}$, $P(a_x, b = \mathbb{LB}(y)|x,y) > 0 \implies P(b = \mathbb{LB}(y)|x,y,a_x)P(a_x|x) > 0$ (by Bayes rule and no-signalling) thus $P(b = \mathbb{LB}(y)|x,y,a_x) > 0 \wedge P(a_x|x) > 0$. Now, using these outcomes for each $x \in \mathsf{X}$, we define the function $\mathbb{LA} : \mathsf{X} \to \mathsf{A}$. Here, $\mathbb{LA}(x) = a_x \in \mathsf{A}$ such that $P(a_x|x) > 0$ and $\Phi^{x,a_x}_{\mathbb{LB}}(P) = \varnothing$. Now consider the local extremal point $P_L = \{P_L(a,b|x,y)\}_{a \in \mathsf{A}, b \in \mathsf{B}, x \in \mathsf{X}, y \in \mathsf{Y}}$ defined by the pair $\mathbb{LA} : \mathsf{X} \to \mathsf{A}$ and $\mathbb{LB} : \mathsf{Y} \to B$, *i.e.*, $P_L(a,b|x,y) = \delta_{a,\mathbb{LA}(x)}\delta_{b,\mathbb{LB}(y)}$. By construction, $\forall x, y$ whenever $P_L(a,b|x,y) = 1$, we also have $P(a,b|x,y) > 0$. Thus, P admits a decomposition as $P = pP_L + (1-p)\tilde{P}$ where $0 < p \leq 1$ and $\tilde{P} \in \mathcal{NS}$. But this contradicts the assumption that P lies on a non-local facet, since such correlations cannot have a convex decomposition with positive weight corresponding to a local extremal point $P_L \in \{P_L^{(i)}\}$. ■

Lemma 5.3. *Given a function $\mathbb{LB} : [d] \to [k]$ and a non-local facet $Face\{P_{NL}^{(i)}\}_{i\in J}$ spanned by non-local extremal points $\{P_{NL}^{(i)}\}_{i\in J}$. There exists an input $x^* \in \mathsf{X}$, independent of the choice of $P_{NL}^{(i)}$ with $i \in J$, such that for every $a \in \mathsf{A}$ there exists a non-empty subset $\Phi_{\mathbb{LB}}^{x^*,a} = \bigcap_{i\in J} \Phi_{\mathbb{LB}}^{x^*,a}(P_{NL}^{(i)}) \subseteq \mathsf{Y}$. For each $a \in \mathsf{A}$ and $y \in \Phi_{\mathbb{LB}}^{x^*,a}$, $P_{NL}^{(i)}(a, b = \mathbb{LB}(y)|x, y) = 0 \;\forall\; i \in J$*

Proof. Consider the correlation $P = \sum_{i\in J} \frac{1}{|J|} P_{NL}^{(i)}$ in non-local facet $Face\{P_{NL}^{(i)}\}_{i\in J}$. From Lemma 5.2, there exists an input $x \in \mathsf{X}$ such that for all $a \in \mathsf{A}$ there exists a non-empty set $\Phi_{\mathbb{LB}}^{x,a}(P) = \{y \in \mathsf{Y} : P(a, b = \mathbb{LB}(y)|x, y) = 0\}$. From the definition of correlation P, $P(a, b = \mathbb{LB}(y)|x, y) = \frac{1}{|J|} \sum_{i\in J} P_{NL}^{(i)}(a, b = \mathbb{LB}(y)|x, y) = 0$. Thus, for $y \in \Phi_{\mathbb{LB}}^{x,a}(P)$, $P_{NL}^{(i)}(a, b = \mathbb{LB}(y)|x, y) = 0$ for all $i \in J$. This implies $y \in \Phi_{\mathbb{LB}}^{x,a}(P_{NL}^{(i)}) \;\forall\; i \in J$ whenever $y \in \Phi_{\mathbb{LB}}^{x,a}(P)$. As a result, with this choice of $x^* = x$ and for all $a \in \mathsf{A}$, $\Phi_{\mathbb{LB}}^{x^*,a} = \bigcap_{i\in J} \Phi_{\mathbb{LB}}^{x^*,a}(P_{NL}^{(i)}) \supseteq \Phi_{\mathbb{LB}}^{x^*,a}(P) \neq \varnothing$. ■

Theorem 5.3. *For the task* **CW**$[d, k]$*, assistance from any correlation lying on a non-local facet of the no-signalling polytope $\mathcal{NS} := \{P = \{P(a, b|x, y)\}_{a\in\mathsf{A},b\in\mathsf{B}=[k],x\in\mathsf{X},y\in\mathsf{Y}=[d]}\}$ achieves the algebraic maximum payoff $\mathcal{S}_W^{alg} = 1$.*

Proof. In the task **CW**$[d, k]$, let the parties share a correlation $P = \{P(a, b|x, y)\}_{a\in\mathsf{A},b\in[k],x\in\mathsf{X},y\in[d]}$ from a non-local facet of the no-signalling polytope $\mathcal{NS} := \{P = \{P(a, b|x, y)\}_{a\in\mathsf{A},b\in[k],x\in\mathsf{X},y\in[d]}\}$. Here $\mathsf{B} = [k], \mathsf{Y} = [d]$. Each input $\mathsf{m} = (\mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \in \mathsf{M} = \{1, \cdots, k\}^d$ can be identified with a function $\mathbb{LB} : [d] \to [k]$ defined by $\mathbb{LB}(i) = \mathsf{m}_{(i)} \forall\; i \in [d]$. From Lemma 5.2, for this function $\mathbb{LB}$ and correlation P, there exists an $x_\mathsf{m} \in \mathsf{X}$ such that for every $a \in \mathsf{A}$ there is a non-empty subset $\Phi_{\mathbb{LB}}^{x_\mathsf{m},a}(P) \subseteq [d]$. For each $y \in \Phi_{\mathbb{LB}}^{x_\mathsf{m},a}(P)$, $P(a, b = \mathbb{LB}(y)|x_\mathsf{m}, y) = 0$. Using this observation, Alice and Bob implement the following protocol for input m.

Encoding: For the input m Alice chooses $x_\mathsf{m} \in \mathsf{X}$ as her input to the correlation P. Obtaining the output $a \in \mathsf{A}$ from P, she sends a fixed element $y_{(\mathsf{m},a)} \in \Phi_{\mathbb{LB}}^{x_\mathsf{m},a}(P) \subseteq [d]$ as message τ.
Decoding: Bob after receiving symbol $\tau' = \tau$ uses it as his input to P, *i.e.* $y = \tau' = y_{(\mathsf{m},a)}$. He uses the outcome $b \in [k]$ from the correlation P as the final output, *i.e.* $\mathsf{n} = b$. The resulting probability $p(\tau', \mathsf{n}|\mathsf{m})$ is

$$p(\tau', \mathsf{n}|\mathsf{m}) = \sum_{x,a,y,b,\tau} \delta_{x,x_\mathsf{m}} \delta_{\tau,y_{(\mathsf{m},a)}} \delta_{\tau,\tau'} \delta_{y,\tau'} P(a, b|x, y) \delta_{\mathsf{n},b} = \sum_a \delta_{\tau',y_{(\mathsf{m},a)}} P(a, \mathsf{n}|x_\mathsf{m}, y_{(\mathsf{m},a)}) \tag{5.31}$$

For a given input m, this protocol guarantees $p(\tau', \mathsf{n}|\mathsf{m}) = 0$ whenever $\tau' \notin \{y_{(\mathsf{m},a)}\}_{a\in\mathsf{A}}$. For $\mathsf{n} = \mathsf{m}_{(\tau')}$, $p(\tau', \mathsf{n}|\mathsf{m}) = 0$ as $\sum_a \delta_{\tau',y_{(\mathsf{m},a)}} P(a, \mathsf{n}|x_\mathsf{m}, y_{(\mathsf{m},a)}) = 0$ in this case from the choice of $y_{(\mathsf{m},a)}$ for each $a \in \mathsf{A}$. As a result, $p(\tau', \mathsf{n} \neq \mathsf{m}_{(\tau')}|\mathsf{m}) = 1$. Thus, the protocol achieves payoff $\sum_{\mathsf{m}\in\mathsf{M}}(\frac{1}{k^d}) = 1$. ■

Corollary 5.1. *Let* $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B}=[k],x\in\mathsf{X},y\in\mathsf{Y}=[d]}$ *belong to a non-local facet of the no-signalling polytope* $\mathcal{NS}$*. For any* p *with* $(1-\frac{1}{k^{d-1}}) < p \leq 1$*, the correlation* $\tilde{P} = pP + (1-p)P_{WN}$ *provides an advantage over shared randomness assistance in the task* $\mathbf{CW}[d,k]$*.*

Proof. Sharing the correlation $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B}=[k],x\in\mathsf{X},y\in\mathsf{Y}=[d]}$ from a non-local facet of the no-signalling polytope $\mathcal{NS}$, the parties can follow the protocol from the proof of Theorem 5.3 and achieve the algebraic maximum $\mathcal{S}_W^{alg} = 1$ in the task $\mathbf{CW}[d,k]$. Now suppose they use the same protocol while sharing the white noise correlation $P_{WN} = \{P_{WN}(a,b|x,y) = \frac{1}{|\mathsf{A}|k}\}_{a\in\mathsf{A},b\in\mathsf{B}=[k],x\in\mathsf{X},y\in\mathsf{Y}=[d]}$. For input $\mathsf{m}\in\mathsf{M}$, Alice chooses $x_{\mathsf{m}}\in\mathsf{X}$, the same as in the protocol for correlation P, as the input to the correlation P_{WN}. Getting output $a\in\mathsf{A}$ from P_{WN}, sends $\tau = y_{(\mathsf{m},a)} \in \Phi_{\mathbb{LB}}^{x_{\mathsf{m}},a}(P) \subseteq [d]$. For m, let $r_{\tau,\mathsf{m}}$ be the number of outcomes $a\in\mathsf{A}$ for which Alice sends τ. Clearly, $\sum_{\tau\in[d]} r_{\tau,\mathsf{m}} = |\mathsf{A}|\ \forall\ \mathsf{m}\in\mathsf{M}$. Bob, after receiving $\tau' = \tau$, queries P_{WN} with $y = \tau'$, and outputs $b\in[k]$ from P_{WN} as $\mathsf{n} = b$. The conditional distribution $\mathbf{p} = \{p(\tau',\mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$ is then

$$p(\tau',\mathsf{n}|\mathsf{m}) = \sum_{a\in\mathsf{A}} \delta_{\tau',y_{(\mathsf{m},a)}} P_{WN}(a,\mathsf{n}|x_{\mathsf{m}},y_{(\mathsf{m},a)}) = \sum_{a\in\mathsf{A}} \delta_{\tau',y_{(m,a)}} \frac{1}{k}\frac{1}{|\mathsf{A}|} = \frac{1}{k|\mathsf{A}|} r_{\tau',\mathsf{m}} \tag{5.32}$$

For each $\mathsf{m}\in\mathsf{M}$ and $\tau'\in[d]$, there is exactly one value $\mathsf{n}\in[k]$ such that $\mathsf{n} = \mathsf{m}_{\tau'}$. From Equation (5.32), $p(\tau',\mathsf{n}\neq\mathsf{m}_{(\tau')}|\mathsf{m}) = \frac{1}{k|\mathsf{A}|} r_{\tau',\mathsf{m}}(k-1)$. The resulting payoff is

$$\mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m}\in\mathsf{M}} [\sum_{\tau'\in[d]} \frac{1}{k^d} p(\tau',\mathsf{n}\neq\mathsf{m}_{(\tau')}|\mathsf{m})] = \sum_{\mathsf{m}\in\mathsf{M}} [\sum_{\tau'\in[d]} \frac{1}{k^d}\frac{1}{k|\mathsf{A}|} r_{\tau',\mathsf{m}}(k-1)] \tag{5.33}$$

$$= \sum_{\mathsf{m}\in\mathsf{M}} \frac{1}{k^{d+1}|\mathsf{A}|}(k-1)[\sum_{\tau'\in[d]} r_{\tau',\mathsf{m}}] = \sum_{\mathsf{m}\in\mathsf{M}} \frac{1}{k^{d+1}}(k-1) = \frac{(k-1)}{k} \tag{5.34}$$

Using linearity of the payoff, sharing the noisy correlation $\tilde{P} = pP + (1-p)P_{WN}$ with $0 \leq p \leq 1$ gives payoff $p(1) + (1-p)\frac{(k-1)}{k}$ with above strategy. This is strictly greater than the shared randomness bound $1 - \frac{1}{k^d}$ whenever $p + (1-p)\frac{(k-1)}{k} > 1 - \frac{1}{k^d} \implies p > (1 - \frac{1}{k^{d-1}})$. ■

Note that local bound on the payoff $s_\Lambda = (1 - \frac{1}{k^d})$ for the task $\mathbf{CW}[d,k]$ approaches the algebraic maximum $\mathcal{S}_W^{alg} = 1$ polynomially in k and exponentially in d. Consider a correlation $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in[2],x\in\mathsf{X},y\in[3]}$ lying on a non-local facet of the polytope $\mathcal{NS}$. In the task $\mathbf{CW}[d=3,k=2]$, following the protocol used in the proof of Corollary 5.1, assistance from any correlation on the isotropic line between P and P_{WN}, namely $\tilde{P} = pP + (1-p)P_{WN}$, does not provide an advantage over shared randomness when $p \leq \frac{3}{4}$. We now present another family of communication tasks. For certain tasks in this family, we show that correlations of the form $pP + (1-p)P_{WN}$ are already advantageous for $\frac{1}{2} < p \leq 1$ (see Corollaries 5.3 and 5.5).

5.2.2 Task: $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$

Task Outline: In task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$[5], Alice receives a randomly chosen relation $\mathcal{R} \subseteq [d] \times [k]$. This relation is a map from $[d]$ to $[k]$ of a particular type, satisfying conditions (i) and (ii) stated below. Alice sends an element $z_1 \in [d]$ to Bob through a noiseless classical channel $\mathscr{T} : [d] \to [d]$. Bob's output $z_2 \in [k]$ must be one of the images of z_1 under the relation $\mathcal{R}$ for them to succeed. The set of possible relations $\mathcal{R}$ is not listed explicitly. Instead, they are defined implicitly by starting with a function $\mathbb{LB} : [d] \to [k]$ and through Lemma 5.2 and Lemma 5.3 stating the compliance conditions (i) and (ii) for the relation. Condition (i) ensures that the payoff under shared randomness assistance is always suboptimal. Condition (ii) guarantees the existence of a protocol that achieves the algebraic maximum payoff whenever the parties are assisted by a correlation P lying on the non-local facet determined by the extremal points $\{P_{NL}^{(i)}\}_{i\in J}$. Unlike a function, a relation $\mathcal{R}$ may assign more than one image in $[k]$ to a given element of $[d]$.

Here, $Face\{P_{NL}^{(i)}\}_{i\in J}$ denotes a non-local facet of the no-signalling polytope $\mathcal{NS} := \{P = \{P(a,b|x,y)\}_{a\in \mathsf{A}, b\in[k], x\in \mathsf{X}, y\in[d]}\}$ (with $\mathsf{B} = [k], \mathsf{Y} = [d]$). This facet is the convex hull of certain non-local extremal points $\{P_{NL}^{(i)}\}_{i\in J}$, where J indexes the non-local extremal correlations that span it. In the task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$, Alice's input set M is a set consisting of some relations $\mathsf{m} \subset [d] \times [k]$, and Bob's output set is $\mathsf{N} = [k]$. To specify the set of relations that form M, we first introduce the following notion. As noted earlier, these relations are not enumerated explicitly. Instead, they are defined implicitly as those satisfying the required conditions (i) and (ii).

From Lemma 5.3, the following is satisfied for the non-local extremal points $\{P_{NL}^{(i)}\}_{i\in J}$ spanning the non-local facet $Face\{P_{NL}^{(i)}\}_{i\in J}$. Consider a function $\mathbb{LB} : [d] \to [k]$. There exists an input $x^* \in \mathsf{X}$, independent of the choice of $P_{NL}^{(i)}$ with $i \in J$, such that for every $a \in \mathsf{A}$ there exists a non-empty subset $\Phi_{\mathbb{LB}}^{x^*,a} = \bigcap_{i\in J} \Phi_{\mathbb{LB}}^{x^*,a}(P_{NL}^{(i)}) \subseteq \mathsf{Y}$. For all $y \in \Phi_{\mathbb{LB}}^{x^*,a}$ and $P_{NL}^{(i)} = \{P_{NL}^{(i)}(a,b|x,y)\}_{a\in\mathsf{A}, b\in[k], x\in\mathsf{X}, y\in[d]}$ with $i \in J$,

$$P_{NL}^{(i)}(a, b = \mathbb{LB}(y)|x^*, y) = 0 \tag{5.35}$$

For each $y \in \Phi_{\mathbb{LB}}^{x^*,a}$ we define the set

$$\Phi_{\mathbb{LB}}^{x^*,a,y} := \{b \in [k] : P_{NL}^{(i)}(a, b|x^*, y) > 0 \text{ for some } i \in J\} \tag{5.36}$$

Note that the above set depends on $\mathbb{LB}$ through $y \in \Phi_{\mathbb{LB}}^{x^*,a}$. Also all $b \in \Phi_{\mathbb{LB}}^{x^*,a,y}$ satisfy $b \neq \mathbb{LB}(y)$.

Remark 5.2. *The set $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} = \varnothing$ if and only if $P_{NL}^{(i)}(a|x^*) = 0$ for all $i \in J$. Indeed, let $\Phi_{\mathbb{LB}}^{x^*,a,y} = \varnothing$ for some $y \in \Phi_{\mathbb{LB}}^{x^*,a}$ then, $\sum_b P_{NL}^{(i)}(a,b|x^*,y) = 0 = P_{NL}^{(i)}(a|x^*)$ for all $i \in J$. It is straightforward to see that if $P_{NL}^{(i)}(a|x^*) = 0$ for all $i \in J$ then $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} = \varnothing$.*

[5] Also discussed in Section 3 of [2]

Consider the sets $\Phi_{\mathbb{LB}}^{x^*,a}$ and $\Phi_{\mathbb{LB}}^{x^*,a,y}$ defined above for the function $\mathbb{LB}$ and each $a \in \mathsf{A}$. Now, consider all relations $\mathcal{R}_{\mathbb{LB}} \subset [d] \times [k]$ that satisfy the following conditions:

(i) For $y \in [d], b \in [k]$, if $(y,b) \in \mathcal{R}_{\mathbb{LB}}$, then $b \neq \mathbb{LB}(y)$.

(ii) For every $a \in \mathsf{A}$ with $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$, there exists $y \in \Phi_{\mathbb{LB}}^{x^*,a}$, such that $(y,b) \in \mathcal{R}_{\mathbb{LB}}$ for all $b \in \Phi_{\mathbb{LB}}^{x^*,a,y}$.

For each function $\mathbb{LB} : [d] \to [k]$, let $\mathsf{M}_{\mathbb{LB}}$ be the set of such relations. We define the set $\mathsf{M} := \cup_{\mathbb{LB}:[d]\to[k]} \mathsf{M}_{\mathbb{LB}}$.

Lemma 5.4. *Given a function $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$, the set $\mathsf{M}_{\mathbb{LB}}$ is non-empty and there are at least two distinct elements in $\mathsf{M} := \cup_{\mathbb{LB}:[d]\to[k]} \mathsf{M}_{\mathbb{LB}}$.*

Proof. Given function $\mathbb{LB} : \mathsf{Y} \to \mathsf{B}$, for the non-local facet $Face\{P_{NL}^{(i)}\}_{i\in J}$ consider the associated input input $x^* \in [X]$ as in Lemma 5.3 (same as discussed above). For all $i \in J$, $\sum_{a\in\mathsf{A}} P_{NL}^{(i)}(a|x^*) = 1$. Thus, for each $i \in J$, there exists at least one $a \in \mathsf{A}$ such that $P_{NL}^{(i)}(a|x^*) > 0$. Consider the set of all $a \in \mathsf{A}$ such that $P_{NL}^{(i)}(a|x^*) > 0$ for some $i \in J$ and denote this set by $\tilde{\mathsf{A}} \subseteq \mathsf{A}$. Using Remark 5.2, for each $a \in \tilde{\mathsf{A}}$ there exists $y \in \Phi_{\mathbb{LB}}^{x^*,a}$ such that $\Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$. Indeed, as the contrary would imply that $P_{NL}^{(i)}(a|x^*) = 0 \; \forall \; i \in J$ for such an a. Also by definition of the set, for $y \in \Phi_{\mathbb{LB}}^{x^*,a}$, all $b \in \Phi_{\mathbb{LB}}^{x^*,a,y}$ satisfy $b \neq \mathbb{LB}(y)$. Consider a simple relation $\mathcal{R} = \{(y,b) \in [d] \times [k] : b \neq \mathbb{LB}(y)\}$. This relation satisfies condition (i) and (ii) for function $\mathbb{LB}$. The reason for satisfying condition (i) is immediate, and for (ii) the argument is as follows. Only for $a \in \tilde{\mathsf{A}}$, $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$. And for each $a \in \tilde{\mathsf{A}}$, take a $y \in \Phi_{\mathbb{LB}}^{x^*,a}$ such that $\Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$. Clearly, for this $y \in \Phi_{\mathbb{LB}}^{x^*,a}$, $(y,b) \in \mathcal{R}$ whenever $b \in \Phi_{\mathbb{LB}}^{x^*,a,y}$.

Now, consider another function $\mathbb{LB}' : \mathsf{Y} \to \mathsf{B}$, using similar argument relation $\mathcal{R}' = \{(y,b) \in [d] \times [k] : b \neq \mathbb{LB}'(y)\}$ satisfies condition (i) and (ii) for function $\mathbb{LB}'$. Now, it is immediate that $\mathcal{R}$ and $\mathcal{R}'$ are two distinct relations and both belong to M. ■

The input set in this task is $\mathsf{M} = \cup_{\mathbb{LB}:[d]\to[k]} \mathsf{M}_{\mathbb{LB}}$. The task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$ is tailored to the non-local facet $Face\{P_{NL}^{(i)}\}_{i\in J}$. In this task, Alice receives a random input $\mathsf{m} \in \mathsf{M}$, *i.e.*, a relation as defined above. Bob outputs $\mathsf{n} \in \mathsf{N} = [k]$. The parties share a noiseless classical channel $\mathscr{T} : [d] \to [d]$ in this task, *i.e.*, $T = [d]$ and $\mathscr{T}(\tau'|\tau) = \delta_{\tau',\tau}$. Let $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$ be the conditional probability distribution describing the probability that Bob receives message $\tau' \in [d]$ and outputs n when Alice's input is m. The payoff function $\mathcal{S}_W$ is then defined as

(5.37)

$$\mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n},\tau'} w^{\mathsf{m}}_{\tau',\mathsf{n}}\, p(\tau',\mathsf{n}|\mathsf{m}) \quad \text{where } w^{\mathsf{m}}_{\tau',\mathsf{n}} = \begin{cases} \mu_{\mathsf{m}} \text{ if } (\tau',\mathsf{n}) \in \mathsf{m} \\ 0 \quad \text{otherwise} \end{cases} ; \mu_{\mathsf{m}} \geq 0, \sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1$$

Additionally, in the task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$ we require the condition that, for each function $\mathbb{LB} : [d] \to [k]$ there must exist some $\mathsf{m} \in \mathsf{M}_{\mathbb{LB}}$ with $\mu_{\mathsf{m}} > 0$. The algebraic maximum of the payoff is $\mathcal{S}_W^{alg} = \sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1$ for this task. If $\mu_{\mathsf{m}} = 0$ for some $\mathsf{m} \in \mathsf{M}$, the input may be removed from the input set without effectively changing the task, since it does not affect the payoff. We now give the local bound on the payoff when Alice and Bob use shared randomness to assist the communication channel.

Theorem 5.4. *In* $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$*, the maximum payoff achievable using shared randomness assistance is*

$$s_\Lambda = \max_{\mathbb{D}:[d]\to[k]} [\sum_{\mathsf{m}\in\mathsf{M}:\exists\alpha\in[d]\wedge(\alpha,\mathbb{D}(\alpha))\in\mathsf{m}} \mu_{\mathsf{m}}] < 1$$

Proof. In the task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$, the payoff function $\mathcal{S}_W$ is linear in the conditional probabilities $p(\tau', \mathsf{n}|\mathsf{m})$. Hence, the maximum payoff with shared randomness can be achieved using some deterministic encoding and decoding strategy. Let Bob's deterministic decoding be given by $\mathbb{D} : [d] \to [k]$, so that on receiving τ', his output is $\mathsf{n} = \mathbb{D}(\tau')$. Consider the set of Alice's inputs $\mathsf{m} \in \mathsf{M}$ such that for all $t \in [d]$, $(t, \mathbb{D}(t)) \notin \mathsf{m}$. We define the set of such inputs as $\mathsf{M}_{\mathbb{D}}$. For example, each $\mathsf{m} \in \mathsf{M}_{\mathbb{LB}}$ where $\mathbb{LB}(t) = \mathbb{D}(t)$ for all $t \in [d]$ also lie in $\mathsf{M}_{\mathbb{D}}$. Alice's deterministic encoding is given by $\mathbb{E} : \mathsf{M} \to [d]$, where she sends $\tau = \mathbb{E}(\mathsf{m})$. For $\mathsf{m} \in \mathsf{M}_{\mathbb{D}}$, the resulting distribution satisfies $p(\tau', \mathsf{n}|\mathsf{m}) = 0$ if $(\tau', \mathsf{n}) \in \mathsf{m}$ (using the definition of $\mathsf{M}_{\mathbb{D}}$). For such inputs $\mathsf{m} \in \mathsf{M}_{\mathbb{D}}$, the payoff is zero when the parties use a deterministic encoding $\mathbb{E}$ and decoding $\mathbb{D}$. Therefore, the total payoff when Bob uses deterministic decoding $\mathbb{D}$ is $\leq \sum_{\mathsf{m}\in\mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}} \mu_{\mathsf{m}}$. If Bob uses the decoding $\mathbb{D}$, this upper bound can be attained if Alice encodes as follows: for each $\mathsf{m} \in \mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}$ she sends $\tau \in [d]$ such that $(\tau' = \tau, \mathbb{D}(\tau')) \in \mathsf{m}$. The existence of $\tau \in [d]$ for each $\mathsf{m} \in \mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}$ follows from the definition of $\mathsf{M}_{\mathbb{D}}$.

Thus, the maximum payoff achievable when Bob uses the deterministic decoding strategy $\mathbb{D} : [d] \to \mathsf{N}$ is $\sum_{\mathsf{m}\in\mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}} \mu_{\mathsf{m}}$. Maximising over all deterministic decoding gives the maximum payoff, which is achievable using shared randomness assistance to the channel, *i.e.*,

$$s_\Lambda = \max_{\mathbb{D}:[d]\to[k]} [\sum_{\mathsf{m}\in\mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}} \mu_{\mathsf{m}}] = \max_{\mathbb{D}:[d]\to[k]} [\sum_{\mathsf{m}\in\mathsf{M}:\exists\alpha\in[d]\wedge(\alpha,\mathbb{D}(\alpha))\in\mathsf{m}} \mu_{\mathsf{m}}] \tag{5.38}$$

For each decoding strategy $\mathbb{D} : [d] \to [k]$ consider function $\mathbb{LB} : [d] \to [k]$ such that $\mathbb{LB}(t) = \mathbb{D}(t)\ \forall t \in [d]$. Every $\mathsf{m} \in \mathsf{M}_{\mathbb{LB}}$ also belong to $\mathsf{M}_{\mathbb{D}}$. Since there always exists some $\mathsf{m} \in \mathsf{M}_{\mathbb{LB}}$ such that $\mu_{\mathsf{m}} > 0$ (from the definition of the task), Alice's deterministic encodings yield zero payoff for such inputs. Thus, $s_\Lambda < \mathcal{S}_W^{alg} = \sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1$. ■

We now show the advantage of assistance from non-local correlations for this task.

Theorem 5.5. *For* $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$*, assistance from any no-signalling correlation* $P \in Face\{P_{NL}^{(i)}\}_{i\in J}$ *leads to the payoff* $\mathcal{S}_W^{alg} = 1$*.*

Proof. In task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$, suppose Alice and Bob share a correlation $P \in Face\{P_{NL}^{(i)}\}_{i\in J}$ on the non-local facet of the no-signalling polytope $\mathcal{NS} := \{P = \{P(a,b|x,y)\}_{a\in\mathsf{A}, b\in[k], x\in\mathsf{X}, y\in[d]}\}$. Each input of Alice $\mathsf{m} \in \mathsf{M}$ is a relation $\mathsf{m} \subset [d] \times [k] \in \mathsf{M}_{\mathbb{LB}}$, where $\mathbb{LB} : [d] \to [k]$ is associated function. For such an input, there exists $x^* \in \mathsf{X}$, independent of the choice of $P_{NL}^{(i)}$ where $i \in J$ such that: for each $a \in \mathsf{A}$ if $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$, then there exists $y \in \Phi_{\mathbb{LB}}^{x^*,a}$, such that $(y,b) \in \mathsf{m}$ for all $b \in \Phi_{\mathbb{LB}}^{x^*,a,y}$.

Alice and Bob use the following strategy. After receiving the input $\mathsf{m} \in \mathsf{M}$, Alice identifies the corresponding function $\mathbb{LB} : [d] \to [k]$ and uses the associated $x^* \in \mathsf{X}$ as her input to the correlation P. If the outcome is $a \in \mathsf{A}$ with $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$, she sends $\tau \in \Phi_{\mathbb{LB}}^{x^*,a}$ such that $(\tau, b) \in \mathsf{m}$ for all $b \in \Phi_{\mathbb{LB}}^{x^*,a,\tau}$. Existence of such τ follows from condition (ii). For outcomes a where $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} = \varnothing$, she send message $\tau = 1$. Note that this branch of encoding will never arise for $P \in Face\{P_{NL}^{(i)}\}_{i\in J}$. Bob, after receiving $\tau' = \tau$ as the message, uses it as his input to the correlation P. Then upon getting outcome $b \in [k]$ from P, he produces it as output, *i.e.*, $\mathsf{n} = b$.

Assume $P_{NL} = P_{NL}^{(i)}$ for some $i \in J$. For input $\mathsf{m} \in \mathsf{M}$, when Alice queries x^* to the shared correlation, if $P_{NL}^{(i)}(a|x^*) > 0$ for some $a \in \mathsf{A}$ then $\cup_{y\in\Phi_{\mathbb{LB}}^{x^*,a}} \Phi_{\mathbb{LB}}^{x^*,a,y} \neq \varnothing$ (see Remark 5.2). According to the protocol, for such output $a \in \mathsf{A}$, Alice sends $\tau \in \Phi_{\mathbb{LB}}^{x^*,a}$ such that $(\tau, b) \in \mathsf{m}$ for all $b \in \Phi_{\mathbb{LB}}^{x^*,a,\tau}$. After receiving the message $\tau' = \tau$, Bob uses it as his input to the shared correlation and obtains an outcome $b' \in [k]$, thus $P_{NL}^{(i)}(b'|x^*, \tau', a) > 0$. By definition, $b' \in \Phi_{\mathbb{LB}}^{x^*,a,\tau'}$. From the protocol and using the fact that m satisfies condition (ii), it implies that always $(\tau', \mathsf{n} = b') \in \mathsf{m}$ for the input $\mathsf{m} \in \mathsf{M}$, leading to a payoff μ_{m}. Thus, the above described strategy for each $\mathsf{m} \in \mathsf{M}$ while sharing correlation $P_{NL}^{(i)}$ ensures that the overall payoff is $\sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1 = \mathcal{S}_W^{alg}$.

This strategy is invariant over shared correlation $P_{NL} = P_{NL}^{(i)}$, where $i \in J$, and always yields to payoff $\mathcal{S}_W^{alg}$. Therefore, it also yields payoff $\mathcal{S}_W^{alg}$ when $P \in Face\{P_{NL}^{(i)}\}_{i\in J}$ using the linearity of the payoff in the conditional probabilities $p(\tau', \mathsf{n}|m)$ and convexity of set $Face\{P_{NL}^{(i)}\}_{i\in J}$. ■

5.2.2.1 Communication task for extremal correlation with dichotomic outputs

We consider some communication tasks corresponding to a no-signalling extremal correlation with dichotomic outputs. In the simplest Bell scenario with $\mathsf{X} = \mathsf{Y} = \mathsf{A} = \mathsf{B} = [2]$, there are 8 non-local extremal correlations, known as PR-boxes, denoted by $P_{NL}^{(\alpha,\beta,\gamma)} = \{P_{NL}^{(\alpha,\beta,\gamma)}(a,b|x,y)\}_{a\in[2],b\in[2],x\in[2],y\in[2]}$ (see Equation (2.19)). These are defined as $P_{NL}^{(\alpha,\beta,\gamma)}(a,b|x,y) = \frac{\delta_{(a-1)\oplus(b-1), f(x,y,\alpha,\beta,\gamma)}}{2}$, where $f(x,y,\alpha,\beta,\gamma) = (x-1)(y-1) \oplus \alpha(x-1) \oplus$

$\beta(y-1) \oplus \gamma$ and $(\alpha, \beta, \gamma) \in \{0,1\}^3$. Now consider a bipartite Bell scenario with dichotomic outputs, *i.e.* $\mathsf{X} = [X], \mathsf{Y} = [d], \mathsf{A} = \mathsf{B} = [2]$ and $X, d \in \mathbb{N}, X \geq 2, d \geq 2$. Let $P_{NL} = \{P_{NL}(a,b|x,y)\}_{a\in[2],b\in[2],x\in[X],y\in[d]} \in \{P_{NL}^{(i)}\}$ denote some non-local extremal correlation of $\mathcal{NS}$ such that for every $y', y'' \in [d]$, there is $x', x'' \in [X]$ for which $\{P_{NL}(a,b|x,y)\}_{a\in[2],b\in[2],x\in\{x',x''\},y\in\{y',y''\}}$ is equivalent to one of the two-input two-output PR-boxes described above.

Tailored to the extremal correlation P_{NL}, we describe the task $\mathbf{CW}[\{P_{NL}\}, d, k=2]$. Alice's input set M consists of the relations $\mathcal{R}_{ijkl} := \{(i,k),(j,l)\} \subset [d] \times [2]$ where $i < j$, $i, j \in [d]$ and $k, l \in [2]$. Each input of Alice is a relation in $\mathsf{m} \in \mathsf{M}$. Bob's output set $\mathsf{N} = [2]$. The parties communicate with a noiseless classical channel $\mathscr{T} : [d] \to [d]$. If $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$ denotes the conditional probability distribution describing the probability that Bob receives message $\tau' \in [d]$ and outputs n when Alice's input is m. The payoff function is defined as

$$\mathcal{S}_W(\mathbf{p}) = \sum_{\mathsf{m},\mathsf{n},\tau'} w^{\mathsf{m}}_{\tau',\mathsf{n}}\, p(\tau', \mathsf{n}|\mathsf{m}) \;\text{ where } w^{\mathsf{m}}_{\tau',\mathsf{n}} = \begin{cases} \mu_{\mathsf{m}} = \frac{1}{2d(d-1)} \text{ if } (\tau', \mathsf{n}) \in \mathsf{m} \\ 0 \quad \text{otherwise} \end{cases} \tag{5.39}$$

We now derive the local bound on the payoff and show that assistance from P_{NL} achieves the algebraic maximum.

Corollary 5.2. *In* $\mathbf{CW}[\{P_{NL}\}, d, k=2]$, (i) *the maximum payoff with shared randomness assistance is* $s_\Lambda = \frac{3}{4}$ *and* (ii) *using the correlation* P_{NL} *leads to the payoff* $\mathcal{S}_W^{alg} = 1$.

Proof. *Proof of* (i): Since the payoff function is linear in the conditional probabilities, the maximum payoff with shared randomness can be obtained using a deterministic encoding-decoding strategy. Let Bob's deterministic decoding be $\mathbb{D} : [d] \to [2]$, so that upon receiving τ', he outputs $\mathsf{n} = \mathbb{D}(\tau')$. Consider Alice's input $\mathsf{m} = \mathcal{R}_{ijkl} \in \mathsf{M}$ with $k \neq \mathbb{D}(i)$ and $l \neq \mathbb{D}(j)$. We denote the set of such inputs by $\mathsf{M}_{\mathbb{D}}$. For each choice of $i, j \in [d]$ where $i < j$, there is one such relation $\mathcal{R}_{ijkl}$. Thus $|M_{\mathbb{D}}| = \binom{d}{2}$. Consider any arbitrary deterministic encoding $\mathbb{E} : M \to [d]$ of Alice, where she sends the message $\tau = \mathbb{E}(\mathsf{m})$ corresponding to the input $\mathsf{m} \in \mathsf{M}$. Using the definition of set $\mathsf{M}_{\mathbb{D}}$, the resulting conditional probability $p(\tau', \mathsf{n}|\mathsf{m}) = 0$ if $\mathsf{m} \in M_{\mathbb{D}}$ and $(\tau', \mathsf{n}) \in \mathsf{m}$. Thus, for $\mathsf{m} \in \mathsf{M}_{\mathbb{D}}$, the payoff is zero while using the deterministic encoding strategy $\mathbb{E}$ and decoding strategy $\mathbb{D}$. The total payoff using this strategy is then $\leq \sum_{\mathsf{m}\in\mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}} \mu_{\mathsf{m}} = [4\binom{d}{2} - \binom{d}{2}]\frac{1}{2d(d-1)} = \frac{3}{4}$. This upper bound is the same for each chosen deterministic decoding strategy $\mathbb{D}$. The upper bound is achieved if Alice's encoding is as follows. For input $\mathsf{m} = \mathcal{R}_{ijkl} \in \mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}$, either $k = \mathbb{D}(i)$ or $l = \mathbb{D}(j)$. Alice sends the message $\tau = \mathbb{E}(\mathsf{m}) = i$ if $k = \mathbb{D}(i)$ else she sends $\tau = \mathbb{E}(\mathsf{m}) = j$ for this input $\mathsf{m} \in \mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}$. Then Bob outputs $\mathsf{n} = \mathbb{D}(\tau') = \mathbb{D}(\tau)$, so $(\tau', \mathbb{D}(\tau')) \in \mathsf{m}$, yielding payoff $\mu_{\mathsf{m}} = \frac{1}{2d(d-1)}$ for the input $m \in M\setminus M_{\mathbb{D}}$. The net payoff is $\sum_{\mathsf{m}\in\mathsf{M}\setminus\mathsf{M}_{\mathbb{D}}} \mu_{\mathsf{m}} = [4\binom{d}{2} - \binom{d}{2}]\frac{1}{2d(d-1)} = \frac{3}{4}$.

Proof of (ii): If the parties share the correlation P_{NL}, the following protocol achieves the algebraic maximum payoff $\mathcal{S}_W^{alg} = 1$. After receiving the message $\tau' = \tau$, Bob uses it as his input to the correlation P_{NL} and, upon obtaining outcome $b \in [2]$ from the correlation. produces it as his output, *i.e.*, $\mathsf{n} = b$. Alice's input $\mathsf{m} = \mathcal{R}_{ijkl} := \{(i,k),(j,l)\} \subset [d] \times [2] \in \mathsf{M}$ where $i < j$, $i,j \in [d]$ and $k,l \in [2]$. For a given $\mathsf{m} = \mathcal{R}_{ijkl}$ consider $y' = i, y'' = j$. From the description of P_{NL}, there exist $x', x'' \in [X]$ (take $x'' > x'$ *wlog*) such that $\{P_{NL}(a,b|x,y)\}_{a\in[2],b\in[2],x\in\{x',x''\},y\in\{y',y''\}}$ is equivalent to a two-input, two-output PR-box (Equation (2.19)). For simplicity, assume it is equivalent to $\{P_{NL}(a,b|x,y)\}_{a\in[2],b\in[2],x\in\{x',x''\},y\in\{y',y''\}} = P_{NL}^{(0,0,0)}$, *i.e.*, $\alpha = 0, \beta = 0, \gamma = 0$. Equivalent encodings can be defined for other (α,β,γ) values, as the different PR-boxes differ only by local labellings. For input relation $m = \mathcal{R}_{ijkl}$, if $k = l$, Alice uses $x'' \in [X]$ as her input to the correlation P_{NL}, and if $k \neq l$, Alice uses $x' \in [X]$ as her input. If $k = 1$, depending on her outcome obtained from P_{NL}, she sends the message $\tau = i$ if $a = 1$ and $\tau = j$ if $a = 2$. Similarly, if $k = 2$, she sends the message $\tau = j$ if $a = 1$ and $\tau = i$ if $a = 2$. For this input $\mathsf{m} = \mathrm{R}_{ijkl}$, using the above strategy along with Bob's decoding, it is straightforward to see that the communicated message τ' and Bob's output n are such that $(\tau', \mathsf{n}) \in \mathsf{m}$. Consequently, it results in a payoff μ_{m} for input $\mathsf{m} = \mathcal{R}_{ijkl}$. Since we did not assume a particular assignment for $i,j \in [d]$, Alice can use a similar encoding for each pair of $i,j \in [d]$. The payoff with the above strategy is μ_{m} for each $\mathsf{m} \in \mathsf{M}$. The net payoff is $\sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1$.

■

Corollary 5.3. *If $\frac{1}{2} < p \leq 1$ then assistance from correlation $pP_{NL} + (1-p)P_{WN}$ is advantageous over shared randomness in the task* $\mathbf{CW}[\{P_{NL}\}, d, k = 2]$.

Proof. Let the shared correlation be $P_{WN} = \{P_{WN}(a,b|x,y) = \frac{1}{4}\}_{a\in[2],b\in[2],x\in[X],y\in[d]}$ to assist the communication channel. Suppose Alice and Bob use the same encoding and decoding strategy as in the proof of (ii) in Corollary 5.2. For each input $\mathsf{m} = \mathcal{R}_{ijkl}$, $p(\tau', \mathsf{n}|\mathsf{m}) = \frac{1}{4}$ if $\tau' \in \{i,j\}$ and $\mathsf{n} \in [2]$ using the strategy. Thus, the payoff for each $\mathsf{m} = \mathcal{R}_{ijkl}$ is $\mu_{\mathsf{m}}(p(\tau' = i, \mathsf{n} = k|\mathsf{m}) + p(\tau' = j, \mathsf{n} = l|\mathsf{m})) = \frac{\mu_{\mathsf{m}}}{2}$. The total payoff with this strategy is therefore $\sum_{\mathsf{m}\in\mathsf{M}} \frac{\mu_{\mathsf{m}}}{2} = \frac{1}{2}$. Now suppose the shared correlation is $\tilde{P} = pP_{NL} + (1-p)P_{WN}$. Using the same strategy as in the proof of (ii) of Corollary 5.2, and noting that the payoff is linear in the conditional probabilities, the resulting payoff is $p(1) + (1-p)\frac{1}{2}$ as the payoff function as the payoff function is linear in conditional probabilities. This payoff is higher than $s_\Lambda = \frac{3}{4}$ if $p + \frac{1-p}{2} > \frac{3}{4} \implies p > \frac{1}{2}$.
■

5.2.2.2 Communication task for extremal correlations violating $I3322$ inequality

We now discuss a task for a non-local facet of the no-signalling polytope $\mathcal{NS} := \{P = \{P(a,b|x,y)\}_{a\in[2],b\in[2],x\in[3],y\in[3]}\}$. We define the extremal correlation $P_{NL}^{(1)} = \{P_{NL}^{(1)}(a,b|x,y)\}_{a\in[2],b\in[2],x\in[3],y\in[3]}$ by $P_{NL}^{(1)}(a,b|x,y) = \frac{\delta_{a,b}}{2}$ if $(x,y) \in \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,2),(3,3)\}$ else $P_{NL}^{(1)}(a,b|x,y) = \frac{1-\delta_{a,b}}{2}$. Similarly extremal correlations $P_{NL}^{(2)} = \{P_{NL}^{(2)}(a,b|x,y)\}_{a\in[2],b\in[2],x\in[3],y\in[3]}$, where $P_{NL}^{(2)}(a,b|x,y) = \frac{\delta_{a,b}}{2}$

if $(x,y) \in \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,2)\}$ else $P^{(2)}_{NL}(a,b|x,y) = \frac{1-\delta_{a,b}}{2}$. Consider non-local facet $Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$. Individually, $P^{(1)}_{NL}$ and $P^{(2)}_{NL}$ maximally violates the $I3322$ Bell inequality [289, 290] given below and using the linearity of the Bell function, correlation on this facet $Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$ also maximally violates the same inequality.

$$\begin{aligned}\mathbf{B}(P) = &- P_A(2|2) - P_B(2|1) - 2P_B(2|2) + P(2,2|1,1) + P(2,2|1,2) + P(2,2|2,1) + \\ &P(2,2|2,2) - P(2,2|1,3) + P(2,2|2,3) - P(2,2|3,1) + P(2,2|3,2)\end{aligned} \tag{5.40}$$

Here, the marginal correlations $P_A(a|x) = \sum_{b\in[2]} P(a,b|x,y)$ for $a \in [2]$ and $x \in [3]$. Similarly, marginal correlations $P_B(b|y) = \sum_{a\in[2]} P(a,b|x,y)$ for $b \in [2]$ and $y \in [3]$. The local bound for the Bell inequality is $\beta_L = 0$, and the no-signalling maximum violation is 1 [289].

The task $\mathbf{CW}[\{P^{(1)}_{NL}, P^{(2)}_{NL}\}, d = 3, k = 2]$ corresponds to the facet $Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$. Alice's input set M consists of relations $\mathcal{R}_{ijkl} := \{(i,k),(j,l)\} \subset [3] \times [2]$ where $i < j$, $i, j \in [3]$ and $k, l \in [2]$. Bob's output set is $\mathsf{N} = [2]$. The parties use a noiseless channel $\mathscr{T} : [3] \to [3]$. Using the conditional probability distribution $\mathbf{p} = \{p(\tau', \mathsf{n}|\mathsf{m})\}_{\tau',\mathsf{n},\mathsf{m}}$ from the task, the payoff function is same as in Equation (5.39) with $d = 3$ and $\mu_{\mathsf{m}} = \frac{1}{2d(d-1)} = \frac{1}{12}$. From Corollary 5.2, the maximum payoff using shared randomness assistance or local bound is $s_\Lambda = \frac{3}{4}$. We now show the advantage of the assistance from no-signalling correlations.

Corollary 5.4. *In* $\mathbf{CW}[\{P^{(1)}_{NL}, P^{(2)}_{NL}\}, d = 3, k = 2]$*, assistance from any correlation* $P \in Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$ *achieves the payoff* $\mathcal{S}^{alg}_W = 1$.

Proof. Let the parties share a correlation $P \in Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$ for assistance. The following protocol achieves the algebraic maximum payoff $\mathcal{S}^{alg}_W = 1$. After receiving the message $\tau' = \tau$, Bob uses it as his input to the correlation P. On obtaining outcome $b \in [2]$ from P he produces it as his output for the task, *i.e.*, $\mathsf{n} = b$. Alice's strategy is as follows. For the inputs $\mathsf{m} = \mathcal{R}_{12kl} \in \mathsf{M}$, if $k = l$ then Alice chooses $x = 3$ as her input to the correlation P, and if $k \neq l$, Alice chooses $x = 1$ as her input. If $l = 2$ in the above input, she sends the message $\tau = 1$ if she gets $a = 1$ from P and $\tau = 2$ if $a = 2$. Similarly, if $l = 1$ in the above input, she sends the message $\tau = 2$ if she gets $a = 1$ from P and $\tau = 1$ if $a = 2$. For other inputs $\mathsf{m} = \mathcal{R}_{ijkl} \in \mathsf{M}$ with $(i,j) \neq (1,2)$, if $k = l$ then Alice chooses $x = 1$ as her input to the correlation P, and if $k \neq l$, she chooses $x = 2$ as her input. For these inputs, if $k = 1$, she sends the message $\tau = i$ if the outcome from P is $a = 1$ and $\tau = j$ if $a = 2$. Similarly, if $k = 2$ in these inputs, then she sends the message $\tau = j$ if the outcome from P is $a = 1$ and $\tau = i$ if $a = 2$. For each input $\mathsf{m} = \mathcal{R}_{ijkl}$, the above protocol ensures that the communicated message τ' and Bob's output n are such that $(\tau', \mathsf{n}) \in \mathsf{m}$. The payoff for each input m is $\mu_{\mathsf{m}} = \frac{1}{12}$. The total payoff is then $\sum_{\mathsf{m}\in\mathsf{M}} \mu_{\mathsf{m}} = 1$. ■

Corollary 5.5. *If* $P \in Face\{P^{(1)}_{NL}, P^{(2)}_{NL}\}$ *and* $\frac{1}{2} < p \leq 1$*, then assistance from* $\tilde{P} = pP + (1-p)P_{WN}$ *is advantageous over shared randomness in the task* $\mathbf{CW}[\{P^{(1)}_{NL}, P^{(2)}_{NL}\}, d = 3, k = 2]$.

Proof. The proof is similar to the proof of Corollary 5.3. Using the protocol given in the proof of Corollary 5.4 with correlation P_{WN} yields a payoff of $\frac{1}{2}$. As the payoff function is linear in conditional probabilities, the resulting payoff is $p(1) + (1-p)\frac{1}{2}$ while using correlation $\tilde{P} = pP + (1-p)P_{WN}$. This payoff is higher than $s_\Lambda = \frac{3}{4}$ if $p + \frac{1-p}{2} > \frac{3}{4} \implies p > \frac{1}{2}$. ■

5.3 Hardy Correlations Advances Classical Communication

We now present communication tasks that demonstrate the advantage of sharing Hardy-type correlations with dichotomic inputs $\mathsf{X} = \mathsf{Y} = \{0, 1\}$, *i.e.*, $x, y \in \{0, 1\}$. Recall the Hardy-type correlations in the Bell scenario with $\mathsf{X} = \mathsf{Y} = [2]$ and $\mathsf{A} = \mathsf{B} = [d]$, where $d \in \mathbb{N}$ and $d \geq 2$, as introduced in Equation (2.21). For convenience, we relabel the inputs by $2 \mapsto 1$ and $1 \mapsto 0$, so that $x, y \in \{0, 1\}$. The outputs remain $a, b \in [d]$. A no-signalling correlation $P = \{P(a, b|x, y)\}_{a\in[d], b\in[d], x\in\{0,1\}, y\in\{0,1\}}$ is said to be of Hardy type [186] if it satisfies:

$$P(a > b|0, 0) = 0, \quad P(a < b|0, 1) = 0, \quad P(a < b|1, 0) = 0, \quad P(a < b|1, 1) = \mathrm{q} > 0 \tag{5.41}$$

where $\mathrm{q} \in \mathbb{R}$ and $P(a > b|x, y) = \sum_{\substack{a\in\mathsf{A}, b\in\mathsf{B}:\\ a>b}} P(a, b|x, y)$. Similarly, $P(a < b|x, y) = \sum_{\substack{a\in\mathsf{A}, b\in\mathsf{B}:\\ a<b}} P(a, b|x, y)$. For quantum-realisable no-signalling correlations $P \in \mathcal{Q}$ satisfying Equation (5.41), let $H^d_{Hardy}(P) = P(a < b|1, 1)$ denote the probability of the non-local event (defined in Equation (2.22)).

Remark 5.3. *From [186], the maximum value of $H^d_{Hardy}(P)$ for quantum realisable no-signalling correlation $P \in \mathcal{Q}$ that satisfy constraints in Equation (5.41), increases with d and is almost* 0.5.

We now introduce a communication task where the payoff achievable with shared randomness is always suboptimal, while Hardy-type correlations with dichotomic inputs provide an advantage. For this purpose, we define parameters $k_{\mathsf{m}}, K_{\mathsf{m}}$ and L_{m} for $(d+1)$ tuples $\mathsf{m} = (\mathsf{m}_{(0)}, \mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \in \{0, 1\}^{d+1}$ whenever they exist. These values are not defined for the trivial tuples $(0, 0, \cdots, 0)$ and $(1, 1, \cdots, 1)$.

$$k_{\mathsf{m}} := \min\{i \in [d] \text{ such that } \mathsf{m}_{(i)} = 0\}, \tag{5.42}$$

$$K_{\mathsf{m}} := \max\{i \in [d] \text{ such that } \mathsf{m}_{(i)} = 0\}, \tag{5.43}$$

$$L_{\mathsf{m}} := \max\{i \in [d] \text{ such that } \mathsf{m}_{(i)} = 1\} \tag{5.44}$$

In the task denoted by $\mathbf{CS}[d, \alpha]$, where $d \geq 2, d \in \mathbb{N}$ and $\alpha \geq 1$ is a real parameter. The input set $\mathsf{M} \subset \{0, 1\}^{d+1}$ is defined in Equation (5.45). Alice's input $\mathsf{m} = (\mathsf{m}_{(0)}, \mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \in \mathsf{M}$ is a $(d+1)$ tuple.

$$\mathsf{M} = \Big\{(0,0,\cdots,0),\ (1,1,\cdots,1)\Big\} \cup \mathsf{M}_1 \cup \mathsf{M}_2 \cup \mathsf{M}_3 \cup \mathsf{M}_4 \quad \text{where} \tag{5.45}$$

$$\mathsf{M}_1 = \Big\{\mathsf{m} : \mathsf{m}_{(0)} = 0,\ 1 < L_{\mathsf{m}} < d,\ \mathsf{m}_{(i)} = 0\ \forall\ i \neq L_{\mathsf{m}}\Big\} \tag{5.46}$$

$$\begin{aligned}\mathsf{M}_2 =& \Big\{\mathsf{m} : \mathsf{m}_{(0)} = 0,\ 2 \leq k_{\mathsf{m}} \leq d-1,\ k_{\mathsf{m}}+1 \leq L_{\mathsf{m}} \leq d,\ \mathsf{m}_{(j)} = 1 \text{ for} \\ & 1 \leq j \leq k_{\mathsf{m}}-1, \mathsf{m}_{(i)} = 0 \text{ for } (k_{\mathsf{m}} \leq i \leq L_{\mathsf{m}}-1), \mathsf{m}_{(i)} = 0 \text{ for } L_{\mathsf{m}}+1 \leq i \leq d\Big\}\end{aligned} \tag{5.47}$$

$$\begin{aligned}\mathsf{M}_3 = \Big\{&\mathsf{m} : \mathsf{m}_{(0)} = 0,\ 2 \leq k_{\mathsf{m}} \leq d,\ L_{\mathsf{m}} = k_{\mathsf{m}}-1,\ \mathsf{m}_{(i)} = 0 \text{ for } k_{\mathsf{m}} \leq i \leq d, \\ & \mathsf{m}_{(j)} = 1 \text{ for } 1 \leq j \leq k_{\mathsf{m}}-1\Big\}\end{aligned} \tag{5.48}$$

$$\mathsf{M}_4 = \Big\{\mathsf{m} : \mathsf{m}_{(0)} = 1, \mathsf{m}_{(1)} = 0, \cdots, \mathsf{m}_{(r)} = 0,\ \mathsf{m}_{(r+1)} = 1, \cdots, \mathsf{m}_{(d)} = 1, 1 \leq r \leq d-1\Big\} \tag{5.49}$$

Bob produces output $\mathsf{n} = (\mathsf{n}_{(0)}, \mathsf{n}_{(1)}) \in \mathsf{N} = \{0,1\} \times [d]$. Alice can use a noiseless one-bit channel $\mathscr{T} : \{0,1\} \to \{0,1\}$. Let $\mathbf{p} = \{p(\mathsf{n}|\mathsf{m})\}_{\mathsf{n},\mathsf{m}}$ is the probability that Bob outputs $\mathsf{n} \in \mathsf{N}$ when Alice receives input $\mathsf{m} \in \mathsf{M}$. This task does not involve wire-reading. The payoff function is defined as

$$\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{m}\in\mathsf{M},\mathsf{n}\in\mathsf{N}} w_{\mathsf{n}}^{\mathsf{m}} p(\mathsf{n}|\mathsf{m}) \quad \text{where} \tag{5.50}$$

$$w_{\mathsf{n}}^{\mathsf{m}} = \begin{cases} +1 & \text{if } (\mathsf{m}_{(0)} = 1, \prod_{i>0} \mathsf{m}_{(i)} = 1, \mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)}) > 1). \\ & \text{if } (\mathsf{m}_{(0)} = 0, \mathsf{n}_{(0)} = 0, \mathsf{n}_{(1)} < k_{\mathsf{m}}) \vee (\mathsf{m}_{(0)} = 0, \mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)} > L_{\mathsf{m}}) \vee \\ -\alpha & \quad (\mathsf{m}_{(0)} = 1, \mathsf{n}_{(0)} = 0, \mathsf{n}_{(1)} > K_{\mathsf{m}}) \vee \mathsf{m}_{(0)} = 1, \mathsf{n}_{(0)} = 0, \prod_{i>0} \mathsf{m}_{(i)} = 1) \vee \\ & \quad (\mathsf{m}_{(0)} = 0, \mathsf{n}_{(0)} = 1, \sum_{i>0} \mathsf{m}_{(i)} = 0). \\ -1 & \text{if } (\mathsf{m}_{(0)} = 1, \mathsf{n}_{(0)} = 1, \mathsf{m}_{(1)} = \cdots = \mathsf{m}_{(r)} = 0, \mathsf{m}_{(r+1)} = \cdots = \mathsf{m}_{(d)} = 1, \mathsf{n}_{(0)} = 1, \\ & \quad \mathsf{n}_{(1)} = r+1) \text{ where } 1 \leq r \leq d-1. \\ 0 & \text{otherwise.} \end{cases} \tag{5.51}$$

Here $\alpha \geq 1$ is the penalty parameter. We refer to this communication task as $\mathbb{C}[d,\alpha]$. Now, we discuss the payoff for the task when the classical communication is assisted by shared randomness. This task $\mathbb{C}[d,\alpha]$ for the case $d = 2$ and $\alpha = -\infty$ resembles the task considered in [165].

Theorem 5.6. *In the task* $\mathbf{CS}[d,\alpha]$*, the maximum payoff achievable with shared randomness assistance to a one-bit channel is* $s_\Lambda \leq 0$*.*

Proof. Since the payoff is linear in conditional probabilities $p(\mathsf{n}|\mathsf{m})$, the maximum payoff with shared randomness assistance s_Λ can be achieved using an optimal deterministic strategy. Let Bob's optimal deterministic decoding is denoted by function $\mathbb{D}:\{0,1\}\rightarrow \mathsf{N}$, so that upon receiving $\tau'\in\{0,1\}$ he outputs $\mathsf{n}=(\mathsf{n}_{(0)},\mathsf{n}_{(1)})=\mathbb{D}(\tau')\in\mathsf{N}$. Similarly, the optimal encoding for Alice can be denoted by the function $\mathbb{E}:\mathsf{M}\rightarrow\{0,1\}$. The strategy results in the conditional probabilities $\{p(\mathsf{n}|\mathsf{m})\}$ where $p(\mathsf{n}|\mathsf{m})=\sum_\tau \delta_{\tau,\mathbb{E}(\mathsf{m})}\delta_{\mathsf{n},\mathbb{D}(\tau)}=\delta_{\mathsf{n},\mathbb{D}(\mathbb{E}(\mathsf{m}))}\in\{0,1\}$ which represents the deterministic encoding and decoding explicitly. Suppose, for contradiction, that a strategy achieves $s_\Lambda>0$. Let us denote $\mathbb{D}(0)=\tilde{\mathsf{n}}=(\tilde{\mathsf{n}}_{(0)},\tilde{\mathsf{n}}_{(1)})$ and $\mathbb{D}(1)=\bar{\mathsf{n}}=(\bar{\mathsf{n}}_{(0)},\bar{\mathsf{n}}_{(1)})$. For each $\mathsf{m}\in\mathsf{M}$, since $\tau=\tau'=\mathbb{E}(\mathsf{m})\in\{0,1\}$ therefore[6],

$$
\begin{aligned}
&p(\tilde{\mathsf{n}}_{(0)},\tilde{\mathsf{n}}_{(1)}|\mathsf{m})+p(\bar{\mathsf{n}}_{(0)},\bar{\mathsf{n}}_{(1)}|\mathsf{m})=1 \quad \text{and}\\
&p(\tilde{\mathsf{n}}_{(0)},\tilde{\mathsf{n}}_{(1)}|\mathsf{m}),p(\bar{\mathsf{n}}_{(0)},\bar{\mathsf{n}}_{(1)}|\mathsf{m})\in\{0,1\}\ \forall\ \mathsf{m}
\end{aligned}
\tag{5.52}
$$

Notice that the only input-output pair for which Bob can obtain a positive payoff is $\mathsf{m}=(1,1,\cdots,1)$ when $\mathsf{n}=(\mathsf{n}_{(0)}=1,\mathsf{n}_{(1)})$ with $\mathsf{n}_{(1)}>1$. For this particular input m, $p(\mathsf{n}_{(0)}=1,\mathsf{n}_{(1)}|\mathsf{m})>0$ for some $\mathsf{n}_{(1)}>1$. Therefore, using Equation (5.52), without loss of generality, we can fix $\tilde{\mathsf{n}}_{(0)}=1$ and $\tilde{\mathsf{n}}_{(1)}>1$. Then from Equation (5.50) and Equation (5.51), $+1$ payoff is obtained for this input when $p(\tilde{\mathsf{n}}_{(0)},\tilde{\mathsf{n}}_{(1)}|\mathsf{m})=1$, *i.e.* if $\mathbb{E}(\mathsf{m})=\tau=0$ in this case.

Next consider input $\mathsf{m}=(0,0,\cdots,0)$. For this input if $p(\tilde{\mathsf{n}}_{(0)}=1,\tilde{\mathsf{n}}_{(1)}|\mathsf{m})=1$ then from Equation (5.50) and (5.51), the payoff is $-\alpha$ (≥ -1) and the overall payoff becomes $s_\Lambda\leq 0$. Using Equation (5.52) and assuming $s_\Lambda>0$ implies $p(\bar{\mathsf{n}}_{(0)},\bar{\mathsf{n}}_{(1)}|\mathsf{m})=1$ and $\bar{\mathsf{n}}_{(0)}=0,\bar{\mathsf{n}}_{(1)}\in[d]$. Now consider the input $\mathsf{m}=(1,\underbrace{0,\ldots,0}_{\{\tilde{\mathsf{n}}_{(1)}-1\text{ times}\}},\underbrace{1,\ldots,1}_{\{d-\tilde{\mathsf{n}}_{(1)}+1\text{ times}\}})\in\mathsf{M}_4$ (see Equation (5.49)). For this input if $p(\tilde{\mathsf{n}}_{(0)}=1,\tilde{\mathsf{n}}_{(1)}|\mathsf{m})=1$ then the payoff is -1 from Equation (5.50) and Equation (5.51) and the overall payoff would be $s_\Lambda\leq 0$. Using Equation (5.52), $p(\bar{\mathsf{n}}_{(0)}=0,\bar{\mathsf{n}}_{(1)}|\mathsf{m})=1$. Now, for the same input $K_\mathsf{m}=\tilde{\mathsf{n}}_{(1)}-1$ and if $\bar{\mathsf{n}}_{(1)}>K_\mathsf{m}$, then again the payoff in this case would be $-\alpha$ (see Equation (5.50) and Equation (5.51)) and the overall payoff would be $s_\Lambda\leq 0$. Thus, $s_\Lambda>0$ implies $\bar{\mathsf{n}}_{(1)}\leq\tilde{\mathsf{n}}_{(1)}-1$.

Next consider the input $\mathsf{m}\in\mathsf{M}_3$ where $\mathsf{m}_{(0)}=0,k_\mathsf{m}=\tilde{\mathsf{n}}_{(1)}$ and $L_\mathsf{m}=\tilde{\mathsf{n}}_{(1)}-1$. For this input if $p(\tilde{\mathsf{n}}_{(0)}=1,\tilde{\mathsf{n}}_{(1)}|\mathsf{m})=1$ then the payoff is $-\alpha$ from Equation (5.50) and Equation (5.51) and the overall payoff would be $s_\Lambda\leq 0$. This is because $\tilde{\mathsf{n}}_{(1)}>L_\mathsf{m}$. Moreover, if $p(\bar{\mathsf{n}}_{(0)}=0,\bar{\mathsf{n}}_{(1)}|\mathsf{m})=1$ then the payoff is $-\alpha$ from Equation (5.50) and Equation (5.51) and the overall payoff would

[6] Alice can send either 0 or 1 as symbol for the input m using her encoding. Then Bob's decoding ensures that the possible outputs for each input of Alice are either $\tilde{\mathsf{n}}$ or $\bar{\mathsf{n}}$.

be $s_\Lambda \leq 0$. This is because $\bar{\mathsf{n}}_1 < k_{\mathsf{m}}$. However, from Equation (5.52), $p(\bar{\mathsf{n}}_{(0)} = 0, \bar{\mathsf{n}}_{(1)}|\mathsf{m}) = 1$ or $p(\tilde{\mathsf{n}}_{(0)} = 1, \tilde{\mathsf{n}}_{(1)}|\mathsf{m}) = 1$. Thus, each deterministic strategy yields $s_\Lambda \leq 0$ and no shared randomness-assisted strategy can achieve a positive payoff. ■

Theorem 5.7. *In the task* $\mathbf{CS}[d, \alpha]$*, assistance from a Hardy-type correlation* $P = \{P(a, b|x, y)\}_{a\in[d],b\in[d],x\in\{0,1\},y\in\{0,1\}}$ *as defined in Equation* (5.41) *yields a payoff* $P(a < b|1, 1) > 0$*.*

Proof. Assume Alice and Bob share a Hardy-type correlation $P = \{P(a, b|x, y)\}_{a\in[d],b\in[d],x\in\{0,1\},y\in\{0,1\}}$ as in Equation (5.41). For each input $\mathsf{m} = (\mathsf{m}_{(1)}, \cdots, \mathsf{m}_{(d)}) \in \mathsf{M}$ they proceed as follows.

Encoding: Alice queries $x = \mathsf{m}_{(0)}$ to the correlation P. Upon obtaining output $a \in [d]$ from the correlation, she sends $\mathsf{m}_{(a)}$ as message τ.

Decoding: Bob on receiving message $\tau' = \tau = \mathsf{m}_{(a)}$, uses $y = \tau$ as input to the correlation P. After getting the output $b \in [d]$ from the correlation, he gives the final output $\mathsf{n} = (\mathsf{n}_{(0)} = \tau', \mathsf{n}_{(1)} = b)$. The above strategy, while sharing correlation P leads to the following conditional probability

$$p(\mathsf{n}|\mathsf{m}) = p(\mathsf{n}_{(0)}, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{x,a,y,b,\tau,\tau'} [\delta_{x,\mathsf{m}_{(0)}} \delta_{\tau,\mathsf{m}_{(a)}} \delta_{\tau,\tau'} \delta_{y,\tau'} P(a, b|x, y) \delta_{\mathsf{n}_{(0)},\tau'} \delta_{\mathsf{n}_{(1)},b}]$$

$$= \sum_a [\delta_{\mathsf{n}_{(0)},\mathsf{m}_{(a)}} P(a, \mathsf{n}_{(1)}|\mathsf{m}_{(0)}, \mathsf{m}_{(a)})] \tag{5.53}$$

The payoff obtained for the input $\mathsf{m} = (1, 1, \cdots, 1)$ using this strategy is $\sum_{\mathsf{n}_{(1)}} -\alpha p(\mathsf{n}_{(0)} = 0, \mathsf{n}_{(1)}|\mathsf{m}) + \sum_{\mathsf{n}_{(1)}>1} p(\mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)}|\mathsf{m}) = \sum_{\mathsf{n}_{(1)}\geq 2} \sum_a [P(a, \mathsf{n}_{(1)}|1, 1)]$ (using Equation (5.53)). The payoff obtained for the input $\mathsf{m} = (0, 0, \cdots, 0)$ using this strategy is $\sum_{\mathsf{n}_{(1)}} -\alpha p(\mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)}|\mathsf{m}) = -\alpha \sum_{\mathsf{n}_{(1)}} \sum_a [\delta_{\mathsf{n}_{(0)}=1,\mathsf{m}_{(a)}} P(a, \mathsf{n}_{(1)}|\mathsf{m}_{(0)}, \mathsf{m}_{(a)})] = 0$. For $\mathsf{m} \in \mathsf{M}_1$, the payoff obtained is

$$-\alpha \Big[\sum_{\mathsf{m}\in\mathsf{M}_1,\mathsf{n}_{(1)}<k_{\mathsf{m}}=1} p(\mathsf{n}_{(0)} = 0, \mathsf{n}_{(1)}|\mathsf{m}) + \sum_{\mathsf{m}\in\mathsf{M}_1,\mathsf{n}_{(1)}>L_{\mathsf{m}}} p(\mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)}|\mathsf{m}) \Big] \tag{5.54}$$

$$= -\alpha \left[\sum_{\mathsf{m}\in\mathsf{M}_1,\mathsf{n}_{(1)}>L_{\mathsf{m}}} \sum_a \delta_{\mathsf{n}_{(0)}=1,\mathsf{m}_{(a)}} P(a, \mathsf{n}_{(1)}|0, \mathsf{m}_{(a)}) \right] \tag{5.55}$$

$$= -\alpha \sum_{m\in\mathsf{M}_1,\mathsf{n}_{(1)}>L_{\mathsf{m}}} P(L_{\mathsf{m}}, \mathsf{n}_{(1)}|0, 1) = 0 \quad \text{(using Equation (5.41))} \tag{5.56}$$

For $\mathsf{m} \in \mathsf{M}_2$, the payoff obtained is

$$-\alpha\Big[\sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}<k_\mathsf{m}} p(\mathsf{n}_{(0)}=0,\mathsf{n}_{(1)}|\mathsf{m}) + \sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}>L_\mathsf{m}} p(\mathsf{n}_{(0)}=1,\mathsf{n}_{(1)}|\mathsf{m})\Big] \tag{5.57}$$

$$= -\alpha\Big[\sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}<k_\mathsf{m}} \sum_a \delta_{\mathsf{n}_{(0)}=0,\mathsf{m}_{(a)}} P(a,\mathsf{n}_{(1)}|0,\mathsf{m}_{(a)}) + \sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}>L_\mathsf{m}} \sum_a \delta_{\mathsf{n}_{(0)}=1,\mathsf{m}_{(a)}} P(a,\mathsf{n}_{(1)}|0,\mathsf{m}_{(a)})\Big] \tag{5.58}$$

$$= -\alpha\Big[\sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}<k_\mathsf{m}} \sum_{\substack{k_\mathsf{m}\le a\le d\\ a\neq L_\mathsf{m}}} P(a,\mathsf{n}_{(1)}|0,0) + \sum_{\mathsf{m}\in\mathsf{M}_2,\mathsf{n}_{(1)}>L_\mathsf{m}} \sum_{\substack{1\le a\le k_\mathsf{m}-1\\ \vee a=L_\mathsf{m}}} P(a,\mathsf{n}_{(1)}|0,1)\Big] \tag{5.59}$$

$$= 0 \qquad (\text{ using Equation (5.41) and } L_\mathsf{m} \geq k_\mathsf{m}+1) \tag{5.60}$$

For $\mathsf{m} \in \mathsf{M}_3$, the payoff obtained is

$$-\alpha\Big[\sum_{\mathsf{m}\in\mathsf{M}_3,\mathsf{n}_{(1)}<k_\mathsf{m}} p(\mathsf{n}_{(0)}=0,\mathsf{n}_{(1)}|\mathsf{m}) + \sum_{\mathsf{m}\in\mathsf{M}_3,\mathsf{n}_{(1)}>L_\mathsf{m}} p(\mathsf{n}_{(0)}=1,\mathsf{n}_{(1)}|\mathsf{m})\Big] \tag{5.61}$$

$$= -\alpha\Big[\sum_{\mathsf{m}\in\mathsf{M}_3,\mathsf{n}_{(1)}<k_\mathsf{m}} \sum_a \delta_{\mathsf{n}_{(0)}=0,\mathsf{m}_{(a)}} P(a,\mathsf{n}_{(1)}|0,\mathsf{m}_{(a)}) + \sum_{\mathsf{m}\in\mathsf{M}_3,\mathsf{n}_{(1)}>L_\mathsf{m}} \sum_a \delta_{\mathsf{n}_{(0)}=1,\mathsf{m}_{(a)}} P(a,\mathsf{n}_{(1)}|0,\mathsf{m}_{(a)})\Big] \tag{5.62}$$

$$= -\alpha\Big[\sum_{m\in\mathsf{M}_3,\mathsf{n}_{(1)}<k_\mathsf{m}} \sum_{k_\mathsf{m}\le a\le d} P(a,\mathsf{n}_{(1)}|0,0) + \sum_{\mathsf{m}\in\mathsf{M}_3,\mathsf{n}_{(1)}>L_\mathsf{m}} \sum_{1\le a\le k_\mathsf{m}-1} P(a,\mathsf{n}_{(1)}|0,1)\Big] \tag{5.63}$$

$$= 0 \qquad (\text{ using Equation (5.41) and } L_\mathsf{m} = k_\mathsf{m}-1) \tag{5.64}$$

For $\mathsf{m} \in \mathsf{M}_4$, the payoff obtained is

(5.65)
$$-\alpha \sum_{\mathsf{m}\in\mathsf{M}_4,\mathsf{n}_{(1)}>K_{\mathsf{m}}} p(\mathsf{n}_{(0)} = 0, \mathsf{n}_{(1)}|\mathsf{m}) - \sum_{\mathsf{m}\in\mathsf{M}_4} p(\mathsf{n}_{(0)} = 1, \mathsf{n}_{(1)} = K_{\mathsf{m}} + 1|\mathsf{m})$$

(5.66)
$$= -\alpha \sum_{\mathsf{m}\in\mathsf{M}_4,\mathsf{n}_{(1)}>K_{\mathsf{m}}} \sum_{a} \delta_{\mathsf{n}_{(0)}=0,\mathsf{m}_{(a)}} P(a, \mathsf{n}_{(1)}|1, \mathsf{m}_{(a)}) - \sum_{\mathsf{m}\in\mathsf{M}_4} \sum_{a} \delta_{\mathsf{n}_{(0)}=1,\mathsf{m}_{(a)}} P(a, K_{\mathsf{m}} + 1|1, \mathsf{m}_{(a)})$$

(5.67)
$$= -\alpha \sum_{\mathsf{m}\in\mathsf{M}_4,\mathsf{n}_{(1)}>K_{\mathsf{m}}} \sum_{1\le a\le K_{\mathsf{m}}} P(a, \mathsf{n}_{(1)}|1, 0) - \sum_{\mathsf{m}\in\mathsf{M}_4} \sum_{K_{\mathsf{m}}+1\le a\le d} P(a, K_{\mathsf{m}} + 1|1, 1)$$

(5.68)
$$= -\sum_{\substack{r+1\le a\le d\\1\le r\le d-1}} P(a, r+1|1, 1) \quad (\text{ using Equation (5.41) and Equation (5.49)})$$

Now the total payoff using this strategy is

$$\mathcal{S}(\mathbf{p}) = \sum_{\mathsf{n}_{(1)}\ge 2} \sum_{a} [P(a, \mathsf{n}_{(1)}|1, 1)] - \sum_{\substack{r+1\le a\le d\\1\le r\le d-1}} P(a, r+1|1, 1) \tag{5.69}$$

$$= \sum_{r\ge 2} \sum_{a} P(a, r|1, 1) - \sum_{\substack{r\le a\le d\\2\le r\le d}} P(a, r|1, 1) = \sum_{\substack{1\le a<r\\2\le r\le d}} P(a, r|1, 1) = P(a < b|1, 1) \tag{5.70}$$

■

Corollary 5.6. *In the task* $\mathbf{CS}[d, \alpha]$*, the gap between the maximum payoff using quantum realisable Hardy-type correlation and the local bound* s_Λ *increases with* d*.*

Proof. From Theorem 5.7, assistance from a Hardy-type correlation $P = \{P(a, b|x, y)\}_{a\in[d],b\in[d],x\in\{0,1\},y\in\{0,1\}}$ yields a payoff $P(a < b|1, 1) > 0$ in the task $\mathbf{CS}[d, \alpha]$. From [186] (see Remark 5.3), the maximum Hardy success probability $H^d_{Hardy}(P) = P(a < b|1, 1)$ achievable by quantum realisable no-signalling correlations increases with d and approaches 0.5 in the limit. Since the classical bound $s_\Lambda \le 0$ is independent of d, while the maximum Hardy success probability $H^d_{Hardy}(P) = P(a < b|1, 1)$ grows with d, the separation between them increases with d. ■

5.4 Bounding quantum advantage in Communication Tasks

In this section, we examine the advantage of quantum non-locality in assisting classical communication for the tasks presented in Sections 5.2.1 and 5.2.2. To estimate the maximum payoff achievable with entanglement assistance, we use the maximal quantum violation of the associated Bell inequality obtained via the wire-cutting technique. We find the dimension-independent upper bound on the maximum quantum violation using semidefinite programming relaxation. We also search for the quantum strategies that achieve this bound. For this, we apply the see-saw algorithm [291] and obtain lower bounds on the maximal quantum violation. For some values of d and k in the task $\mathbf{CW}[d,k]$, we numerically compute the quantum correlation-assisted payoff by using the see-saw algorithm with two-qubit entangled states and observe values exceeding the classical (local) bound (see Table 5.2). In the case of the task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$ described in Section 5.2.2.2, the see-saw algorithm with two-qubit entanglement yields a payoff of 0.85355, which is higher than the local bound of 0.75 for the task (See Section 3.3 of [2]).

We use the NPA hierarchy [284, 285] to compute upper bounds on the maximal quantum violation of the associated Bell inequality for several cases listed in Table 5.3. For the task $\mathbf{CW}[d,k]$ with $d=2, k=2$, the upper bound from the level-2 NPA hierarchy is 0.85355, which matches the lower bound obtained using two-qubit states up to machine precision. This confirms the maximal quantum violation of the Bell inequality and, hence, the maximum entanglement-assisted payoff in the task. Interestingly, for $d=2, k=3$ and $d=3, k=2$, the upper bounds at level $1+AB$ of the NPA hierarchy differ from the qubit lower bounds obtained via the see-saw algorithm. We examine these cases in more detail.

For the task $\mathbf{CW}[d=2, k=3]$, we found that a two-qutrit strategy with non-projective local measurements achieves a higher payoff than any two-qubit quantum strategy. To bound the payoff achievable with two qubits, we applied the Navascués-Vértesi method [1], which yields an upper bound of 0.93491 for two-qubit entanglement with rank 1 projectors. This bound was saturated by the see-saw lower bound we obtained. Using the see-saw method with two-qutrit entanglement, we achieved a higher payoff that matches the upper bound 0.93883 from level-2 of the NPA hierarchy. Thus, this value is the maximal quantum correlation-assisted payoff for the task. For the case $d=3, k=2$, the level-2 NPA bound of 0.95209 is tighter than lower levels but does not match the see-saw lower bounds. Moreover, the see-saw lower bounds did not increase up to local dimension 8. For the task $\mathbf{CW}[\{P_{NL}^{(i)}\}_{i\in J}, d, k]$ discussed in Section 5.2.2.2, the upper bound 0.85355 from the level-2 NPA hierarchy coincides with the two-qubit see-saw lower bound up to machine precision. Thus, we recovered the maximal quantum correlation-assisted payoff for the task. Interestingly, the task $\mathbf{CW}[d=2, k=3]$ enables both device-independent and semi-device-independent certification of the local dimension of the shared entangled state.

Summary: We studied correlation-assisted classical communication in a one-way PM scenario with bounded communication, where Bob has no input. Using wire-reading, we constructed Bell inequalities tailored to these tasks for linear payoff functions. We showed that the advantage of entanglement assistance is equivalent to a violation of the associated inequality, and vice versa. We discussed how *wire-reading* can reveal the advantage of non-local correlations in scenarios where no such advantage is otherwise possible. We then introduced families of

(d, k)	Quantum payoff	Local Bound	Advantage
$(2, 2)$	$0.85355^{(i)}$	0.75	0.10355
$(3, 2)$	0.94975	0.875	0.07475
$(4, 2)$	0.97487	0.9375	0.03737
$(5, 2)$	0.99235	0.96875	0.02360
$(6, 2)$	0.99371	0.98437	0.00934
$(2, 3)$	$0.93491^{(ii)}$	0.88888	0.04603
$(2, 4)$	0.96338	0.9375	0.02588
$(2, 5)$	0.97656	0.96	0.01656
$(2, 6)$	0.98373	0.97222	0.01151
$(3, 3)$	0.98511	0.96296	0.02215

Table 5.2: Advantage from quantum correlation assistance in communication tasks. The table shows the payoff obtained in the task $\mathbf{CW}[d, k]$ using *two-qubit* entanglement from the see-saw algorithm for different values of d, k. In (i), the payoff matches the upper bound on the entanglement-assisted payoff from the level-2 NPA hierarchy (see Table 5.3). In (ii), the payoff matches the upper bound from the Navascués-Vértési method [1] for two-qubit entanglement with rank-1 projectors. (See Section 3.3 of [2])

(d, k)	Maximum Quantum Payoff (upper bound)
$(2, 2)$	$0.85355^{(i)}$ (NPA level 2)
$(3, 2)$	0.96231 (NPA $1 + AB$)
$(3, 2)$	0.95209 (NPA level 2)
$(2, 3)$	$0.93883^{(ii)}$ (NPA level 2)

Table 5.3: Upper bounds on the maximum quantum correlation-assisted payoff in the task $\mathbf{CW}[d, k]$ obtained from the NPA hierarchy for different d, k. In case (i), the upper-bound matches the lower bound achieved with two-qubit entanglement (see Table 5.2). In case (ii), the bound matches the see-saw lower bound for two-qutrit entanglement, while two-qubit strategies give a smaller payoff (see Table 5.2). Also discussed in Section 3.3 of [2].

bounded communication tasks in the Bob-without-input PM scenario. In these tasks, non-local correlations yield an advantage, while shared randomness gives strictly suboptimal payoffs. In the first family, correlations from any non-local facet of the no-signalling polytope optimally assist the bounded channel by achieving the maximal payoff. We next described tasks tailored to a non-local facet of the polytope. For correlations with dichotomic outputs, we showed that all correlations on the isotropic line between one such facet/extremal point and white noise remain resourceful as long as the noise fraction is below 0.5. Finally, we presented a third family of tasks where Hardy-type non-local correlations with dichotomic inputs. Here, the maximum payoff with shared randomness is zero, while Hardy correlations achieve a payoff equal to the Hardy violation. The payoff with quantum realisable Hardy-type correlation assistance increases with the task parameter. We also gave explicit examples of quantum advantage for some of these tasks.

Chapter 6

Conclusion

In this thesis, we investigated the advantages of non-classical resources within quantum theory and beyond in various information-processing tasks involving communication and non-local scenarios. Utilising such advantages, we design detection methods for some non-classical features in quantum theory. The key contributions are in three operational scenarios: (i) we demonstrate unbounded advantages of quantum communication and shared resources, (ii) we explore the interplay between non-local correlations and the advantages they provide in correlation-assisted classical communication, and (iii) we present quantum correlation-based advantages for randomness-free detection of non-projective-simulable measurements. These results contribute to our understanding of the cost of simulating of some quantum features and how some quantum behaviours are impossible to simulate classically under some operational constraints. Such non-classical resources potentially improve the performance of information-processing tasks.

In Chapter 3, we studied relations defined by orthogonality graphs and analysed the classical and quantum communication required for one-way zero-error *distributed computation* and *relation reconstruction* without shared resources. We showed that the distributed computation of the relation cannot demonstrate a quantum advantage for certain graphs. In contrast, relation reconstruction reveals a clear separation. The classical communication cost without shared and private randomness grows logarithmically with the order of the graph n. This classical cost without shared randomness (but with local randomness) is at least $\log_2 K(\mathcal{G})$ bit, where $K(\mathcal{G})$ is the disjointness number of $\mathcal{G}$ (see Definition 3.4) and is lower bounded by $\max\{\omega(\mathcal{G}),\ \log_2 n + \frac{1}{2}\log_2\log_2 n + O(1)\}$. On the other hand, the quantum communication cost is $\log_2 \omega(\mathcal{G})$ qubit whenever $\mathcal{G}$ admits a faithful orthogonal representation in dimension $\omega(\mathcal{G})$. We identified families of graphs where the order n grows without bound while the faithful orthogonal range remains constant. For these graphs, the task can be completed with a fixed quantum cost of $\log_2 d$ qubit, independent of n. In these cases, the classical cost grows without bound and is at least $\max\{\log_2 d, \log_2(\log_2 n + \frac{1}{2}\log_2\log_2 n + c)\}$ bit, for some constant c. This establishes an unbounded quantum advantage for relation reconstruction. The input size in this task scales logarithmically with n. Although this unbounded advantage disappears in the presence of shared resources, we quantified the necessary shared randomness for the task while using bounded classical communication. For some graphs, the lower bound on the shared randomness required scales as $\log_2(\lceil\log_2 \alpha\rceil + 1)$ bit with the number of maximal cliques α. We also showed that the gap between entanglement assistance and shared randomness assistance to a limited classical communication can be arbitrarily large and scales as $\log_2(\lceil\log_2 \alpha\rceil + 1)$ bit with the number of maximal cliques α.

Some previously studied communication problems based on orthogonality graphs [292] showed a quantum advantage in one-way communication whenever the graphs belonged to the class of state-independent contextuality graphs. In contrast, several graphs we considered are binary colourable. Thus, the quantum advantage in the communication task we discussed can be observed independent of the usefulness of the graphs in demonstrating state-independent contextuality. This leaves open a foundational question about the source of the quantum advantage in these tasks. It also remains open whether there exist tasks with binary colourable graphs that give rise to a large separation, for instance, an exponential, between classical and quantum communication in the presence/absence of shared resources. Another open direction could be exploring whether unbounded separations can also appear beyond the zero-error setting considered here. We also derived the necessary amount of shared randomness for some specific graphs. Establishing such lower bounds for more general classes of graphs remains open. We can interpret the relation reconstruction task as a qualitative simulation of quantum statistics, where spatially separated parties are given a set of favourable events, and the requirement is that these events are simulated with non-zero probability using classical communication, as they are in the quantum case. Seen this way, the protocol resembles the distribution of conditional randomness with classical communication. This motivates the question of whether the task or its variations share a relation to discrete analogues of boson sampling [293] or similar computational problems.

In Chapter 4, we introduced a *randomness-free* detection scheme for *non-projective simulable measurements* in a scenario with separate parties sharing systems with bounded local operational dimension. We proved that all correlations achievable in this setting using a local projective measurement on a d-dimensional quantum system, together with classical post-processing, can also be reproduced by shared classical systems of the same local dimension, and vice versa. We then presented detection schemes in a bipartite setting for three- and four-outcome qubit non-projective simulable measurements, some of which remain robust under arbitrary depolarising noise. We discussed the detection schemes for five-outcome qutrit non-projective simulable measurements. In the bipartite case with identical devices, we discussed such a scheme and provided numerical evidence of its robustness against arbitrary depolarising noise. Relaxing the identical-device assumption, we constructed detection schemes in both bipartite and tripartite scenarios. For these tasks, we demonstrated violations of the bounds achievable by qutrit projective-simulable measurements, using a shared two-qutrit state and a qutrit POVM. Finally, we extended the notion of non-projective simulable measurements for GPTs. We also presented a randomness-free test to show that the square-bit model or box world theory is unphysical.

Several questions remain open for future work. First, it is unknown whether the detection schemes discussed in Section 4.2.1 for three and four-outcome qubit non-projective-simulable measurements are robust against noise. Second, while some of the target correlations in the noisy detection schemes in Section 4.2.2.1 and Section 4.2.2.2 can be simulated with qubit POVMs and two-qubit shared states, it remains open whether all correlations in these sets with mutual information less than 1 can be reproduced in this way (see region $R3$ in Figure 4.2 and Figure 4.3). Third, it remains open if it is possible to analytically derive a non-trivial upper bound on the projective-simulable payoff discussed in Section 4.3.2 and Section 4.3.3. Fourth, in Section 4.3.1 we showed the impossibility of generating certain large but finite target correlations with qutrit projective-simulable measurements by discretising parameters. It remains unexplored whether all correlations in this set are beyond qutrit projective-simulable strategies. A conclusive result would establish the robustness of the detection scheme against arbitrary white noise. Finally, an open question is whether randomness-free tasks, such as discussed here, can also rule out a class of GPTs such as *polygon models* [294].

In Chapter 5, we presented families of tasks in a minimal PM scenario where non-local correlations enhance bounded classical communication. Using wire-cutting, we derived a Bell inequality tailored to any correlation-assisted classical communication task with a linear payoff. We showed that violation of this inequality implies an advantage in the corresponding task, and conversely, that any non-local correlation providing an advantage in the task must violate it. We also introduced wire-reading and showed that it reveals the advantage of non-local correlation assistance in scenarios where no such advantage exists otherwise. With wire-reading, we presented two families of tasks in the minimal PM scenario where shared randomness gives a strictly suboptimal payoff. In the first family, we showed that correlations from any non-local facet of the no-signalling polytope lead to the optimal payoff. For some of these tasks, we explicitly demonstrated quantum advantage. In the second family, each task is tailored to a specific non-local facet, and we showed that correlations from that facet achieve the optimal payoff. As a concrete example, we consider tasks tailored to non-local extremal correlations with dichotomic outputs, where the shared randomness-assisted payoff is bounded from above by 0.75. Any correlation on the isotropic line connecting the extremal point to white noise provides an advantage as long as the noise fraction is below 0.5. We then introduced a third family of tasks demonstrating advantageous assistance from Hardy-type correlations with dichotomic inputs to a one-bit channel. Here, the payoff with shared randomness is bounded by zero from above, while Hardy correlations yield a positive payoff. The maximum quantum payoff using Hardy-type correlations grows with the task parameter. Finally, we also identified a communication task from the first family where the maximum quantum payoff is achieved by a two-qutrit entangled state. In contrast, two-qubit entangled states achieve payoffs that are higher than the local bound but strictly less than the quantum maximum. This gap offers an operational method to witness the local dimension of the entangled system.

Non-local correlations provide an advantage in tasks involving classical communication. It is not known, however, whether every non-local correlation is advantageous in some classical communication task. We approached the broader question of whether non-locality is both necessary and sufficient for a quantum communication advantage. The wire-cutting method answers the necessity part, but the sufficiency question, whether every non-local correlation can yield an advantage, is still open. The wire-reading technique, as we showed, allows the construction of communication tasks closely tailored to a given non-local resource. This may serve as a useful tool for probing the sufficiency of non-locality. So far, we have only presented a few families of tasks based on a non-local facet of the no-signalling polytope. Many similar possibilities remain unexplored. Exploring extensions of these constructions and studying their behaviour when the communication channel is potentially noisy could be promising directions for future work.

We demonstrated quantum advantage in several information-processing tasks, including cases with unbounded separation between classical and quantum resources. These results establish a fundamental operational distinction between classical and quantum theories. The work reinforces connections between graph theory and quantum communication advantage, and also proposes ways to explore the role of non-local correlation in assisting communication. We have also provided a new approach to certify non-classical measurement resources. The open questions outlined here invite further research in quantum foundations and quantum information. Some of the theoretical results presented may also guide future applications in the growing field of quantum technologies.

Appendix A

Relaxed Relation Reconstruction

Consider graphs $\mathcal{G}$ satisfying conditions **(G0)** and **(G1)** with faithful orthogonal range $\omega(\mathcal{G})$. From Theorem 3.3 we have quantum advantage in relation reconstruction of $\mathcal{R}(\mathcal{G})$ without shared randomness exists whenever $K(\mathcal{G}) > \omega(\mathcal{G})$. If Alice has no private randomness, quantum advantage already exists whenever $|\mathbf{V}(\mathcal{G})| > \omega(\mathcal{G})$. A similar quantum advantage can be shown for a larger class of graphs by considering a task inspired by relation reconstruction.

Consider a one-way PM scenario as in Chapter 3. When parties do not have shared resources, quantum advantage can also be shown for another task whenever $\mathcal{G}$ satisfies conditions **(G0)** and **(G1)** with faithful orthogonal representation in dimension $d < K(\mathcal{G})$. If Alice additionally has no private randomness in this task, then quantum advantage already exists whenever $|\mathbf{V}(\mathcal{G})| > d$. We refer to this task as *relaxed relation reconstruction*. It will be defined for relations specified by a variant of the distributed clique labelling problem. We will denote the modification of the distributed clique labelling problem shortly as m-CLP for a graph $\mathcal{G}$ and the relation it specifies by $\tilde{\mathcal{R}}(\mathcal{G})$.

The relation $\tilde{\mathcal{R}}(\mathcal{G})$ is defined by the distributed *modified clique labelling problem* (m-CLP). As in the case of CLP, this problem is also specified using a simple graph $\mathcal{G}$ (see Definition 2.15) with vertex set $\mathbf{V}(\mathcal{G})$ and edge set $\mathbf{E}(\mathcal{G})$. Each vertex of $\mathcal{G}$ belongs to at least one maximum clique in the graph. If the graph has α maximum cliques and clique number $\omega(\mathcal{G})$, we denote the set of maximum cliques as $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$. The set of clique labels for each maximum clique in $\mathcal{G}$ is $\mathcal{O} = \{0, 1, \cdots, \omega(\mathcal{G}) - 1\}$. Each label corresponds to a binary colouring of the clique C_i according to Definition 2.27. We extend the set of clique labels to $\tilde{\mathcal{O}} = \mathcal{O} \cup \{\varkappa\}$, by adding a trivial label $\varkappa$ that corresponds to a colouring (not binary colouring) which assigns colour 0 to all the vertices in the corresponding maximum clique.

The distributed m-CLP is defined in a PM scenario where Alice and Bob know the orthogonality graph $\mathcal{G}$. The input and output sets for the problem are: $\mathsf{M}_A = \mathcal{C} \times \mathcal{O}$, $\mathsf{M}_B = \mathcal{C}$ and $\mathsf{N} = \tilde{\mathcal{O}}$. The inputs are chosen randomly following a uniform distribution. Alice receives $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) \in \mathsf{M}_A$, *i.e.*, a maximum clique $\mathsf{m_{a(1)}} \in \mathcal{C}$ and a clique label $\mathsf{m_{a(2)}} \in \mathcal{O}$ associated with the maximum clique, both chosen randomly. Bob receives a maximum clique $\mathsf{m_b} \in \mathsf{M}_B = \mathcal{C}$. Bob's output $\mathsf{n} \in \mathsf{N} = \tilde{\mathcal{O}}$ is a clique label for his input maximum clique. The output must satisfy the constraints described below. Each clique label except $\varkappa$ is mapped to a particular binary colouring of the vertices in the maximum clique through definition 2.27. Clique label $\varkappa$ corresponds to a colouring which assigns colour 0 to all the vertices in the corresponding maximum clique.

(i) If Alice's and Bob's input maximum cliques share some vertices, then Bob's colouring (either 0 or 1 by output clique label) of each shared vertex must be identical to Alice's colouring (from her input clique label) of these vertices.

(ii) If their input maximum cliques are distinct and a vertex from Alice's input maximum clique is adjacent to some vertex from Bob's input maximum clique, then both the vertices sharing the edge must not be simultaneously assigned colour 1 (by the input-output clique labels of Alice and Bob).

(iii) In the remaining cases, Bob can output any valid clique label for his input maximum clique, independent of Alice's input clique label.

If Alice and Bob receive the same maximum clique, *i.e.* $\mathsf{m_{a(1)}} = \mathsf{m_b}$, then by condition (i) the output clique label of Bob n must be the same as the input clique label of Alice $\mathsf{m_{a(2)}}$. If their input cliques are different and some vertex of Alice's input maximum clique is adjacent to a vertex of Bob's input maximum clique in the graph $\mathcal{G}$, then their labels must ensure that such adjacent vertices are not both assigned the colour 1 at the same time. Note that these rules are the same as CLP except that Bob can sometimes output $\varkappa$ as his output clique label whenever condition (i) is not violated.

The relation $\tilde{\mathcal{R}}(\mathcal{G}) \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N} = (\mathcal{C} \times \mathcal{O}) \times \mathcal{C} \times \tilde{\mathcal{O}}$ can be directly defined from the distributed m-CLP for the graph $\mathcal{G}$. The tuple $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) = ((\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}), \mathsf{m_b}, \mathsf{n})$ belongs to $\tilde{\mathcal{R}}(\mathcal{G}) \subseteq \mathsf{M}_A \times \mathsf{M}_B \times \mathsf{N}$ if Alice's and Bob's input maximum clique, Alice's input clique label and output clique label of Bob satisfy the constraints described above. We will denote the tuples as $(\mathsf{m_a}, \mathsf{m_b}, \mathsf{n}) = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}, \mathsf{m_b}, \mathsf{n}) \in \tilde{\mathcal{R}}(\mathcal{G})$ here for convenience.

Remark A.1. *For a graph $\mathcal{G}$, consider the relation $\mathcal{R}(\mathcal{G})$ specified by CLP and relation $\tilde{\mathcal{R}}(\mathcal{G})$ specified by m-CLP. If a tuple $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \mathcal{R}(\mathcal{G})$ for $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$ and $k, l \in \mathcal{O}$ then $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \tilde{\mathcal{R}}(\mathcal{G})$. Also, if $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \mathcal{R}(\mathcal{G})$ for $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$ and $k, l \in \mathcal{O}$ then $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \tilde{\mathcal{R}}(\mathcal{G})$. This can be directly seen by comparing the constraints of CLP and m-CLP.*

In the relaxed relation reconstruction of $\tilde{\mathcal{R}}(\mathcal{G})$, given Alice's input maximum clique $\mathsf{m}_{\mathsf{a}(1)}$ with clique label $\mathsf{m}_{\mathsf{a}(2)}$, Bob must output clique label n for his input maximum clique m_{b} such that $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \in \tilde{\mathcal{R}}(\mathcal{G})$. Additionally, for Alice's input $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$, Bob's output clique label n for his maximum clique m_{b} in different rounds should additionally span all valid labels $\{\mathsf{n} \in \mathsf{N} \setminus \{\varkappa\} = \mathcal{O} : (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \in \tilde{\mathcal{R}}(\mathcal{G})\}$. Note that Bob may not output $\mathsf{n} = \varkappa$ and still accomplish the task. Thus, we refer to this task as relaxed relation reconstruction.

For relaxed relation reconstruction, Alice and Bob follow some protocol. Following a protocol, suppose Bob outputs a clique label n for his input maximum clique m_{b} with conditional probability $p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$ where $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$ is the input of Alice. The set $\{p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})\}_{\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}}$ can be represented as the entries of a table. The rows are indexed by Alice's inputs $\mathsf{m}_{\mathsf{a}} = (\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)})$, and the columns by Bob's input-output pairs $(\mathsf{m}_{\mathsf{b}}, \mathsf{n})$. For rows, the index first runs over the clique labels $\mathcal{O} = \{0, 1, \ldots, \omega(\mathcal{G}) - 1\}$ in increasing order and then moves to the next maximum clique in $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$. Similarly, for columns, the index first runs over the clique labels $\tilde{\mathcal{O}} = \{0, 1, \ldots, \omega(\mathcal{G}) - 1, \varkappa\}$ in the order listed and then moves to the next maximum clique in $\mathcal{C} = \{\mathrm{C}_1, \mathrm{C}_2, \cdots, \mathrm{C}_\alpha\}$. We can denote such a table by $\mathbf{W}$. In terms of elements $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}), (\mathsf{m}_{\mathsf{b}}, \mathsf{n})} := p(\mathsf{n}|\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}})$ of this table, a protocol accomplishes relation reconstruction if the table satisfies both **(T0)** and **(T1)**.

(T0') : $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}), (\mathsf{m}_{\mathsf{b}}, \mathsf{n})} = 0$ if $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \notin \tilde{\mathcal{R}}(\mathcal{G})$

(T1') : $\mathbf{W}_{(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}), (\mathsf{m}_{\mathsf{b}}, \mathsf{n})} > 0$ if $(\mathsf{m}_{\mathsf{a}(1)}, \mathsf{m}_{\mathsf{a}(2)}, \mathsf{m}_{\mathsf{b}}, \mathsf{n}) \in \tilde{\mathcal{R}}(\mathcal{G})$ and $n \neq \varkappa$

Now, consider the same class of orthogonality graph $\mathcal{G}$ considered in Section 3.2.2. We can calculate the classical communication cost of relaxed relation reconstruction.

Theorem A.1. *For $\tilde{\mathcal{R}}(\mathcal{G})$, where $\mathcal{G}$ satisfies conditions **(G0)** and **(G1)**, the classical communication necessary and sufficient for zero-error relaxed relation reconstruction without shared resources and without private randomness for Alice is $\log_2 |\mathbf{V}(\mathcal{G})|$ bits. More generally, the classical communication necessary for zero-error relaxed relation reconstruction without shared resources is $\log_2 K(\mathcal{G})$ bit.*

Proof. The proof of the theorem follows an exactly the same argument as presented in the proof of Theorem 3.2. Intuitively, from Remark A.1, if a tuple $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \mathcal{R}(\mathcal{G})$ for $\mathrm{C}_i, \mathrm{C}_j \in \mathcal{C}$ and $k, l \in \mathcal{O}$ then $(\mathrm{C}_i, k, \mathrm{C}_j, l) \in \tilde{\mathcal{R}}(\mathcal{G})$ for a graph $\mathcal{G}$. Also if $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \mathcal{R}(\mathcal{G})$ then $(\mathrm{C}_i, k, \mathrm{C}_j, l) \notin \tilde{\mathcal{R}}(\mathcal{G})$. Thus, relation reconstruction for $\mathcal{R}(\mathcal{G})$ is necessary for achieving relaxed relation reconstruction of $\mathcal{R}(\mathcal{G})$. Following a classical communication protocol that ensures relation reconstruction for $\mathcal{R}(\mathcal{G})$ trivially ensures relaxed relation reconstruction of $\tilde{\mathcal{R}}(\mathcal{G})$ ■

Theorem A.2. *Let $\mathcal{G}$ be a graph satisfying **(G0)** and **(G1)** with faithful orthogonal range d, then $\log_2 d$ qubit communication is sufficient for zero-error relaxed relation reconstruction of $\tilde{\mathcal{R}}(\mathcal{G})$ without shared resources (and private randomness).*

Proof. Alice and Bob know the graph $\mathcal{G}$ and a faithful orthogonal representation (FOR) $\phi : \mathbf{V}(\mathcal{G}) \to \mathbb{C}^d$ of $\mathcal{G}$. The FOR ϕ assigns unit vector $\phi(v) \in \mathbb{C}^d$ to each vertex $v \in \mathbf{V}(\mathcal{G})$ such that $\langle \phi(v)|\phi(v')\rangle = 0$ if and only if v and v' are adjacent in $\mathcal{G}$. Say, Alice is given input $\mathsf{m_a} = (\mathsf{m_{a(1)}}, \mathsf{m_{a(2)}}) = (\mathrm{C}_i, k)$, where $k \in \mathcal{O}$ and $\mathrm{C}_i \in \mathcal{C}$. The clique label k corresponds to the binary colouring that assigns 1 to vertex v of the maximum clique $\mathrm{C}_i \in \mathcal{C}$. Alice encodes it using the quantum state $\rho_{(\mathrm{C}_i,k)} = |\phi(v)\rangle \langle\phi(v)| \in D(\mathbb{C}^d)$ and sends it to Bob. If Bob receives input $\mathsf{m_b} = \mathrm{C}_j$, then he performs the measurement $\{E = |\phi(v')\rangle \langle\phi(v')| : v' \in \mathbf{V}(\mathrm{C}_j)\} \cup \{I_d - \sum_{v'\in\mathbf{V}(\mathrm{C}_j)} |\phi(v')\rangle \langle\phi(v')|\}$. On obtaining outcome corresponding to the projector $|\phi(v')\rangle \langle\phi(v')|$, Bob outputs $\mathsf{n} = l$, where the clique label $l \in \mathcal{O}$ for C_j assigns colour 1 to the vertex v'. And obtaining outcome corresponding to the the projector $I_d - \sum_{v'\in\mathbf{V}(\mathrm{C}_j)} |\phi(v')\rangle \langle\phi(v')|$, Bob outputs $\mathsf{n} = \varkappa$. This quantum strategy always satisfies the consistency condition. This is so because $\mathrm{Tr}[(|\phi(v)\rangle \langle\phi(v)|)(|\phi(v')\rangle \langle\phi(v')|)]$ is 1 if $v = v'$ and 0 if v is adjacent to v'. In all other cases, it is positive. Thus, the protocol accomplishes relation reconstruction. ■

Theorem A.3. *Let $\mathcal{G}$ be a graph satisfying conditions **(G0)** and **(G1)** with faithful orthogonal range d. Then quantum advantage in relation reconstruction of $\tilde{\mathcal{R}}(\mathcal{G})$ without shared randomness exists whenever $K(\mathcal{G}) > d$. If Alice also does not have private randomness, quantum advantage already exists whenever $|\mathbf{V}(\mathcal{G})| > d$.*

Proof. The proof of the claim follows from Theorem A.1 and Theorem A.2. For $\tilde{\mathcal{R}}(\mathcal{G})$, the classical communication necessary for zero-error relaxed relation reconstruction without shared resources is at least $\log_2 K(\mathcal{G})$ bit, and without private randomness for Alice, it is $\log_2 |\mathbf{V}(\mathcal{G})|$ bit. In contrast, the $\log_2 d$ qubit communication in both scenarios is sufficient for relaxed relation reconstruction. ■

As an example, consider Paley graphs $\mathcal{G}_{Paley(q)}$ with q vertices, where $q = 4k + 1 = p^m$ for some prime p and positive $m \in \mathbb{N}$. These graphs have order q and have FOR in dimension $\frac{q+1}{2}$ (see Theorem 4 in [184]), so they provided quantum communication advantage in the sense of the Theorem A.3. The graphs $\mathcal{G}$ satisfying conditions **(G0)** and **(G1)** with faithful orthogonal range d, where $d < |\mathbf{V}(\mathcal{G})|$ were claimed to show advantage in relation reconstruction task in [184]. However, it turns out that the original relation reconstruction task (also see Chapter 3) requires a minor correction - the relaxation presented here - to reveal quantum advantage for such graph as demonstrated above.

Appendix B

Bounding Quantum Non-local Correlations

Recall that a no-signalling correlation $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},y\in\mathsf{Y}} \in \mathcal{NS}$ is quantum realisable if there is a shared quantum state $\rho_{AB} \in D(\mathbb{C}^{d_A}\otimes\mathbb{C}^{d_B})$ and POVMs $\mathcal{M}_A^x = \{E_a^x \in \mathcal{B}(\mathbb{C}^{d_A}) : E_a^x \geq 0, \sum_{a\in\mathsf{A}} E_a^x = I_{d_A}\}\ \forall x \in \mathsf{X}$ for Alice and POVMs $\mathcal{M}_B^y = \{E_b^y \in \mathcal{B}(\mathbb{C}^{d_B}) : E_b^y \geq 0, \sum_{b\in\mathsf{B}} E_b^y = I_{d_B}\}\ \forall y \in \mathsf{Y}$ for Bob such that

$$P(a,b|x,y) = \mathrm{Tr}\big[\rho_{AB}(E_a^x \otimes E_b^y)\big] \ \ \forall \ \ a,b,x,y \tag{B.1}$$

We denote by $\mathcal{Q} \subset \mathcal{NS}$ the set of all no-signalling correlations $P \in \mathcal{NS}$ which are quantum realisable. Deciding membership in $\mathcal{Q}$ is, however, an undecidable problem in general [219, 220]. A systematic approach to approximate $\mathcal{Q}$ is provided by the Navascués-Pironio-Acín (NPA) hierarchy. This is a sequence of semidefinite programming relaxations of the quantum set, denoted by $\{\mathcal{Q}_L\}$ where $L \in \mathbb{N}$ is the order of the relaxation, which offer a way of addressing the membership problem. Each level $\mathcal{Q}_L$ gives an outer approximation of $\mathcal{Q}$, with higher levels yielding tighter approximations.

The key tool behind this technique is the construction of a positive semidefinite moment matrix Γ. Its entries are defined using a positive semidefinite variable $\rho \geq 0$ and an operator list $\mathcal{F} = \{F_1, F_2, \cdots, F_n\}$. The entries of the moment matrix are given by $\Gamma_{F_i,F_j} = \mathrm{Tr}\Big[\rho(F_i^\dagger F_j)\Big]$. Now, the following is satisfied by a moment matrix Γ:

Lemma B.1. *Let $\mathcal{F} = \{F_1, F_2, \cdots, F_n\}$ be a set of operators, then the orthogonal matrix Γ with entries $\Gamma_{F_i,F_j} = \mathrm{Tr}\Big[\rho(F_i^\dagger F_j)\Big]$ is positive semi-definite. [[218]].*

Here $\mathcal{F}$ may be any set of linear operators. We choose them to be Alice's projective measurement operators $\mathcal{M}_A^x = \{E_a^x \in \mathcal{B}(\mathbb{C}^{d_A}) : E_a^x \geq 0, \sum_{a\in\mathsf{A}} E_a^x = I_{d_A}\}\ \forall x \in \mathsf{X}$ and Bob's projective measurement operators $\mathcal{M}_B^y = \{E_b^y \in \mathcal{B}(\mathbb{C}^{d_B}) : E_b^y \geq 0, \sum_{b\in\mathsf{B}} E_b^y = I_{d_B}\}\ \forall y \in \mathsf{Y}$ in a bipartite

Bell scenario $\mathscr{B}_{\mathsf{AB}}^{\mathsf{XY}}$. Then the operator list $\mathcal{F}_1 = \{I_{d_A d_B}\} \bigcup (\cup_{x\in\mathsf{X},a\in\mathsf{A}}\{E_a^x\}) \bigcup (\cup_{y\in\mathsf{Y},b\in\mathsf{B}}\{E_b^y\})$. These operators satisfy the projectivity conditions (normalisation and orthogonality) as well as commutation relations. In this case, certain entries of the moment matrix correspond directly to probabilities of the correlation P. For example, $[\Gamma]_{I_{d_A d_B},E_a^x} = P_A(a|x)$, $\Gamma_{I_{d_A d_B},E_b^y} = P_B(b|y)$, $\Gamma_{E_a^x,E_b^y} = P(a,b|x,y)$. Let $\mathcal{Q}_1$ be the set of correlations $P = \{P(a,b|x,y)\}_{a\in\mathsf{A},b\in\mathsf{B},x\in\mathsf{X},y\in\mathsf{Y}} \in \mathcal{NS}$ for which such $\Gamma \geq 0$ exist while satisfying these constraints. The associated semi-definite programme then asks whether such a matrix Γ can be found. If such Γ does not exist or the programme is infeasible, then $P \notin \mathcal{Q}_1$ and hence $P \notin \mathcal{Q}$. If the programme is feasible and $P \in \mathcal{Q}_1$, we cannot conclude that $P \in \mathcal{Q}$, since $\mathcal{Q}_1$ is only an outer approximation of the quantum set.

For $L = 2$, the operator list includes all strings of operators in $\mathcal{F}_1$ up to length two, *i.e.* $\mathcal{F}_2 = \mathcal{F}_1 \bigcup (\cup_{x,x'\in\mathsf{X},a,a'\in\mathsf{A}}\{E_a^x E_{a'}^{x'}\}) \bigcup (\cup_{y,y'\in\mathsf{Y},b,b'\in\mathsf{B}}\{E_b^y E_{b'}^{y'}\}) \bigcup (\cup_{x\in\mathsf{X},y\in\mathsf{Y},a\in\mathsf{A},b\in\mathsf{B}}\{E_a^x E_b^y\})$. Let $\mathcal{Q}_2$ be the set of no-signalling correlations $P = P(a,b|x,y)$ for which there exists a moment matrix $\Gamma \geq 0$ built from $\mathcal{F}_2$ that satisfies the same constraints as before. More generally, for each $L \in \mathbb{N}$, the operator list $\mathcal{F}_L$ contains all strings of operators in $\mathcal{F}_1$ of length at most L, and $\mathcal{Q}_L$ denotes the corresponding set of feasible correlations. These sets satisfy $\mathcal{Q}_1 \supseteq \mathcal{Q}_2 \cdots \supseteq Q_L \supseteq Q_{L+1}$. It was shown in [284, 285] that $\lim_{L\to\infty} \mathcal{Q}_L \equiv \mathcal{Q}$. Hence, this scheme provides a hierarchy of strengthening conditions for deciding the membership problem. Since the membership problem can be converted to the problem of maximising the value of a Bell violation, the hierarchy returns a converging sequence of dimension-independent upper bounds on the maximum quantum Bell violation. Denoting these bounds by $\beta_{\mathcal{Q}_1} \geq \beta_{\mathcal{Q}_2} \geq \cdots \geq \beta_{\mathcal{Q}_L}$, we have $\lim_{L\to\infty} \beta_{\mathcal{Q}_L} \equiv \beta_{\mathcal{Q}}$. Here, $\beta_{\mathcal{Q}}$ is the maximum violation of the Bell inequality and $\beta_{\mathcal{Q}_L}$ is the upper bound obtained at level L.

Whether a given correlation is realisable with finite-dimensional quantum systems remains a complicated problem. However, this can also be addressed similarly. In particular, [1, 295] observed that for a fixed Bell scenario, moment matrices arising from states in finite-dimensional Hilbert spaces form a strict subspace of all feasible moment matrices. This subspace is equipped with the Hilbert-Schmidt inner product and admits an orthogonal basis. Their method constructs such a basis by sampling finite-dimensional moment matrices at random and applying Gram-Schmidt orthogonalisation. A detailed introduction to the topic can be found in reviews such as [291].

Bibliography

[1] M. Navascués and T. Vértesi, Phys. Rev. Lett. **115**, 020501 (2015).

[2] S. Rout, A. Chaturvedi, S. S. Bhattacharya, and P. Horodecki, Facets of non-locality and advantage in entanglement-assisted classical communication tasks (2025), arXiv:2507.10830 [quant-ph] .

[3] R. P. Feynman, International Journal of Theoretical Physics **21**, 467–488 (1982).

[4] P. Benioff, Phys. Rev. Lett. **48**, 1581 (1982).

[5] D. Deutsch, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97 (1985).

[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[7] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[8] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[9] D. Deutsch and R. Jozsa, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences **439**, 553 (1992).

[10] G. Brassard and P. Hoyer, in *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems* (1997) pp. 12–23.

[11] P. W. Shor, SIAM Journal on Computing **26**, 1484–1509 (1997).

[12] D. R. Simon, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, SFCS '94 (IEEE Computer Society, USA, 1994) p. 116–123.

[13] L. K. Grover, Phys. Rev. Lett. **79**, 325 (1997).

[14] M. Boyer, G. Brassard, P. Høyer, and A. Tapp, Fortschritte der Physik **46**, 493–505 (1998).

[15] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation (2002).

[16] H. Buhrman, R. Cleve, and A. Wigderson, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98 (Association for Computing Machinery, New York, NY, USA, 1998) p. 63–68.

[17] R. Raz, in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99 (Association for Computing Machinery, New York, NY, USA, 1999) p. 358–367.

[18] D. Gavinsky, in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08 (Association for Computing Machinery, New York, NY, USA, 2008) p. 95–102.

[19] A. Montanaro, Quantum Information & Computation **11**, 574 (2011).

[20] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).

[21] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[22] M. Steiner, Physics Letters A **270**, 239–244 (2000).

[23] T. Maudlin, PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association **1992**, 404–417 (1992).

[24] S. Massar, D. Bacon, N. J. Cerf, and R. Cleve, Phys. Rev. A **63**, 052305 (2001).

[25] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).

[26] M. J. Renner, A. Tavakoli, and M. T. Quintino, Phys. Rev. Lett. **130**, 120801 (2023).

[27] N. J. Cerf, N. Gisin, and S. Massar, Phys. Rev. Lett. **84**, 2521 (2000).

[28] C. Perry, R. Jain, and J. Oppenheim, Phys. Rev. Lett. **115**, 030504 (2015).

[29] Z.-W. Liu, C. Perry, Y. Zhu, D. E. Koh, and S. Aaronson, Phys. Rev. A **93**, 012347 (2016).

[30] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, IEEE Transactions on Information Theory **48**, 2637 (2002).

[31] C. Vieira, C. de Gois, L. Pollyceno, and R. Rabelo, New Journal of Physics **25**, 113004 (2023).

[32] H. Buhrman, R. Cleve, and W. van Dam, SIAM Journal on Computing **30**, 1829 (2001).

[33] R. Cleve and H. Buhrman, Phys. Rev. A **56**, 1201 (1997).

[34] H. Buhrman, W. van Dam, P. Høyer, and A. Tapp, Phys. Rev. A **60**, 2737 (1999).

[35] M. Pawłowski and M. Żukowski, Phys. Rev. A **81**, 042326 (2010).

[36] A. Tavakoli, J. Pauwels, E. Woodhead, and S. Pironio, PRX Quantum **2**, 040357 (2021).

[37] J. Pauwels, A. Tavakoli, E. Woodhead, and S. Pironio, New Journal of Physics **24**, 063015 (2022).

[38] J. Pauwels, S. Pironio, E. Z. Cruzeiro, and A. Tavakoli, Phys. Rev. Lett. **129**, 120504 (2022).

[39] Č. Brukner, M. Żukowski, and A. Zeilinger, Phys. Rev. Lett. **89**, 197901 (2002).

[40] Č. Brukner, T. Paterek, and M. Żukowski, International Journal of Quantum Information **01**, 519–525 (2003).

[41] Č. Brukner, M. Żukowski, J.-W. Pan, and A. Zeilinger, Physical Review Letters **92** (2004).

[42] A. Tavakoli and M. Żukowski, Phys. Rev. A **95**, 042305 (2017).

[43] A. Tavakoli, M. Żukowski, and Č. Brukner, Quantum **4**, 316 (2020).

[44] J. Ho, G. Moreno, S. Brito, F. Graffitti, C. L. Morrison, R. Nery, A. Pickston, M. Proietti, R. Rabelo, A. Fedrizzi, and R. Chaves, npj Quantum Information **8** (2022).

[45] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, Phys. Rev. Lett. **104**, 230503 (2010).

[46] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, IEEE Transactions on Information Theory **57**, 5509 (2011).

[47] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, SIAM Journal on Computing **38**, 366–384 (2008).

[48] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).

[49] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).

[50] B. D. Clader, B. C. Jacobs, and C. R. Sprouse, Physical Review Letters **110**, 10.1103/physrevlett.110.250504 (2013).

[51] A. M. Childs, E. Farhi, and S. Gutmann, Quantum Information Processing **1**, 35–43 (2002).

[52] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, STOC03 (ACM, 2003) p. 59–68.

[53] J. Kempe, Probability Theory and Related Fields **133**, 215–235 (2005).

[54] A. Ambainis, A. M. Childs, B. W. Reichardt, R. Špalek, and S. Zhang, SIAM Journal on Computing **39**, 2513–2530 (2010).

[55] M. Santha, in *Theory and applications of models of computation*, Lecture Notes in Comput. Sci., Vol. 4978 (Springer, Berlin, 2008) pp. 31–46.

[56] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson, SIAM Journal on Computing **32**, 1570–1585 (2003).

[57] M. Szegedy, in *45th Annual IEEE Symposium on Foundations of Computer Science* (2004) pp. 32–41.

[58] F. Magniez, A. Nayak, J. Roland, and M. Santha, SIAM Journal on Computing **40**, 142–164 (2011).

[59] S. Lloyd, Science **273**, 1073–1078 (1996).

[60] D. W. Berry, A. M. Childs, and R. Kothari, in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, 2015) p. 792–809.

[61] I. Buluta and F. Nori, Science **326**, 108–111 (2009).

[62] K. L. Brown, W. J. Munro, and V. M. Kendon, Entropy **12**, 2268–2307 (2010).

[63] I. Georgescu, S. Ashhab, and F. Nori, Reviews of Modern Physics **86**, 153–185 (2014).

[64] S. K. Leyton and T. J. Osborne, A quantum algorithm to solve nonlinear differential equations (2008).

[65] D. W. Berry, Journal of Physics A: Mathematical and Theoretical **47**, 105301 (2014).

[66] N. Wiebe, D. Braun, and S. Lloyd, Phys. Rev. Lett. **109**, 050505 (2012).

[67] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).

[68] A. Montanaro, npj Quantum Information **2**, 10.1038/npjqi.2015.23 (2016).

[69] B. Paredes, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **95**, 140501 (2005).

[70] A. B. Watts, R. Kothari, L. Schaeffer, and A. Tal, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019 (Association for Computing Machinery, New York, NY, USA, 2019) p. 515–526.

[71] S. Bravyi, D. Gosset, and R. König, Science **362**, 308 (2018).

[72] S. Zhang and L. Li, CCF Transactions on High Performance Computing **4**, 53–62 (2022).

[73] L. Gyongyosi and S. Imre, Comput. Sci. Rev. **31**, 51 (2019).

[74] D. Bacon and W. van Dam, Communications of the ACM **53**, 84–93 (2010).

[75] A. M. Childs and W. van Dam, Rev. Modern Phys. **82**, 1 (2010).

[76] I. L. Chuang, N. Gershenfeld, and M. Kubinec, Phys. Rev. Lett. **80**, 3408 (1998).

[77] E. Martín-López, A. Laing, T. Lawson, R. Alvarez, X.-Q. Zhou, and J. L. O'Brien, Nature Photonics **6**, 773–776 (2012).

[78] L. M. K. Vandersypen, M. Steffen, M. H. Sherwood, C. S. Yannoni, G. Breyta, and I. L. Chuang, Applied Physics Letters **76**, 646–648 (2000).

[79] J. King, S. Yarkoni, M. M. Nevisi, J. P. Hilton, and C. C. McGeoch, Benchmarking a quantum annealing processor with the time-to-target metric (2015).

[80] X.-D. Cai, C. Weedbrook, Z.-E. Su, M.-C. Chen, M. Gu, M.-J. Zhu, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Phys. Rev. Lett. **110**, 230501 (2013).

[81] S. Barz, I. Kassal, M. Ringbauer, Y. O. Lipp, B. Dakić, A. Aspuru-Guzik, and P. Walther, Scientific Reports **4**, 10.1038/srep06115 (2014).

[82] J. Pan, Y. Cao, X. Yao, Z. Li, C. Ju, H. Chen, X. Peng, S. Kais, and J. Du, Phys. Rev. A **89**, 022313 (2014).

[83] S. Somaroo, C. H. Tseng, T. F. Havel, R. Laflamme, and D. G. Cory, Phys. Rev. Lett. **82**, 5381 (1999).

[84] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, Science **334**, 57–61 (2011).

[85] A. Politi, J. C. F. Matthews, and J. L. O'Brien, Science **325**, 1221 (2009).

[86] R. Blatt and C. F. Roos, Nature Physics **8**, 277–284 (2012).

[87] T. C. Ralph, Nature Photonics **7**, 514–515 (2013).

[88] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen,

A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Nature **574**, 505–510 (2019).

[89] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Science **370**, 1460–1463 (2020).

[90] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, Nature **414**, 883–887 (2001).

[91] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, R. Cleve, and I. L. Chuang, Phys. Rev. Lett. **85**, 5452 (2000).

[92] M. Steffen, M. Lieven, Vandersypen, and I. Chuang, IEEE Micro **21**, 24 (2001).

[93] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Nature **464**, 45–53 (2010).

[94] M. Liu, R. Shaydulin, P. Niroula, M. DeCross, S.-H. Hung, W. Y. Kon, E. Cervero-Martín, K. Chakraborty, O. Amer, S. Aaronson, A. Acharya, Y. Alexeev, K. J. Berg, S. Chakrabarti, F. J. Curchod, J. M. Dreiling, N. Erickson, C. Foltz, M. Foss-Feig, D. Hayes, T. S. Humble, N. Kumar, J. Larson, D. Lykov, M. Mills, S. A. Moses, B. Neyenhuis, S. Eloul, P. Siegfried, J. Walker, C. Lim, and M. Pistoia, Nature **640**, 343–348 (2025).

[95] T. Lunghi, J. B. Brask, C. C. W. Lim, Q. Lavigne, J. Bowles, A. Martin, H. Zbinden, and N. Brunner, Phys. Rev. Lett. **114**, 150501 (2015).

[96] R. Colbeck and A. Kent, Journal of Physics A: Mathematical and Theoretical **44**, 095305 (2011).

[97] M. Herrero-Collantes and J. C. Garcia-Escartin, Rev. Mod. Phys. **89**, 015004 (2017).

[98] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, npj Quantum Information **2**, 10.1038/npjqi.2016.21 (2016).

[99] A. Acín and L. Masanes, Nature **540**, 213–219 (2016).

[100] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, Bangalore, India, 1984) pp. 175–179.

[101] A. K. Ekert, Physical Review Letters **67**, 661–663 (1991).

[102] D. Mayers and A. Yao, in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, FOCS '98 (IEEE Computer Society, USA, 1998) p. 503.

[103] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).

[104] P. Eraerds, N. Walenta, M. Legré, N. Gisin, and H. Zbinden, New Journal of Physics **12**, 063027 (2010).

[105] G. Brassard, in *IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security, 2005.* (2005) pp. 19–23.

[106] A. R. Dixon, Z. L. Yuan, J. F. Dynes, A. W. Sharpe, and A. J. Shields, Optics Express **16**, 18790 (2008).

[107] P. A. Hiskett, D. Rosenberg, C. G. Peterson, R. J. Hughes, S. Nam, A. E. Lita, A. J. Miller, and J. E. Nordholt, New Journal of Physics **8**, 193–193 (2006).

[108] W.-Y. Hwang, Phys. Rev. Lett. **91**, 057901 (2003).

[109] B. Korzh, C. C. W. Lim, R. Houlmann, N. Gisin, M. J. Li, D. Nolan, B. Sanguinetti, R. Thew, and H. Zbinden, Nature Photonics **9**, 163–168 (2015).

[110] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in Optics and Photonics **12**, 1012 (2020).

[111] Y.-A. Chen, Q. Zhang, T.-Y. Chen, W.-Q. Cai, S.-K. Liao, J. Zhang, K. Chen, J. Yin, J.-G. Ren, Z. Chen, S.-L. Han, Q. Yu, K. Liang, F. Zhou, X. Yuan, M.-S. Zhao, T.-Y. Wang, X. Jiang, L. Zhang, W.-Y. Liu, Y. Li, Q. Shen, Y. Cao, C.-Y. Lu, R. Shu, J.-Y. Wang, L. Li, N.-L. Liu, F. Xu, X.-B. Wang, C.-Z. Peng, and J.-W. Pan, Nature **589**, 214–219 (2021).

[112] M. Pittaluga, M. Minder, M. Lucamarini, M. Sanzaro, R. I. Woodward, M.-J. Li, Z. Yuan, and A. J. Shields, Nature Photonics **15**, 530–535 (2021).

[113] H.-Y. Liu, X.-H. Tian, C. Gu, P. Fan, X. Ni, R. Yang, J.-N. Zhang, M. Hu, J. Guo, X. Cao, X. Hu, G. Zhao, Y.-Q. Lu, Y.-X. Gong, Z. Xie, and S.-N. Zhu, Phys. Rev. Lett. **126**, 020503 (2021).

[114] C. M. Knaut, A. Suleymanzade, Y.-C. Wei, D. R. Assumpcao, P.-J. Stas, Y. Q. Huan, B. Machielse, E. N. Knall, M. Sutula, G. Baranes, N. Sinclair, C. De-Eknamkul, D. S. Levonian, M. K. Bhaskar, H. Park, M. Lončar, and M. D. Lukin, Nature **629**, 573–578 (2024).

[115] J.-L. Liu, X.-Y. Luo, Y. Yu, C.-Y. Wang, B. Wang, Y. Hu, J. Li, M.-Y. Zheng, B. Yao, Z. Yan, D. Teng, J.-W. Jiang, X.-B. Liu, X.-P. Xie, J. Zhang, Q.-H. Mao, X. Jiang, Q. Zhang, X.-H. Bao, and J.-W. Pan, Nature **629**, 579–585 (2024).

[116] A. J. Stolk, K. L. van der Enden, M.-C. Slater, I. te Raa-Derckx, P. Botma, J. van Rantwijk, J. J. B. Biemond, R. A. J. Hagen, R. W. Herfst, W. D. Koek, A. J. H. Meskers, R. Vollmer, E. J. van Zwet, M. Markham, A. M. Edmonds, J. F. Geus, F. Elsen, B. Jungbluth, C. Haefner, C. Tresp, J. Stuhler, S. Ritter, and R. Hanson, Science Advances **10**, 10.1126/sciadv.adp6442 (2024).

[117] J. Preskill, Quantum **2**, 79 (2018).

[118] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Müller, W. Riess, G. Salis, J. Smolin, I. Tavernelli, and K. Temme, Quantum Science and Technology **3**, 030503 (2018).

[119] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, Journal of the Physical Society of Japan **90**, 032001 (2021).

[120] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nature Communications **5**, 10.1038/ncomms5213 (2014).

[121] J. Allcock and S. Zhang, National Science Review **6**, 26–28 (2018).

[122] V. Dunjko and H. J. Briegel, Rep. Progr. Phys. **81**, 074001, 67 (2018).

[123] J. Lin, Z. Y. Lai, and X. Li, Phys. Rev. A **101**, 052327 (2020).

[124] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, Quantum Inf. Comput. **15**, 1 (2015).

[125] S. P. Jordan, K. S. M. Lee, and J. Preskill, Science **336**, 1130–1133 (2012).

[126] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Phys. Rev. A **90**, 022305 (2014).

[127] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning (2013).

[128] S. Aaronson, Nature Physics **11**, 291–293 (2015).

[129] A. Aspuru-Guzik and P. Walther, Nature Physics **8**, 285–291 (2012).

[130] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Nature Physics **8**, 325–330 (2012).

[131] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Chemical Reviews **120**, 12685–12717 (2020).

[132] S. Barzanjeh, S. Pirandola, D. Vitali, and J. M. Fink, Science Advances **6**, 10.1126/sciadv.abb0451 (2020).

[133] J. S. Bell, Physics Physique Fizika **1**, 195–200 (1964).

[134] E. Schrödinger, Mathematical Proceedings of the Cambridge Philosophical Society **31**, 555–563 (1935).

[135] E. Schrödinger, Mathematical Proceedings of the Cambridge Philosophical Society **32**, 446–452 (1936).

[136] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[137] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).

[138] A. Einstein, B. Podolsky, and N. Rosen, Physical Review **47**, 777–780 (1935).

[139] S. KOCHEN and E. P. SPECKER, Journal of Mathematics and Mechanics **17**, 59 (1967).

[140] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).

[141] J. S. BELL, Rev. Mod. Phys. **38**, 447 (1966).

[142] S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).

[143] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[144] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, Nature **403**, 515–519 (2000).

[145] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Nature **526**, 682–686 (2015).

[146] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[147] D. Mayers and A. Yao, Quantum Information and Computation **4**, 273–286 (2004).

[148] M. McKague, T. H. Yang, and V. Scarani, Journal of Physics A: Mathematical and Theoretical **45**, 455304 (2012).

[149] F. Buscemi, Phys. Rev. Lett. **108**, 200401 (2012).

[150] E. Verbanis, A. Martin, D. Rosset, C. C. W. Lim, R. T. Thew, and H. Zbinden, Phys. Rev. Lett. **116**, 190501 (2016).

[151] T. Vértesi and E. Bene, Phys. Rev. A **82**, 062115 (2010).

[152] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Science Advances **6**, 10.1126/sciadv.aaw6664 (2020).

[153] N. Brunner, M. Navascués, and T. Vértesi, Phys. Rev. Lett. **110**, 150501 (2013).

[154] S. Wehner, M. Christandl, and A. C. Doherty, Phys. Rev. A **78**, 062112 (2008).

[155] M. M. Wolf and D. Perez-Garcia, Phys. Rev. Lett. **102**, 190504 (2009).

[156] J. Bowles, M. T. Quintino, and N. Brunner, Phys. Rev. Lett. **112**, 140407 (2014).

[157] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, 1996).

[158] J. Hromkovič, *Communication Complexity and Parallel Computing* (Springer-Verlag, New York, 1997).

[159] A. S. Holevo, Problems of Information Transmission **9**, 177 (1973), originally published in *Probl. Peredachi Inf.* **9**(3), 3–11 (1973).

[160] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07 (Association for Computing Machinery, New York, NY, USA, 2007) p. 516–525.

[161] J. Barrett, Phys. Rev. A **75**, 032304 (2007).

[162] D. Leung, L. Mancinska, W. Matthews, M. Ozols, and A. Roy, Communications in Mathematical Physics **311**, 97–111 (2012).

[163] P. E. Frenkel and M. Weiner, Quantum **6**, 662 (2022).

[164] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[165] M. Alimuddin, A. Chakraborty, G. L. Sidhardh, R. K. Patra, S. Sen, S. R. Chowdhury, S. G. Naik, and M. Banik, Phys. Rev. A **108**, 052430 (2023).

[166] I. Ivanovic, Physics Letters A **123**, 257 (1987).

[167] A. Peres, Physics Letters A **128**, 19 (1988).

[168] M. T. DiMario and F. E. Becerra, npj Quantum Information **8**, 84 (2022).

[169] J. Bae and L.-C. Kwek, Journal of Physics A: Mathematical and Theoretical **48**, 083001 (2015).

[170] J. Shang, A. Asadian, H. Zhu, and O. Gühne, Phys. Rev. A **98**, 022309 (2018).

[171] R. Derka, V. Buzˇek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).

[172] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Journal of Mathematical Physics **45**, 2171–2180 (2004).

[173] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, IEEE Transactions on Information Theory **63**, 5628 (2017).

[174] R. Stricker, M. Meth, L. Postler, C. Edmunds, C. Ferrie, R. Blatt, P. Schindler, T. Monz, R. Kueng, and M. Ringbauer, PRX Quantum **3**, 040310 (2022).

[175] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne, Phys. Rev. Lett. **129**, 220502 (2022).

[176] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, Phys. Rev. A **56**, 1163 (1997).

[177] S. Ishizaka and T. Hiroshima, Phys. Rev. Lett. **101**, 240501 (2008).

[178] M. Studziński, S. Strelchuk, M. Mozrzymas, and M. Horodecki, Scientific Reports **7**, 10871 (2017).

[179] M. Mozrzymas, M. Studziński, S. Strelchuk, and M. Horodecki, New Journal of Physics **20**, 053006 (2018).

[180] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Phys. Rev. A **94**, 052108 (2016), arXiv:1608.02634 [quant-ph] .

[181] M. Szczykulska, T. Baumgratz, and A. Datta, Advances in Physics: X **1**, 621–639 (2016).

[182] F. Albarelli and R. Demkowicz-Dobrzański, Phys. Rev. X **12**, 011039 (2022).

[183] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, Phys. Rev. A **93**, 040102 (2016).

[184] S. Rout, N. Sakharwade, S. S. Bhattacharya, R. Ramanathan, and P. Horodecki, Phys. Rev. Res. **7**, 023104 (2025).

[185] S. Rout, S. Sankar Bhattacharya, and P. Horodecki, New Journal of Physics **27**, 033024 (2025).

[186] J.-L. Chen, A. Cabello, Z.-P. Xu, H.-Y. Su, C. Wu, and L. C. Kwek, Phys. Rev. A **88**, 062116 (2013).

[187] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer Science & Business Media, New York, 2012) p. 676.

[188] J. J. Sakurai, *Advanced Quantum Mechanics*, 1st ed. (Addison–Wesley, Reading, MA, 1967).

[189] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2nd ed. (Cambridge University Press, Cambridge, 2017).

[190] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2012).

[191] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, Journal of Physics A: Mathematical and General **38**, 5979 (2005).

[192] G. Chiribella, G. M. D'Ariano, and D. Schlingemann, Phys. Rev. Lett. **98**, 190403 (2007).

[193] M. A. Naimark, Izv. Akad. Nauk USSR Ser. Mat. **4**, 277 (1940).

[194] F. Riesz and B. Szőkefalvi-Nagy, *Functional Analysis* (Dover Publications, New York, 1990).

[195] V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, 2020).

[196] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (Ungar, New York, 1963).

[197] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Phys. Rev. Lett. **119**, 190501 (2017).

[198] M. Oszmaniec, F. B. Maciejewski, and Z. Puchała, Phys. Rev. A **100**, 012351 (2019).

[199] M. Kotowski and M. Oszmaniec, Pretty-good simulation of all quantum measurements by projective measurements (2025).

[200] T. Singal, F. B. Maciejewski, and M. Oszmaniec, npj Quantum Information **8**, 10.1038/s41534-022-00589-1 (2022).

[201] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[202] M. Horodecki, P. Horodecki, and R. Horodecki, Physics Letters A **223**, 1 (1996).

[203] B. M. Terhal, Physics Letters A **271**, 319–326 (2000).

[204] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Physical Review A **62** (2000).

[205] O. Gühne and G. Tóth, Physics Reports **474**, 1–75 (2009).

[206] M. Plávala, Physics Reports **1033**, 1–64 (2023).

[207] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, PRX Quantum **2**, 010331 (2021).

[208] J. H. Selby, M. E. Stasinou, M. Wilson, and B. Coecke, Generalised process theories (2025), arXiv:2502.10368 [math.CT] .

[209] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **84**, 012311 (2011).

[210] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).

[211] G. Chiribella, Electronic Proceedings in Theoretical Computer Science **172**, 1–14 (2014).

[212] P. Janotta, Electronic Proceedings in Theoretical Computer Science **95**, 183–192 (2012).

[213] S. Popescu and D. Rohrlich, Foundations of Physics **24**, 379–385 (1994).

[214] D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951).

[215] J. S. Bell and A. Aspect, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, 2004).

[216] V. Scarani, *Bell Nonlocality* (Oxford University PressOxford, 2019).

[217] A. Fine, Physical Review Letters **48**, 291–295 (1982).

[218] V. Scarani, arXiv preprint arXiv:1303.3081 (2013), arXiv:1303.3081 [quant-ph] .

[219] W. SLOFSTRA, Forum of Mathematics, Pi **7**, 10.1017/fmp.2018.3 (2019).

[220] H. Fu, C. A. Miller, and W. Slofstra, Communications in Mathematical Physics **406**, 10.1007/s00220-024-05229-7 (2025).

[221] B. S. Cirel'son, Letters in Mathematical Physics **4**, 93–100 (1980).

[222] T. Cope and R. Colbeck, Phys. Rev. A **100**, 022114 (2019).

[223] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, American Journal of Physics **58**, 1131–1143 (1990).

[224] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).

[225] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).

[226] S. Goldstein, Phys. Rev. Lett. **72**, 1951 (1994).

[227] N. D. Mermin, American Journal of Physics **62**, 880–887 (1994).

[228] P. G. Kwiat and L. Hardy, American Journal of Physics **68**, 33–36 (2000).

[229] M. Ferrero and A. van der Merwe, eds., *New Developments on Fundamental Problems in Quantum Physics* (Springer Netherlands, 1997).

[230] D. Boschi, S. Branca, F. De Martini, and L. Hardy, Phys. Rev. Lett. **79**, 2755 (1997).

[231] S. Kunkri and S. K. Choudhary, Phys. Rev. A **72**, 022348 (2005).

[232] K. P. Seshadreesan and S. Ghosh, Journal of Physics A: Mathematical and Theoretical **44**, 315305 (2011).

[233] R. Rabelo, L. Y. Zhi, and V. Scarani, Phys. Rev. Lett. **109**, 180401 (2012).

[234] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Quantum **4**, 280 (2020).

[235] D. Schmid, D. Rosset, and F. Buscemi, Quantum **4**, 262 (2020).

[236] L. Lovasz, IEEE Transactions on Information Theory **25**, 1 (1979).

[237] A. Solís-Encina and J. R. Portillo, Orthogonal representation of graphs (2015), arXiv:1504.03662 [math.CO] .

[238] L. Lovász, M. Saks, and A. Schrijver, Linear Algebra and its Applications **114–115**, 439–454 (1989).

[239] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-A. Larsson, Rev. Mod. Phys. **94**, 045007 (2022).

[240] A. Cabello, J. Estebaranz, and G. García-Alcaine, Physics Letters A **212**, 183–187 (1996).

[241] A. Chi-Chih Yao, in *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science* (1993) pp. 352–361.

[242] I. Kremer, *Quantum Communication*, Master's thesis, Hebrew University, Computer Science Department (1995).

[243] C. E. Shannon, Bell System Technical Journal **27**, 379–423 (1948).

[244] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).

[245] B. Kalyanasundaram and G. Schnitger, SIAM Journal on Discrete Mathematics **5**, 545 (1992).

[246] E. Sperner, Mathematische Zeitschrift **27**, 544–548 (1928).

[247] Y. L. Luke, *The Special Functions and Their Approximations*, Vol. 1 (Academic Press, New York, NY, USA, 1969).

[248] H. Maehara and V. Rödl, Graphs and Combinatorics **6**, 365–367 (1990).

[249] D. Goyeneche and K. Życzkowski, Phys. Rev. A **90**, 022316 (2014).

[250] D. Goyeneche, D. Alsina, J. I. Latorre, A. Riera, and K. Życzkowski, Phys. Rev. A **92**, 032316 (2015).

[251] D. Goyeneche, J. Bielawski, and K. Życzkowski, Phys. Rev. A **94**, 012346 (2016).

[252] S. Pang, H. Xu, and M. Chen, Entropy **24**, 1000 (2022).

[253] A. K. Pati, Phys. Rev. A **63**, 014302 (2000).

[254] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, Phys. Rev. Lett. **87**, 077902 (2001).

[255] B.-G. Englert and Y. Aharonov, Physics Letters A **284**, 1 (2001).

[256] P. K. Aravind, Zeitschrift für Naturforschung A **58**, 85 (2003).

[257] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **92**, 067902 (2004).

[258] M. A. Ballester and S. Wehner, Phys. Rev. A **75**, 022319 (2007).

[259] W. K. Wootters and B. D. Fields, Annals of Physics **191**, 363 (1989).

[260] I. D. Ivonovic, Journal of Physics A: Mathematical and General **14**, 3241 (1981).

[261] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).

[262] J. Ahrens, P. Badziąg, M. Pawłowski, M. Żukowski, and M. Bourennane, Phys. Rev. Lett. **112**, 140401 (2014).

[263] Y. Cai, J.-D. Bancal, J. Romero, and V. Scarani, Journal of Physics A: Mathematical and Theoretical **49**, 305301 (2016).

[264] A. M. Childs and W. van Dam, Rev. Mod. Phys. **82**, 1 (2010).

[265] D. Bacon, A. M. Childs, and W. v. Dam, Chicago Journal of Theoretical Computer Science **2006** (2006).

[266] P. Sen, in *21st Annual IEEE Conference on Computational Complexity (CCC'06)* (2006) pp. 14 pp.–287.

[267] J. Barrett, Phys. Rev. A **65**, 042302 (2002).

[268] J. S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. L. Pregnell, C. Silberhorn, T. C. Ralph, J. Eisert, M. B. Plenio, and I. A. Walmsley, Nature Physics **5**, 27 (2009).

[269] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, Science Advances **6**, eaaw6664 (2020), https://www.science.org/doi/pdf/10.1126/sciadv.aaw6664 .

[270] D. Martínez, E. S. Gómez, J. Cariñe, L. Pereira, A. Delgado, S. P. Walborn, A. Tavakoli, and G. Lima, Nature Physics **19**, 190 (2023).

[271] T. Vértesi and E. Bene, Phys. Rev. A **82**, 062115 (2010).

[272] E. S. Gómez, S. Gómez, P. González, G. Cañas, J. F. Barra, A. Delgado, G. B. Xavier, A. Cabello, M. Kleinmann, T. Vértesi, and G. Lima, Phys. Rev. Lett. **117**, 260401 (2016).

[273] J. Barrett and N. Gisin, Phys. Rev. Lett. **106**, 100406 (2011).

[274] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, Phys. Rev. Lett. **113**, 190402 (2014).

[275] M.-O. Renou, E. Bäumer, S. Boreiri, N. Brunner, N. Gisin, and S. Beigi, Phys. Rev. Lett. **123**, 140401 (2019).

[276] T. Guha, M. Alimuddin, S. Rout, A. Mukherjee, S. S. Bhattacharya, and M. Banik, Quantum **5**, 569 (2021).

[277] Z. Ma, M. Rambach, K. Goswami, S. S. Bhattacharya, M. Banik, and J. Romero, Phys. Rev. Lett. **131**, 130201 (2023).

[278] A. Chaturvedi, M. Pawlowski, and K. Horodecki, Phys. Rev. A **96**, 022125 (2017).

[279] V. M, R. k. Patra, M. Janpandit, S. Sen, M. Banik, and A. Chaturvedi, Phys. Rev. A **104**, 012420 (2021).

[280] A. Hameedi, D. Saha, P. Mironowicz, M. Pawłowski, and M. Bourennane, Phys. Rev. A **95**, 052345 (2017).

[281] P. E. Frenkel and M. Weiner, Communications in Mathematical Physics **340**, 563–574 (2015).

[282] S. A. Yadavalli and R. Kunjwal, Quantum **6**, 839 (2022).

[283] K. Agarwal, S. G. Naik, A. Chakraborty, S. Sen, P. Ghosal, B. Paul, M. Banik, and R. K. Patra, Nonlocality-assisted enhancement of error-free communication in noisy classical channels (2025), arXiv:2412.04779 [quant-ph] .

[284] M. Navascués, S. Pironio, and A. Acín, New Journal of Physics **10**, 073013 (2008).

[285] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98**, 010401 (2007).

[286] P. J. Cavalcanti, J. H. Selby, J. Sikora, T. D. Galley, and A. B. Sainz, npj Quantum Information **8**, 10.1038/s41534-022-00574-8 (2022).

[287] T. Van Himbeeck, J. Bohr Brask, S. Pironio, R. Ramanathan, A. B. Sainz, and E. Wolfe, Quantum **3**, 186 (2019).

[288] A. B. Sainz, M. J. Hoban, P. Skrzypczyk, and L. Aolita, Phys. Rev. Lett. **125**, 050404 (2020).

[289] D. Collins and N. Gisin, Journal of Physics A: Mathematical and General **37**, 1775–1787 (2004).

[290] R. Horodecki, P. Horodecki, and M. Horodecki, Physics Letters A **200**, 340–344 (1995).

[291] P. Mironowicz, Journal of Physics A: Mathematical and Theoretical **57**, 163002 (2024).

[292] D. Saha, P. Horodecki, and M. Pawłowski, New Journal of Physics **21**, 093057 (2019).

[293] B. T. Gard, K. R. Motes, J. P. Olson, P. P. Rohde, and J. P. Dowling, An introduction to boson-sampling, in *From Atomic to Mesoscale* (WORLD SCIENTIFIC, 2015) p. 167–192.

[294] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, New Journal of Physics **13**, 063024 (2011).

[295] M. Navascués, A. Feix, M. Araújo, and T. Vértesi, Phys. Rev. A **92**, 042117 (2015).