

Characterizing the projective simulability of quantum measurements and instruments

by Julia Strömberg



LUND
UNIVERSITY

M.Sc. Thesis

Project duration: 4 months

Supervisor: Armin Tavakoli

Co-supervisors: Gabriele Cobucci, Shishir Khandelwal

Examiner: Peter Samuelsson

Division of Mathematical Physics

Department of Physics

Lund University, Faculty of Engineering (LTH)

Lund, Sweden

May 2026

Abbreviations and symbols

PSD	Positive Semidefinite
QM	Quantum Measurement
QC	Quantum Channel
CPTP	Complete Positive Trace Preserving
QI	Quantum Instrument
POVM	Positive Operator-Valued Measure
PVM	Projective Operator-Valued Measure
PM	Projective Measurement
PI	Projective Instrument
SDP	Semidefinite Program
$\mathcal{H}^d$	d -dimensional Hilbert space.
$\mathcal{D}(\mathcal{H})$	Space of density matrices.
$\mathcal{L}_+(\mathcal{H})$	Hilbert-Schmidt space of (bounded, finite-norm) positive semidefinite linear operators.
$\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$	Space of linear operators taking a Hilbert space $\mathcal{H}_A$ to a Hilbert space $\mathcal{H}_B$.
$\mathcal{P}_E$	Set of projectively simulable measurements.
$\mathcal{P}_I$	Set of projectively simulable instruments.

Abstract

In the heart of every physical theory lies measurement theory. While classical measurements are passive probes, QMs can be viewed as active physical processes that shape the available information. Depending on the measurement type, different aspects of the information about a system are retrieved. All possible measurements in quantum physics are described by mathematical objects called Positive Operator-Valued Measures (POVMs). The most familiar standard measurements are called Projective-Valued Measures (PVMs) and make up a subset of POVMs. The complementary subset of POVMs is non-projective. However, some non-projective POVMs can be expressed as a probabilistic mixture of PVMs. Such POVMs are called projectively simulable, while non-simulable POVMs are genuinely non-projective. These can be made projectively simulable by adding noise to the measurements. If a genuinely non-projective POVM requires a large amount of noise before it becomes projectively simulable, it is more non-projective than a POVM that requires less noise. Hence, the degree of non-projectiveness is determined by the required noise strength.

The act of measuring a quantum system changes the quantum state. While quantum measurements only take into account the readable classical output value, quantum instruments account for both the classical outcome and the post-measurement state. Hence, instruments are associated with measurements, and in an analogous way, projective and projectively simulable instruments can be defined.

For many purposes, non-projective POVMs are more useful than PVMs, and the more non-projective a POVM is, the greater the advantage of using them over PVMs. Therefore, the aim of the thesis is to investigate the genuinely non-projective POVMs. More precisely, the degree of non-projectiveness in qubit POVMs is studied. This is done by convex optimization, specifically by solving semidefinite programs. Furthermore, the thesis addresses the open problem of finding the most non-projective qubit quantum instrument by a witness-based see-saw algorithm.

Popular science summary

Is it not beautiful to throw oneself into the abyss of unknowingness? In her poem, *Utopia*, Wisława Szymborska wrote: “As if all you can do here is leave and plunge, never to return, into the depths. Into unfathomable life.” The conventional interpretation is of a political nature; however, I would like to stretch it to the extreme and read it through the eyes of a physicist. Through these new lenses, the isolated island called *Utopia*, where “The Tree of Understanding, dazzlingly straight and simple, sprouts by the spring called *Now I Get It*”, is an allegorical description of classical physics. Seemingly disentangled and complete, it is nevertheless merely a small part of the story, part of something far bigger. The realm of quantum mechanics stretches across the island and far into the sea surrounding it, where physics becomes a conundrum.

In an attempt to tangent the tiny world we know very little about, quantum measurements and instruments play a key role. In a classical physics picture, the attributes of physical objects are well-defined and revealed by performing a measurement. In contrast, the characteristics of quantum physical systems can be viewed as emerging in the course of the measurement process itself. Depending on the measurement procedure used, different aspects of the information are obtained.

In quantum mechanics, the general measurement procedures that can be performed on a quantum system are described by mathematical objects called Positive Operator-Valued Measures (POVMs). A special class of such measurements is the Projective-Valued Measures (PVMs). The POVMs that do not belong to this set are complementarily called non-projective measurements. PVMs are generally easy to implement in an experiment and are considered standard measurements in quantum mechanics. However, for many purposes, they do not give a good picture of the characteristics. Instead, it is more advantageous to use non-projective POVMs to reveal more useful information. However, such measurements are generally difficult to implement experimentally as they require more resources.

It is known that, in the set of non-projective POVMs, there exist measurements that can be simulated by a mixture of PVMs. These POVMs offer no advantage over PVMs, as they reveal the same informational aspects and can therefore be implemented using the simpler projective schemes. On the other hand, for the POVMs that cannot be simulated projectively, we cannot resort to simpler experimental implementations, and the data retrieved might be more interesting. Such measurements are genuinely non-projective. It is, however, known that some of these genuinely non-projective POVMs can be easily made projectively simulable by perturbing them with noise. Some POVMs require more noise and some less. If a POVM requires a large amount of noise, it is more non-projective than one that needs less noise. Very non-projective POVMs are of greater interest as their statistics are very far from resembling the statistics of PVMs. Therefore, it might be worth the trouble of implementing them.

In the spirit of this, this thesis is about drawing the border between projectively simulable POVMs and genuinely non-projective POVMs. To this end, we characterize the degree of non-projectiveness among genuinely non-projective POVMs and ask which quantum measurements are the most non-projective.

To extend the picture of measurements, the thesis also treats quantum instruments. Imagine an electron moving in space. When we measure its properties, e.g., position, we register a classical value. Hence, a quantum measurement device takes in a quantum state and outputs a readable classical outcome. However, the act of measuring perturbs the state of the electron, leaving it in a changed state after the measurement has been performed. A quantum instrument is a mathematical model that describes both the classical outcome and the post-measurement state of the electron. Thus, a quantum instrument has an associated measurement part and can therefore be seen as an extension of the measurement. Consequently, the same questions can be asked in the context of instruments. For this purpose, a search algorithm is developed for finding the most non-projective instruments.

Acknowledgements

I would first like to express my great gratitude to Armin for introducing me to the research group and for giving me the opportunity of writing my thesis there and on this topic. Then, I want to give my deepest thanks to Gabriele for your supervision and advice throughout the thesis. Thank you for always finding the time to answer my questions (trivial or not) and for showing understanding and patience. I would also like to deeply thank Shishir for your feedback and advice. Thank you for always taking the time to answer questions and for the inspiring discussions we have had.

I would also like to thank all the people I shared the office with for being so kind and for making the writing process more fun and relaxing. Special thanks to Oscar, Tim, William, and Aria, for all the discussions we have had and for your comradeship.

I would then like to thank my family from the bottom of my heart for your unconditional love and support throughout my university studies. Thank you for believing in me when I could not.

Lastly, I would like to thank the QIT Scania group as well as the division for giving me insight into the incredible work underway at the corner of Fysicum. Keep on inspiring the world around you.

I acknowledge the use of ChatGPT for grammatical refinement and structural improvement in preparation of the thesis, clarifying code syntax and theoretical mathematical formalism.

Contents

Popular scientific summary (English)	i
Acknowledgements	iii
1 Introduction	1
1.1 Scope of the work	3
1.2 Thesis outline	3
2 Quantum information theory	4
2.1 Qubits	4
2.1.1 The density operator	4
2.2 Bloch sphere representation and state geometry	5
2.3 Quantum measurements	6
2.3.1 Projection-valued measures (PVMs)	6
2.3.2 Positive operator-valued measures (POVMs)	7
2.4 Quantum channels	8
2.5 Representations of quantum channels	9
2.5.1 Choi representation	9
2.5.2 Kraus operators	10
2.6 Quantum instruments	10
2.7 Semidefinite programs	13
3 Projective simulability of quantum measurements	14
3.1 Protocol	14
3.2 Quantification of projective simulability	15
3.3 Condition for simulability of POVMs	18
3.4 Three-outcome POVM	19
3.4.1 White noise	20
3.4.2 Antipodal noise	20
3.4.3 Worst-case noise	20
3.5 Four-outcome equal-weight POVM	23
3.5.1 White noise, antipodal noise, and worst-case noise	24
4 Projective simulability of quantum instruments	27
4.1 Protocol	27
4.2 Condition for simulability of instruments	29
4.3 The most non-projective instruments	32
5 Conclusions and outlook	38

6	Appendix	44
A	Why Eq. (3.4) uses minimization	44
B	Proof of redundancy of classical post-processing of outcomes in the PVM-simulation	44
C	Derivation of $v_{\text{white}}(\theta)$	46
D	Derivation of $v_{\text{ap}}(\theta)$	48
E	Derivation of $v_{\text{wc}}(\theta)$	50
F	Proof of why $\theta + \psi \geq \pi$	58
G	Derivation of $v_{\text{white}}(\theta, \psi)$	59
	G.1 Preliminaries	59
	G.2 Minimization problem	60
	G.3 Solution to the minimization problem	64
H	Derivation of $v_{\text{wc}}(\theta, \psi)$	68
	H.1 Expression for $T(\vec{r})$	69
	H.2 The auxiliary variable $t = 1/v$	72
	H.3 Case I: Balanced branch, $R_1 \leq R_2 + R_3$	73
	H.4 Case II: Unbalanced branch, $R_1 > R_2 + R_3$	74
	H.5 Tangency condition in the unbalanced branch	75
	H.6 Parametrization by λ	76
	H.7 Stationary equation for λ_*	77
	H.8 Algebraic root form	78
	H.9 Final formula	79
I	Proof of redundancy of classical post-processing of outcomes in PI-simulation	79
J	Derivation of Eq. 4.3	82
K	Proof of Theorem 1	82
L	Proof of Theorem 2	85
M	Motivation for using the positive partial transpose criterion for qubit-qubit instruments	89
N	Unsharp observables	89
O	Dual SDP for fixed noise	90
P	Dual SDP for worst-case noise	92
Q	Derivation of $v_{\text{wc}}^{(3)}(\zeta)$	95
	Q.1 Preliminaries	95
	Q.2 Positivity condition	97
	Q.3 Expression for critical visibility, v_{max}	103
	Q.4 Solution to the t -optimization	105
R	Derivation of $v_{\text{wc}}^{(4)}(\zeta)$	107
	R.1 Preliminaries	107
	R.2 The value of μ	108
	R.3 Expression of $v(\zeta, t)$	110
S	Computational cost	111

Chapter 1

Introduction

The theory of quantum measurement (QM) is a wide-ranging field and a building block in quantum theory. QMs provide the link between quantum systems and classical information. Fundamentally, a measurement device takes a quantum system as input and produces a classical outcome, which can then be recorded and processed. Although similar to classical physics, the quantum case is fundamentally different: the act of measuring can change the system being measured, and the measurement procedure used determines what information is obtained and how strongly the system is disturbed. This makes QMs essential in almost every area of quantum science, from state tomography and quantum sensing to cryptography, quantum computing, and foundational tests of quantum theory. [1]

The field of quantum measurements contains a wide range of regimes. Some measurements are strong, i.e., they extract sharp information and typically cause a large disturbance of the state [2]. Standard measurements belong to this familiar category. Other measurements are weak, meaning that they extract only partial information in a single run and therefore disturb the system less. Such measurements play an important role in sequential protocols and continuous monitoring of quantum systems [3, 4]. Measurements can also have either discrete outcomes, such as spin up or spin down, or continuous outcomes, such as time, position, and phase [5, 6]. This diversity is mathematically captured by generalized quantum measurements, and by quantum instruments when the post-measurement state is also included [7].

The standard and most familiar measurements in quantum mechanics are called projective-valued measures (PVMs). Such measurements are mathematically described as complete sets of orthogonal projectors. In the common situation of qubits, a PVM has at most two non-trivial outcomes, which intuitively correspond to asking the system a binary question.

However, the full scope of quantum measurements (QMs) is described by a more general notion of generalized measurements, known as positive operator-valued measures (POVMs). Unlike PVMs, they do not need to be described by orthogonal projectors and can also have more outcomes than the dimension of the system. In the qubit case, POVMs can have up to four irreducible outcomes [1]. In this sense, POVMs extend the measurement possibilities beyond projective measurements, which are a special case of POVMs.

The generality of POVMs stretches beyond mathematics. The advantages of generalized measurements have been recognized since the late 1980s. They have been shown to

offer advantages in tasks such as quantum state discrimination [8–11], entanglement detection [12], quantum tomography [13–19], and quantum metrology [20–22]. More recently, non-projective measurements have also been found to outperform projective measurements in information-processing protocols, including randomness certification [23] and quantum cryptography [24]. Further applications in quantum computing [25–27] and in foundational studies of quantum information [28, 29] have been established.

While generalized measurements may reveal intrinsically inaccessible information and are more powerful compared to projective ones, their physical implementation is far more complex, often requiring additional degrees of freedom or ancillary systems beyond probing and controlling the system in question [30–32]. A practical example is provided by the implementation with polarised photons, in which a PVM can be implemented using a simple polarized beam-splitter, whereas a POVM requires coupling the polarization to other degrees of freedom [33], e.g., orbital angular momentum or spatial modes. For these reasons, the ability to detect and characterize non-projective measurements is of both theoretical and practical relevance.

Interestingly, there exist non-projective POVMs that can be realized via probabilistic mixtures of PVMs, with classical post-processing applied to the outcomes [30, 34]. More precisely, such POVMs can be implemented with sole access to three resources: classical randomness, standard projective measurements on the system, and classical post-processing of the measurement outcomes. POVMs that can be simulated by an experimenter with sole access to these three resources are called projectively simulable or PM-simulable. Hence, simulable POVMs offer no operational advantage over projective schemes, as such a simulation excludes the non-projective features. This raises the interest in characterizing which POVMs can be simulated projectively or not.

As of today, determining the conditions under which POVMs outperform projective measurements remains a challenging and open problem [32]. The structure of the set of projectively simulable POVMs is not yet fully understood. However, in practice, POVMs are not noise-free, and it is known that adding sufficient noise to any POVM eventually makes it simulable [30]. The main idea is to treat non-projectiveness as a resource. If a measurement requires a lot of noise to become PM-simulable, then it is strongly non-projective. If it becomes simulable after only a little noise, then it is closer to the projective world. Hence, a second question emerges: which are the most non-projective POVMs, i.e., which POVMs require the most amount of noise to simulate projectively?

While POVMs specify the probabilities of the classical measurement outcomes, they do not describe how the quantum state changes after the measurement is performed. A full description of a measurement must therefore include both the classical outcome and the corresponding quantum post-measurement output state. The object entailing this hybrid output is called a quantum instrument (QI) [35]. Hence, QIs are central to quantum theory since they capture not only the information obtained from a system, but also the disturbance caused by the measurement [7].

In analogy to POVMs, the same questions can be asked for QIs. Instruments associated with PVMs are called projective instruments (PIs). A QI that can be simulated by an experimenter with sole access to projective measurements on the system, classical

randomness, classical post-processing of outcomes, and quantum processing of the post-measurement state is referred to as a projectively simulable or PI-simulable QI. Non-simulable QIs can, in a fully analogous way, be made PI-simulable by mixing the instrument with noise, and the noise level determines the degree of non-projectiveness of the QI. To this end, it is highly relevant to seek the QI that most strongly defies projectivity. This framework is an extension of projective simulability for POVMs, now accounting for the complete measurement process rather than only the outcome statistics.

1.1 Scope of the work

The purpose of the thesis is for it to serve as a step in the largely unresolved problem of characterizing non-projective QMs and the more recent corresponding problem for QIs. We restrict to the practical case of qubit measurements and instruments. Following the framework developed in [7, 34], we consider simulation protocols based on projective measurements, classical randomness and processing, and quantum post-processing. The protocols are used to distinguish measurements and instruments that are genuinely non-projective from those that are projectively simulable. By introducing a measure called visibility, the extent of non-projectivity can be quantified using convex optimization. More precisely, the so-called critical visibility quantifies the minimum amount of noise required for a QM or QI to become simulable. The critical visibility is optimized by solving a semidefinite program (SDP). Such a program is defined for selected families of POVMs and QIs, for different types of noise. For robustness, analytical results are derived for specific cases. Furthermore, an algorithmic approach for determining the most non-projective measurements was developed in [34]. This is extended for the purpose of the instruments in this thesis.

1.2 Thesis outline

The thesis is structured as follows. In Chapter 2, we introduce the fundamentals of quantum information theory. In Chapter 3, we present building blocks of the framework of projective simulability of qubit POVMs, including the simulation protocol and the necessary condition for PM-simulability. From this, an SDP is formulated for various noise types and two different POVM families: three-outcome symmetric POVMs and four-outcome equal-weight POVMs, defined by a set of parameters. The critical visibility is obtained as a function of these parameters, both numerically and analytically, from the SDP. In Chapter 4, a protocol for the projective simulability of QIs is described, and necessary and sufficient criteria for qubit instruments are formulated as an SDP. Furthermore, an iterative search algorithm to find the most non-projective QIs is built and run for qubit systems with varying numbers of outcomes and noise types. The results are also presented and discussed at the end of the chapter. Lastly, Chapter 5 provides a conclusion and future prospects of this research topic.

Chapter 2

Quantum information theory

2.1 Qubits

The quantum analog of a classical bit is the qubit, the fundamental unit of quantum information theory (QIT). While the classical bit can only be either in the state 0 or 1, the qubit obeys the laws of quantum mechanics and will generally be in a linear superposition of the two eigenstates $|0\rangle$ and $|1\rangle$. Qubit states reside in $\mathcal{H}^2 = \mathbb{C}^2$, which is spanned by the computational basis $\{|i\rangle\}_{i=0}^1$. Consequently, an arbitrary qubit state $|\psi\rangle$ has a complete vector representation given by [36]:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad \alpha, \beta \in \mathbb{C}, \quad (2.1)$$

where α and β are the probability amplitudes, which satisfy the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Performing a measurement in the computational basis on $|\psi\rangle$ causes it to collapse into one of the eigenstates with a probability given by Born's rule $|\langle a|\psi\rangle|^2 = |c|^2$, where $c = \alpha$ if $\langle a| = \langle 0|$ and β when $\langle a| = \langle 1|$ [37]. In this way, it is not possible to extract the full information a qubit stores.

2.1.1 The density operator

The qubit state in Eq. 2.1 is completely characterized by α and β . States with this property are called pure states [37]. In operator form, a pure state can be written as a projector $\rho = P_\psi = |\psi\rangle\langle\psi|$. A generalization of this is the so-called density operator, which provides a means of defining states that are not fully known. Such states are called mixed states and are described by a statistical mixture of pure states [36]. More precisely, if a quantum system is given by the pure state $|\psi_j\rangle$ with probability p_j , the ensemble $\{p_j, |\psi_j\rangle\}$ is described by

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \quad \sum_j p_j = 1. \quad (2.2)$$

Here, each $|\psi_j\rangle\langle\psi_j|$ is a projector P_{ψ_j} that projects ρ onto the subspace of $|\psi_j\rangle$ with a certain probability p_j . Hence, the pure state is a special case of a mixed state.

All density operators are positive semidefinite (PSD), i.e., $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle$, and normalized, $\text{Tr}(\rho) = 1$. The PSD property implies that ρ is also Hermitian, $\rho^\dagger = \rho$. Hermiticity further implies that ρ is normal (diagonalizable with an orthonormal eigenbasis).

Therefore, it possesses a spectral decomposition as in Eq. (2.2), where the eigenvectors $|\psi_j\rangle$ satisfy $\langle\psi_j|\psi_k\rangle = \delta_{j,k}$ with corresponding eigenvalues p_j [38]. Hence, the form in Eq. (2.2). It can be noted that for a pure state, ρ has only one non-zero eigenvalue, therefore $\rho = |\psi\rangle\langle\psi|$. [39]

2.2 Bloch sphere representation and state geometry

Since the coefficients α and β in Eq. (2.1) are complex, the state above is specified by four real variables. On the other hand, the quantum states are defined by a normalized vector up to a global phase factor, reducing the number of real variables to two. In polar coordinates, the state in Eq. (2.1) can be written as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad (2.3)$$

where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$. Eq. (2.3) is referred to as the Bloch sphere representation and can be visualized as in Fig. 2.1a. The projector for this state is [40]

$$\begin{aligned} \psi = |\psi\rangle\langle\psi| &= \begin{bmatrix} \cos^2\left(\frac{\theta}{2}\right) & e^{-i\varphi}\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) \\ e^{i\varphi}\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{bmatrix} \\ &= \frac{1}{2}(\mathbb{1} + \sin(\theta)\cos(\varphi)\sigma_x + \sin(\theta)\sin(\varphi)\sigma_y + \cos(\theta)\sigma_z). \end{aligned} \quad (2.4)$$

Since the Pauli matrices $\{\mathbb{1}, \sigma_X, \sigma_Y, \sigma_Z\}$ constitute a basis of the qubit Hilbert-Schmidt space $\mathcal{H}^d$, any single-qubit state has a Bloch decomposition [39]:

$$\rho = \frac{1}{2}(\mathbb{1}_2 + \vec{n} \cdot \vec{\sigma}) = \frac{1}{2} \left(\mathbb{1}_2 + \sum_{i=x,y,z} n_i \sigma_i \right), \quad (2.5)$$

where the position vector $\vec{n} = \lambda(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is called the Bloch vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. States with $|\lambda| = 1$ are pure and reside on the surface of the Bloch sphere, while states for which $|\lambda| < 1$ are mixed states, lying in the interior of the Bloch sphere. Consequently, the maximally mixed state ($|\lambda| = 0$) lives in the center of the sphere. Hence, from a geometrical point of view, the mixed states are convex mixtures of pure states, occupying the boundary of the state space in Fig. 2.1b. [41]

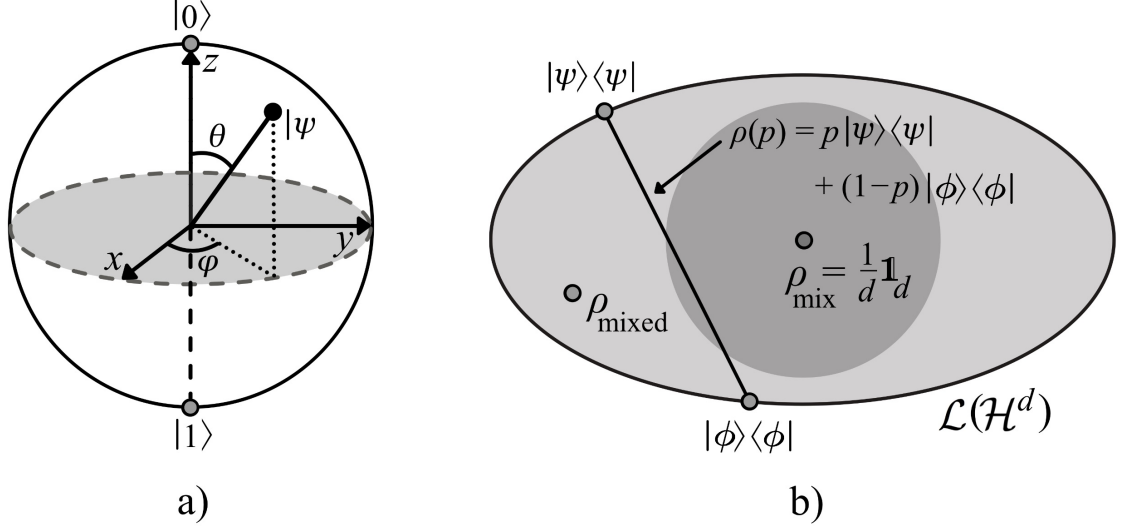


Figure 2.1: (a) An arbitrary state $|\psi\rangle$ represented in the Bloch sphere. (b) Quantum state space geometry. The Hilbert-Schmidt space $\mathcal{L}(\mathcal{H}^d)$ forms a convex subset of the Hilbert $\mathcal{H}$ space. Pure states $|\psi\rangle, |\phi\rangle$ compose the boundary, and mixed states $\rho(p)$ reside inside the set. The solid line represents a mixed state formed by a probabilistic mixture of two pure states. The maximally mixed state $\rho_{\text{mix}} = \frac{1}{d}\mathbb{1}_d$ lies at the center of the set (dark grey circle), where all states of the same purity are equidistant from it.

2.3 Quantum measurements

Predictions about measurement outcomes make up the core of every physical theory. In QIT, quantum states encode all available information about a physical system, including the probabilities of measurement outcomes. Quantum measurements are used to access the classical outcome. Up to this point, the focus has been on quantum states. This section provides a mathematical description of all physically realizable quantum measurements.

2.3.1 Projection-valued measures (PVMs)

In quantum theory, standard measurements are represented by a complete set of mutually orthogonal projection operators [34]. These measurements are therefore called projective and can mathematically be described by a set of measurement operators $\{M_n\}_{n=1,\dots,k}$. This is a complete set of orthogonal projectors, i.e., satisfying [39]

$$M_m M_n = \delta_{mn} M_n \quad \text{orthogonality} \quad (2.6)$$

$$E_n := M_n^\dagger M_n = M_n^2 = M_n \quad \text{projectivity} \quad (2.7)$$

$$\sum_n M_n^\dagger M_n = \sum_{n=1}^k E_n = \mathbb{1} \quad \text{completeness}, \quad (2.8)$$

where E_a are called POVM effects and $n \leq \dim(\mathcal{H})$ is the number of distinct outcomes labeled by $n = 1, 2, \dots, k$.

Assume we are given a quantum state $\rho \in \mathcal{L}(\mathcal{H}^d)$. While the probability of obtaining outcome a is given by the Born rule $p(a) = \text{Tr}(M_a^\dagger M_a \rho) = \text{Tr}(E_a \rho)$, the post-measurement

state is given by the Bayesian Lüders update rule [39]

$$\rho(a) = \frac{M_a \rho M_a^\dagger}{p(a)}. \quad (2.9)$$

This is a conditional state, given that measurement outcome a has been obtained. Furthermore, the factor $1/p(a)$ enforces unit trace of all valid states (normalization).

In the non-degenerate case where $k = \dim(\mathcal{H}^d)$, the set $\{M_n\}$ consists of d number of one-dimensional projectors given by $M_a = |a\rangle \langle a|$ for all a . Moreover, the states $|a\rangle$ form an orthonormal basis of $\mathcal{H}$. This is due to the fact that the projectors are orthogonal and sum to the identity. Hence, the sum of their ranks must be d [7, 42]. If there are already d non-zero outcomes, each one must have rank one. Hence, given a quantum state $\rho \in \mathcal{L}(\mathcal{H}^d)$, the probability of obtaining outcome a reduces to [39]

$$p(a) = \text{Tr}(E_a \rho) = \text{Tr}(M_a \rho) = \text{Tr}(|a\rangle \langle a| \rho) = \langle a | \rho | a \rangle, \quad (2.10)$$

where projectivity ensures $E_a = M_a^2 = M_a$. Physically, rank-one PVMs can be understood as measuring an observable O [39], i.e., a measurement in the eigenbasis $\{|a\rangle\}_a$ of [37]:

$$O = \sum_a \lambda_a |a\rangle \langle a|. \quad (2.11)$$

Consequently, the measurement outcome a indicates that the measured value of the observable is the corresponding eigenvalue λ_a . Moreover, the update rule in Eq. (2.9) simplifies to $\rho(a) = |a\rangle \langle a|$.

2.3.2 Positive operator-valued measures (POVMs)

Open quantum theory requires measurements more general than complete sets of projectors [41]. The fundamental requirement for measurement operators M_a is to ensure reasonable outcome probabilities. That is, the requirement of $p(a) \geq 0$ with $\sum_a p(a) = 1$ is satisfied independently of the initial state ρ for any set of positive operators E_a , given by [39]:

$$\{E_a | E_a \geq 0, \quad \sum_a E_a = \mathbb{1}\}_a, \quad (2.12)$$

where $E_a = M_a^\dagger M_a$. This is called a positive operator-valued measure. It can be noted that this definition is less restrictive than that of PVMs. More precisely, since orthogonal projection matrices are always positive semidefinite, the only difference between the two measurement types is the extra requirement of projectiveness for PVMs. Hence, non-orthogonal POVMs provide a more sensible way of extracting information about incompatible observables [43]. Moreover, $\{E_a\}$ fixes the outcome probabilities via Born's rule $p(a) = \text{Tr}(E_a \rho)$.

As for qubit states, qubit measurement operators can be described by vectors in the Bloch sphere. More precisely, POVM (including PVM) elements are given by the Bloch form

$$E_a = \lambda_a (\mathbb{1} + \vec{n}_a \cdot \vec{\sigma}), \quad (2.13)$$

where $\vec{n}_a$ is the Bloch vector associated to outcome a and $\|\vec{n}_a\| \leq 1$. This comes from the fact that the Pauli matrices $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ constitute a basis of the Hermitian qubit

Hilbert space. Furthermore, the form above satisfies the positivity and normalization conditions [44]

$$\lambda_a > 0, \quad \sum_a \lambda_a = 1, \quad \sum_a \lambda_a \vec{n}_a = 0. \quad (2.14)$$

When the non-vanishing effects of a d -dimensional POVM are linearly independent, the POVM is known as extremal. Extremal POVMs for qubits have at most four elements and are made of rank-one operators. For such POVMs, $\|\vec{n}_a\| = 1$. Since non-extremal POVMs can therefore be written as a convex combination of extremal ones, we will only study extremal POVMs.

Symmetric informationally complete quantum measurements (SIC-POVMs)

On a fixed Hilbert space, there are POVMs that are intrinsically non-projective in the sense that they cannot be represented in terms of projective measurements alone [44]. A seminal class of such measurements is formed by symmetric informationally complete POVMs, or SIC-POVMs [16, 19]. For a d -dimensional system, a SIC-POVM has d^2 outcomes, which is the largest possible number of irreducible outcomes. Each effect, E_a , is rank-one, and the effects are distributed such that all pairwise Hilbert-Schmidt inner products are equal. Equivalently, the d^2 Bloch vectors are equiangular in the Bloch sphere, hence they are symmetric. Furthermore, a POVM is called informationally complete when its measurement statistics are sufficient to uniquely reconstruct the quantum state. [16, 34]

For state reconstruction to be maximally efficient, the POVM effects must be rank-one, i.e., each effect is a positive multiple of a projector onto a pure state [16, 44]. The additional symmetry condition requires the pairwise overlaps between different POVM elements ($\text{Tr}(E_a E_b)$ for all $a \neq b$) to be identical [16]. Hence, the SIC-POVMs can mathematically be constructed in the following way [39]

$$\left\{ E_i = \frac{1}{d} M_i \middle| M_i^2 = M_i, \text{Tr}(M_i M_j) = \frac{d\delta_{ij} + 1}{d+1} \right\}_{i=1}^{d^2}, \quad (2.15)$$

where M_i, M_j for all $i, j \in (1, \dots, d^2)$ are rank-one projectors.

2.4 Quantum channels

The most general quantum evolution of a quantum state is described by a so-called quantum channel (QC). Adopting a black-box view of the quantum physical evolutions, we assume that the experimenter can prepare any state (pure, mixed, or even one share of an entangled state) before the evolution begins [39, 41]. To mathematically define a QC, specific criteria and definitions need to be imposed in order for the channel to describe a valid physical evolution that a quantum state can undergo. First, the input and output of quantum channels are restricted to density operators, since quantum evolutions preserve density operators. Therefore, a quantum channel on a Hilbert space $\mathcal{H}_A$, can be represented by the map

$$\Lambda : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B), \quad \rho \mapsto \Lambda[\rho], \quad (2.16)$$

which brings linear operators in $\mathcal{L}(\mathcal{H}_A)$ to those in $\mathcal{L}(\mathcal{H}_B)$. Second, a map defined by Eq. (2.16) must be linear in order to be defined for all states. Third, for the output states to be

normalized, Λ must also be trace preserving (TP) ($\text{Tr}(\Lambda[\rho]) = \text{Tr}(\rho) = 1$), and to ensure the objects are valid density operators, $\Lambda[\rho]$ must also be positive ($\Lambda[\rho] \succeq 0, \quad \forall \rho \succeq 0$) [40]. The latter criterion, however, is not sufficient when accounting for input states that describe a subsystem of a larger system in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ [41]. Applying a positive map to only one part of a system does not guarantee that the total state remains positive. Instead, we demand complete positivity (CP) [44]:

$$(\Lambda_{A \rightarrow A'} \otimes \mathbb{1}_B)[X_{AB}] \succeq 0, \quad X_{AB} \succeq 0, \quad (2.17)$$

for a reference system B of arbitrary dimension.

The noise in a quantum channel can be interpreted as the loss of a measurement outcome. To this end, consider a third party performing a measurement with operators $\{M_k\}$ on the input state ρ . Hence, the overall evolution is captured in the QC described by

$$\Lambda(\rho) = \sum_k M_k \rho M_k^\dagger, \quad (2.18)$$

where the index k may be interpreted as a measurement outcome that is not revealed, see Fig. 2.2.

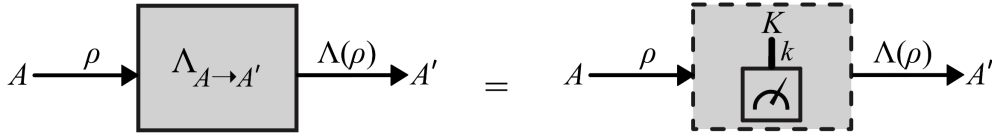


Figure 2.2: The left diagram shows a quantum channel $\Lambda_{A \rightarrow A'}$ taking a state ρ in system A to a state $\Lambda_{A \rightarrow A'}(\rho)$ in system A' . The diagram on the right shows an interpretation of the channel in which a third party performs a measurement on the input system without informing the receiver of the measurement outcome k .

2.5 Representations of quantum channels

2.5.1 Choi representation

Consider two isomorphic Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_R$ with orthonormal bases $\{|i\rangle_A\}$ and $\{|i\rangle_R\}$, respectively. More precisely, $\mathcal{H}_A$ denotes the input system and $\mathcal{H}_R$ is a reference system. Let $\Lambda : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_{A'})$ be a linear map, where $\mathcal{H}_{A'}$ is another Hilbert space. The reference system is a copy of the input system A , which is not acted upon by the channel. Then, the Choi operator, $\eta_{A'A}$,¹ corresponding to the map after relabeling² R as A is [41]:

$$\eta_{A'A} = (\Lambda_{A \rightarrow A'} \otimes \mathbb{1}_A)[\phi_{AA}^+] = \sum_{i,j=0}^{d_A-1} \Lambda_{A \rightarrow A'}(|i\rangle \langle j|_A) \otimes |i\rangle \langle j|_A, \quad (2.19)$$

where AA means the input system A and a reference copy of A , and, ϕ^+ is the maximally entangled state of local dimension $d_A \equiv \dim(\mathcal{H}_A)$.

¹The convention is to put the output system first, $(A' \otimes A)$, where A' is the output system and A the input system.

²Without using the convention of renaming R as A , the Choi operator would be written as $\eta_{A'R} = (\Lambda_{A \rightarrow A'} \otimes \mathbb{1}_R)[\phi_{AR}^+] = \sum_{i,j=0}^{d_A-1} \Lambda_{A \rightarrow A'}(|i\rangle \langle j|_A) \otimes |i\rangle \langle j|_R$.

The action of a quantum channel on a state $\rho \in \mathcal{L}_+(\mathcal{H}_A)$ of dimension d has a so-called Choi state representation given by [44]:

$$\Lambda_{A \rightarrow A'}(\rho) = d \operatorname{Tr}_A \{ (\mathbb{1}_{A'} \otimes \rho_A^T) \eta_{A'A} \}, \quad (2.20)$$

where $\eta \in \mathcal{L}(\mathcal{H}_{A'} \otimes \mathcal{H}_A)$. Therefore, the Choi operator represents a quantum operation Λ as an unnormalized quantum state of a larger system [40]. The reason one can associate a state given by Eq. (2.20) to the channel in Eq. (2.19) is that there is a one-to-one correspondence between CP maps from $\mathcal{H}_A$ to $\mathcal{H}_R$ and PSD states on $\mathcal{H}_A \otimes \mathcal{H}_R$, referred to as Choi-Jamiołkowski isomorphism [39]. The channel-state duality holds only for Choi states, which constitute a subset of $\mathcal{L}_+(\mathcal{H}^\Gamma)$. The defining properties of a Choi state are:

$$\eta \succeq 0, \quad \operatorname{Tr}(\eta) = 1, \quad \operatorname{Tr}_{A'}(\eta_{A'A}) = \frac{\mathbb{1}}{d}, \quad (2.21)$$

where the last property is unique for Choi states. The Choi state further implies the properties: (i) Λ is CP $\iff \eta \succeq 0$, and (ii) Λ is TP $\iff \operatorname{Tr}_{A'}(\eta_{A'A}) = \frac{\mathbb{1}}{d}$ [40, 41]. Hence, Choi operators provide a method for determining whether a map is CPTP.

2.5.2 Kraus operators

Another way to represent a channel is via the Kraus representation. A map $\Lambda_{A \rightarrow A'}$ is a QC if and only if it admits a Kraus decomposition [39]:

$$\Lambda_{A \rightarrow A'}(\rho) = \sum_a K_a \rho K_a^\dagger, \quad \text{s.t.} \quad \sum_a K_a K_a^\dagger = \mathbb{1}_A, \quad (2.22)$$

where K_a are Kraus operators.

Given that an operator Λ possesses a Kraus representation defined in Eq. (2.22), the corresponding Choi representation can be constructed directly from the Kraus operators [7, 40]

$$\eta = (\Lambda \otimes \mathbb{1})[\phi^+] = \sum_{i=1}^d (K_i \otimes \mathbb{1}) \phi^+ (K_i^\dagger \otimes \mathbb{1}). \quad (2.23)$$

2.6 Quantum instruments

As mentioned above, a quantum measurement can be viewed as a channel that maps a quantum state to a classical state. This definition can be extended to a more general object, a quantum instrument. Such objects consider both the classical outcome and the post-measurement quantum state. Consequently, it provides the most general evolution, transforming an input quantum system into a composite quantum-classical state. Hence, in a scenario where Alice wants to transmit both classical and quantum information, Bob can decode both using a QI. [40, 41]

A QI is composed of a set of completely positive trace-non-increasing maps $\{\Lambda_j\}_j$ such that the sum map $\sum_j \Lambda_j$ is trace-preserving. Trace non-increasing maps are defined by the following property: $\operatorname{Tr}\{\mathcal{M}(X)\} \leq \operatorname{Tr}\{X\}$ for all PSD $X \in \mathcal{L}(\mathcal{H})$. The action of a QI on a

density operator $\rho \in \mathcal{D}(\mathcal{H})$ is described by a quantum channel featuring both a quantum and classical output [41]:

$$\Lambda_{A \rightarrow JA'} : \quad \rho \rightarrow \sum_j \Lambda_j(\rho) \otimes |j\rangle \langle j|_J, \quad \rho \mapsto \{[\rho_j, p(j|\rho)]\} \quad (2.24)$$

where $\{|j\rangle\}$ is an orthonormal basis for the Hilbert space $\mathcal{H}_J$ and j indexes the classical outcome.

As noted earlier, the noise from a quantum channel can be interpreted as the loss of a measurement outcome. To this end, consider now a third party performs a two-outcome measurement with outcomes j and k , and measurement operators $\{M_{j,k}\}_{j,k}$. Suppose the measurement is performed on a system with density operator ρ and both outcomes are made accessible by the third party. The Kraus map for the overall quantum evolution is given by

$$\sum_{j,k} M_{j,k} \rho M_{j,k}^\dagger \otimes |j\rangle \langle j|_J \otimes |k\rangle \langle k|_K, \quad (2.25)$$

where $M_{j,k}$ play the role of Kraus operators and $\{|j\rangle, |k\rangle\}$ are orthonormal bases. Suppose that only outcome j is accessible. The state is then obtained by tracing out the classical register K , leading to the following Kraus map:

$$\sum_{j,k} M_{j,k} \rho M_{j,k}^\dagger \otimes |j\rangle \langle j|_J. \quad (2.26)$$

This map corresponds to a QI and describes a general noisy quantum evolution producing both a quantum and a classical output, see Fig. 2.3. Comparing Eq. (2.24) to Eq. (2.26), each j -dependent map Λ_j can be defined as $\Lambda_j(\rho) \equiv \sum_k M_{j,k} \rho M_{j,k}^\dagger$. This is a CP trace-non-increasing map, since

$$\text{Tr} [\Lambda_j(\rho)] = p_J(j) \leq 1, \quad (2.27)$$

where $p_J(j)$ is a marginal probability, which must sum to unity. To see this, consider

$$\text{Tr} [\Lambda_j(\rho)] = \sum_k \text{Tr} (M_{j,k} \rho M_{j,k}^\dagger) = \sum_k \text{Tr} (M_{j,k}^\dagger M_{j,k} \rho) = \text{Tr}(E_j \rho) = p_J(j), \quad (2.28)$$

where in the first equality, cyclicity of trace was used, and, in the penultimate equality, the POVM effect $E_j = \sum_k M_{j,k}^\dagger M_{j,k}$ was defined. The last equality used Born's rule $p_J(j) = \text{Tr}(E_j \rho)$.

Then, the flagged (outcome-labeled) instrument branch shown in Fig. 2.3 is

$$\mathcal{I}_j(\rho) := \Lambda_j(\rho) \otimes |j\rangle \langle j|_J. \quad (2.29)$$

The output quantum state for this branch, $\Lambda_j(\rho)$, is sub-normalized and given by

$$\Lambda_j(\rho) = p_J(j) \frac{\Lambda_j(\rho)}{\text{Tr}(\Lambda_j(\rho))}, \quad (2.30)$$

where $\frac{\Lambda_j(\rho)}{\text{Tr}(\Lambda_j(\rho))}$ is the actual post-measurement state. The quantum output of the full QI is retrieved by tracing over the classical register J :

$$\text{Tr}_J \left\{ \sum_j \Lambda_j(\rho) \otimes |j\rangle \langle j|_J \right\} = \sum_j \Lambda_j(\rho). \quad (2.31)$$

This sum map is trace-preserving:

$$\text{Tr} \left\{ \sum_j \Lambda_j(\rho) \right\} = \sum_j \text{Tr} \{ \Lambda_j(\rho) \} = \sum_j p_J(j) = 1. \quad (2.32)$$

The last equality follows from the fact that marginal probabilities $p_J(j)$ must sum to one. Hence, the sum map $\sum_j \Lambda_j$ is trace-preserving.

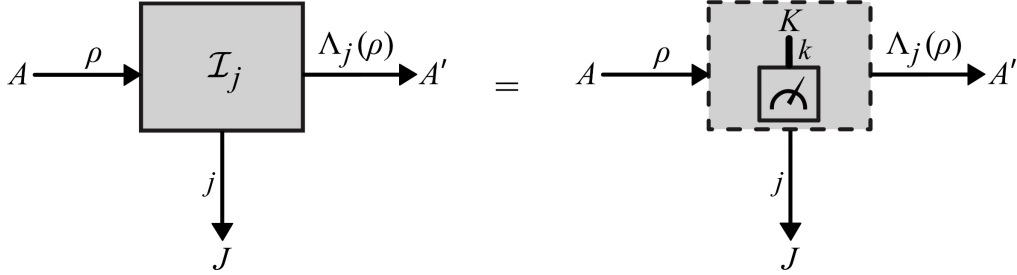


Figure 2.3: The diagram on the left shows a QI, which transforms an incoming state ρ into both a classical outcome j and an updated quantum state $\Lambda_j(\rho)$. The diagram to the right displays the internal workings of the QI, indicating only partial access to the measurement outcomes.

A special form of an instrument is the so-called Lüders instrument form, given by [45]

$$I_a(\rho) = \sqrt{E_a} \rho \sqrt{E_a}. \quad (2.33)$$

Not all QIs can be reduced to this form, but if they can, it is very useful and insightful, as it gives an explicit relationship between an instrument and its corresponding measurement. This form will be used to analyze results in later chapters.

The POVM $\{E_a\}_a$ that is realized by the instrument is obtained from its Choi state [7]:

$$E_a = d \text{Tr}_{A'}(\eta_a)^T. \quad (2.34)$$

The possible associated post-measurement states given by Eq. (2.9) are not unique. To see this, we note that since $E_a = M_a^\dagger M_a$ is PSD, it is diagonalizable [42, 46, 47]. Consequently, there exists a unitary U_a such that $E_a = U_a D_a U_a^\dagger$ where D_a is a diagonal matrix. Hence, one can directly build at least one measurement operator M_a via $\sqrt{E_a}$ [39]:

$$M_a := \sqrt{E_a} = U_a \sqrt{D_a} U_a^\dagger. \quad (2.35)$$

On the other hand, the M_a operators defined by $M_a := V_a \sqrt{E_a}$ also satisfies $M_a^\dagger M_a = \sqrt{E_a} V_a^\dagger V_a \sqrt{E_a} = E_a$, for any set of unitaries $\{V_a\}_a$ with $V_a^\dagger V_a = \mathbb{1}$. Therefore, the reconstruction in Eq. (2.35) is not unique. Hence, while the probabilities $p(a)$ for any outcome a are the same for any realization of a given POVM, the post-measurement states associated with a fixed POVM are not uniquely determined by the POVM as they can differ depending on the choice of M_a in Eq. (2.9) [39]. This means that there is a many-to-one correspondence between QIs and QMs. Instruments corresponding to the same measurement are not necessarily equal as they can have different post-measurement states and two measuring processes are statistically equivalent if they define the same instrument [48].

2.7 Semidefinite programs

For many purposes in QIT and as the building block of this thesis, a semidefinite program optimizes a real linear function, called the objective function, over a convex set of Hermitian matrices subjected to linear constraints. Quantum states are semipositive operators subjected to the linear constraints $\rho \geq 0$ and $\text{Tr}(\rho) = 1$, making semidefinite programming a powerful tool in optimization problems in quantum information science. [49], [50]

For an operator variable X , the primal SDP formulation is given by [49]:

$$\begin{aligned} \max_X \quad & \text{Tr}(AX) \\ \text{subject to} \quad & \Phi_i(X) = B_i, \quad \forall i \\ & \Gamma_j(X) \leq C_j, \quad \forall j, \end{aligned} \tag{2.36}$$

where A is a Hermitian operator, $\Phi_i(\cdot) = B_i$ and $\Gamma_j(\cdot)$ are Hermiticity-preserving linear maps, and B_i and C_j are Hermitian operators of arbitrary finite dimension. The corresponding so-called dual SDP formulation is [49]:

$$\begin{aligned} \min_{Y_i, Z_j} \quad & \sum_{i=1}^m \text{Tr}(Y_i B_i) + \sum_{j=1}^n \text{Tr}(Z_j C_j) \\ \text{subject to} \quad & A - \sum_{i=1}^m \Phi_i^\dagger(Y_i) - \sum_{j=1}^n \Gamma_j^\dagger(Z_j) = 0, \\ & Z_j \geq 0 \quad \forall j, \end{aligned} \tag{2.37}$$

where the dual variables Y_i and Z_j are arbitrary but Hermitian [50].

Any dual feasible point gives the tightest upper bound to the primal optimal value while every primal feasible point provides the tightest lower bound of the dual optimal value. This relation is always satisfied and is called weak duality. In addition, when the values of the primal and dual SDP coincide, i.e., the bounds of the SDP are optimal, strong duality holds.

Variables that arise when switching from a primal to a dual problem, which are not present in the objective function are called slack variables. Such variables can be removed from the LP, which has the effect of converting equality constraints into inequality constraints [49].

Chapter 3

Projective simulability of quantum measurements

Building on the framework of projective simulability of POVMs developed in [34], this chapter presents the protocol that defines PM-simulability and a necessary condition for it. Together with the critical visibility as a notion for quantifying projectiveness, an SDP can be defined. The SDP will later be used to express the critical visibility as a function of the angle parameters that characterize two highly symmetric rank-one POVM models: a three-outcome POVM and a four-outcome POVM.

3.1 Protocol

Consider that an experimenter has access to three resources: (i) classical randomness, represented by a stochastic variable λ subject to a probability distribution $\{q_\lambda\}_\lambda$. (ii) A set of PVMs, $\{P_{a'|\lambda}\}_{a'=1}^n \in \mathcal{H}^d$, where a' denotes the measurement outcome. Here, the variable λ is interpreted as the selection of measurement. (iii) Classical post-processing of a' and λ , through a probabilistic rule $p(a|a', \lambda)$, generating the final output a . An n -outcome POVM $E = \{E_a\}_a^n \in \mathcal{H}^d$ is then PM-simulable if it can be realized by this restricted class of resources, in a way shown in Fig. 3.1 below.

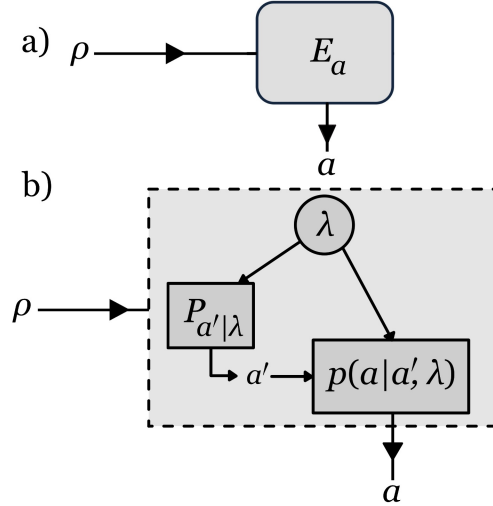


Figure 3.1: (a) A general POVM transforming an input state ρ to a classical output a . (b) The internal workings of a PM-simulable POVM which uses the classical random variable λ to select a PVM, $\{P_{a'|\lambda}\}$, yielding an internal outcome a' . Lastly, a classical post-processing rule $p(a|a', \lambda)$ is applied to a' to determine the classical output.

This means that the POVM effect E_a can be modeled by [34]:

$$E_a = \sum_{\lambda} q_{\lambda} \sum_{a'} p(a|a', \lambda) P_{a'|\lambda} = \sum_{\lambda} q_{\lambda} P_{a|\lambda}. \quad (3.1)$$

In the last equality, post-processing was neglected without loss of generality, as illustrated in Fig. 3.2 (see Appendix B for proof of why post-processing can be neglected).

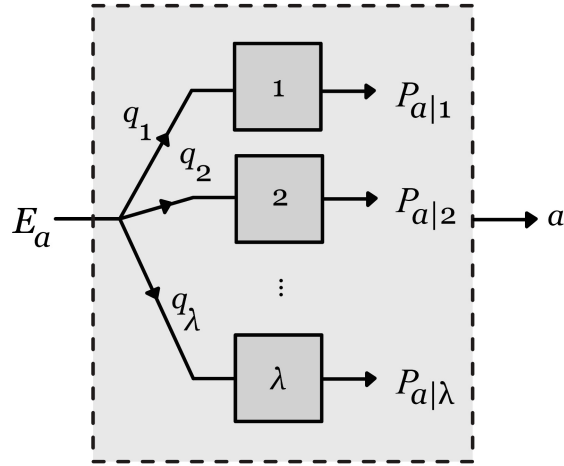


Figure 3.2: No post-processing.

3.2 Quantification of projective simulability

Let $\mathcal{P}_E$ denote the set of PM-simulable POVMs, i.e., POVMs given by Eq. (3.1). Further let $\mathbf{E} = \{E_a\}_{a=1}^n$ be a POVM that is not projectively simulable ($\mathbf{E} \notin \mathcal{P}_E$). Then, the extent of non-projectivity of a POVM can be quantified by estimating the amount of noise a measurement can tolerate before it becomes PM-simulable. More precisely, by mixing a

POVM $\mathbf{E} \notin \mathcal{P}_E$ with noise so that it becomes PM-simulable. This is represented by the map Φ_v , given by

$$\Phi_v(E_a) := vE_a + (1-v)E_a^{\text{noise}}, \quad \forall a, \quad (3.2)$$

where $v \in [0, 1]$. The output state $\Phi_v(\mathbf{E})$ defined above can be viewed as the noisy implementation of the original POVM $\mathbf{E}$, in which every effect E_a is mixed with an arbitrary noise measurement¹ $\{E_a^{\text{noise}}\}_a$. Since the noise is a POVM itself, it must fulfill the same constraints as a valid POVM does. The factor $1-v$ can be viewed as the noise strength of Φ_v .

From Eq. (3.2), we can define the critical (maximal) visibility for which $\Phi_v(\mathbf{E})$ is still PM-simulable [30]:

$$v^*(\mathbf{E}) := \max \{v | \{\Phi_v(E_a)\}_a \in \mathcal{P}(d, n)\}. \quad (3.3)$$

Therefore, $\Phi_v(\mathbf{E})$ given by Eq. (3.2) is PM-simulable for all $v \leq v^*(\mathbf{E})$ and loses its simulability for $v > v^*(\mathbf{E})$. This means that for a small $v^*(\mathbf{E})$ (i.e., large $1 - v^*(\mathbf{E})$), it is difficult to simulate $\mathbf{E}$ projectively using Φ_v , since a large amount of noise must be added to each of E_a before $\Phi_v(\mathbf{E})$ becomes PM-simulable. Meanwhile, a large $v^*(\mathbf{E})$ makes it easier to simulate. Hence, the maximum visibility is sought, as any non-simulable POVM $\mathbf{E}$ is guaranteed to become simulable after acting with Φ_v on every E_a with $v \leq v^*(\mathbf{E})$.

Since PM-simulable POVMs can be reproduced using only PVMs and classical randomness (see Eq. (3.1)), they form a subset of all POVMs, see Fig. 3.3. Fig. 3.3 further indicates that POVMs on the boundary of the set are far from being projectively simulable, while other POVMs lie closer to the simulable set. Naturally, this means that more noise needs to be added to the POVMs farthest away from the PM-simulable set, compared to those lying closer to the set.

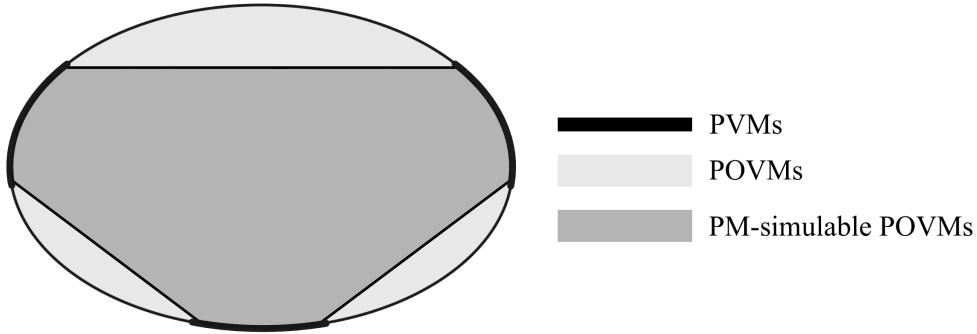


Figure 3.3: schematic illustration of the geometry of POVMs, the most general class of QMs, for a given dimension. Solid lines represent extremal projective measurements. Their convex hull (shaded in dark gray) represents the set of PM-simulable measurements.

Furthermore, determining which POVM most strongly defies a projective simulation corresponds to finding the specific $\mathbf{E}$ that solves:

$$v_d \equiv \min_{\mathbf{E}} \{v^*(\mathbf{E}) | \mathbf{E} \in \mathcal{P}_E\}, \quad (3.4)$$

where v_d is the visibility associated with the most non-projective measurement in $\mathcal{H}^d$. The minimization is over the set of d^2 -output measurements on $\mathcal{H}^d$ and is justified by

¹For the qubit case: $E_a^{\text{noise}} = \lambda_a^{\text{noise}}(\mathbb{1} + \vec{n}_a^{\text{noise}} \cdot \vec{\sigma})$.

the concavity of the function $\mathbf{E} \rightarrow v(\mathbf{E})$, causing the minimum to be achieved for some extremal POVM, which at most has d^2 outcomes [44, 51], see Appendix A. Hence, for the purpose of minimizing $v(E)$, it is enough to search among extremal POVMs.

A natural noise model is given by the depolarizing channel [30]:

$$\Phi_v(E_a) = vE_a + (1-v)\frac{\text{tr}(E_a)}{d}\mathbb{1}, \quad (3.5)$$

where the POVM $\{E_a^{\text{noise}}\}_{a=1}^n = \{\text{Tr}(E_a)\frac{\mathbb{1}}{d}\}_{a=1}^n$ is called white noise. This noise can be realized by mixing deterministic dprojective measurements $(\mathbb{1}, 0, \dots, 0), \dots, (0, \dots, 0, \mathbb{1})$. If $v = 1$, the original POVM $\mathbf{E}$ is kept, while if $v = 0$, $\mathbf{E}$ is completely discarded. The white noise $\left\{\frac{\text{Tr}(E_a)}{d}\mathbb{1}\right\}_{a=1}^n$ can be interpreted as a completely random measurement as it carries no state-dependent information. The depolarizing map is illustrated in Fig. 3.4 where it becomes evident that the smaller the v , the larger the factor $1-v$ must be to compensate in order to maintain PM-simulability.

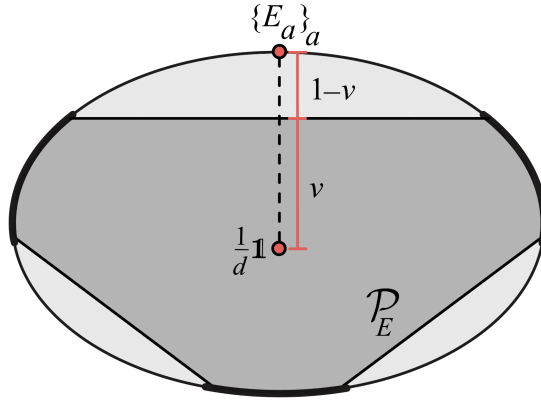


Figure 3.4: Geometric illustration of the depolarizing channel quantifying the projectiveness of a POVM $\{E_a\}_{a=1}^n$.

Other interesting noise models include the so-called antipodal noise and worst-case noise. For the former, $\{E_a^{\text{noise}}\}_a$ in Eq. (3.2) corresponds to a measurement whose Bloch vectors always oppose the Bloch vectors of $\{E_a\}_{a=1}^n$. The latter is the noise most effective at destroying non-projectiveness. Compared with the other two noise types mentioned, the worst-case noise is a variable. Analogous to the depolarizing noise, the following noise map can be defined as

$$\tilde{\Phi}_v(X, X_{\text{noise}}) = vX + (1-v)X_{\text{noise}}, \quad (3.6)$$

with the corresponding worst-case critical visibility

$$\tilde{v}^*(\mathbf{E}) = \max_{\mathbf{E}_{\text{noise}}} \{v | \tilde{\Phi}_v(E_a, E_{\text{noise}}^a)\}_a \in \mathcal{P}_E, \quad (3.7)$$

which is additionally maximized over all valid POVMs² $\{E_{\text{noise}}^a\}_{a=1}^n$. For the special case of the fixed white noise, $E_{\text{noise}}^a = \text{Tr}(E_a)\mathbb{1}/2$, it can be seen that $v^*(\mathbf{E}) \leq \tilde{v}^*(\mathbf{E})$. Hence, the worst-case noise is optimized such that the amount of it required before $\Phi_v(E_a)$ becomes PM-simulable is smaller than for any other noise type. The noise types will be discussed in more detail in sections 3.4 and 3.5.

²That have the same number of outcomes as $\mathbf{E}$ so that each POVM element is affected by the noise.

3.3 Condition for simulability of POVMs

Now we turn to the question of how we can determine if a given POVM $\mathbf{E}$ is projective-simulable. While a full solution to the problem of characterizing the entire set of PM-simulable POVMs is unknown, it is possible to identify a necessary condition in the form of an SDP [34]. The main idea is to partition the space of PVMs over $\mathcal{H}^d$ by the ranks of the measurement operators. To this end, consider a vector $\vec{r} = (r_1, \dots, r_n)$, where r_a are non-negative integers satisfying $\sum_{a=1}^n r_a = d$. To each such vector we can associate a distinct class of PVMs, $\{P_a\}_a$, where $r_a = \text{rank}(P_a) \geq 0$, such that the union of all such classes forms the set of all PVMs. Hence, $\vec{r}$ is called a rank-vector, which can be used for indexing the partitioning of the space of PVMs. This enables a separate treatment of the measurements associated to different $\vec{r}$ -vectors. Consequently, the variable λ in Eq. (3.1) can be substituted by a pair of variables $(\vec{r}, \chi)$, where $\vec{r}$ gives the rank-vector and χ is a random variable serving as an index specifying the PVM within that $\vec{r}$ -specific class.

With all this, Eq. (3.1) becomes:

$$E_a = \sum_{\vec{r}, \chi} q(\vec{r}, \chi) P_{a|\vec{r}, \chi} = \sum_{\vec{r}} \left(\sum_{\chi} q_{\vec{r}, \chi} P_{a|\vec{r}, \chi} \right) = \sum_{\vec{r}} F_{a|\vec{r}}, \quad (3.8)$$

where $F_{a|\vec{r}} \succeq 0$ by construction. In addition, from conditional probability, it follows that $q_{\vec{r}, \chi} = q_{\vec{r}} q_{\chi|\vec{r}}$. Therefore, we have

$$\sum_a F_{a|\vec{r}} = \sum_a \sum_{\chi} q_{\vec{r}, \chi} P_{a|\vec{r}, \chi} = q_{\vec{r}} \sum_a \sum_{\chi} q_{\chi|\vec{r}} P_{a|\vec{r}, \chi} = q_{\vec{r}} \sum_a P_{a|\vec{r}, \chi} = q_{\vec{r}} \mathbb{1}, \quad (3.9)$$

where measurement completeness was used in the last equality. Furthermore, since projectors have eigenvalues that are either 0 or 1 [42], the trace gives the rank for a projector, $\text{Tr}(P_{a|\vec{r}, \chi}) = r_a$. Then, it follows that

$$\text{Tr}(F_{a|\vec{r}}) = \sum_{\chi} q_{\vec{r}, \chi} \text{Tr}(P_{a|\vec{r}, \chi}) = q_{\vec{r}} r_a, \quad (3.10)$$

where, again, $q_{\vec{r}, \chi} = q_{\vec{r}} q_{\chi|\vec{r}}$. The equation above implies that the rank-vector distribution is given by $q_{\vec{r}} = \frac{1}{d} \sum_a \text{Tr}(F_{a|\vec{r}})$.

Putting this together, the necessary condition for $E \in \mathcal{P}$ is given by the following SDP [7]:

$$\begin{aligned} & \text{find} \quad \{F_{a|\vec{r}}\}_{a, \vec{r}} \\ & \text{s.t.} \quad \sum_{\vec{r}} F_{a|\vec{r}} = E_a, \quad \text{Tr}(F_{a|\vec{r}}) = q_{\vec{r}} r_a \\ & \quad \sum_a F_{a|\vec{r}} = q_{\vec{r}} \mathbb{1}, \quad \text{and} \quad F_{a|\vec{r}} \succeq 0. \end{aligned} \quad (3.11)$$

The necessary condition for PM-simulation defined in the primal SDP above is also sufficient in the case of qubits [34]. This is justified by the fact that PVMs cannot have more than d non-zero outcomes. All extremal dichotomic measurements are projective [52], and d -outcome measurements constitute a strict superset of measurements through Eq.

(3.11), implying the criterion is exact for qubits.

The SDP in Eq. (3.11) can be adapted to compute upper bounds on the critical visibility $v^*(E)$. For this purpose, E_a is substituted with its noisy implementation $\Phi_v(E_a)$, and the objective is to maximize the visibility v (see Eq. (3.3)) (assuming fixed noise):

$$\begin{aligned}
& \max_{v, \{F_{a|\vec{r}}\}} v \\
& \text{s.t.} \quad vE_a + (1-v)E_a^{\text{noise}} = \sum_{\vec{r}} F_{a|\vec{r}}, \\
& \quad \sum_a F_{a|\vec{r}} = q(\vec{r})\mathbb{1}, \quad \forall \vec{r}, \\
& \quad \text{Tr}\{F_{a|\vec{r}}\} = q(\vec{r})r_a, \quad \forall a, \vec{r} \\
& \quad F_{a|\vec{r}} \succeq 0.
\end{aligned} \tag{3.12}$$

The goal of the primal SDP above is to compute the largest possible visibility $v \in [0, 1]$ for which a projective simulation is tangible.

3.4 Three-outcome POVM

In this section, a three-outcome qubit POVM parameterized by an angle θ , see Fig. 3.5, is treated. The POVM has the form (see Eq. (2.13))

$$E_a = \lambda_a(\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma}), \quad a = 1, 2, 3, \tag{3.13}$$

and defined such that the $\lambda_2 = \lambda_3$. Given the POVM in Fig. 3.5, we impose the SDP formulation given by Eq. (3.12) for white noise, antipodal noise, and worst-case noise. The SDP is run for all angles in the range $(\pi/2, \pi)$, allowing for the visibility to be plotted as a function of the parameter θ , both numerically and analytically.

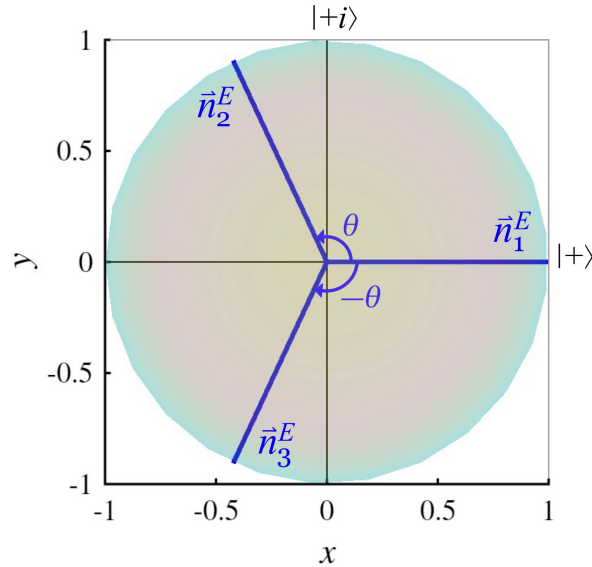


Figure 3.5: Trine POVM model Bloch vectors symmetrically separated by an angle parameter $\theta \in (\pi/2, \pi)$ plotted in the xy -plane of the Bloch sphere. The blue lines represent the Bloch vectors $\vec{n}_a^E$ and point outward from the origin.

3.4.1 White noise

A natural noise model used to quantify the magnitude of non-projectivity is through applying the depolarizing channel given by Eq. (3.5). This corresponds to replacing E_a^{noise} in Eq. (3.12) with the white noise measurement $\text{Tr}(E_a)\frac{1}{d}$. Physically, it can be interpreted as adding a random measurement to the POVM. The closed-form of the white-noise visibility as a function of the angle θ is given by:

$$v_{\text{white}}(\theta) = \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \quad (3.14)$$

where

$$\begin{aligned} A &:= (\lambda_1^F)^2 + (\lambda_2^E \sin(\theta))^2 \\ B &:= 2\lambda_2^E \sin(\theta)\lambda_2^F \\ C &:= (\lambda_2^F)^2 - (\lambda_1^F)^2, \end{aligned} \quad (3.15)$$

and

$$\lambda_1^F = -\frac{1}{2} \left(\frac{\cos(\theta)}{1 - \cos(\theta)} \right), \quad \lambda_2^E = \frac{1}{2} \left(\frac{1}{1 - \cos(\theta)} \right), \quad \lambda_2^F = \frac{1}{2} \left(\frac{1 + \cos(\theta)}{1 - \cos(\theta)} \right). \quad (3.16)$$

The $v_{\text{white}}(\theta)$ -curve is plotted in Fig. 3.6. The minimum critical visibility is analytically given by $v_{\text{white}}^{\min}(\theta) = \sqrt{3}/2$ and numerically from the SDP by 0.8660. It was verified both numerically and analytically that $v_{\text{white}}^{\min}$ corresponds to $\theta = 120^\circ$. For the derivation, see Appendix C. The POVM with $\theta = 120^\circ$ is therefore the most non-projective one for the white-noise model in the POVM family studied here.

3.4.2 Antipodal noise

The Bloch vectors of the antipodal noise POVM point in the opposite direction of the corresponding POVM Bloch vectors given in Fig. 3.5. As a result, the left-hand side of the objective function in the SDP (Eq. (3.12)) can be written as

$$v_{\text{ap}} E_a + (1 - v_{\text{ap}}) E_a^\perp = v_{\text{ap}} \lambda_a^E (\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma}) + (1 - v_{\text{ap}}) \lambda_a^E (\mathbb{1} - \vec{n}_a^E \cdot \vec{\sigma}) = \lambda_a^E (\mathbb{1} + (2v_{\text{ap}} - 1) \vec{n}_a^E \cdot \vec{\sigma}), \quad (3.17)$$

where the Bloch representation and $E_a^\perp = 2\mathbb{1} - E_a$ were used in the second equality. Comparing the last form of Eq. (3.17) with the corresponding expression for white noise (Eq. (3.5)) implies that

$$\lambda_a^E (\mathbb{1} + (2v_{\text{ap}}(\theta) - 1) \vec{n}_a^E \cdot \vec{\sigma}) = \lambda_a^E (\mathbb{1} + v_{\text{white}}(\theta) \vec{n}_a^E \cdot \vec{\sigma}) \iff v_{\text{ap}} = \frac{1 + v_{\text{white}}(\theta)}{2}, \quad (3.18)$$

which is plotted in Fig. 3.6. An alternative derivation of v_{ap} is given in Appendix D. Moreover, the smallest critical visibility, which corresponds to $\theta = 120^\circ$, is given by $v_{\text{ap}}^{\min}(\theta) = (1 + v_{\text{white}}^{\min}(\theta))/2 = (\sqrt{3} + 2)/4 \approx 0.9330$. Hence, the symmetric POVM with $\theta = 120^\circ$ is the most non-projective POVM in the class studied for all three noise types.

3.4.3 Worst-case noise

The worst-case noise uses a variable noise, N_a . The objective function in Eq. (3.12) is given by

$$E_a v + (1 - v) N_a = \sum_{\vec{r}} F_{a|\vec{r}}. \quad (3.19)$$

This means that the noise is optimized together with the critical visibility. To avoid quadratic terms such as vN_a , we define $\tilde{N}_a = (1 - v)N_a$ so that the equation above becomes

$$E_a v + \tilde{N}_a = \sum_{\vec{r}} F_{a|\vec{r}}, \quad (3.20)$$

where $\tilde{N}_a$ fulfills the constraints

$$\frac{1}{d} \sum_a \text{tr}(\tilde{N}_a) = (1 - v), \quad \sum_a \tilde{N}_a = (1 - v) \mathbb{1}. \quad (3.21)$$

In order for the noise N_a to be a valid POVM, it must satisfy the following properties

$$\sum_a N_a = \mathbb{1}, \quad \frac{1}{d} \sum_a \text{tr}(N_a) = \sum_a \lambda_a = \frac{1}{d} \text{tr}(\mathbb{1}) = 1. \quad (3.22)$$

Hence, additional constraints in Eq. (3.21) must be included to the SDP in Eq. (3.12), which is now given by

$$\begin{aligned} & \max_{v, \{F_{a|\vec{r}}\}, \{\tilde{N}_a\}} v \\ & \text{s.t.} \quad vE_a + \tilde{N}_a = \sum_{\vec{r}} F_{a|\vec{r}}, \quad \forall a, \\ & \quad F_{a|\vec{r}} \succeq 0, \quad \forall a, \vec{r}, \\ & \quad \text{tr}(F_{a|\vec{r}}) = q_{\vec{r}} r_a, \quad \forall a, \vec{r}, \\ & \quad \sum_a F_{a|\vec{r}} = q_{\vec{r}} \mathbb{1}, \quad \forall \vec{r}, \\ & \quad \tilde{N}_a \succeq 0, \quad \forall a, \\ & \quad \tilde{N}_a = (1 - v) \mathbb{1}. \end{aligned} \quad (3.23)$$

The worst-case noise visibility is given indirectly by an auxiliary parameter $t \in (0, 1)$ (see Appendix E for derivation)

$$t_*(x) = \begin{cases} -\frac{(x+1) + 2\sqrt{\Delta_0} \cos\left(\frac{\Phi - 2\pi}{3}\right)}{3(3x-1)}, & 0 < x < \frac{1}{3}, \\ \frac{\sqrt{17}-1}{8}, & x = \frac{1}{3}, \\ -\frac{(x+1) + 2\sqrt{\Delta_0} \cos\left(\frac{\Phi + 2\pi}{3}\right)}{3(3x-1)}, & \frac{1}{3} < x < 1, \end{cases} \quad (3.24)$$

where $\Phi = \arccos\left(\frac{\Delta_1}{2\Delta_0^{3/2}}\right)$ and $\Delta_1 := -19x^2 + 13x - 4$ and $\Delta_0 = -187x^3 + 177x^2 - 57x + 11$, with $x := \cot(\theta/2)$. The closed form of the visibility is

$$v_{\text{wc}}(\theta) = \frac{(1 + 2t_*)^2(1 - t_*^2)}{(1 + 2t_*)^2(1 - t_*^2) + t_*^2(1 - x^2)}, \quad x = \cot \frac{\theta}{2}. \quad (3.25)$$

The minimum visibility is analytically given by $v_{\text{wc}}^{\min}(\theta) = (\sqrt{3} + 2)/4 \approx 0.9330$, with $\theta = 120^\circ$.

The curve is plotted in Fig. 3.6, which also indicates that the v_{wc} -curve lies above all other curves. Since the worst-case noise is most effective at destroying non-projectiveness, the highest critical visibilities are expected for this noise. The worst-case noise follows closely the curve of the antipodal noise. This is due to the fact that the POVM, which affects the measurement most strongly, has Bloch vectors pointing in opposite directions in the qubit case. It was also observed from the SDP that the optimal worst-case noise had Bloch vectors pointing in directions very similar to those of the antipodal noise. The shape of the curves in Fig. 3.6 highlights that measurements with either small or large θ -values are easy to simulate projectively.

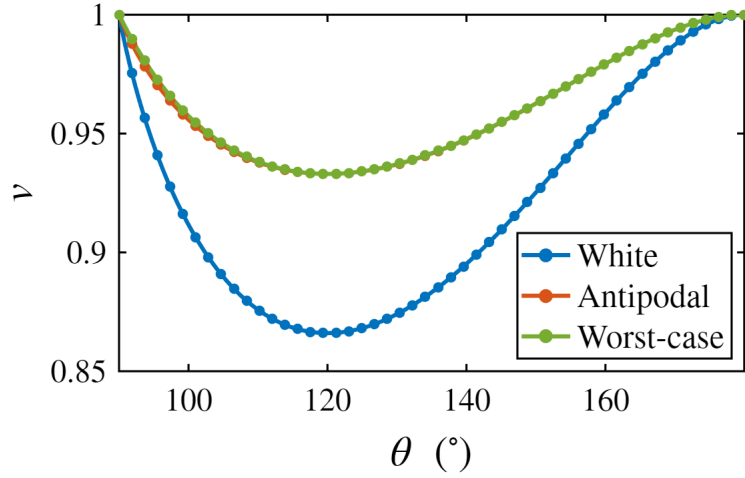


Figure 3.6: Visibility v as a function of $\theta \in (90, 180)$ (°) for the three noise models obtained numerically from the SDP (dotted) and analytically (solid line).

3.5 Four-outcome equal-weight POVM

In this section, the four-outcome equal-weight ($\lambda_a = \frac{1}{4}$ for all a) POVM model shown in Fig. 3.7 is treated for the white noise, antipodal noise, and worst-case noise.

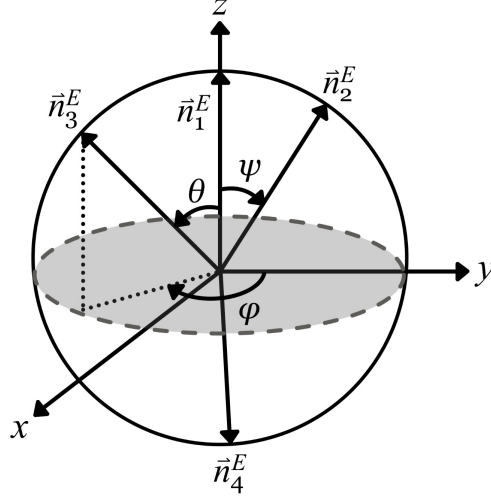


Figure 3.7: Three-angle parameterization of the four-outcome equal-weight POVM. The angle φ depends on the input angles $(\theta, \psi) \in (0, \pi)$.

Since all POVM elements belonging to any valid POVM sum to identity, the following holds for the POVM defined in Fig. 3.7, working in the Bloch representation in Eq. (3.13) but with $a = 1, 2, 3, 4$:

$$\mathbb{1} = \sum_a \lambda_a^E (\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma}) = \mathbb{1} + \frac{1}{4} \sum_{a=1}^4 \vec{n}_a^E \cdot \vec{\sigma} \iff \sum_a \vec{n}_{a,i}^E = 0, \quad i = x, y, z.. \quad (3.26)$$

Hence, the Bloch vectors defined in Fig. 3.7, are given by

$$\begin{aligned} \vec{n}_1^E &= (0, 0, 1) \\ \vec{n}_2^E &= (0, \sin(\psi), \cos(\psi)), \\ \vec{n}_3^E &= (\sin(\theta) \cos(\varphi), \sin(\theta) \sin(\varphi), \cos(\theta)), \\ \vec{n}_4^E &= -\vec{n}_1^E - \vec{n}_2^E - \vec{n}_3^E, \end{aligned} \quad (3.27)$$

where $\|\vec{n}_a^E\| = 1$ for all a , since the rank-one extremal POVMs are enough to consider. The normalization of $\vec{n}_4^E$ gives rise to a relationship between φ and the input angles (θ, ψ) in the following way:

$$\begin{aligned} \|\vec{n}_4^E\|^2 &= \|\vec{n}_1^E + \vec{n}_2^E + \vec{n}_3^E\|^2 = 3\|\vec{n}_i^E\|^2 + 2(\vec{n}_1^E \cdot \vec{n}_2^E + \vec{n}_1^E \cdot \vec{n}_3^E + \vec{n}_2^E \cdot \vec{n}_3^E) = 1 \\ \iff \cos(\psi) + \cos(\theta) + \sin(\psi) \sin(\theta) \sin(\varphi) + \cos(\psi) \cos(\theta) &= -1 \\ \iff \sin(\varphi) &= \frac{-1 - \cos(\psi) - \cos(\theta) - \cos(\psi) \cos(\theta)}{\sin(\psi) \sin(\theta)} \end{aligned} \quad (3.28)$$

The natural choice in the yz -plane, $y > 0$, restricts the input arguments to $(0, \pi)$. Moreover, $\varphi \in \mathbb{R}$ and exists if and only if $|\sin(\varphi)| \leq 1$. This existence constraint leads to the angle restriction, $\theta + \psi \geq \pi$, that the input angles must satisfy. See Appendix F for details. The

extra angle constraint is indicated by the staircase-like behavior in Fig. (3.8).

An interesting special case of this POVM class occurs when $\theta = \psi = \arccos(-1/3)$. This corresponds to the SIC-POVM. To see this, consider Eq. (2.15), from which

$$\text{Tr}(M_a M_b) = d^2 \text{Tr}(E_a E_b) = \frac{1}{3}, \quad a \neq b. \quad (3.29)$$

Moreover,

$$d^2 \text{Tr}(E_a E_b) = \text{Tr} \left[\frac{1}{4} (\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma})(\mathbb{1} + \vec{n}_b^E \cdot \vec{\sigma}) \right] = \frac{1}{2} (1 + \vec{n}_a^E \cdot \vec{n}_b^E), \quad (3.30)$$

where $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$ was used in the last equality, together with the fact that the Pauli operators are traceless. This means that

$$\frac{1}{2} (1 + \vec{n}_a^E \cdot \vec{n}_b^E) = \frac{1}{3} \iff \vec{n}_a^E \cdot \vec{n}_b^E = -\frac{1}{3}. \quad (3.31)$$

The Bloch vectors are of unit length since the POVMs considered are of rank one. Consequently, the SIC-POVM for the POVM family defined in Fig. 4a is given by the special case $\theta = \psi = \arccos(-1/3)$.

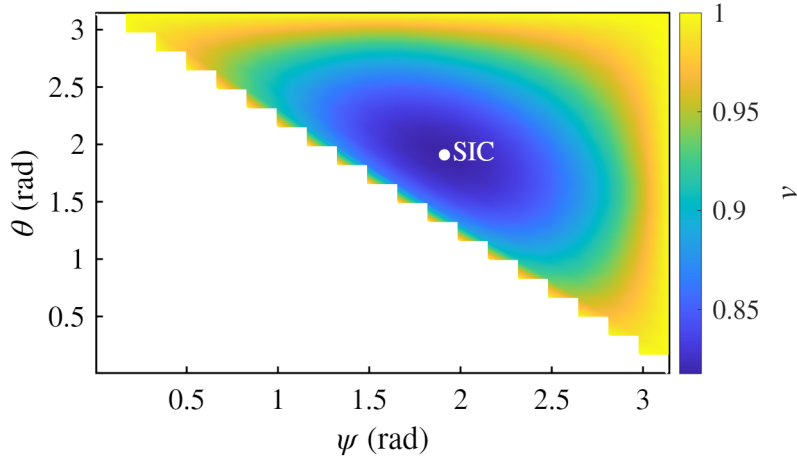


Figure 3.8: Two-dimensional $v_{\text{white}}(\theta, \psi)$ map over the simultaneous sweep over θ and ψ . The minimum visibility is marked by the white dot corresponding to SIC-POVM, with $\theta = \psi = \arccos(-1/3)$.

3.5.1 White noise, antipodal noise, and worst-case noise

The SDPs used to obtain the SDP critical visibility curves as functions of the input angles θ and ψ for the three noise models are the same as discussed in the previous section, but with more outcome variables. The closed-form visibility for white noise is given by

$$v_{\text{white}}(\theta, \psi) = \begin{cases} \frac{\sqrt{2}}{x + y + z}, & m \leq \frac{x + y + z}{2}, \\ \frac{1}{\sqrt{m^2 + (x + y + z - m)^2}}, & m > \frac{x + y + z}{2}, \end{cases} \quad (3.32)$$

with

$$m = \max\{x, y, z\}, \text{ and} \quad (3.33)$$

$$x = \cos\left(\frac{\psi}{2}\right), \quad y = \cos\left(\frac{\theta}{2}\right), \quad z = \sqrt{-\frac{\cos(\psi) + \cos(\theta)}{2}}.$$

See Appendix H for derivation. The smallest critical visibility for this noise is given by $v_{\text{white}}^{\max} = \sqrt{\frac{2}{3}} \approx 0.8165$. This value is achieved for the SIC angles.

Furthermore, as earlier discussed, the antipodal noise visibility is given by $v_{\text{ap}}(\theta, \psi) = \frac{1+v_{\text{white}}}{2}$, with $v_{\text{ap}}^{\min}(\theta, \psi) = v_{\text{ap}}(\theta_{\text{SIC}}, \psi_{\text{SIC}}) = \frac{\sqrt{2/3+1}}{2} \approx 0.9082$.

Lastly, the worst-case noise visibility can analytically be reduced to (see Appendix H)

$$v_{\text{wc}}(\theta, \psi) = \begin{cases} \frac{\sqrt{3} + \sqrt{2}}{\sqrt{3} + s}, & a(\sqrt{6} - 1) \leq B + \sqrt{3} - \sqrt{2}, \\ \frac{2}{1 + p_*^2 + 2Bp_* + a\sqrt{1 - 6p_*^2 + p_*^4}}, & a(\sqrt{6} - 1) > B + \sqrt{3} - \sqrt{2}, \end{cases} \quad (3.34)$$

where

$$s := x + y + z, \quad a := \max\{x, y, z\}, \quad B := s - a \quad (3.35)$$

$$x = \cos\left(\frac{\psi}{2}\right), \quad y = \cos\left(\frac{\theta}{2}\right), \quad z = \sqrt{-\frac{\cos(\psi) + \cos(\theta)}{2}}.$$

In Eq. (3.34), p_* is the unique root of

$$(B + p)^2(p^4 - 6p^2 + 1) - a^2p^2(3 - p^2)^2 = 0, \quad 0 < p_* < \sqrt{3} - \sqrt{2}, \quad (3.36)$$

which needs to be solved numerically. Hence, the piece-wise function in Eq. (3.34) is semi-analytical as it cannot be reduced to a closed-form expression of elementary functions and radicals. The minimum critical visibility is $v_{\text{wc}}(\theta_{\text{SIC}}, \psi_{\text{SIC}}) = v_{\text{ap}}^{\min}(\theta, \psi)$, which was verified numerically. Together with the other visibilities, $v_{\text{wc}}(\theta, \psi)$ is plotted in Fig. 3.9 below. Due to reasons discussed above, the worst-case noise largely coincides with the antipodal noise and yields the highest visibilities. The shape of the curves in Fig. 3.9 emphasizes that POVMs of very small values of (θ, ψ) are very easily PM-simulable. The most non-projective POVM in the POVM family considered corresponds to the SIC-POVM in all noise cases. For the qubit case, this result is expected from the algorithm for finding the most non-projective qubit POVM developed in [34].

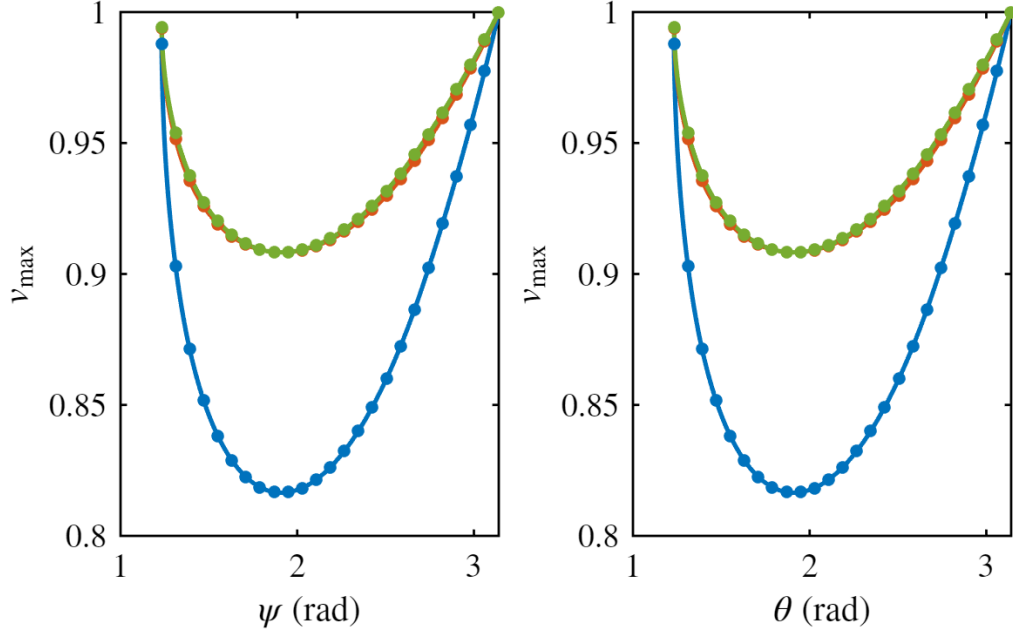


Figure 3.9: Analytical³ and SDP visibilities for white (blue), antipodal (red), and worst-case (green) noise as a function of θ and ψ for the four-outcome equal-weight POVM model. To the left, $\theta = \arccos(-1/3)$ is fixed, while ψ is changed and to the right, θ is changed while $\psi = \arccos(-1/3)$ is held fixed.

³Semi-analytical for v_{wc} .

Chapter 4

Projective simulability of quantum instruments

The formalism of projective-simulability of quantum instruments was developed in [7]. In this chapter, we provide the relevant details needed. This includes a protocol and a necessary condition for the projective simulability of QIs, which is also sufficient in the qubit case. More precisely, an SDP is formulated to compute the critical visibility for the white and worst-case noise. Later on, in 4.3, we propose an algorithm based on the approach for the measurement case in [34] to address the open problem of characterizing the most non-projective QIs. The algorithm is used for qubit QIs with varying numbers of outcomes and noise types.

4.1 Protocol

The protocol for simulating a QI, $\mathcal{I}$, is closely related to the measurement protocol in Fig. 3.1. The difference is that the protocol for a QI transforms a state, $\rho \in \mathcal{D}(\mathcal{H}_A)$, into both a classical outcome, a , and a corresponding post-measurement state, $\omega_a \in \mathcal{D}(\mathcal{H}_{A'})$, see Fig. 4.1a. Accounting for the post-measurement state requires quantum processing, represented by the additional box labeled $\Lambda_{a',\lambda}$ in Fig. 4.1b.

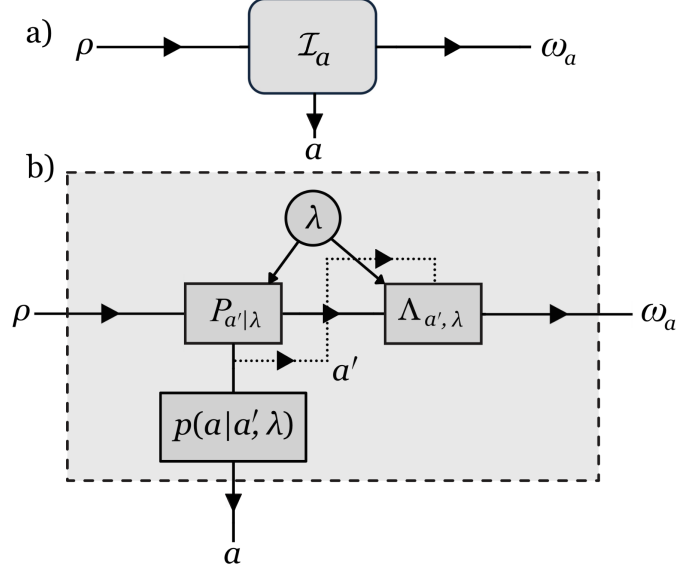


Figure 4.1: (a) A general QI transforming an input state ρ to a quantum updated state ω_a , and a classical output a . (b) The internal workings of a PI, which uses the classical random variable λ to select a PVM, $\{P_{a'|\lambda}\}$, then sends the post-projection state through a quantum channel, $\Lambda_{a',\lambda}$, and uses a rule $p(a|a', \lambda)$ to determine the classical output.

A QI can be represented by the set $\mathcal{I} = \{\mathcal{I}_a\}_{a=1}^N$, where each $\mathcal{I}_a : \mathcal{H}_A \rightarrow \mathcal{H}_{A'}$ is a CP and trace non-increasing map, such that $\sum_a \mathcal{I}_a$ is CPTP. The output of the QI branch $\mathcal{I}_a$ is given by the subnormalized state $\mathcal{I}_a(\rho)$ where $\text{Tr}(\mathcal{I}(\rho)) = p(a) \leq 1$; see Eq. (2.27). However, the actual post-measurement state ω_a in Fig. 4.1b is normalized ($\text{Tr}(\omega_a) = 1$) and given by $\omega_a = \frac{\mathcal{I}_a(\rho)}{\text{tr}(\mathcal{I}_a(\rho))}$. The normalization factor is the probability of outcome a , $p(a)$, which ensures trace-preservation of the map $\sum_a \mathcal{I}_a$, in accordance with Eq. (2.32).

In the spirit of channel-state duality, we can assign a Choi state to an QI, meaning that each $\mathcal{I}_a$ can be associated with a bipartite operator $\eta_a \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathcal{H}_A)$ defined as $\eta_a = (\mathcal{I}_a \otimes \mathbb{1})[\phi^+]$, where $\phi^+ = |\phi^+\rangle \langle \phi^+|$, and $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ is the maximally entangled state of dimension $d = \dim(\mathcal{H}_A)$. Moreover, the CP and normalization conditions are fulfilled by $\eta_a \geq 0$ and $\text{tr}_{A'} \sum_a \eta_a = \frac{\mathbb{1}}{d}$, respectively, see Eq. (2.21). Lastly, the sub-normalized quantum output of the instrument branch is given by $\mathcal{I}_a(\rho) = d \text{Tr}_A((\mathbb{1}_{A'} \otimes \rho_A^T) \eta_a)$ (comes from Eq. (2.20)). Furthermore, the instrument's reduced Choi operator realizes the POVM $\{M_a\}_a$ as $M_a = d \text{Tr}_{A'}(\eta_a)^T$, see Eq. (2.34).

Analogously to the POVM case, Fig. 4.1b defines the class of instruments that are PI-simulable. More precisely, if there exist $P_{a'|\lambda}$ and $\Lambda_{a',\lambda}$ such that the instrument outputs outcomes a and ω_a , $\mathcal{I} \in \mathcal{P}_{\mathcal{I}}$. An experimenter can implement such instruments with sole access to three resources: (i) classical randomness and classical post-processing, (ii) projective measurements on $\mathcal{H}_A$, and (iii) quantum post-processing. Upon drawing a classical random variable λ from some distribution $\{q_\lambda\}_\lambda$, the experimenter measures the input state ρ using a projective measurement $\{P_{a'|\lambda}\}_{a'}$ with outcome a' ($P_{a'|\lambda}^2 = P_{a'|\lambda}$ and $\sum_{a'} P_{a'|\lambda} = \mathbb{1}$). The sub-normalized post-projection state is given by $P_{a'|\lambda} \rho P_{a'|\lambda}$, which is, in general, not equal to the QI's post-measurement state, ω_a . The projected state is then passed through a CPTP quantum channel $\Lambda_{a',\lambda} : \mathcal{H}_A \rightarrow \mathcal{H}_{A'}$, which gives the freedom to possibly obtain the quantum output state $\Lambda_{a',\lambda}[P_{a'|\lambda} \rho P_{a'|\lambda}] = \omega_a$. The post-processing rule

$p(a|a', \lambda)$ determines the classical outcome a . This defines a projective instrument in the following way [7]:

$$\mathcal{I}_a(\rho) = \sum_{\lambda} q_{\lambda} \sum_{a'} p(a|a', \lambda) \Lambda_{a'}[E_{a'|\lambda} \rho E_{a'|\lambda}] \in \mathcal{P}_{\mathcal{I}}, \quad (4.1)$$

where $\mathcal{P}_{\mathcal{I}}$ is the set of PIs. Conditioned on a given λ and internal outcome a' , the sub-normalized quantum output is $\Lambda_{a',\lambda}[P_{a'|\lambda} \rho P_{a'|\lambda}]$, and classical outcome a , are announced with probability $p(a|a', \lambda)$. Hence, the sub-normalized output assigned to a is given by $p(a|a', \lambda) \Lambda_{a',\lambda}[P_{a'|\lambda} \rho P_{a'|\lambda}]$. By summing over all internal projective outcomes¹ a' , and averaging over the random choice of λ , Eq. (4.1) is obtained.

4.2 Condition for simulability of instruments

How can we determine whether a QI is projectively simulable? To address this question, we start with a general and necessary condition for PI-simulability. Recall from Section 3.3, that we can associate a rank-vector $\vec{r} = (r_1, r_2, \dots, r_N)$, to every N -outcome PVM $\{P_a\}_{a=1}^N$, where $r_a = \text{rank}(P_a)$ due to completeness ($\sum_a P_a = \mathbb{1}$). Since the measurements associated with each such $\vec{r}$ can be treated separately, we write $\lambda = (\chi, \vec{r})$, where χ is a random variable for choosing a PVM with rank vector $\vec{r}$.

The Choi representation in Eq. (2.19) of a QI is given by

$$\mathcal{I}_a \rightarrow \eta_a = (\mathcal{I}_a \otimes \mathbb{1})[\phi^+], \quad (4.2)$$

where post-processing has been neglected, i.e., $p(a|a', \lambda) = \delta_{a,a'}$ and $\Lambda_{a',\lambda} = \Lambda_{\lambda}$, without loss of generality (proven in Appendix I). Consequently, the Choi state of a generic PI can be written as

$$n_a = \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} (\Lambda_{\chi, \vec{r}} \otimes \mathbb{1}) [\nu_{a|\chi, \vec{r}}], \quad (4.3)$$

where $\nu_{a|\chi, \vec{r}} = (E_{a|\chi, \vec{r}} \otimes \mathbb{1})\phi^+(E_{a|\chi, \vec{r}} \otimes \mathbb{1})$. The equation above is derived in Appendix J. This relates to the following theorem.

Theorem 1. (*Instrument simulation*). *Consider any N -outcome PI and let $\{\eta_a\}_{a=1}^N$ be its Choi representation. There exists a decomposition:*

$$\begin{aligned} \eta_a &= \sum_{\vec{r}} \sigma_{a|\vec{r}}, \quad \text{where } \sigma_{a|\vec{r}} \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathcal{H}_A) \\ \text{tr}(\sigma_{a|\vec{r}}) &= q_{\vec{r}} \frac{r_a}{d}, \quad \sum_a \text{tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{\mathbb{1}}{d}, \quad SN(\sigma_{a|\vec{r}}) \leq r_a, \end{aligned} \quad (4.4)$$

where $\vec{r}$ runs over all rank vectors and SN denotes the Schmidt number.

See Appendix K for the proof of Theorem 1. Theorem 1 can be interpreted as: if $\mathcal{I}$ cannot be written as a convex combination over the Choi representations associated with the different rank-vectors, then $\mathcal{I} \notin \mathcal{P}_{\mathcal{I}}$. The operators $\{\sigma_{a|\vec{r}}\}_a$ correspond to a sub-normalized Choi representation of a PI using only measurements with rank-vector $\vec{r}$ and play the

¹Since a can arise from any a' via post-processing.

same role as the $F_{a|\vec{r}}$ -operators in Eq. (3.12) in Section 3.3. The sub-normalization, $q_{\vec{r}} = \sum_a \text{tr}(\sigma_{a|\vec{r}})$, corresponds to the probability of selecting $\vec{r}$, since

$$\sum_a \text{tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \sum_a \frac{r_a}{d} = q_{\vec{r}} \frac{1}{d} \sum_a r_a = q_{\vec{r}} \leq 1, \quad (4.5)$$

by definition. In the last equality, $\sum_a r_a = d$ was used.

Theorem 1 is necessary but not sufficient for PI-simulability, since not all Choi operators with a Schmidt number as in Theorem 1 can be associated with PIs [7]. However, in the practical case of qubits ($d = 2$), there is an exact characterization.

Theorem 2. (*Qubit instrument simulation*): *For instruments with qubit input, Theorem 1 is both necessary and sufficient.*

In particular, for a qubit-qubit instrument, an exact SDP, analogous to Eq. (3.12), can be formulated for any fixed noise model as

$$\begin{aligned} & \max_{v, \{\sigma_{a|\vec{r}}\}} v \\ & \text{s.t.} \quad \sum_{\vec{r}} \sigma_{a|\vec{r}} = v\eta_a + (1-v)\eta_a^{\text{noise}}, \quad \forall a \\ & \quad \text{Tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{r_a}{d}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \\ & \quad \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{\mathbb{1}}{d}, \quad \forall \vec{r} \\ & \quad \sigma_{a|\vec{r}}^{T_{A'}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1 \\ & \quad \sigma_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \end{aligned} \quad (4.6)$$

where the three sector sets are defined as $\mathcal{R}_0 := \{(a, \vec{r}) : r_a = 0\}$, $\mathcal{R}_1 = \{(a, \vec{r}) : r_a = 1\}$, $\mathcal{R}_2 := \{(a, \vec{r}) : r_a = 2\}$, and $\mathcal{R}_{1,2} := \mathcal{R}_1 \cup \mathcal{R}_2$. Here, the Schmidt number constraint in Theorem 1, $SN \leq 1$ has been replaced by the positive partial transpose (PPT) constraint $\sigma_{a|\vec{r}}^{T_A} \geq 0$ [53] (see Appendix M for motivation of the PPT criterion and Appendix L for the proof of Theorem 2).

Considering the noise contribution to be a variable² $\tilde{\eta}_a^{\text{noise}} := (1-v)\eta_a^{\text{noise}}$, requires changes to the SDP above. It can be noted that $\tilde{\eta}_a^{\text{noise}} \succeq 0$ and since $\{\eta_a^{\text{noise}}\}_a$ corresponds to an instrument, we further have $\sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) = \frac{1-v}{d} \mathbb{1}_A$. Hence, the primal SDP for the

²This choice avoids a variable multiplied by a variable.

worst-case noise is given by

$$\begin{aligned}
& \max_{v, \{\sigma_{a|\vec{r}}\}, \{\tilde{\eta}_a^{\text{noise}}\}} v \\
& \text{s.t.} \quad \sum_{\vec{r}} \sigma_{a|\vec{r}} = q_{\vec{r}} \frac{r_a}{d}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \\
& \quad \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{1_A}{d}, \quad \forall \vec{r}, \\
& \quad \sigma_{a|\vec{r}}^{T_{A'}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& \quad \sigma_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \\
& \quad \tilde{\eta}_a^{\text{noise}} \succeq 0, \quad \forall a, \\
& \quad \sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) = \frac{1-v}{d} \mathbb{1}_A.
\end{aligned} \tag{4.7}$$

The numerical³ visibility curves for the different noise models are plotted in Fig. 4.2 for the unsharp Pauli-Z instruments, which are defined by the following Kraus operators [7]:

$$K_a = \sqrt{\frac{1 - (-1)^a \gamma}{2}} |0\rangle \langle 0| + \sqrt{\frac{1 + (-1)^a \gamma}{2}} |1\rangle \langle 1|, \quad a = \{1, 2\}. \tag{4.8}$$

See Appendix N for details. The minimum critical visibility occurs at $\gamma = 1/\sqrt{2}$ with a value of 0.8162 for the worst-case noise, while for the white noise it occurs at a somewhat smaller γ with a visibility of 0.6928 [7]. The shape of the curves in Fig. 4.2 underlines that instruments close to sharp measurements ($\gamma = 1$) are more difficult to simulate projectively than those that are close to non-interacting measurements ($\gamma = 0$).

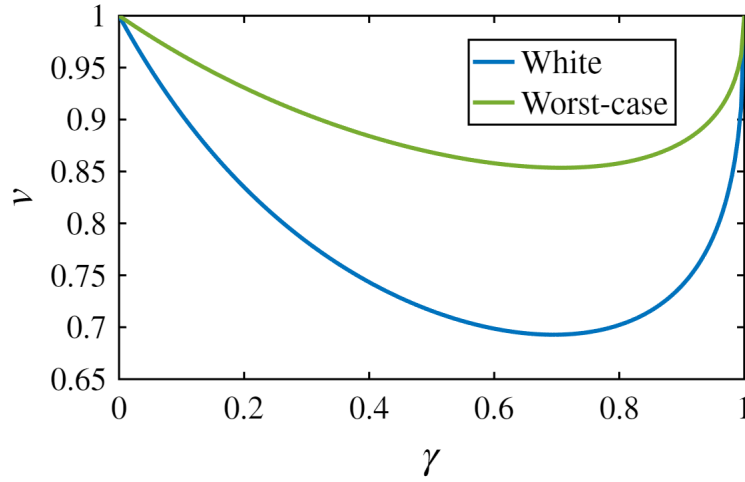


Figure 4.2: Weak-Z instrument numerical visibility v as a function of the sharpness parameter $\gamma \in [0, 1]$ for the isotropic white noise and worst-case noise.

We distinguish between two types of white noise. The white noise in Fig. 4.2 is isotropic, and is, unlike the weighted white noise in Section 3.2 (Eq. (3.5)), given by $\eta_a^{\text{noise}} = \frac{1}{d^2 N} \mathbb{1}$. This means that the white noise is equally distributed for all POVM effects. However, since trace is homogeneous for the weak-Z measurements (see Appendix N), the visibility curves for the two types of white noise coincide.

³Analytical forms can be found in [7].

4.3 The most non-projective instruments

The problem of finding the most non-projective quantum measurement has recently been solved, where a numerical see-saw algorithm was developed for this task in [34]. In this section, we propose an adaptation of the algorithm to fit quantum instruments, as an approach to the unresolved problem of characterizing the most non-projective QIs. As for QMs [34], the algorithm iteratively searches over the space of non-projective instruments by utilizing the SDP criterion. Adapted to qubit instruments, the method consists of the following steps describing the first iteration (see Fig. 4.3): (i) Choose a non-projective Choi state, η_a^0 , to start the search from and compute a bound on its visibility from the primal SDP. (ii) Extract a witness of non-projectivity, W_1 , from the dual SDP. The witness is a hyperplane separating projective POVMs from some of the non-projective ones, as shown in Fig. 4.3. (iii) Compute the QI, η_a^1 , that gives the largest witness violation by a new SDP. (iv) Return to step (i), using the instrument η_a^1 from (iii) as the input. Repeat the process until the desired convergence is reached.

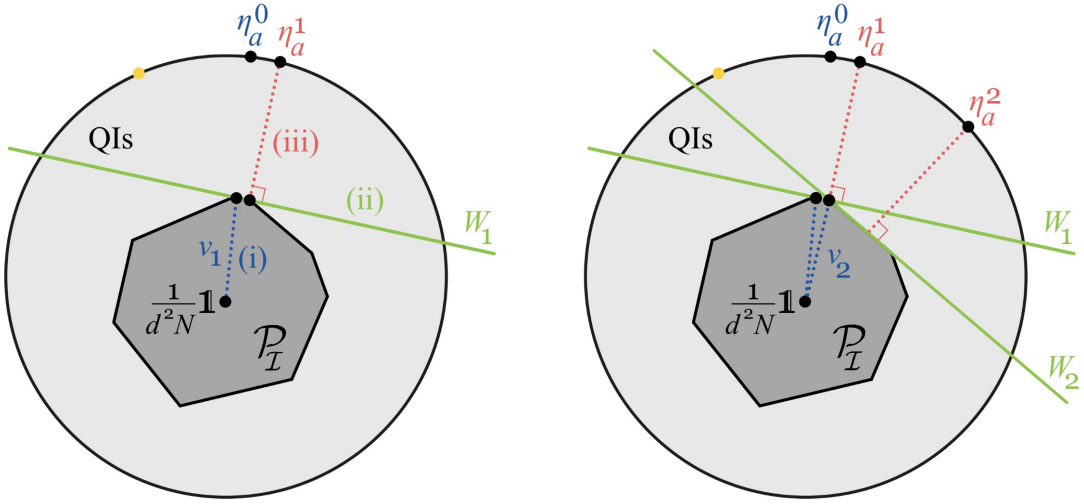


Figure 4.3: Geometrical visualization of the search algorithm, illustrated for isotropic white noise. To the left, one iteration; to the right, two iterations. Step (i) (blue), step (ii) (green), and step (iii) (red). The most non-projective instrument is the final instrument represented by the yellow dot.

For fixed noise, the dual SDP used in the algorithm is given by

$$\begin{aligned}
& \min_{\{\Gamma_a\}, \{Y_{\vec{r}}\}, \{y_{a|\vec{r}}\}, \{R_{a|\vec{r}}\}_{\mathcal{R}_1}} & 1 + \sum_a \text{tr}(\Gamma_a \eta_a) \\
& \text{s.t.} & 1 + \sum_a \text{tr}(\Gamma_a (\eta_a - \eta_a^{\text{noise}})) = 0, \\
& & \text{tr}(Y_{\vec{r}}) + \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\
& & \Gamma_a - R_{a|\vec{r}}^{T_{A'}} - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& & \Gamma_a - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\
& & R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1,
\end{aligned} \tag{4.9}$$

and for the worst-case noise, the reduced dual is modified to

$$\begin{aligned}
(\text{D}) \quad & \min_{\{\Gamma_a\}, \{Y_{\vec{r}}\}, \{y_{a|\vec{r}}\}, \{R_{a|\vec{r}}\} \mathcal{R}_1} 1 + \sum_a \text{tr}(\Gamma_a \eta_a) \\
\text{s.t.} \quad & 1 + \sum_a \text{tr}(\Gamma_a(\eta_a)) + \frac{1}{d} \text{Tr}(\beta) = 0, \\
& \text{tr}(Y_{\vec{r}}) + \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\
& \Gamma_a - R_{a|\vec{r}}^{T_{A'}} - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& \Gamma_a - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\
& -\Gamma_a - \mathbb{1}_{A'} \otimes \beta \succeq 0, \quad \forall a, \\
& R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1.
\end{aligned} \tag{4.10}$$

The dual SDPs above are derived in Appendix O and P, respectively.

To see how the dual SDP can be used to extract a witness for non-projectiveness, consider the following. The feasible solutions $\{\Gamma_a\}_a$ of the dual problem above provide a witness to detect $\mathcal{I} \notin \mathcal{P}$. More precisely, the witness can be defined as [34]:

$$W(\eta) = \sum_a \text{Tr}(\Gamma_a \eta_a). \tag{4.11}$$

Due to strong duality, the optimal primal value equals the optimal dual value. Hence, the critical visibility is given by $v^* = 1 + \sum_a \text{Tr}(\Gamma_a \eta_a)$ for the optimal $\{\Gamma_a\}_a$. From the dual constraint in Eq. (4.9), we have

$$\sum_a \text{Tr}(\Gamma_a \eta_a^{\text{noise}}) = 1 + \sum_a \text{Tr}(\Gamma_a \eta_a) = v^*, \tag{4.12}$$

at the optimum. Consequently, for any $\eta_a^{(v)}$:

$$\begin{aligned}
W(\{\eta_a^{(v)}\}) &= \sum_a \text{Tr}(\Gamma_a \eta_a^{(v)}) = v \underbrace{\sum_a \text{Tr}(\Gamma_a \eta_a)}_{v^* - 1} + (1 - v) \underbrace{\sum_a \text{Tr}(\Gamma_a \eta_a^{\text{noise}})}_{v^*} \\
&= v(v^* - 1) + (1 - v)v^* = v^* - v \geq 0.
\end{aligned} \tag{4.13}$$

From this we note that for $v < v^*$, the noisy instrument $\{\eta_a^{(v)}\}_a$ lies inside the PI-simulable region and $W(\{\eta_a^{(v)}\}) > 0$. Moreover, for the original instrument, for which $v = 1$, $W(\{\eta_a\}) = v^* - 1$. Hence, $W(\{\eta_a\}) < 0$ for $v^* < 1$. This means that a negative witness value certifies non-projectivity since the inequality in Eq. (4.13) is satisfied by all $\eta \notin \mathcal{P}$, meaning that finding $W(\eta) \not\geq 0$ implies $\mathcal{I} \notin \mathcal{P}_{\mathcal{I}}$.

The table below shows the results from the algorithm for different outcomes and noise types.

Table 4.1: Critical visibilities corresponding to the most non-projective instruments obtained from the search algorithm for $N = 2, 3, 4$ number of outcomes and for the isotropic white, weighted white, and the worst-case noise.

$N \backslash$ noise	Isotropic white	Weighted white ⁴	Worst-case
2	0.4810	0.6928	0.8536
3	0.4810	0.6928	0.8273
4	0.4810	0.6928	0.8162

To explain the results in the table above, the obtained instrument Choi state η_{opt} was investigated by extracting its associated measurement part by using $E_a = d \text{Tr}_{A'}(\eta_a)^T$ [7]. More precisely, by investigating trace properties of the measurement operators, we can understand the reason behind the obtained visibility values in Table 4.1.

First, it can be noted that the visibility of 0.8162 for the worst-case noise and $N = 2$ is the same as that of the weak- Z instruments in Fig. 4.2, where the corresponding γ -value is $1/\sqrt{2}$. This suggests that the unsharp instrument given $\gamma = 1/\sqrt{2}$ is the most non-projective two-outcome instrument for the worst-case noise. This is further supported by [48], which states that weak- Z instruments are extremal two-outcome instruments. Second, for $N = 3$ and the worst-case noise, it was noted that the extracted measurement part of η_{opt} has equal traces $\text{Tr}(E_a)$, equal pairwise overlaps $\text{Tr}(E_a E_b)$ for all $a \neq b$, and equal Bloch radius lengths, less than unity. These properties strongly suggest that the measurement operators associated with η_{opt} can be expressed as

$$E_a = \frac{1}{3}(\mathbb{1} + \zeta \vec{n}_a^E \cdot \vec{\sigma}), \quad a = 1, 2, 3, \quad (4.14)$$

where $\|\vec{n}_a\| = 1$ and ζ is the Bloch radius⁵. POVMs of the form defined above are called trine POVMs and are illustrated in Fig. 4.4a. Opposite to the POVM class studied in Section 3.4, the geometry is fixed, and the radius is varied.

⁴For this noise, the algorithm achieved a value of 0.7003 for all N . However, by running the primal SDP many times with random instruments of varying numbers of outcomes, a lower value of 0.6928 was obtained. This indicates that the algorithm got stuck at a local minimum.

⁵This can be seen as a sharpness factor indicating how sharply the trine-POVM is implemented.

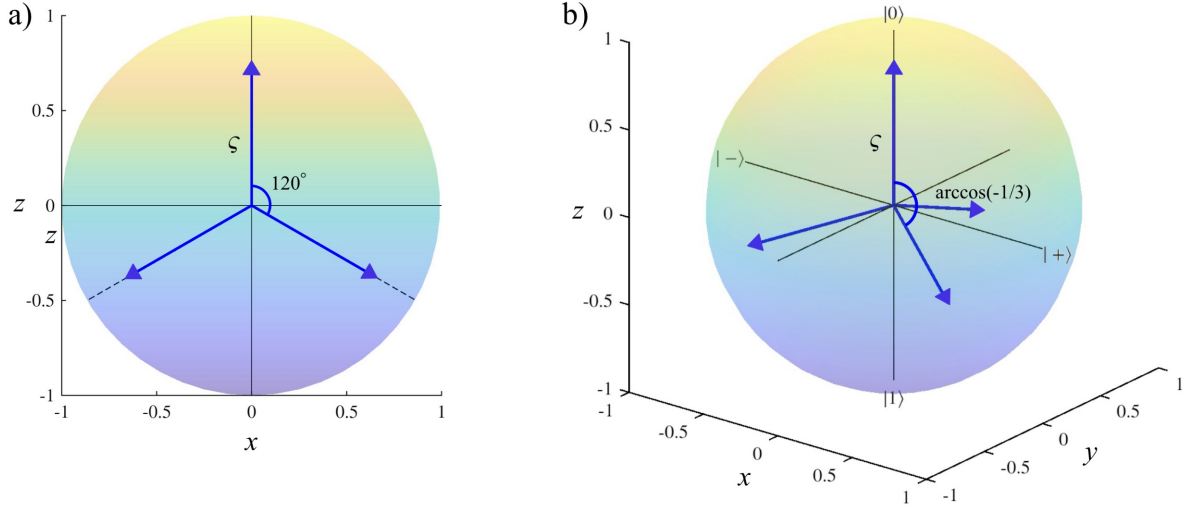


Figure 4.4: a) Bloch-vector representation of the trine POVM family in the xz -plane of the Bloch sphere. This POVM class has effects with equal weights λ_a and equal Bloch radii, ζ . All the pairwise angles between the Bloch vectors are equal to 120° . b) Illustration of the Bloch vectors of the tetrahedral POVM family in the Bloch sphere. Such POVMs have equal weights λ_a , equal Bloch radius lengths given by $\zeta \in [0, 1]$, and equal pairwise angles of $\arccos(-1/3)^\circ$.

The Bloch radii were extracted in the following way. An arbitrary qubit effect is given by

$$E_a = \lambda_a(\mathbb{1} + \vec{r}_a \cdot \vec{\sigma}), \quad (4.15)$$

where $\lambda_a = \text{Tr}(E_a)/2$ and $\|\vec{r}_a\| \leq 1$. From this,

$$\text{Tr}(E_a E_b) = 2\lambda_a \lambda_b (\mathbb{1} + \vec{r}_a \cdot \vec{r}_b) \iff \vec{r}_a \cdot \vec{r}_b = \frac{\text{Tr}(E_a E_b)}{2\lambda_a \lambda_b} - 1 = \frac{2 \text{Tr}(E_a E_b)}{\text{Tr}(E_a) \text{Tr}(E_b)} - 1. \quad (4.16)$$

Using this formula, it was found that $r_a = \zeta \approx 0.7559$ for all $a = 1, 2, 3$. Hence, the most non-projective instrument associated with the visibility of 0.8273 in Table 4.1 corresponds to a trine measurement of $\zeta \approx 0.7559$.

The values for v and ζ associated with η_{opt} can be explained by the analytical form of the instrument visibility as a function of ζ . To this end, we use the following procedure. First, the measurement part of η_{opt} , denoted E_{opt} was extracted using [7]

$$E_{\text{opt}} = d \text{Tr}_{A'}(\eta_{\text{opt}})^T. \quad (4.17)$$

Second, given E_{opt} , the corresponding Lüders form is given by [7]

$$\eta_{\text{lüders}, a} = (K_{\text{opt}, a} \otimes \mathbb{1}) |\phi^+\rangle \langle \phi^+| (K_{\text{opt}, a}^\dagger \otimes \mathbb{1}), \quad (4.18)$$

where $K_{\text{opt}, a} = \sqrt{E_{\text{opt}, a}}$ and $|\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$, was constructed. Running the SDP on the Lüders instrument constructed from E_{opt} resulted in the same visibility value as in Table 4.1. This means that η_{opt} has a Lüders form. Consequently, a family of Lüders instruments corresponding to the trine POVM class defined by Eq. (4.14) is indirectly parameterized by the shrinking factor ζ . Running the primal instrument SDP for this family of instruments for $\zeta \in [0, 1]$ allows us to plot the instrument visibility as a function

of ζ , as Fig. 4.5 shows. The shape of the $v(\zeta)$ -curves in Fig. 4.5 indicates that instruments of larger shrinking factors are harder to simulate compared to those of lower values.

The closed-form visibility is given by (see Appendix Q for derivation):

$$v_{\text{wc}}^{(3)}(\zeta) = \begin{cases} 1, & \zeta = 0, \\ \frac{(1 - t_*(\zeta))(1 + 7t_*(\zeta))}{8 \left[1 - A(\zeta)(1 - t_*(\zeta)) - B(\zeta)t_*(\zeta) \right]}, & 0 < \zeta < \zeta_c, \\ \frac{1}{8(1 - A(\zeta))}, & \zeta_c \leq \zeta \leq 1, \end{cases} \quad (4.19)$$

where

$$t_*(\zeta) = \frac{2(6 - 7A + B)}{14(1 - A) + \sqrt{28(1 - B)(7 + B - 8A)}}, \quad (4.20)$$

with $A := \frac{\sqrt{3}}{2}\zeta$ and $B = \sqrt{1 - \zeta^2}$. Moreover, the transition point ζ_c is

$$\zeta_c = \frac{84\sqrt{3} + 2\sqrt{7}}{151} \approx 0.9986. \quad (4.21)$$

The minimum $v_{\text{wc}}^{(3),\min}(\zeta) = (7 + \sqrt{21})/14 \approx 0.8273$ with $\zeta = 2/\sqrt{7} \approx 0.7559$, coincides with the value given in Table 4.1 and the extracted corresponding Bloch radius in Eq. (4.16), respectively.

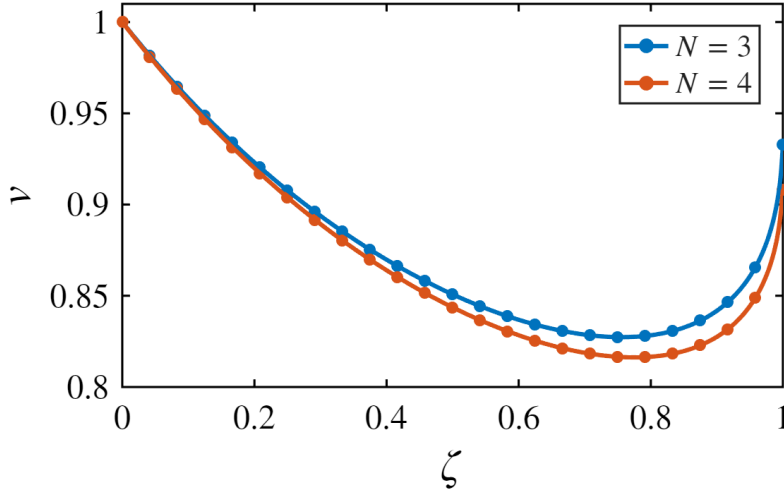


Figure 4.5: Worst-case noise instrument visibility v as a function of Bloch radius length ζ corresponding to the trine POVM instrument (blue) and the tetrahedral POVM instrument (red). Analytical visibility is plotted as a solid line, and SDP as a dotted line.

Furthermore, for $N = 4$, η_{opt} corresponds to a uniformly⁶ shrunken equal-weight ($\lambda_a = 1/4 = \text{Tr}(M_a)/2$ for all a) four-outcome POVM of equal pairwise trace overlaps. Hence, the data strongly suggests that the measurement part of the obtained instrument η_{opt} is of the form of a tetrahedral POVM of Bloch radius ζ (see Fig. 4.4b):

$$E_a = \frac{1}{4}(\mathbb{1} + \zeta \vec{n}_a^E \cdot \vec{\sigma}), \quad a = 1, 2, 3, 4, \quad (4.22)$$

⁶The Bloch radii lengths were calculated in the same way as for the three-outcome instrument above.

where $\|\vec{n}_a\| = 1$. This POVM family has the same angle geometry as the SIC-POVM in Section 3.5, where $\theta = \psi = \arccos(-1/3)$, but the Bloch vector length is variable instead of unity. By following the same method as for the $N = 3$ case, i.e., converting η_{opt} into E_{opt} by Eq. (4.17) and then transforming E_{opt} back into an instrument on the Lüders form by Eq. (4.18), the instrument visibility for the restricted Lüders instrument family corresponding to the tetrahedral POVM family can be plotted numerically as a function of ζ . This is plotted in Fig. 4.5, which includes the closed-form instrument visibility, given by

$$v_{\text{wc}}^{(4)}(\zeta) = \begin{cases} 1, & \zeta = 0, \\ \frac{(1 - t_\star)(1 + 5t_\star)}{6[1 - A(1 - t_\star) - Bt_\star]}, & 0 < \zeta < \zeta_c, \\ \frac{1}{6(1 - A)}, & \zeta_c \leq \zeta \leq 1, \end{cases} \quad (4.23)$$

where

$$t_\star(\zeta) = \frac{2(4 - 5A(\zeta) + B(\zeta))}{10(1 - A(\zeta)) + \sqrt{20(1 - B(\zeta))(5 + B(\zeta) - 6A(\zeta))}}, \quad (4.24)$$

$A := \sqrt{\frac{2}{3}}\zeta$, and $B := \sqrt{1 - \zeta^2}$. The transition point is given by

$$\zeta_c = \frac{20\sqrt{6} + \sqrt{15}}{53} \approx 0.9974. \quad (4.25)$$

See Appendix R for details. The minimum of the piecewise function is $v_{\text{wc}}^{(4), \min} = \frac{1 + \sqrt{2/5}}{2} \approx 0.8162$, which agrees with the value in Table 4.1. Moreover, the analytical value of $\zeta = \sqrt{3/5} \approx 0.7746$, corresponding to $v_{\min}$, also matches the value of the uniform shrinking factor extracted from E_{opt} .

For the weighted white noise, the same value of 0.6928 is obtained for all N , where η_{opt} is a two-outcome instrument on Lüders form⁷. This value is equal to the minimum v_{white} which was obtained for unsharp instruments in Fig. 4.2, with a γ -value slightly smaller than $1/\sqrt{2}$. Due to the unitary freedom of instruments, several instruments can, in principle, yield the same critical visibility. However, since no other examples of an instrument giving this visibility value and no other examples of lower visibility values were found with the algorithm, we can conclude that the weak- Z instrument with this γ is the most non-projective instrument for the weighted white noise.

For the isotropic white noise, the instrument reduces to a two-outcome antipodal POVM of equal weights for all N of visibility 0.4810, see Table 4.1. Since higher values of N dilute the noise more, a smaller (larger) v ($1 - v$) is required to make $\eta_a^{(v)}$ PI-simulable. Therefore, the visibility for the isotropic noise is significantly smaller than for the weighted white noise.

⁷Since the primal SDP was only run for random instruments defined on Lüders form.

Chapter 5

Conclusions and outlook

In this thesis, a three-resource protocol for projective-simulability has been studied for qubit-qubit QIs and rank-one qubit QMs. SDPs for different noise models were used to estimate the critical visibilities for both objects. Particularly for QMs, two POVM classes were characterized and studied: a three-outcome model defined by an angle θ and a four-outcome model characterized by two input angles θ and ψ . From the noise-dependent SDP, visibility curves were produced both numerically and analytically, as functions of the angle parameters. It was found that the most non-projective POVMs for all noise types are highly symmetric: an equal weight rank-one trine-POVM with $\theta = 120^\circ$, and the SIC-POVM, as expected.

For QIs, a search algorithm to find the most non-projective instrument was developed and run across different noise models and outcome numbers. Since, for the worst-case noise and weighted white noise, we found one example of an instrument for each yielding the critical visibilities in Table 4.1 under the weak- Z measurement, we can conjecture that the weak- Z instrument is the least projective among these noise types. Furthermore, it was observed that the most non-projective instruments for the worst-case noise are highly symmetric. It was analytically verified that the most non-projective POVMs in the trine and tetrahedral POVM families correspond to the instruments obtained using the algorithm. Although the results were strongly justified, the algorithm has the drawback of occasionally getting stuck in local minima.

A natural next step is to employ an asymmetric model for the three-outcome POVM. Instead of using one angle parameter θ , which symmetrically separates two of the Bloch vectors in Fig. 3.5, two arbitrary angles can be used, accounting for a larger class of three-outcome POVMs. To further develop the characterization of non-projectiveness in QMs and QIs, it is of high interest to study higher dimensions. However, it would require a different method. The convex optimization problems used in the thesis can handle low-dimensional systems, but lack control over the number of optimization variables, making them experimentally flexible but computationally insufficient. Particularly, the computational cost of the SDPs used in the thesis grows exponentially with d (see Appendix S for details). Common tools such as YALMIP [54], along with built-in optimizers like self-dual-minimization (SeDuMi) and MOSEK, primarily use a semidefinite-quadratic-linear programming approach [55] and are typically employed to solve such problems. However, these tools become insufficient for large-scale systems, often requiring many hours of computation on a moderately configured laptop [56].

Bibliography

- [1] S. Rout, S. Sankar Bhattacharya, and P. Horodecki, “Randomness-free detection of non-projective measurements: qubits & beyond,” *New Journal of Physics*, vol. 27, p. 033024, Mar. 2025.
- [2] J.-P. Pellonpää, “Quantum instruments: II. Measurement theory,” *Journal of Physics A: Mathematical and Theoretical*, vol. 46, p. 025303, Jan. 2013.
- [3] T. Heinosaari and M. M. Wolf, “Nondisturbing quantum measurements,” *Journal of Mathematical Physics*, vol. 51, p. 092201, Sept. 2010.
- [4] P. Busch and P. J. Lahti, “The standard model of quantum measurement theory: History and applications,” *Foundations of Physics*, vol. 26, pp. 875–893, July 1996.
- [5] Y. Aharonov, D. Z. Albert, and L. Vaidman, “How the result of a measurement of a component of the spin of a spin- $1/2$ particle can turn out to be 100,” *Physical Review Letters*, vol. 60, pp. 1351–1354, Apr. 1988.
- [6] K. Jacobs and D. A. Steck, “A Straightforward Introduction to Continuous Quantum Measurement,” *Contemporary Physics*, vol. 47, pp. 279–303, Sept. 2006. arXiv:quant-ph/0611067.
- [7] S. Khandelwal and A. Tavakoli, “Simulating quantum instruments with projective measurements and quantum post-processing,” *Physical Review Letters*, vol. 135, p. 040202, July 2025. arXiv:2503.00956 [quant-ph].
- [8] S. M. Barnett and S. Croke, “Quantum state discrimination,” *Advances in Optics and Photonics*, vol. 1, p. 238, Apr. 2009.
- [9] A. Peres, “How to differentiate between non-orthogonal states,” *Physics Letters A*, vol. 128, no. 19, 1988.
- [10] M. T. DiMario and F. E. Becerra, “Demonstration of optimal non-projective measurement of binary coherent states with photon counting,” *npj Quantum Information*, vol. 8, p. 84, July 2022.
- [11] J. Bae and L.-C. Kwek, “Quantum state discrimination and its applications,” *Journal of Physics A: Mathematical and Theoretical*, vol. 48, p. 083001, Feb. 2015. arXiv:1707.02571 [quant-ph].
- [12] J. Shang, A. Asadian, H. Zhu, and O. Gühne, “Enhanced entanglement criterion via symmetric informationally complete measurements,” *Physical Review A*, vol. 98, p. 022309, Aug. 2018.

- [13] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, “Sample-optimal tomography of quantum states,” in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, (Cambridge MA USA), pp. 913–925, ACM, June 2016.
- [14] H. C. Nguyen, J. L. Bönsel, J. Steinberg, and O. Gühne, “Optimizing Shadow Tomography with Generalized Measurements,” *Physical Review Letters*, vol. 129, p. 220502, Nov. 2022.
- [15] R. Stricker, M. Meth, L. Postler, C. Edmunds, C. Ferrie, R. Blatt, P. Schindler, T. Monz, R. Kueng, and M. Ringbauer, “Experimental Single-Setting Quantum State Tomography,” *PRX Quantum*, vol. 3, p. 040310, Oct. 2022.
- [16] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, June 2004.
- [17] R. Derka, V. Buzek, and A. K. Ekert, “Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement,” *Physical Review Letters*, vol. 80, pp. 1571–1575, Feb. 1998.
- [18] L.-T. Feng, X.-M. Hu, M. Zhang, Y.-J. Cheng, C. Zhang, Y. Guo, Y.-Y. Ding, Z. Hou, F.-W. Sun, G.-C. Guo, D.-X. Dai, A. Tavakoli, X.-F. Ren, and B.-H. Liu, “Higher-dimensional symmetric informationally complete measurement via programmable photonic integrated optics,” *Optica*, vol. 12, p. 1014, July 2025.
- [19] A. J. Scott, “Tight informationally complete quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 39, p. 13507, Oct. 2006.
- [20] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, “Compatibility in multiparameter quantum metrology,” *Physical Review A*, vol. 94, p. 052108, Nov. 2016.
- [21] F. Albarelli and R. Demkowicz-Dobrzański, “Probe Incompatibility in Multiparameter Noisy Quantum Metrology,” *Physical Review X*, vol. 12, p. 011039, Mar. 2022.
- [22] M. Szczykulska, T. Baumgratz, and A. Datta, “Multi-parameter quantum metrology,” *Advances in Physics: X*, vol. 1, pp. 621–639, July 2016.
- [23] A. Acín, S. Pironio, T. Vértesi, and P. Wittek, “Optimal randomness certification from one entangled bit,” *Physical Review A*, vol. 93, p. 040102, Apr. 2016.
- [24] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, “Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy,” *Physical Review A*, vol. 56, pp. 1163–1172, Aug. 1997.
- [25] A. M. Childs and W. Van Dam, “Quantum algorithms for algebraic problems,” *Reviews of Modern Physics*, vol. 82, pp. 1–52, Jan. 2010.
- [26] D. Bacon and A. M. Childs, “OPTIMAL MEASUREMENTS FOR THE DIHEDRAL HIDDEN SUBGROUP PROBLEM,” *Chicago Journal of Theoretical Computer Science*, vol. 2006, no. 3, pp. 1–19, 2006.

- [27] P. Sen, “Random measurement bases, quantum state distinction and applications to the hidden subgroup problem,” *Proceedings of the 21st Annual IEEE Conference on Computational Complexity (CCC 2006)*, pp. 14–28, 2006.
- [28] J. Barrett, “Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality,” *Physical Review A*, vol. 65, p. 042302, Mar. 2002.
- [29] T. Vértesi and E. Bene, “Two-qubit Bell inequality for which positive operator-valued measurements are relevant,” *Physical Review A*, vol. 82, p. 062115, Dec. 2010.
- [30] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, “Simulating positive-operator-valued measures with projective measurements,” *Physical Review Letters*, vol. 119, p. 190501, Nov. 2017. arXiv:1609.06139 [quant-ph].
- [31] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, “Experimental characterization of qutrits using SIC-POVMs,” *Physical Review A*, vol. 83, p. 051801, May 2011. arXiv:1006.4905 [quant-ph].
- [32] R. Brinster, P. Tirler, S. Khandelwal, M. Meth, H. Kampermann, D. Bruß, R. Blatt, M. Ringbauer, A. Tavakoli, and N. Wyderka, “Robust certification of non-projective measurements: theory and experiment,” Nov. 2025. arXiv:2511.04446 [quant-ph].
- [33] Y. Ota, S. Ashhab, and F. Nori, “Implementing general measurements on linear optical and solid-state qubits,” 2012.
- [34] G. Cobucci, R. Brinster, S. Khandelwal, H. Kampermann, D. Bruß, N. Wyderka, and A. Tavakoli, “Maximally non-projective measurements are not always symmetric informationally complete,” Aug. 2025. arXiv:2508.03652 [quant-ph].
- [35] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum Measurement. Theoretical and Mathematical Physics*, Cham: Springer International Publishing, 2016.
- [36] M. Nielsen and I. Chuang, “Quantum Computation and Quantum Information,” vol. 10th edition, 2010.
- [37] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Cambridge University Press, 3 ed., Sept. 2020.
- [38] B. C. Hall, *Quantum Theory for Mathematicians*, vol. 267 of *Graduate Texts in Mathematics*. New York, NY: Springer New York, 2013.
- [39] R. Bertlmann and N. Friis, *Modern Quantum Theory: From Quantum Mechanics to Entanglement and Quantum Information*. Oxford University Press Oxford, 1 ed., Sept. 2023.
- [40] A. Hashim, L. B. Nguyen, N. Goss, B. Marinelli, R. K. Naik, T. Chistolini, J. Hines, J. Marceaux, Y. Kim, P. Gokhale, T. Tomesh, S. Chen, L. Jiang, S. Ferracin, K. Rudinger, T. Proctor, K. C. Young, I. Siddiqi, and R. Blume-Kohout, “Practical Introduction to Benchmarking and Characterization of Quantum Computers,” *PRX Quantum*, vol. 6, p. 030202, Aug. 2025.

- [41] M. Wilde, *Quantum information theory*. Cambridge, UK ; New York: Cambridge University Press, second edition ed., 2017.
- [42] A. Holst, *Matrix theory*. Lund: Studentlitteratur, 1. ed ed., 2014.
- [43] A. Peres, *Quantum Theory: Concepts and Methods*. No. 57 in Fundamental Theories of Physics, Dordrecht: Imprint: Springer, 1st ed. 2002 ed., 2002.
- [44] G. M. D’Ariano, P. L. Presti, and P. Perinotti, “Classical randomness in quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 38, pp. 5979–5991, July 2005.
- [45] A. Łuczak and H. Podsiadkowska, “Lüders Instruments, Generalised Lüders Theorem, and Some Aspects of Sufficiency,” *International Journal of Theoretical Physics*, vol. 54, pp. 4283–4292, Dec. 2015.
- [46] L. N. Trefethen, *Numerical linear algebra*. Philadelphia, Pa: Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1997.
- [47] N. Wheeler, “Generalized Quantum Measurement,” 2012.
- [48] J.-P. Pellonpää, “Quantum instruments: I. Extreme instruments,” *Journal of Physics A: Mathematical and Theoretical*, vol. 46, p. 025302, Jan. 2013.
- [49] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*. IOP Publishing, Mar. 2023.
- [50] A. Tavakoli, A. Pozas-Kerstjens, P. Brown, and M. Araújo, “Semidefinite programming relaxations for quantum correlations,” *Reviews of Modern Physics*, vol. 96, p. 045006, Dec. 2024. arXiv:2307.02551 [quant-ph].
- [51] G. Chiribella, G. M. D’Ariano, and D. Schlingemann, “How Continuous Quantum Measurements in Finite Dimensions Are Actually Discrete,” *Physical Review Letters*, vol. 98, p. 190403, May 2007.
- [52] E. B. Davies and J. T. Lewis, “An operational approach to quantum probability,” 1970.
- [53] P. Horodecki, “Separability criterion and inseparable mixed states with positive partial transposition,” *Physics Letters A*, vol. 232, pp. 333–339, Aug. 1997.
- [54] J. Lofberg, “YALMIP : a toolbox for modeling and optimization in MATLAB,” in *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, (Taipei, Taiwan), pp. 284–289, IEEE, 2004.
- [55] A. Gaikwad, M. S. Torres, S. Ahmed, and A. F. Kockum, “Gradient-descent methods for fast quantum state tomography,” *Quantum Science and Technology*, vol. 10, p. 045055, Dec. 2025.
- [56] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Full reconstruction of a 14-qubit state within four hours,” *New Journal of Physics*, vol. 18, p. 083036, Aug. 2016.

- [57] K. E. Lestari, U. S. Pasaribu, S. W. Indratno, and H. Garminia, “Generating roots of cubic polynomials by Cardano’s approach on correspondence analysis,” *Heliyon*, vol. 6, p. e03998, June 2020.
- [58] I. Devnath, “Trigonometric Solution to the General Cubic Equation,” 2024.
- [59] H. Chen, “Schmidt numbers of low-rank bipartite mixed states,” *Physical Review A*, vol. 67, p. 062301, June 2003.
- [60] B. M. Terhal and P. Horodecki, “Schmidt number for density matrices,” *Physical Review A*, vol. 61, p. 040301, Mar. 2000.

Chapter 6

Appendix

A Why Eq. (3.4) uses minimization

Consider a non-extremal instrument $\mathbf{E}$, which can be written as

$$\mathbf{E} = \lambda \mathbf{E}^{(1)} + (1 - \lambda) \mathbf{E}^{(2)}, \quad 0 < \lambda < 1 \quad (1)$$

with $\mathbf{E}^{(1)} \neq \mathbf{E}^{(2)}$. Due to the concavity of v , we have

$$v(\mathbf{E}) \geq \lambda v(\mathbf{E}^{(1)}) + (1 - \lambda) v(\mathbf{E}^{(2)}). \quad (2)$$

Since the convex mixture is at least the minimum of the two $\mathbf{E}^{(1)}$ and $\mathbf{E}^{(2)}$,

$$v(\mathbf{E}) \geq \min\{v(\mathbf{E}^{(1)}), v(\mathbf{E}^{(2)})\}. \quad (3)$$

Hence, at least one of $\mathbf{E}^{(1)}$ or $\mathbf{E}^{(2)}$ satisfies $v(\mathbf{E}^{(i)}) \leq v(\mathbf{E})$. This means that any non-extremal POVM can be decomposed and moved to one of its components without increasing the objective value. Iterating the argument pushes the minimization all the way to an extremal POVM

$$\min_{\mathbf{E} \in \mathcal{P}_E} v(\mathbf{E}) = \min_{\mathbf{E} \text{ extremal}} v(\mathbf{E}). \quad (4)$$

B Proof of redundancy of classical post-processing of outcomes in the PVM-simulation

Proof. ¹ As seen in Eq. (3.1), a PVM-simulable POVM $\mathbf{E} = \{E_a\}_a$ can be decomposed as:

$$E_a = \sum_{\lambda} q_{\lambda} \sum_k p(a|k, \lambda) P_{k|\lambda}, \quad (5)$$

where λ is a random variable with probability distribution q_{λ} , $P_{k|\lambda}$ is a PVM for a given λ , and $p(a|k, \lambda)$ is a probabilistic post-processing rule. In this section, we show that the classical post-processing step in the simulation equation above can be neglected without loss of generality.

¹This proof is based on [34].

First, consider the case of deterministic post-processing, i.e., $p(a|k, \lambda) = \delta_{a,f(k,\lambda)}$, where f is some function. Then, the simulation equation is given by

$$E_a = \sum_{\lambda} q_{\lambda} \sum_k P_{k|\lambda} \delta_{a,f(k,\lambda)}. \quad (6)$$

We can define

$$F_{a|\lambda} := \sum_k P_{k|\lambda} \delta_{a,f(k,\lambda)}, \quad (7)$$

where $F_{a|\lambda}$ is implicitly dependent on f .

To verify that $\{F_{a|\lambda}\}_a$ is a projective measurement for any choice of (λ, f) , fix (λ, f) and define the set of original outcomes that are relabeled to a as

$$S_a^{(\lambda,f)} := \{k : f(k, \lambda) = a\}. \quad (8)$$

Consequently,

$$F_{a|\lambda} = \sum_{k \in S_a^{(\lambda,f)}} P_{k|\lambda}. \quad (9)$$

Since $\{P_{k|\lambda}\}_k$ is a projective measurement, we have

$$P_{k|\lambda} P_{\ell|\lambda} = \delta_{k\ell} P_{k|\lambda}, \quad \sum_k P_{k|\lambda} = \mathbb{1}. \quad (10)$$

We can now check the PVM conditions. First, each $F_{a|\lambda}$ is a projector:

$$F_{a|\lambda}^2 = \left(\sum_{k \in S_a} P_{k|\lambda} \right) \left(\sum_{\ell \in S_a} P_{\ell|\lambda} \right) = \sum_{k, \ell \in S_a} P_{k|\lambda} P_{\ell|\lambda} = \sum_{k \in S_a} P_{k|\lambda} = F_{a|\lambda}. \quad (11)$$

It is also Hermitian because it is a sum of Hermitian projectors:

$$F_{a|\lambda}^\dagger = \left(\sum_{k \in S_a} P_{k|\lambda} \right)^\dagger = \sum_{k \in S_a} P_{k|\lambda}^\dagger = \sum_{k \in S_a} P_{k|\lambda} = F_{a|\lambda}. \quad (12)$$

Second, distinct outcomes are orthogonal. If $a \neq b$, then

$$S_a \cap S_b = \emptyset. \quad (13)$$

Hence,

$$F_{a|\lambda} F_{b|\lambda} = \sum_{k \in S_a} \sum_{\ell \in S_b} P_{k|\lambda} P_{\ell|\lambda} = 0. \quad (14)$$

Third, the measurement is complete:

$$\sum_a F_{a|\lambda} = \sum_a \sum_k P_{k|\lambda} \delta_{a,f(k,\lambda)} = \sum_k P_{k|\lambda} \sum_a \delta_{a,f(k,\lambda)} = \sum_k P_{k|\lambda} = \mathbb{1}. \quad (15)$$

Hence, $\{F_{a|\lambda}\}_a$ is a projective measurement.

With the definition in Eq. (7)

$$E_a = \sum_{\lambda} q_{\lambda} F_{a|\lambda}. \quad (16)$$

Here, the post-processing is fully removed and the POVM is written as a stochastic mixture of the projectors $F_{a|\lambda}$.

We now allow a general classical post-processing step. Any stochastic post-processing map can be decomposed as a convex combination of deterministic relabelings. Hence, for suitable deterministic functions g and probabilities q_g , we may write

$$p(a|k, \lambda) = \sum_g q_g \delta_{a,g(k,\lambda)}. \quad (17)$$

Substituting this expression into the simulation equation gives

$$E_a = \sum_{\lambda} q_{\lambda} \sum_k P_{k|\lambda} p(a|k, \lambda) = \sum_{\lambda} q_{\lambda} \sum_{k,g} P_{k|\lambda} q_g \delta_{a,g(k,\lambda)} = \sum_{\lambda,g} q_{\lambda} q_g \sum_k P_{k|\lambda} \delta_{a,g(k,\lambda)}. \quad (18)$$

Define

$$q_{\lambda,g} := q_{\lambda} q_g, \quad G_{a|\lambda,g} := \sum_k P_{k|\lambda} \delta_{a,g(k,\lambda)}. \quad (19)$$

Then the simulation can be written as

$$E_a = \sum_{\lambda,g} q_{\lambda,g} G_{a|\lambda,g}. \quad (20)$$

For each fixed pair (λ, g) , the collection $\{G_{a|\lambda,g}\}_a$ is again a projective measurement, since it is obtained from $\{P_{k|\lambda}\}_k$ only by grouping outcomes according to the deterministic rule g . Finally, by absorbing g into the classical variable and defining

$$\lambda' := (\lambda, g), \quad (21)$$

we obtain

$$E_a = \sum_{\lambda'} q_{\lambda'} G_{a|\lambda'}. \quad (22)$$

Thus, the stochastic post-processing can be incorporated into the classical randomization over projective measurements. Consequently, the explicit post-processing step may be omitted without loss of generality. $\square$

C Derivation of $v_{\text{white}}(\theta)$

The Bloch decomposition of a PVM is given by [36]:

$$E_a = \lambda_a (\mathbb{1} + \vec{n}_a \cdot \hat{\sigma}), \quad (23)$$

where $|\vec{n}_a| \leq 1$ and $\lambda_a \geq 0$ to ensure operator positivity [36]. For normalization, it is required that:

$$\mathbb{1} = \sum_a E_a = \mathbb{1} \sum_a \lambda_a + \hat{\sigma} \cdot \sum_a \lambda_a \vec{n}_a. \quad (24)$$

The above conditions lead to the following POVM constraints:

$$\sum_i \lambda_i = 1, \quad (25)$$

$$\sum_a \lambda_a \vec{n}_a = 0. \quad (26)$$

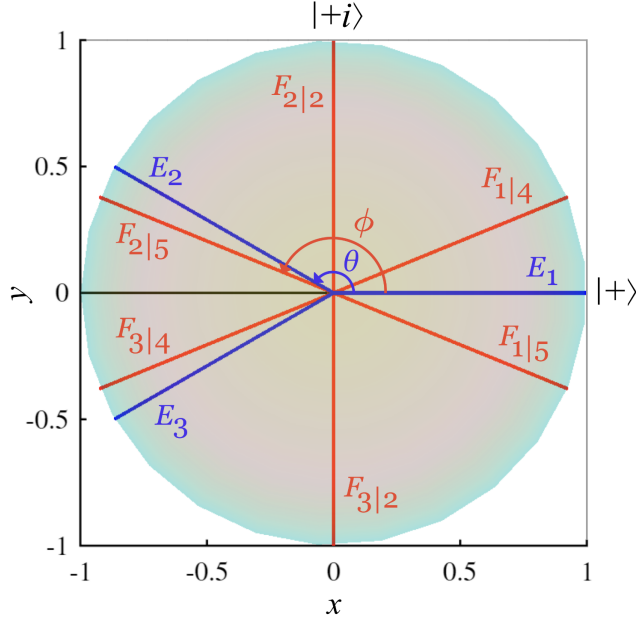


Figure 1: Single-angle parameterization of $F_{a|r_a}$ -vectors in the xy -plane, defining the angles between the POVM (blue) Bloch vectors and the (red) F -vectors. The Bloch vectors are normalized to unit vectors for simplicity. M_2 and M_3 have equal weight λ_2^M .

The POVM to be simulated is a three-outcome POVM defined by an angle θ , as shown in Fig. 1. With this parameterization, we assign the weights for the POVM E and PVMs F as λ_a^M and λ_a^F , respectively. Due to symmetry, it can be assumed that the Bloch vectors for E_2 and E_3 have equal weights, denoted λ_2 . Moreover, the weights for $F_{1|4}, F_{1|5}, F_{2|5}$ and $F_{3|4}$ share a weight λ_1^F , while $F_{2|2}$ and $F_{3|2}$ share a different weight λ_2^F , as indicated by the SDP results. It can also be noted from the SDP that

$$\sum_{\vec{r}} q(\vec{r}) = 1 \implies \lambda_2^F = \frac{1 - 4\lambda_1^F}{2}. \quad (27)$$

Furthermore, Eq. (25) implies

$$\lambda_2^E = \frac{1}{2}(1 - \lambda_1^E) \quad (28)$$

Together with Eq. (26), λ_1^E can be written in terms of θ in the following way²

$$\lambda_1^E \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1 - \lambda_1^E}{2} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} + \frac{1 - \lambda_1^E}{2} \begin{bmatrix} \cos(\theta) \\ -\sin(\theta) \end{bmatrix} = 0 \iff \lambda_1^E = -\frac{\cos(\theta)}{1 - \cos\theta}. \quad (29)$$

Drawing inspiration from the objective function of the SDP in Eq. (3.12), combined with Eq. (23), the following equation system holds

$$v\lambda_a^E(\mathbb{1} + \vec{n}_a^E \cdot \hat{\sigma}) + (1 - v)\lambda_a^E\mathbb{1} = \sum_{\vec{r}} \lambda_a^F(\mathbb{1} + \vec{n}_a^F \cdot \hat{\sigma}) \quad (30)$$

For $a = 1$, this becomes

$$\begin{bmatrix} \lambda_1^M & v\lambda_1^M \\ v\lambda_1^M & \lambda_1^M \end{bmatrix} = \begin{bmatrix} 2\lambda_1^F & -2\cos(\phi)\lambda_1^F \\ -2\cos(\phi)\lambda_1^F & 2\lambda_1^F \end{bmatrix}, \quad (31)$$

²The sign convention is the following: $+\cos(\theta)$ in the left half-plane and $+\sin(\theta)$ in the upper half-plane.

which is equivalent to

$$\lambda_1 = 2\lambda'_1 \quad (32)$$

$$v\lambda_1 = -\cos(\phi)\lambda_1 \iff v = -\cos\phi. \quad (33)$$

For simplicity, the Bloch vectors fulfill the following equation for all outcomes:

$$v\lambda_a^E \vec{n}_a^E = \sum_{\vec{r}} \vec{n}_a^F \lambda_a^F, \quad (34)$$

which for $a = 2$ leads to

$$v\lambda_2^E \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = \lambda_2^F \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_1^F \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix} = \lambda_2^F \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_1^F \begin{bmatrix} -v \\ \sqrt{1-v^2} \end{bmatrix}, \quad (35)$$

where Eq. (33) is used in the last equality³. The equation system above is equivalent to

$$\lambda_2^E \cos(\theta) = -\lambda_1^F \quad (36)$$

$$v\lambda_2^E \sin(\theta) = \lambda_2^F + \lambda_1^F \sqrt{1-v^2}. \quad (37)$$

The Bloch vector equation for $a = 3$ provides no additional information. Only v is unknown in Eq. (37), and can thereby be solved for. The idea is to reduce the equation into a quadratic by variable substitutions:

$$v\lambda_2^E \sin \theta = \lambda_1^F \sqrt{1-v^2} + \lambda_2^F \quad (38)$$

$$\iff \left(\underbrace{(\lambda_2^E \sin \theta)}_b v - \underbrace{\lambda_2^F}_c \right)^2 = \left(\underbrace{\lambda_1^F}_a \sqrt{1-v^2} \right)^2 \quad (39)$$

$$\iff \underbrace{(a^2 + b^2)}_A v^2 - \underbrace{2bc}_B v + \underbrace{c^2 - a^2}_C = 0 \quad (40)$$

Applying the quadratic formula to find the roots of Eq. (40) yields

$$v_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \quad (41)$$

where v_1 gives the correct solution. This is plotted in Fig. 3.6. Moreover, the minimum visibility is obtained at $\theta = 120^\circ$, forming a Mercedes-Benz POVM. The fact that all vectors are equally spaced by 30° gives $v_{\min} = \cos(\pi/6) = \sqrt{3}/2 \approx 0.8860$.

D Derivation of $v_{\text{ap}}(\theta)$

As the name suggests, the Bloch vectors of the antipodal noise point are antipodal to the corresponding Bloch vectors of the POVM, symmetrically separated by an angle θ , see Fig. 2. The weights for the Bloch vectors involved follow the same rules as the white-noise case.

³ $\sin(\phi) = +\sqrt{1-v^2}$ in the upper half of the unit circle.

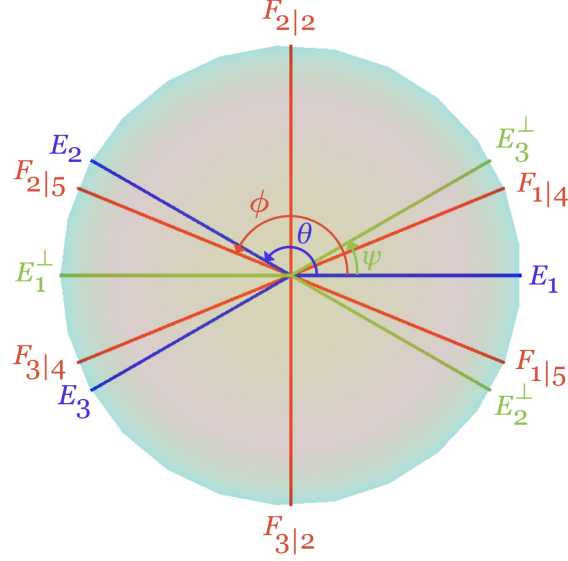


Figure 2: Single-angle parameterization of $F_{a|r_a}$ -vectors in the xy -plane, defining the angles between the POVM (blue) Bloch vectors and the (red) F -vectors. The Bloch vectors are normalized to unit vectors for simplicity. M_2 and M_3 have equal weight λ_2^M .

The term involving E_a^{noise} in the SDP in Eq. (3.12) is replaced by $E_a^\perp = 2\lambda_a^E \mathbb{1} - E_a$, leading to the following objective function

$$vE_a + (1-v)(2\lambda_a^E \mathbb{1} - E_a) = \sum_{\vec{r}} F_{a|\vec{r}} \quad (42)$$

$$\iff (2v-1)E_a + 2(1-v)\lambda_a^E \mathbb{1} = \sum_{\vec{r}} F_{a|\vec{r}}. \quad (43)$$

Analogously to the white noise case, the last equation above leads to an equation system, which for $a = 1$ is

$$\begin{bmatrix} \lambda_1^E(2v-1) & \lambda_1^E(2v-1) \\ \lambda_1^E(2v-1) & \lambda_1^E(2v-1) \end{bmatrix} + \begin{bmatrix} 2\lambda_1^E(1-v) & 0 \\ 0 & 2\lambda_1^E(1-v) \end{bmatrix} = \begin{bmatrix} 2\lambda_1^F & -2\lambda_1^F \cos \phi \\ -2\lambda_1^F \cos \phi & 2\lambda_1^F \end{bmatrix}, \quad (44)$$

which is equivalent to

$$2v\lambda_1^E - \lambda_1^E + 2\lambda_1^E - 2\lambda_1^E v = 2\lambda_1^F \iff \lambda_1^E = 2\lambda_1^F \quad (45)$$

$$(2v-1)\lambda_1^E = -2\lambda_1^F \cos \phi \quad (46)$$

Substituting Eq. (45) into (46) leads to

$$\cos \phi = 1 - 2v. \quad (47)$$

Furthermore, the Bloch vector equation for $a = 2$ is given by

$$(2v-1)\lambda_2^E \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = \lambda_2^F \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_1^F \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix}, \quad (48)$$

which leads to (one more independent equation?)

$$(2v-1)\lambda_2^E \cos(\theta) = \lambda_1^F \cos(\phi) \quad (49)$$

$$(2v-1)\lambda_2^E \sin(\theta) = \lambda_2^F + \lambda_1^F \sin(\phi). \quad (50)$$

Substituting Eq. (47) and $\sin(\phi) = \sqrt{1 - (1 - 2c)^2}$ into Eq. (50), a quadratic equation can be obtained through variable substitutions as for the case of white noise:

$$(2v - 1)\lambda_2^E \sin(\theta) = \lambda_2^F + \lambda_1^F \sqrt{1 - (1 - 2v)^2} \quad (51)$$

$$\iff \left(\frac{2v\lambda_2^E \sin(\theta) - \lambda_2^E \sin \theta - \lambda_2^F}{\lambda_1^F} \right)^2 = 1 - (1 - 2v)^2 \quad (52)$$

$$\iff \left(\underbrace{\frac{2\lambda_2^E \sin \theta}{\lambda_1^F}}_a v - \underbrace{\frac{\lambda_2^E \sin \theta + \lambda_2^F}{\lambda_1^F}}_b \right)^2 = 1 - (1 - 4v + 4v^2) \quad (53)$$

$$\iff a^2 v^2 - 2abv + b^2 = 4v - 4v^2 \quad (54)$$

$$\iff \underbrace{(a^2 + 4)}_A v^2 + \underbrace{(-2ab - 4)}_B v + \underbrace{b^2}_C = 0. \quad (55)$$

Using the quadratic formula gives

$$v_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \quad (56)$$

where v_1 gives the correct visibility curve as a function of θ .

E Derivation of $v_{wc}(\theta)$

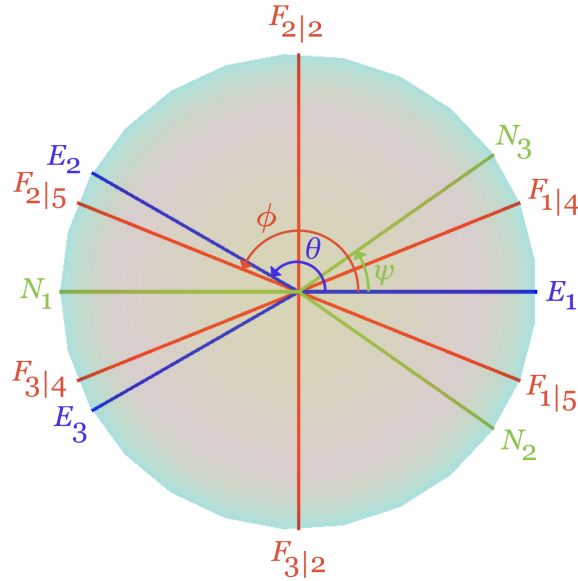


Figure 3: Single-angle parameterization of $F_{a|ra}$ -vectors in the xy -plane, defining the angles between the POVM (blue) Bloch vectors and the (red) F -vectors. The Bloch vectors are normalized to unit vectors for simplicity. M_2 and M_3 have equal weight λ_2^M .

Due to symmetries, $F_{1|4}$ and $F_{1|5}$ in Fig. 3 have equal weight, labeled λ_1^F while the remaining F -operators share the weight λ_2^F . Moreover, E_2 and E_3 share a common weights λ_2^E , while the weight for E_1 is denoted λ_1^E . The same argument holds for the weights for the worst-case noise Bloch vectors, but denoted with a superscript N .

Following Fig. 3, for the POVM (blue) and the noise Bloch vectors (green), the following prevails:

$$\sum_a \lambda_a = 1 \iff \lambda_2 = \frac{1 - \lambda_1}{2} \quad (57)$$

$$\sum_a \lambda_a \vec{n} = 0 \quad (58)$$

For the POVM vectors, n_1^E , Eqs (57)-(58) implies that

$$\begin{aligned} \lambda_1^E \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1 - \lambda_1^E}{2} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} + \frac{1 - \lambda_1^E}{2} \begin{bmatrix} \cos(\theta) \\ -\sin(\theta) \end{bmatrix} &= 0 \\ \iff \lambda_1 + (1 - \lambda_1^E) \cos \theta &\iff \lambda_1^E (1 - \cos(\theta)) = -\cos(\theta) \iff \lambda_1^E = -\frac{\cos(\theta)}{1 - \cos \theta}. \end{aligned} \quad (59)$$

Hence, λ_1^E and λ_2^E are determined by the angle θ in Fig. 3. Similarly, for the n_2^N Bloch vectors:

$$\begin{aligned} \lambda_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \frac{1 - \lambda_1}{2} \begin{bmatrix} \cos(\psi) \\ \sin(\psi) \end{bmatrix} + \frac{1 - \lambda_1}{2} \begin{bmatrix} \cos(\psi) \\ -\sin(\psi) \end{bmatrix} &= 0 \\ \iff -\lambda_1^N + (1 - \lambda_1) \cos(\psi) &\iff \lambda_1^N (1 + \cos(\psi)) = \cos(\psi) \iff \lambda_1^N = \frac{\cos(\psi)}{1 + \cos(\psi)}. \end{aligned} \quad (60)$$

As noted from the SDP results, we further have

$$\lambda_2^F = \frac{1 - 4\lambda_1^F}{2}. \quad (61)$$

Using the Bloch vector equation for $a = 1$ gives

$$v\lambda_1^E \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1 - v)\lambda_1^N \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \lambda_1^F \begin{bmatrix} -\cos(\phi) \\ \sin(\phi) \end{bmatrix} + \lambda_1^F \begin{bmatrix} -\cos(\phi) \\ -\sin(\phi) \end{bmatrix}, \quad (62)$$

which leads to the equation

$$v\lambda_1^E - (1 - v)\lambda_1^N \cos(\psi) = -2\cos(\phi). \quad (63)$$

Analogously, for $a = 2$

$$v\lambda_2^E \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} + (1 - v)\lambda_2^N \begin{bmatrix} \cos \psi \\ -\sin \psi \end{bmatrix} = \lambda_2^F \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_1^F \begin{bmatrix} \cos \phi \\ \sin \phi \end{bmatrix}, \quad (64)$$

leads to the set of equations:

$$\lambda_2^E v \cos \theta + \lambda_2^N (1 - v) \cos \psi = \lambda_1^F \cos \phi \quad (65)$$

$$\lambda_2^E v \sin \theta - \lambda_2^N (1 - v) \sin \psi = \lambda_2^F + \lambda_1^F \sin(\phi). \quad (66)$$

The Bloch vector equation for $a = 3$ does not yield any new information. Since Eqs (63) and (65) are equivalent, there are three independent equations but four unknowns $(\psi, \lambda_1^F, v, \phi)$. However, using optimality, a fourth equation will later appear. The Bloch

vector equation for $a = 3$ gives no extra information.

To make expressions cleaner, we start by making the following definitions:

$$A := v\lambda_1^E + (1-v)\lambda_1^N = 2\lambda_1^F \quad (67)$$

$$X := v\lambda_1^E - (1-v)\lambda_1^N = -2\lambda_1^F \cos \phi \quad (68)$$

$$Y := v\lambda_2^E \sin \theta - (1-v)\lambda_2^N \sin \psi = \lambda_2^F + \lambda_1^F \sin \phi, \quad (69)$$

where:

$$\lambda_2^F = \frac{1 - 4\lambda_1^F}{2}. \quad (70)$$

Eliminating $\lambda_1^F, \lambda_2^F, \phi$ gives:

$$\lambda_1^F = \frac{A}{2}, \quad \cos \phi = -\frac{X}{A}. \quad (71)$$

Consequently,

$$\sin \phi = \sqrt{1 - \frac{X^2}{A^2}}, \quad (72)$$

which, substituted into Eq. (69) gives:

$$Y = \frac{1 - 4\lambda_1^F}{2} + \lambda_1^F \sin \phi = \frac{1}{2} - A + \frac{A}{2} \sqrt{1 - \frac{X^2}{A^2}} = \frac{1}{2} - A + \frac{1}{2} \sqrt{A^2 - X^2}. \quad (73)$$

Next, we compute $A^2 - X^2$. From Eq. (67)-(68):

$$A^2 - X^2 = (v\lambda_1^E + (1-v)\lambda_1^N)^2 - (v\lambda_1^E - (1-v)\lambda_1^N)^2 = 4v(1-v)\lambda_1^E \lambda_1^N. \quad (74)$$

Hence, Eq. (69) becomes:

$$v\lambda_2^E \sin \theta - (1-v)\lambda_2^N \sin \psi = \frac{1}{2} - (v\lambda_1^E + (1-v)\lambda_1^N) + \sqrt{v(1-v)\lambda_1^E \lambda_1^N}. \quad (75)$$

Now, define:

$$x := \cot \frac{\theta}{2}, \quad t := \tan \frac{\psi}{2}. \quad (76)$$

Furthermore, $\frac{\pi}{2} < \theta < \pi$ implies $0 < x < 1$. From half-angle identities:

$$\cos \theta = \frac{x^2 - 1}{x^2 + 1}, \quad \sin \theta = \frac{2x}{x^2 + 1}, \quad (77)$$

$$\cos \psi = \frac{1 - t^2}{1 + t^2}, \quad \sin \psi = \frac{2t}{1 + t^2} \quad (78)$$

substitutions can be made for the purpose of simplifying the expressions for the weights involved.

Starting with λ_a^E . Combining

$$\lambda_1^E = -\frac{\cos \theta}{1 - \cos \theta}. \quad (79)$$

with

$$1 - \cos \theta = 1 - \frac{x^2 - 1}{x^2 + 1} = \frac{2}{x^2 + 1}, \quad (80)$$

λ_1^E can be rewritten as:

$$\lambda_1^E = -\frac{(x^2 - 1)/(x^2 + 1)}{2/(x^2 + 1)} = -\frac{x^2 - 1}{2} = \frac{1 - x^2}{2}. \quad (81)$$

Then,

$$\lambda_2^E = \frac{1 - \lambda_1^E}{2} = \frac{1 - (1 - x^2)/2}{2} = \frac{1 + x^2}{4}. \quad (82)$$

Moreover,

$$\lambda_2^E \sin \theta = \frac{1 + x^2}{4} \cdot \frac{2x}{1 + x^2} = \frac{x}{2}. \quad (83)$$

Then, the weights λ_a^N can be treated analogously. Using

$$\lambda_1^N = \frac{\cos \psi}{1 + \cos \psi}, \quad (84)$$

with

$$1 + \cos \psi = 1 + \frac{1 - t^2}{1 + t^2} = \frac{2}{1 + t^2}, \quad (85)$$

the weight λ_1^N becomes:

$$\lambda_1^N = \frac{(1 - t^2)/(1 + t^2)}{2/(1 + t^2)} = \frac{1 - t^2}{2}. \quad (86)$$

Consequently,

$$\lambda_2^N = \frac{1 - \lambda_1^N}{2} = \frac{1 + t^2}{4}. \quad (87)$$

Furthermore:

$$\lambda_2^N \sin \psi = \frac{1 + t^2}{4} \cdot \frac{2t}{1 + t^2} = \frac{t}{2}. \quad (88)$$

When substituting everything into Eq. 75, the LHS becomes:

$$v\lambda_2^E \sin \theta - (1 - v)\lambda_2^N \sin \psi = \frac{vx}{2} - \frac{(1 - v)t}{2}, \quad (89)$$

while the RHS is given by:

$$\begin{aligned} & \frac{1}{2} - (v\lambda_1^E + (1 - v)\lambda_1^N) + \sqrt{v(1 - v)\lambda_1^E\lambda_1^N} \\ &= \frac{1}{2} - \frac{v(1 - x^2)}{2} - \frac{(1 - v)(1 - t^2)}{2} + \sqrt{\frac{v(1 - v)(1 - x^2)(1 - t^2)}{4}}. \end{aligned} \quad (90)$$

By multiplying both sides by 2, Eq. 75 becomes:

$$vx - (1 - v)t = 1 - v(1 - x^2) - (1 - v)(1 - t^2) + \sqrt{v(1 - v)(1 - x^2)(1 - t^2)}. \quad (91)$$

Simplifying the non-square-root part:

$$1 - v(1 - x^2) - (1 - v)(1 - t^2) = 1 - v + vx^2 - (1 - v) + (1 - v)t^2 = vx^2 + (1 - v)t^2, \quad (92)$$

yields

$$vx - (1 - v)t = vx^2 + (1 - v)t^2 + \sqrt{v(1 - v)(1 - x^2)(1 - t^2)}. \quad (2)$$

After rearranging, a reduced scalar equation is obtained:

$$vx(1 - x) - (1 - v)t(1 + t) = \sqrt{v(1 - v)(1 - x^2)(1 - t^2)}. \quad (93)$$

Replacing v by $s := \sqrt{\frac{v}{1-v}}$ leads to the following variable substitutions:

$$v = \frac{s^2}{1 + s^2}, \quad 1 - v = \frac{1}{1 + s^2}, \quad \sqrt{v(1 - v)} = \frac{s}{1 + s^2}. \quad (94)$$

Inserting this into Eq. (93) and multiplying through by $1 + s^2$ gives:

$$x(1-x)s^2 - t(1+t) = s\sqrt{(1-x^2)(1-t^2)} \iff x(1-x)s^2 - \sqrt{(1-x^2)(1-t^2)}s - t(1+t) = 0, \quad (95)$$

i.e., a quadratic equation in s . Furthermore, defining

$$S := \sqrt{(1-x^2)(1-t^2)}, \quad (96)$$

leads to

$$x(1-x)s^2 - Ss - t(1+t) = 0. \quad (97)$$

The positive root of the quadratic formula (95) is

$$s = \frac{S + \sqrt{S^2 + 4x(1-x)t(1+t)}}{2x(1-x)}, \quad (98)$$

and

$$v = \frac{s^2}{1 + s^2}, \quad (99)$$

which means that the visibility is a function of $x(\theta)$ and $t(\psi)$.

To obtain a solution of v that is only a function of v , we can recall that the maximum visibility is the objective of the SDP so we optimize v over t . The reason for this is that $v = \frac{s^2}{1+s^2}$ is strictly increasing in $s > 0$. Therefore, maximizing v is the same as maximizing s . When searching for stationary points of $s(t)$, differentiating Eq. (97) implicitly instead of differentiating Eq. (98) directly is a much cleaner approach. To this end, let

$$F(s, t) := x(1-x)s^2 - Ss - t(1+t) = 0 \quad (100)$$

and differentiate it with respect to t :

$$2x(1-x)ss' - S's - Ss' - (1+2t) = 0. \quad (101)$$

At a stationary point, $s' = 0$, so

$$-S's - (1+2t) = 0. \quad (102)$$

Since

$$S = \sqrt{(1-x^2)(1-t^2)}, \quad (103)$$

and x is fixed, we have

$$S' = \frac{1}{2} \frac{(1-x^2)(-2t)}{\sqrt{(1-x^2)(1-t^2)}} = \frac{t(1-x^2)}{S}. \quad (104)$$

Inserting the expression above into Eq. (102) gives:

$$\frac{t(1-x^2)}{S} s - (1+2t) = 0 \iff s = \frac{(1+2t)S}{t(1-x^2)} = \frac{(1+2t)\sqrt{(1-x^2)(1-t^2)}}{t(1-x^2)}. \quad (105)$$

Then,

$$s^2 = \frac{(1+2t)^2(1-t^2)}{t^2(1-x^2)} \quad (106)$$

can be substituted into $v = s^2/(1+s^2)$, giving:

$$v(x, t) = \frac{(1+2t)^2(1-t^2)}{(1+2t)^2(1-t^2) + t^2(1-x^2)}. \quad (107)$$

Furthermore, a cubic form can be derived for t . When substituting Eq. (105) into (95), we first compute:

$$Ss = S \cdot \frac{(1+2t)S}{t(1-x^2)} = \frac{(1+2t)(1-t^2)}{t}, \quad (108)$$

where Eq. (105) is used in the first equality and (103) in the last. Moreover,

$$x(1-x)s^2 = x(1-x) \frac{(1+2t)^2(1-t^2)}{t^2(1-x^2)} = \frac{x(1+2t)^2(1-t^2)}{t^2(1+x)}, \quad (109)$$

where Eq. (106) is used. Inserting both equations above into Eq. (95) gives:

$$\frac{x(1+2t)^2(1-t^2)}{t^2(1+x)} - \frac{(1+2t)(1-t^2)}{t} - t(1+t) = 0. \quad (110)$$

Multiplying this by $t^2(1+x)$ and factorizing out $-(1+t)$ leads to:

$$\begin{aligned} x(1+2t)^2(1-t^2) - t(1+x)(1+2t)(1-t^2) - t^3(1+x)(1+t) &= 0 \\ \iff -(1+t) [-x(1+2t)^2(1-t) + t(1+x)(1+2t)(1-t) + t^3(1+x)] &= 0 \\ \iff -(1+t) [(3x-1)t^3 + (x+1)t^2 + (1-2x)t - x] &= 0. \end{aligned} \quad (111)$$

Since $t \geq 0$, $1+t \neq 0$. Therefore, the stationary point satisfies

$$(3x-1)t^3 + (x+1)t^2 + (1-2x)t - x = 0, \quad (112)$$

which defines a cubic equation.

To solve the cubic, Cardano roots can be used. Eq. (112) can be solved in closed form. For cleaner expressions, the coefficients in Eq. 112 can be relabeled in the following way:

$$\underbrace{(3x-1)}_a t^3 + \underbrace{(x+1)}_b t^2 + \underbrace{(1-2x)}_c t - \underbrace{x}_d, \quad (113)$$

which defines the cubic equation as:

$$at^3 + bt^2 + ct + d = 0 \quad (114)$$

For the special point $x = \frac{1}{3}$ which makes $a = 3x - 1 = 0$, the cubic degenerates to a quadratic equation

$$\frac{4}{3}t^2 + \frac{1}{3}t - \frac{1}{3} = 0 \iff 4t^2 + t - 1 = 0, \quad (115)$$

whose positive root is given by

$$t = \frac{\sqrt{17} - 1}{8}. \quad (116)$$

For all other points ($x \neq \frac{1}{3}$), the cubic in Eq. (114) divided by 3, can be reduced to a depressed cubic by setting [57]:

$$t = u - \frac{b}{3a}, \quad (117)$$

which removes the quadratic term, i.e., the cubic takes on the form

$$u^3 + pu + q = 0, \quad (118)$$

where

$$p = \frac{3ac - b^2}{3a^2}, \quad q = \frac{27a^2d - 9abc + 2b^3}{27a^3}. \quad (119)$$

To compute p and q , we first compute

$$\begin{aligned} 3ac - b^2 &= 3(3x - 1)(1 - 2x) - (x + 1)^2 = -19x^2 + 13x - 4 =: \Delta_0 \\ \implies p &= -\frac{\Delta_0}{3a^2}. \end{aligned} \quad (120)$$

Similarly for q :

$$\begin{aligned} 27a^2d - 9abc + 2b^3 &= -187x^3 + 177x^2 - 57x + 11 =: \Delta_1 \\ \implies q &= \frac{\Delta_1}{27a^3}. \end{aligned} \quad (121)$$

Furthermore, the discriminant for Eq. 114 is given by:

$$\Delta = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 = \frac{\Delta_1^2 - 4\Delta_0^3}{2916a^6} = \frac{(x+1)(31x^3 - 51x^2 + 21x - 5)}{108(3x-1)^4}. \quad (122)$$

Since, $x+1 > 0$, $(3x-1)^4 > 0$ and $31x^3 - 51x^2 + 21x - 5 < 0$ for $0 < x < 1$, it follows that $\Delta < 0$. The negative sign of the discriminant implies that the equation has three distinct real roots, and thereby a trigonometric method to solving the cubic is appropriate [58]. To this end, it can be noted that the trigonometric identity

$$4\cos^3(\xi) + 3\cos(\xi) = \cos(3\xi), \quad (123)$$

also takes on the form of a depressed cubic. Therefore, u can be set to

$$u = 2\sqrt{-\frac{p}{3}}\cos(\xi). \quad (124)$$

To see this, we seek a substitution of the form:

$$u = A \cos(\xi). \quad (125)$$

Substituting the ansatz above into the depressed cubic leads to

$$A^3 \cos^3 \xi + pA \cos(\xi) + q = 0 \iff \frac{A^3}{4} \cos^3 \xi + pA \cos(\xi) = \frac{A^3}{4} \left(4 \cos^3 \xi + \frac{4p}{A^2} \cos(\xi) \right). \quad (126)$$

Comparing to the trigonometric identity

$$4 \cos^3(\xi) + 3 \cos(\xi) = \cos(3\xi), \quad (127)$$

we require

$$\frac{4p}{A^2} = -3 \implies A^2 = -\frac{4p}{3} \iff A = \pm 2\sqrt{-\frac{p}{3}}, \quad (128)$$

in order to make the bracket in Eq. (126) equal to $4 \cos^3 \xi - 3 \cos(\xi)$. Hence, $u = 2\sqrt{-\frac{p}{3}} \cos(\xi) = -\frac{2\sqrt{\Delta_0}}{3a} \cos(\xi)$. Inserting this into Eq. 118 yields

$$\begin{aligned} & -\frac{8\Delta_0^{3/2}}{27a^3} \cos^3(\xi) + \frac{6\Delta_0^{3/2}}{27a^3} \cos(\xi) + \frac{\Delta_1}{27a^3} = 0 \\ \iff & \frac{1}{27a^3} \left[\left(-2\Delta_0^{3/2} \right) (4 \cos^3(\xi) - 3 \cos(\xi)) + \Delta_1 \right] = 0 \\ \iff & -2\Delta_0^{3/2} \cos(3\xi) + \Delta_1 = 0 \\ \iff & \cos(3\xi) = \frac{\Delta_1}{2\Delta_0^{3/2}}. \end{aligned} \quad (129)$$

By defining $\Phi = \arccos\left(\frac{\Delta_1}{2\Delta_0^{3/2}}\right)$, the last equality above has the solution

$$\xi_k = \frac{\Phi + 2\pi k}{3}, \quad k = 0, \pm 1, \quad (130)$$

are the three real solutions.

Putting all of this together, the solution for t in Eq. 112 is

$$t = u - \frac{b}{3a} = -\frac{b + 2\sqrt{\Delta_0} \cos\left(\frac{\Phi + 2\pi k}{3}\right)}{3a}, \quad k = 0, \pm 1, \quad (131)$$

where $\Phi = \arccos\left(\frac{\Delta_1}{2\Delta_0^{3/2}}\right)$. Since the root in question should satisfy $t = \tan \frac{\psi}{2} \in [0, 1]$. In the physical domain $0 < x < 1$, t becomes a piecewise function:

$$t_*(x) = \begin{cases} -\frac{(x+1) + 2\sqrt{\Delta_0} \cos\left(\frac{\Phi - 2\pi}{3}\right)}{3(3x-1)}, & 0 < x < \frac{1}{3}, \\ \frac{\sqrt{17}-1}{8}, & x = \frac{1}{3}, \\ -\frac{(x+1) + 2\sqrt{\Delta_0} \cos\left(\frac{\Phi + 2\pi}{3}\right)}{3(3x-1)}, & \frac{1}{3} < x < 1. \end{cases} \quad (132)$$

Hence, the closed form of the visibility is

$$v(\theta) = \frac{(1 + 2t_*)^2(1 - t_*^2)}{(1 + 2t_*)^2(1 - t_*^2) + t_*^2(1 - x^2)}, \quad x = \cot \frac{\theta}{2}. \quad (133)$$

Moreover, the minimum visibility is obtained for the Mercedes-Benz POVM ($\theta = 120^\circ$). It is obtained from Bloch vector equation

$$vM_a + (1 - v)N_a = \sum_{\vec{r}} F_a|\vec{r} \quad (134)$$

For $a = 1$:

$$\frac{1}{3}v \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{3}(1 - v) \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} \cos \frac{\pi}{6} \\ \sin \frac{\pi}{6} \end{bmatrix} + \frac{1}{6} \begin{bmatrix} \cos \frac{\pi}{6} \\ -\sin \frac{\pi}{6} \end{bmatrix} \quad (135)$$

$$\iff v = \frac{\cos \frac{\pi}{6} + 1}{2} \iff v_{\min} = \frac{\sqrt{3} + 2}{4}. \quad (136)$$

F Proof of why $\theta + \psi \geq \pi$

Recall Eq. (3.28):

$$\sin(\varphi) = \frac{-1 - \cos(\psi) - \cos(\theta) - \cos(\psi) \cos(\theta)}{\sin(\psi) \sin(\theta)}. \quad (137)$$

After factorizing the numerator, the equation can be rewritten as

$$\sin(\varphi) = -\frac{(1 + \cos(\psi))(1 + \cos(\theta))}{\sin(\psi) \sin(\theta)}. \quad (138)$$

Using the half-angle identities

$$1 + \cos x = 2 \cos^2 \left(\frac{x}{2} \right), \quad \sin x = 2 \sin \left(\frac{x}{2} \right) \cos \left(\frac{x}{2} \right), \quad (139)$$

gives

$$\frac{1 + \cos x}{\sin x} = \frac{2 \cos^2(x/2)}{2 \sin(x/2) \cos(x/2)} = \cot \left(\frac{x}{2} \right). \quad (140)$$

Applying this for both θ and ψ gives

$$\sin(\varphi) = -\cot \left(\frac{\psi}{2} \right) \cot \left(\frac{\theta}{2} \right) = -\frac{1}{\tan \left(\frac{\psi}{2} \right) \tan \left(\frac{\theta}{2} \right)}. \quad (141)$$

Since $\cot(\psi/2), \cot(\theta/2), \tan(\psi/2), \tan(\theta/2) > 0$ in the domain $(0, \pi)$, we have the following equivalences

$$|\sin \varphi| \leq 1 \iff \cot \left(\frac{\psi}{2} \right) \cot \left(\frac{\theta}{2} \right) \leq 1 \iff \tan \left(\frac{\psi}{2} \right) \tan \left(\frac{\theta}{2} \right) \geq 1. \quad (142)$$

From the last equivalence, we have

$$\tan \left(\frac{\theta}{2} \right) \geq \frac{1}{\tan(\psi/2)} = \cot(\psi/2) = \tan \left(\frac{\pi}{2} - \frac{\psi}{2} \right), \quad (143)$$

where $\cot x = \tan(\pi/2 - x)$ has been used. Finally,

$$\tan \left(\frac{\theta}{2} \right) \geq \tan \left(\frac{\pi}{2} - \frac{\psi}{2} \right). \quad (144)$$

Since $\tan x$ is a strictly increasing function on the interval $(0, \pi/2)$ and $(\theta, \psi) \in (0, \pi)$, we can conclude that $\theta + \psi \geq \pi$.

G Derivation of $v_{\text{white}}(\theta, \psi)$

G.1 Preliminaries

As earlier discussed, for white noise, the objective function in the SDP in Eq. (3.12) is for the qubit case given by

$$vE_a + (1-v)\frac{\text{Tr}(E_a)}{2}\mathbb{1} = \frac{1}{4}(\mathbb{1} + v\vec{w}_a \cdot \hat{\sigma}), \quad (145)$$

where $E_a = \frac{1}{4}(\mathbb{1} + \vec{w}_a \cdot \hat{\sigma})$ and $\text{Tr}(E_a) = \frac{1}{2}$ were used in the second equality. Hence, white noise shrinks the Bloch vectors by a factor of v .

As discussed earlier, the constraint $|v_4| = 1$ leads to $\sin(\phi) = -\cot(\psi/2)\cot(\theta/2)$ where the non-degenerate extremal region is $\psi + \theta > \pi$. To proceed, it is convenient to reduce the POVM defined in Fig. 3.7 to a standard tetrahedron form. Any centered equal-weight rank-one four-outcome qubit POVM forms a tetrahedron whose opposite edges are equal, i.e., the four Bloch vectors defined in Eq. (3.27) can be written as

$$\vec{w}_1 = (x, y, z), \quad \vec{w}_2 = (x, -y, -z), \quad \vec{w}_3 = (-x, y, -z), \quad \vec{w}_4 = (-x, -y, z), \quad (146)$$

up to rotations. The four vectors defined above have the same POVM geometry as the original parameterization and satisfy

$$\sum_{a=1}^4 \vec{w}_a = 0. \quad (147)$$

The Bloch vectors in Eq. (146) are simply rotated into a more convenient coordinate system. Since unitary transformations such as rotations preserve scalar products [42], we have

$$\vec{w}_a \cdot \vec{w}_b = \vec{v}_a \cdot \vec{v}_b, \quad \forall (a, b) = 1, \dots, 4. \quad (148)$$

By comparing inner products, we can relate x, y , and z to θ and ψ . To this end, the inner product

$$\vec{w}_1 \cdot \vec{w}_2 = x^2 - y^2 - z^2 = 2x^2 - 1, \quad (149)$$

where $x^2 + y^2 + z^2 = 1$ is used in the second equality, can be compared with

$$\vec{v}_1 \cdot \vec{v}_2 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \sin(\psi) & \cos(\psi) \end{bmatrix}^T = \cos(\psi). \quad (150)$$

Hence, the vectors $\vec{w}_1$ and $\vec{w}_2$ corresponds to an angle ψ

$$\vec{w}_1 \cdot \vec{w}_2 = \cos(\psi) = 2x^2 - 1 \iff x^2 = \frac{1 + \cos(\psi)}{2}. \quad (151)$$

From the half-angle identity $\cos^2\left(\frac{\psi}{2}\right) = \left(\frac{1 + \cos(\psi)}{2}\right)$ we get that

$$x^2 = \cos^2\left(\frac{\psi}{2}\right) \implies x = \cos\left(\frac{\psi}{2}\right), \quad (152)$$

where the positive root is chosen since $\psi/2 \in (0, \pi/2)$. The same arguments apply to the inner product $(\vec{w}_1, \vec{w}_3) = (\vec{v}_1, \vec{v}_3)$, which leads to the following relationship between y and θ

$$y = \cos\left(\frac{\theta}{2}\right). \quad (153)$$

Lastly, to get a relationship between z and the angles in Fig. 3.7, the normalization condition

$$x^2 + y^2 + z^2 = 1, \quad (154)$$

can be used. Substituting $x^2 = \frac{1+\cos(\psi)}{2}$ and $y^2 = \frac{1+\cos(\theta)}{2}$ into Eq. (154) gives

$$z^2 = 1 - \frac{1+\cos(\psi)}{2} - \frac{1+\cos(\theta)}{2} \implies z = \sqrt{-\frac{\cos(\psi) + \cos(\theta)}{2}}, \quad (155)$$

where the positive root is chosen since the sign of z can be absorbed into the choice of coordinate axes w.l.o.g. (here, we simply choose $z \geq 0$ by convention). This transformation simplifies the geometry; instead of working directly with θ, ψ and φ , the tetrahedron is described fully by the positive numbers x, y, z subject to normalization.

G.2 Minimization problem

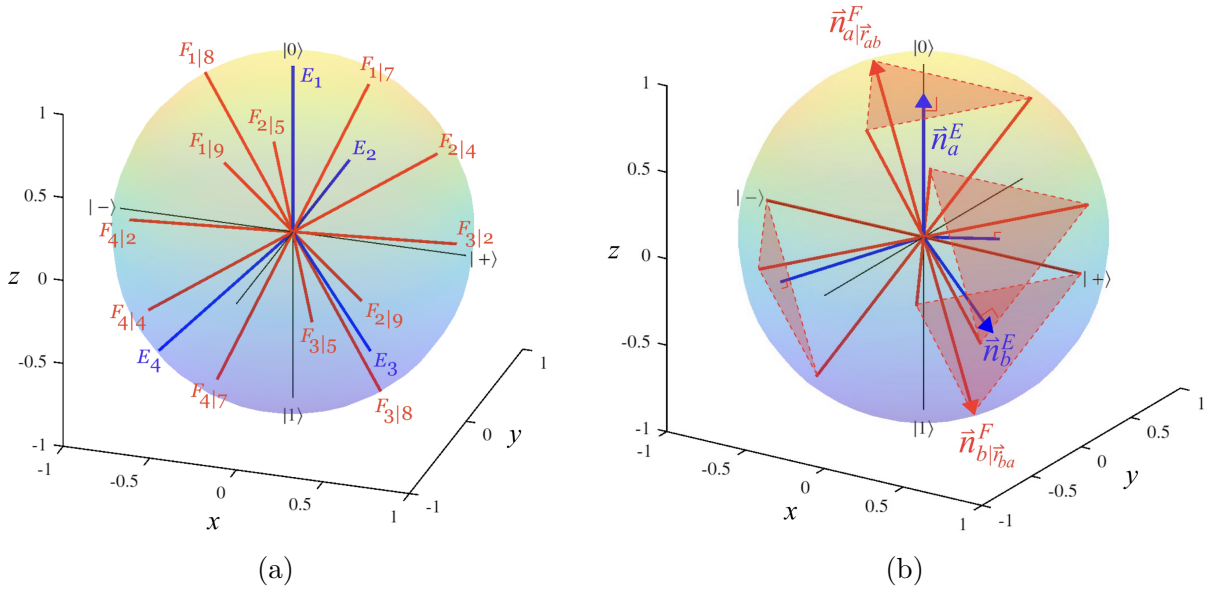


Figure 4: a) Bloch vectors corresponding to the POVM elements E_a (blue) and non-zero projective Bloch vectors contributions corresponding to $F_{a|\vec{r}}$ (red), plotted in the Bloch sphere. b) Convex hull boundary (shaded red) defined by the PVMs used to simulate each POVM outcome. Shows selected Bloch vectors. $\vec{r}_{ab} = \vec{r}_{ba}$.

It is helpful to think about what projective simulation means geometrically. In a four-outcome simulation, a PVM contributes to two POVM outcomes at a time, as Fig. 4a shows. We can further assign to each $F_{a|\vec{r}}$ Bloch vector, a vector g_{ab} , which is the contribution coming from a 2-outcome PVM acting on outcomes a and b , where $g_{ba} = -g_{ab}$. As Fig. 4a indicates, for a given rank-vector $\vec{r}$, there are two antipodal F -Bloch vectors acting on two different outcomes. The following mapping can be made:

$$n_{a|\vec{r}}^F \leftrightarrow g_{ab}, \quad n_{b|\vec{r}}^F \leftrightarrow g_{ba}, \quad (156)$$

where $\vec{r}_{ab}$ is the rank vector for the outcome pair (a, b) , see Fig. 4b. Hence,

$$F_{a|\vec{r}_{ab}} = \lambda_{a|\vec{r}_{ab}}^F (\mathbb{1} + \vec{n}_{a|\vec{r}_{ab}}^F \cdot \hat{\sigma}), \quad \vec{g}_{ab} := \lambda_{a|\vec{r}_{ab}}^F \vec{n}_{a|\vec{r}_{ab}}^F. \quad (157)$$

Fig. 4b shows the projective Bloch vector pair marked out in the simulation of outcome a and b . At outcome a , all incoming pair contributions must add up to the desired noisy Bloch vector:

$$\sum_{b \neq a} \vec{g}_{ab} = \frac{v}{4} \vec{w}_a, \quad (158)$$

where the factor $1/4$ comes from comparing to the effect defined in Eq. (145)

$$E_a^{(v)} = \frac{1}{4}(\mathbb{1} + v \vec{w}_a \cdot \vec{\sigma}). \quad (159)$$

Hence, for outcome $a = 1$, the equation above becomes

$$g_{12} + g_{13} + g_{14} = \frac{v}{4} \vec{w}_1. \quad (160)$$

A qubit effect of the form

$$F_{a|\vec{r}_{ab}} = \alpha_{ab} \mathbb{1} \pm \vec{g}_{ab} \cdot \hat{\sigma}, \quad b \neq a \quad (161)$$

is positive semidefinite if and only if

$$\alpha_{ab} \geq \|\vec{g}_{ab}\|. \quad (162)$$

To see this, consider the following. Noting that

$$\vec{g}_{ab} \cdot \hat{\sigma} = g_x \sigma_x + g_y \sigma_y + g_z \sigma_z \quad (163)$$

has eigenvalues $\pm \|\vec{g}_{ab}\|$ ($(\vec{g}_{ab} \cdot \hat{\sigma})^2 = \|\vec{g}_{ab}\|^2 \mathbb{1}$). Consequently, the operator

$$A := \alpha_{ab} \mathbb{1} + \vec{g}_{ab} \cdot \hat{\sigma} \quad (164)$$

has eigenvalues $\alpha_{ab} \pm \|\vec{g}_{ab}\|$. Since A is positive semidefinite, both eigenvalues must be nonnegative. Therefore, the smaller eigenvalue must satisfy

$$\alpha_{ab} - \|\vec{g}_{ab}\| \geq 0 \iff \alpha_{ab} \geq \|\vec{g}_{ab}\|. \quad \square \quad (165)$$

Hence, every vector contribution $\vec{g}_{ab}$ requires a minimum of $\|\vec{g}_{ab}\|$ of scalar weight:

$$F_{a|\vec{r}_{ab}} = \alpha_a \mathbb{1} + \vec{g}_{ab} \cdot \vec{\sigma} \succeq 0 \iff \alpha_{ab} \geq \|\vec{g}_{ab}\|. \quad (166)$$

The minimum possible scalar weight is obtained when $\alpha_{ab} = \|\vec{g}_{ab}\|$. Then, one of the eigenvalues is zero, hence $F_{a|\vec{r}_{ab}}$ is of rank one.

From the SDP, we have

$$\sum_{b \neq a} F_{a|\vec{r}_{ab}} = \left(\sum_{b \neq a} \alpha_{ab} \right) \mathbb{1} + \left(\sum_{b \neq a} \vec{g}_{ab} \right) \cdot \vec{\sigma}. \quad (167)$$

Comparing with:

$$E_a^{(v)} = \frac{1}{4} \mathbb{1} + \frac{v}{4} \vec{w}_a \cdot \vec{\sigma}, \quad (168)$$

the scalar parts must match

$$\sum_{b \neq a} \alpha_{ab} = \frac{1}{4}. \quad (169)$$

Using that $F_{a|\vec{r}_{ab}}$ is positive semidefinite, we get that

$$\alpha_{ab} \geq \|\vec{g}_{ab}\| \implies \sum_{b \neq a} \|\vec{g}_{ab}\| \leq \sum_{b \neq a} \alpha_{ab} = \frac{1}{4}. \quad (170)$$

Making a change of variables

$$\vec{h}_{ab} := \frac{4}{v} \vec{g}_{ab} \iff \vec{g}_{ab} := \frac{v}{4} \vec{h}_{ab} \quad (171)$$

and substituting into Eq. (158) leads to

$$\sum_{b \neq a} \frac{v}{4} \vec{h}_{ab} = \frac{v}{4} \vec{w}_a. \quad (172)$$

Further assuming $v > 0$, and dividing both sides by $v/4$ yields the Bloch vector balance equations (these tell you how the Bloch vector part of each noisy POVM effect is assembled from the pairwise projective pieces)

$$\sum_{b \neq a} \vec{h}_{ab} = \vec{w}_a, \quad (173)$$

and the scalar constraint

$$\frac{v}{4} \sum_{b \neq a} \|\vec{h}_{ab}\| \leq \frac{1}{4} \iff v \sum_{b \neq a} \|\vec{h}_{ab}\| \leq 1. \quad (174)$$

In the above equation, we can define the minimal total amount of scalar weight needed in order to realize the desired Bloch vectors by pairwise projective contributions in the following way

$$T := \min \sum_{b \neq a} \|\vec{h}_{ab}\|. \quad (175)$$

This is the total length of the rescaled branch vectors that are used to build the vectors $\vec{w}_a$. Then, from Eq. (174), the largest possible visibility is given by

$$v = \frac{1}{T}. \quad (176)$$

Therefore, the problem reduces to finding the smallest value of T needed to build the tetrahedron vectors $\vec{w}_a$, which maximizes the visibility, v .

By symmetry of the POVM family, it is sufficient⁴ to only consider outcome $a = 1$, where $\vec{w}_1 = (x, y, z)$. The three edges for outcome 1 are $(1, 2)$, $(1, 3)$, and $(1, 4)$. For $a = 1$, the corresponding projective rescaled Bloch vectors (see Eq. (173)) are given by the ansatz

$$\vec{h}_{12} = (0, \alpha, \beta), \quad \vec{h}_{13} = (\gamma, 0, \delta), \quad \vec{h}_{14} = (\varepsilon, \zeta, 0). \quad (177)$$

Figure 5 shows how these vectors are oriented in the Bloch sphere.

⁴This implies that $T_1 = T_2 = T_3 = T_4$. Since the full visibility is $v_{\text{total}} = \min_a \frac{1}{T_a} = \frac{1}{\max_a T_a}$, for a nonsymmetric POVM, we have $v_{\text{total}} = \frac{1}{T_1}$ here, since all T_a are equal.

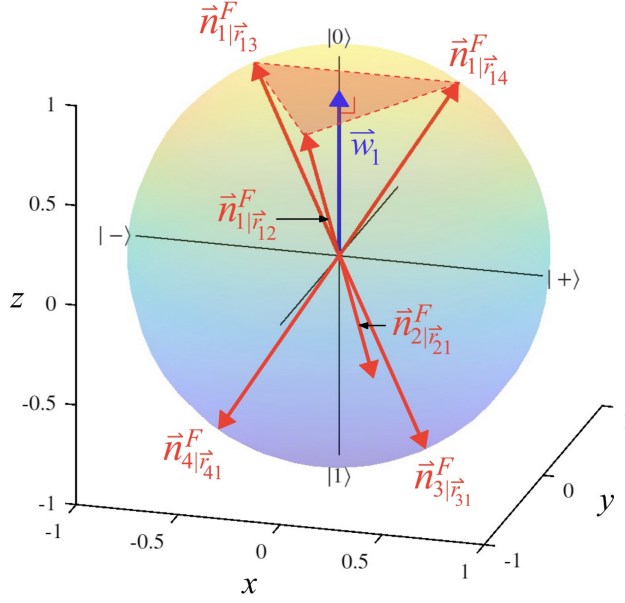


Figure 5: Orientation of vectors $\vec{n}_{1|\vec{r}_{12}}^F \parallel \vec{h}_{12}$, $\vec{n}_{1|\vec{r}_{13}}^F \parallel \vec{h}_{13}$, and $\vec{n}_{1|\vec{r}_{14}}^F \parallel \vec{h}_{14}$.

To see this, consider the rescaled Bloch vector $\vec{h}_{12}$ of fixed rank-vector $\vec{r}_{12}$ connecting outcomes 1 and 2 so that at outcome 1, it contributes $+\vec{h}_{12}$ and at outcome 2, the same branch contributes the opposite vector $\vec{h}_{21} = -\vec{h}_{12}$. So, the rank-vector $\vec{r}_{12}$ -branch creates opposite Bloch-vector contributions at outcomes 1 and 2. It can be noted that the two target Bloch vectors $\vec{w}_1 = (x, y, z)$ and $\vec{w}_2 = (x, -y, -z)$ have the same x -component. That means that the branch connecting outcomes 1 and 2 does not need to create any x -difference between them. The same logic holds for $\vec{h}_{13}$. Since $\vec{w}_1 = (x, y, z)$ and $\vec{w}_3 = (-x, y, -z)$ have the same y -component, we choose

$$\vec{h}_{13} = (\gamma, 0, \delta), \quad (178)$$

and for the remaining vector we have

$$\vec{h}_{14} = (\varepsilon, \zeta, 0). \quad (179)$$

by the same argument.

The condition

$$\vec{h}_{12} + \vec{h}_{13} + \vec{h}_{14} = \vec{w}_1 = (x, y, z) \quad (180)$$

leads to a set of equations:

$$\begin{aligned} \gamma + \varepsilon &= x, \\ \alpha + \zeta &= y, \\ \beta + \delta &= z. \end{aligned} \quad (181)$$

Moreover, the total length (see Eq. (175)) at outcome 1 is

$$T = \sqrt{\alpha^2 + \beta^2} + \sqrt{\gamma^2 + \delta^2} + \sqrt{\varepsilon^2 + \zeta^2}. \quad (182)$$

Hence, the problem in Eq. (175) is

$$T = \min \left[\sqrt{\alpha^2 + \beta^2} + \sqrt{\gamma^2 + \delta^2} + \sqrt{\varepsilon^2 + \zeta^2} \right], \quad (183)$$

subject to

$$\gamma + \varepsilon = x, \quad \alpha + \zeta = y, \quad \beta + \delta = z. \quad (184)$$

G.3 Solution to the minimization problem

To proceed, let

$$U := \sqrt{\alpha^2 + \beta^2}, \quad V := \sqrt{\gamma^2 + \delta^2}, \quad W := \sqrt{\varepsilon^2 + \zeta^2}. \quad (185)$$

The method of Lagrange multipliers can be employed. This method is used to minimize a function subject to some constraints [49]. The function to be minimized is

$$T = \sqrt{\alpha^2 + \beta^2} + \sqrt{\gamma^2 + \delta^2} + \sqrt{\varepsilon^2 + \zeta^2}, \quad (186)$$

with the constraints in Eq. (184). Define

$$\mathcal{L} = T + \lambda_x(\gamma + \varepsilon - x) + \lambda_y(\alpha + \zeta - y) + \lambda_z(\beta + \delta - z), \quad (187)$$

where $\lambda_x, \lambda_y, \lambda_z$ are auxiliary variables called Lagrange multipliers. At the optimum derivative of $\mathcal{L}$ with respect to every variable must vanish. Then, the complete Lagrange system is:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha} &= 0, & \frac{\partial \mathcal{L}}{\partial \beta} &= 0, & \frac{\partial \mathcal{L}}{\partial \gamma} &= 0, \\ \frac{\partial \mathcal{L}}{\partial \delta} &= 0, & \frac{\partial \mathcal{L}}{\partial \varepsilon} &= 0, & \frac{\partial \mathcal{L}}{\partial \zeta} &= 0, \end{aligned} \quad (188)$$

together with

$$\frac{\partial \mathcal{L}}{\partial \lambda_x} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda_y} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda_z} = 0. \quad (189)$$

With the definitions above, we have

$$T = U + V + W. \quad (190)$$

Now compute the derivatives

$$\begin{aligned} \frac{\partial U}{\partial \alpha} &= \frac{\alpha}{U}, & \frac{\partial U}{\partial \beta} &= \frac{\beta}{U}, \\ \frac{\partial V}{\partial \gamma} &= \frac{\gamma}{V}, & \frac{\partial V}{\partial \delta} &= \frac{\delta}{V}, \\ \frac{\partial W}{\partial \varepsilon} &= \frac{\varepsilon}{W}, & \frac{\partial W}{\partial \zeta} &= \frac{\zeta}{W}. \end{aligned} \quad (191)$$

Differentiating $\mathcal{L}$ for α and ζ :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha} &= \frac{\alpha}{U} + \lambda_y = 0, \\ \frac{\partial \mathcal{L}}{\partial \zeta} &= \frac{\zeta}{W} + \lambda_y = 0, \end{aligned} \quad (192)$$

both contain the same multiplier λ_y as both α and ζ appear in the same constraint $\alpha + \zeta = y$. From this

$$\frac{\alpha}{U} = \frac{\zeta}{W}. \quad (193)$$

Similarly,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \beta} &= \frac{\beta}{U} + \lambda_z = 0, \\ \frac{\partial \mathcal{L}}{\partial \delta} &= \frac{\delta}{V} + \lambda_z = 0, \end{aligned} \quad (194)$$

hence,

$$\frac{\beta}{U} = \frac{\delta}{V}. \quad (195)$$

Lastly,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \gamma} &= \frac{\gamma}{V} + \lambda_x = 0, \\ \frac{\partial \mathcal{L}}{\partial \varepsilon} &= \frac{\varepsilon}{W} + \lambda_x = 0, \end{aligned} \quad (196)$$

thus,

$$\frac{\gamma}{V} = \frac{\varepsilon}{W}. \quad (197)$$

In total, the lagrange equations are:

$$\frac{\alpha}{U} = \frac{\zeta}{W}, \quad \frac{\beta}{U} = \frac{\delta}{V}, \quad \frac{\gamma}{V} = \frac{\varepsilon}{W}. \quad (198)$$

Defining:

$$p := \frac{\alpha}{U} = \frac{\zeta}{W}, \quad q := \frac{\beta}{U} = \frac{\delta}{V}, \quad r := \frac{\gamma}{V} = \frac{\varepsilon}{W}. \quad (199)$$

Since each vector is normalized by its own length⁵, we also have

$$p^2 + q^2 = 1, \quad q^2 + r^2 = 1, \quad p^2 + r^2 = 1. \quad (203)$$

These three equations imply:

$$p = q = r = \frac{1}{\sqrt{2}}. \quad (204)$$

Hence, at the optimum:

$$\alpha = \beta = \frac{U}{\sqrt{2}}, \quad \gamma = \delta = \frac{V}{\sqrt{2}}, \quad \varepsilon = \zeta = \frac{W}{\sqrt{2}}. \quad (205)$$

Inserting this into the constraints gives

$$\begin{aligned} \gamma + \varepsilon = x &\implies \frac{V + W}{\sqrt{2}} = x \iff V + W = \sqrt{2}x, \\ \alpha + \zeta = y &\implies U + W = \sqrt{2}y, \\ \beta + \delta = z &\implies U + V = \sqrt{2}z. \end{aligned} \quad (206)$$

⁵Recall that

$$\vec{h}_{12} = (0, \alpha, \beta), \quad \vec{h}_{13} = (\gamma, 0, \delta), \quad \vec{h}_{14} = (\varepsilon, \zeta, 0). \quad (200)$$

Their lengths are therefore

$$U = \|\vec{h}_{12}\| = \sqrt{\alpha^2 + \beta^2}, \quad V = \|\vec{h}_{13}\| = \sqrt{\gamma^2 + \delta^2}, \quad W = \|\vec{h}_{14}\| = \sqrt{\varepsilon^2 + \zeta^2}. \quad (201)$$

Hence, the normalized version of $\vec{h}_{12}$ is

$$\frac{\vec{h}_{12}}{\|\vec{h}_{12}\|} = \frac{(0, \alpha, \beta)}{U} = \left(0, \frac{\alpha}{U}, \frac{\beta}{U}\right). \quad (202)$$

Solving this system of equations leads to

$$U = \frac{y + z - x}{\sqrt{2}}, \quad V = \frac{x + z - y}{\sqrt{2}}, \quad W = \frac{x + y - z}{\sqrt{2}}. \quad (207)$$

For this to be a valid solution we need

$$U \geq 0, \quad V \geq 0, \quad W \geq 0, \quad (208)$$

since lengths cannot be negative. This means

$$x \leq y + z, \quad y \leq x + z, \quad z \leq x + y. \quad (209)$$

These are triangle inequalities. When they hold,

$$T = U + V + W = \frac{y + z - x + x + z - y + x + y - z}{\sqrt{2}} = \frac{x + y + z}{\sqrt{2}}. \quad (210)$$

Thus,

$$v_{\text{white}}(\theta, \psi) = \frac{1}{T} = \frac{\sqrt{2}}{x + y + z}, \quad (211)$$

gives the an interior solution to the problem defined in Eq. (183).

To find the boundary solution to Eq. (183), we do the following. Suppose one of the coordinates x, y, z is much larger than the others, e.g., $z > x + y$. Then, the triangle inequality from Eq. (209)

$$z \leq x + y \quad (212)$$

fails. In the solution given by Eq. (207), it was found that

$$W = \frac{x + y - z}{\sqrt{2}}. \quad (213)$$

If $z > x + y$, then $W < 0$, which is unphysical. Therefore, the optimum must lie on the boundary $W = 0$, which means that

$$\vec{h}_{14} = (\varepsilon, \zeta, 0) = 0. \quad (214)$$

Hence, the minimization problem in Eq. (183) reduces to

$$T = \min_{\beta + \delta = z} \left[\sqrt{y^2 + \beta^2} + \sqrt{x^2 + \delta^2} \right], \quad (215)$$

where

$$\alpha = y, \quad \gamma = x, \quad \varepsilon = \zeta = 0. \quad (216)$$

has been used. The above equalities come from substituting $\varepsilon = \zeta = 0$ into Eq. (184). Hence, the minimization occurs over (β, δ) subject to

$$\beta + \delta = z. \quad (217)$$

For this, set

$$\delta = z - \beta. \quad (218)$$

Then

$$T(\beta) = \sqrt{y^2 + \beta^2} + \sqrt{x^2 + (z - \beta)^2}. \quad (219)$$

At the optimum, the derivative of T with respect to β must vanish:

$$\frac{dT}{d\beta} = \frac{\beta}{\sqrt{y^2 + \beta^2}} - \frac{z - \beta}{\sqrt{x^2 + (z - \beta)^2}} = 0 \iff \frac{\beta}{\sqrt{y^2 + \beta^2}} = \frac{\delta}{\sqrt{x^2 + \delta^2}}. \quad (220)$$

Squaring both sides and simplifying gives

$$\frac{\beta}{y} = \frac{\delta}{x}. \quad (221)$$

Together with

$$\beta + \delta = z, \quad (222)$$

we obtain

$$\beta = \frac{yz}{x + y}, \quad \delta = \frac{xz}{x + y}. \quad (223)$$

Substituting these back into Eq. (215), gives

$$\begin{aligned} T &= \sqrt{y^2 + \left(\frac{yz}{x + y}\right)^2} + \sqrt{x^2 + \left(\frac{xz}{x + y}\right)^2} = y\sqrt{1 + \frac{z^2}{(x + y)^2}} + x\sqrt{1 + \frac{z^2}{(x + y)^2}} \\ &= (x + y)\sqrt{1 + \frac{z^2}{(x + y)^2}} = \sqrt{(x + y)^2 + z^2}. \end{aligned} \quad (224)$$

Since the critical visibility is given by $v = 1/T$, we conclude that

$$v = \frac{1}{\sqrt{z^2 + (x + y)^2}}. \quad (225)$$

Hence, if z dominates, the white-noise visibility is

$$v_{\text{white}}(\theta, \psi) = \frac{1}{\sqrt{z^2 + (x + y)^2}}, \quad (226)$$

where $z > x + y$.

Instead of treating, x , y , and z separately, all dominated cases can be written compactly by defining:

$$s := x + y + z, \quad m := \max\{x, y, z\}. \quad (227)$$

If one coordinate dominates, then

$$m > s - m, \quad (228)$$

where $s - m$ is the sum of the other two coordinates. Consequently, we have

$$T = \sqrt{m^2 + (s - m)^2}, \quad (229)$$

and

$$v = \frac{1}{T} = \frac{1}{\sqrt{m^2 + (s - m)^2}}. \quad (230)$$

In the same spirit, the triangle inequality for the interior solution can be rewritten as:

$$m \leq s - m \iff m \leq \frac{s}{2}. \quad (231)$$

So the final closed form is

$$v_{\text{white}}(\theta, \psi) = \begin{cases} \frac{\sqrt{2}}{x + y + z}, & m \leq \frac{x + y + z}{2}, \\ \frac{1}{\sqrt{m^2 + (x + y + z - m)^2}}, & m > \frac{x + y + z}{2}, \end{cases} \quad (232)$$

with

$$m = \max\{x, y, z\}. \quad (233)$$

Furthermore, at the SIC POVM point, we have

$$\theta = \psi = \arccos(-1/3), \quad (234)$$

which gives

$$x = y = z = \frac{1}{\sqrt{3}}. \quad (235)$$

Hence,

$$x + y + z = \sqrt{3}. \quad (236)$$

Since all three are equal, the triangle inequalities hold. Therefore

$$v_{\text{white}}^{\text{SIC}} = \sqrt{\frac{2}{3}} \approx 0.8165, \quad (237)$$

which confirms the qubit SIC visibility under white noise analytically.

H Derivation of $v_{\text{wc}}(\theta, \psi)$

The noisy implementation of E_a for the worst-case noise is given by

$$E_a^{(v)} = vE_a + (1 - v)N_a. \quad (238)$$

With

$$E_a = \frac{1}{4}(\mathbb{1} + \vec{w}_a \cdot \vec{\sigma}), \quad N_a = \frac{1}{4}(\mathbb{1} + \vec{n}_a \cdot \vec{\sigma}), \quad (239)$$

then

$$vE_a + (1 - v)N_a = \frac{1}{4}[\mathbb{1} + (v\vec{w}_a + (1 - v)\vec{n}_a) \cdot \vec{\sigma}]. \quad (240)$$

We can make the definition

$$\vec{r}_a = v\vec{w}_a + (1 - v)\vec{n}_a. \quad (241)$$

It can be noted that $\vec{n}_a = 0$ in the white-noise case. Hence, Eq. (238) becomes

$$E_a^{(v)} = \frac{1}{4}(\mathbb{1} + \vec{r}_a \cdot \vec{\sigma}) \in \mathcal{P}, \quad (242)$$

where $\vec{r}_1 = (r_x, r_y, r_z)$ is the Bloch vector of $E_a^{(v)}$. For the worst-case noise, $\vec{r}$ is a free optimization variable describing a general PM-simulable candidate POVM⁶ and can be chosen w.l.o.g. to be

$$\begin{aligned}\vec{r}_1 &= (r_x, r_y, r_z), \\ \vec{r}_2 &= (r_x, -r_y, -r_z), \\ \vec{r}_3 &= (-r_x, r_y, -r_z), \\ \vec{r}_4 &= (-r_x, -r_y, r_z),\end{aligned}\tag{243}$$

since $E_a^{(v)}$ follows the same rotational symmetry as E_a . It also satisfies normalization $r_x^2 + r_y^2 + r_z^2 \leq 1$.

A POVM is PM-simulable if it is a convex mixture of projective measurements. For qubits, every nontrivial projective measurement is binary, hence a 4-outcome POVM must be simulated by mixing binary projective measurements whose two outcomes are assigned to pairs of labels. For the family treated here, the six outcome pairs reduce to three coordinate-plane types: xy, xz, yz . Hence, analogously to the previous section, we ask, how much total projective weight is needed to build the vector $\vec{r} = (r_x, r_y, r_z)$? Since a convex mixture of PVMs has a total probability weight of at most unity, we are looking for the most efficient projective decomposition with a total weight of ≤ 1 . The most efficient decomposition is the one of minimal projective weight. Again, the minimum total weight is denoted $T(\vec{r})$.

H.1 Expression for $T(\vec{r})$

Following the arguments in the previous section, $\vec{r}$ can be decomposed as

$$\vec{r} = \vec{h}_{xy} + \vec{h}_{xz} + \vec{h}_{yz},\tag{244}$$

where

$$\vec{h}_{xy} = (h_x, h_y, 0), \quad \vec{h}_{xz} = (k_x, 0, k_z), \quad \vec{h}_{yz} = (0, \ell_y, \ell_z).\tag{245}$$

Each $\vec{h}$ -vector corresponds to a family of binary projective measurements, whose length is the projective weight needed for that branch. This is the same as for the white noise case. Thus, the minimum total projective-measurement weight needed to simulate the POVM is

$$T(\vec{r}) := \min_{\vec{r} = \vec{h}_{xy} + \vec{h}_{xz} + \vec{h}_{yz}} \left(\|\vec{h}_{xy}\| + \|\vec{h}_{xz}\| + \|\vec{h}_{yz}\| \right).\tag{246}$$

Then

$$E_a^{(v)} \in \mathcal{P} \iff T(\vec{r}) \leq 1.\tag{247}$$

Sorting the coordinates $r_1 > r_2 > r_3 > 0$ leads to

$$\begin{aligned}m_r = r_1 = \max\{r_x, r_y, r_z\} &\implies \vec{r} = \vec{h}_{12} + \vec{h}_{13} + \vec{h}_{23}, \\ s_r = r_1 + r_2 + r_3 &= \sum_i r_i, \quad i = x, y, z.\end{aligned}\tag{248}$$

From this, two cases can be identified.

⁶For white noise, the general vector is forced to be $\vec{r} = v\vec{w} = (vx, vy, vz)$.

Case I: Balanced branch, $r_1 \leq r_2 + r_3$.

Suppose

$$r_1 \leq r_2 + r_3 \iff m_r \leq \frac{s_r}{2}. \quad (249)$$

Using the decomposition

$$\vec{h}_{12} = a_{12}(1, 1, 0), \quad \vec{h}_{13} = a_{13}(1, 0, 1), \quad \vec{h}_{23} = a_{23}(0, 1, 1), \quad (250)$$

we need

$$\begin{cases} r_1 = a_{12} + a_{13} \\ r_2 = a_{12} + a_{23} \\ r_3 = a_{13} + a_{23} \end{cases} \quad (251)$$

Solving this linear system gives

$$\begin{cases} a_{12} = \frac{r_1 + r_2 - r_3}{2} \\ a_{13} = \frac{r_1 + r_3 - r_2}{2} \\ a_{23} = \frac{r_2 + r_3 - r_1}{2} \end{cases} \quad (252)$$

These coefficients are nonnegative when $r_1 \leq r_2 + r_3$. Furthermore, the cost is⁷

$$T(\vec{r}) = \min_{\vec{r} = \vec{h}_{xy} + \vec{h}_{xz} + \vec{h}_{yz}} \left(\|\vec{h}_{12}\| + \|\vec{h}_{13}\| + \|\vec{h}_{23}\| \right). \quad (253)$$

Since $\|a(1, 1, 0)\| = \sqrt{2}a$,

$$T(\vec{r}) = \sqrt{2}(a_{12} + a_{13} + a_{23}), \quad (254)$$

but

$$a_{12} + a_{13} + a_{23} = \frac{r_1 + r_2 + r_3}{2} = \frac{s_r}{2}. \quad (255)$$

Therefore,

$$T(\vec{r}) = \sqrt{2} \frac{s_r}{2} \leq \frac{s_r}{\sqrt{2}}. \quad (256)$$

To show that the upper bound is also the minimum, let $\vec{q} := \frac{1}{\sqrt{2}}(1, 1, 1)$. Using the Cauchy-Schwarz on each branch gives

$$\vec{q} \cdot \vec{h}_{ij} \leq \|\vec{h}_{ij}\|, \quad i = 1, 2, \quad j = 2, 3, \quad (257)$$

where $\|\vec{q}\| = 1$ and $\vec{h}_{ij} \in \text{span}\{e_i, e_j\}$ has been used. Hence, from Eq. (248)

$$\vec{q} \cdot \vec{r} \leq \|\vec{h}_{12}\| + \|\vec{h}_{13}\| + \|\vec{h}_{23}\|. \quad (258)$$

Moreover,

$$\vec{q} \cdot \vec{r} = \frac{r_1 + r_2 + r_3}{\sqrt{2}} = \frac{s_r}{\sqrt{2}}. \quad (259)$$

Every valid decomposition satisfies

$$\|\vec{h}_{12}\| + \|\vec{h}_{13}\| + \|\vec{h}_{23}\| \geq \frac{s_r}{\sqrt{2}}. \quad (260)$$

Hence

$$\begin{cases} T(\vec{r}) \geq \frac{s_r}{\sqrt{2}} \\ T(\vec{r}) \leq \frac{s_r}{\sqrt{2}} \end{cases} \implies T(\vec{r}) = \frac{s_r}{\sqrt{2}}. \quad (261)$$

⁷Minimizing T maximizes v .

Case II: Unbalanced branch, $r_1 > r_2 + r_3$

Assume

$$r_1 \geq r_2 + r_3 \iff m_r \geq \frac{s_r}{2}. \quad (262)$$

This indicates that the largest coordinate r_1 is too large to use the symmetric three-branch decomposition with all positive weights above. Instead, we use only the two branches that contain the dominant coordinate r_1 , which is the 12- and 13-branches. Hence,

$$\vec{r} = (r_1, r_2, r_3) = \vec{h}_{12} + \vec{h}_{13}, \quad (263)$$

where

$$\vec{h}_{12} = (u, r_2, 0), \quad \vec{h}_{13} = (r_1 - u, 0, r_3). \quad (264)$$

The only freedom⁸ is how the dominant coordinate r_1 is split between the two branches. Consequently, the minimization problem in Eq. (246) becomes

$$T(\vec{r}) = \min_u \left[\sqrt{u^2 + r_2^2} + \sqrt{(r_1 - u)^2 + r_3^2} \right], \quad (265)$$

This is the cost of using only the two branches $\vec{h}_{12}$ and $\vec{h}_{13}$. To get the minimum, we differentiate with respect to u :

$$\frac{d}{du} \left[\sqrt{u^2 + r_2^2} + \sqrt{(r_1 - u)^2 + r_3^2} \right] = \frac{u}{\sqrt{u^2 + r_2^2}} - \frac{r_1 - u}{\sqrt{(r_1 - u)^2 + r_3^2}}, \quad (266)$$

which at the minimum (vanishing derivative with respect to u) is

$$\begin{aligned} \frac{u}{\sqrt{u^2 + r_2^2}} &= \frac{r_1 - u}{\sqrt{(r_1 - u)^2 + r_3^2}} \iff \frac{u^2}{u^2 + r_2^2} = \frac{(r_1 - u)^2}{(r_1 - u)^2 + r_3^2} \iff u^2 r_3^2 = (r_1 - u)^2 r_2^2 \\ \implies u r_3 &= (r_1 - u) r_2 \iff u = \frac{r_1 r_2}{r_2 + r_3}, \quad r_1 - u = \frac{r_1 r_3}{r_2 + r_3}. \end{aligned} \quad (267)$$

Further let $L := r_2 + r_3$. We use only two plane branches

$$\vec{h}_{12} = \left(\frac{r_1 r_2}{L}, r_2, 0 \right), \quad \vec{h}_{13} = \left(\frac{r_1 r_3}{L}, 0, r_3 \right). \quad (268)$$

Moreover, since $L = r_2 + r_3$, we have

$$\frac{r_1 r_2}{L} + \frac{r_1 r_3}{L} = \frac{r_1 (r_2 + r_3)}{L} = r_1. \quad (269)$$

Thus,

$$\vec{h}_{12} + \vec{h}_{13} = \left(\frac{r_1 r_2}{L} + \frac{r_1 r_3}{L}, r_2, r_3 \right) = (r_1, r_2, r_3). \quad (270)$$

To compute the cost,

$$\begin{aligned} \|\vec{h}_{12}\| &= \sqrt{\left(\frac{r_1 r_2}{L} \right)^2 + r_2^2} = \frac{r_2}{L} \sqrt{r_1^2 + L^2}, \\ \|\vec{h}_{13}\| &= \frac{r_3}{L} \sqrt{r_1^2 + L^2}. \end{aligned} \quad (271)$$

⁸Since we are not using the 23-branch, the r_2 -component can only come from the 12-branch, and the r_3 -component can only come from the 13-branch. Hence, the only remaining degree of freedom is how to split r_1 between the two branches used.

Then,

$$T(\vec{r}) = \left(\frac{r_2 + r_3}{L} \right) \sqrt{r_1^2 + L^2} = \sqrt{r_1^2 + L^2} = \sqrt{r_1^2 + (r_2 + r_3)^2}, \quad (272)$$

and since $r_1 = m_r$ and $r_2 + r_3 = s_r - m_r$, this becomes

$$T(\vec{r}) = \sqrt{m_r^2 + (s_r - m_r)^2}. \quad (273)$$

Putting the two cases together gives

$$T(\vec{r}) = \begin{cases} \frac{s_r}{\sqrt{2}}, & m_r \leq \frac{s_r}{2}, \\ \sqrt{m_r^2 + (s_r - m_r)^2}, & m_r \geq \frac{s_r}{2}, \end{cases} \quad (274)$$

where

$$s_r = r_x + r_y + r_z, \quad m_r = \max\{r_x, r_y, r_z\}. \quad (275)$$

The worst-case noise asks whether there exists some POVM N such that

$$E_a^{(v)} = vE_a + (1 - v)N_a, \quad \forall a \quad (276)$$

is PM-simulable. Again, due to the symmetry of the POVM family studied, only one outcome needs to be considered. Equivalently,

$$E_a^{(v)} - vE_a = (1 - v)N_a \succeq 0. \quad (277)$$

Hence, we need

$$E^{(v)} \in \mathcal{P}, \quad E_a^{(v)} - vE_a \succeq 0. \quad (278)$$

For outcome 1

$$E_1^{(v)} - vE_1 = \frac{1}{4}[(1 - v)\mathbb{1} + (\vec{r} - v\vec{w}) \cdot \vec{\sigma}], \quad (279)$$

using $E_a^{(v)} = \frac{1}{4}(\mathbb{1} + \vec{r}_a \cdot \vec{\sigma})$. From Eq. (162), a qubit operator

$$\alpha\mathbb{1} + \vec{q} \cdot \vec{\sigma} \succeq 0 \iff \alpha \geq \|\vec{q}\|. \quad (280)$$

Therefore,

$$E_1^{(v)} - vE_1 \succeq 0 \iff \|\vec{r} - v\vec{w}\| \leq 1 - v. \quad (281)$$

Consequently, the worst-case visibility is

$$v_{\text{wc}} = \max_{\vec{r}} \{v : T(\vec{r}) \leq 1, \quad \|\vec{r} - v\vec{w}\| \leq 1 - v\}. \quad (282)$$

H.2 The auxiliary variable $t = 1/v$

Letting

$$t := \frac{1}{v}, \quad \vec{R} := \frac{\vec{r}}{v} \iff \vec{r} := v\vec{R}. \quad (283)$$

Since T is homogeneous⁹,

$$T(\vec{r}) = T(v\vec{R}) = vT(\vec{R}). \quad (284)$$

⁹ $T(\vec{r}) = \left(\|\vec{h}_{xy}\| + \|\vec{h}_{xz}\| + \|\vec{h}_{yz}\| \right)$ and $\|c\vec{h}\| = c\|\vec{h}\|$. Therefore, $T(c\vec{r}) = cT(\vec{r})$.
 $\vec{r} = \vec{h}_{xy} + \vec{h}_{xz} + \vec{h}_{yz}$

The PM-simulability condition $T(\vec{r}) \leq 1$ thereby becomes

$$vT(\vec{R}) \leq 1 \iff T(\vec{R}) \leq t, \quad (285)$$

since $t = 1/v$. The positivity condition in Eq. (281) becomes

$$\|v\vec{R} - v\vec{w}\| \leq 1 - v \iff v\|\vec{R} - \vec{w}\| \leq 1 - v. \quad (286)$$

Dividing by v , yields

$$\|\vec{R} - \vec{w}\| \leq \frac{1}{v} - 1 = t - 1. \quad (287)$$

Hence, the inverse worst-case visibility can be written as

$$t_{\text{wc}} = \min_{\vec{R}} \max \left\{ T(\vec{R}), \quad 1 + \|\vec{R} - \vec{w}\| \right\}. \quad (288)$$

At the optimum, $T(\vec{R}) = t$ and $\|\vec{R} - \vec{w}\| = t - 1$.

Furthermore, letting (a, b, c) be the sorted version of (x, y, z) such that

$$a \geq b \geq c \geq 0. \quad (289)$$

Further defining $B := b + c$, which leads to $s = a + B = x + y + z$, we have

$$a = \max\{x, y, z\}, \quad B = s - a. \quad (290)$$

The two branches of T are

$$T(\vec{R}) = \begin{cases} \frac{R_1 + R_2 + R_3}{\sqrt{2}}, & R_1 \leq R_2 + R_3, \\ \sqrt{R_1^2 + (R_2 + R_3)^2}, & R_1 > R_2 + R_3 \end{cases}. \quad (291)$$

We are now in a position to solve the two cases directly.

H.3 Case I: Balanced branch, $R_1 \leq R_2 + R_3$

For this branch,

$$T(\vec{R}) = \frac{R_1 + R_2 + R_3}{\sqrt{2}}. \quad (292)$$

The surface $T(\vec{R}) = t$ is equivalent to the plane

$$R_1 + R_2 + R_3 = \sqrt{2}t. \quad (293)$$

The point on this plane closest to $\vec{w} = (a, b, c)$ is obtained by shifting all coordinates equally:

$$R_i = w_i - d, \quad i = 1, 2, 3. \quad (294)$$

The direction normal to the plane is $(1, 1, 1)/\sqrt{3}$, which means that

$$\|\vec{R} - \vec{w}\| = \sqrt{3}d. \quad (295)$$

Since the distance is maximally equal to $t - 1$,

$$\sqrt{3}d = t - 1 \iff d = \frac{t - 1}{\sqrt{3}}. \quad (296)$$

Moreover,

$$R_1 + R_2 + R_3 = (a + b + c) - 3d = s - 3d. \quad (297)$$

Then, the plane condition gives

$$s - 3d = \sqrt{2}t. \quad (298)$$

Substituting Eq. (296) into the equation above gives

$$s - \frac{3(t - 1)}{\sqrt{3}} = \sqrt{2}t \iff s - \sqrt{3}(t - 1) = \sqrt{2}t \iff t = \frac{s + \sqrt{3}}{\sqrt{2} + \sqrt{3}}. \quad (299)$$

Hence,

$$v = \frac{\sqrt{2} + \sqrt{3}}{s + \sqrt{3}}. \quad (300)$$

We also need a condition for the closest point to lie in the region $R_1 \leq R_2 + R_3$. Since

$$R_1 = a - d, \quad R_2 + R_3 = B - 2d, \quad (301)$$

we demand

$$a - d \leq B - 2d \iff a + d \leq B. \quad (302)$$

Using Eq. (296) and (299) gives

$$d = \frac{1}{\sqrt{3}} \left(\frac{s + \sqrt{3}}{\sqrt{2} + \sqrt{3}} - 1 \right) = \frac{1}{\sqrt{3}} \left(\frac{s + \sqrt{3} - \sqrt{2} - \sqrt{3}}{\sqrt{2} + \sqrt{3}} \right) = \frac{s - \sqrt{2}}{\sqrt{3}(\sqrt{2} + \sqrt{3})}. \quad (303)$$

Hence, the balanced condition is

$$a + \frac{s - \sqrt{2}}{\sqrt{3}(\sqrt{2} + \sqrt{3})} \leq B \iff a(\sqrt{6} - 1) \leq B + \sqrt{3} - \sqrt{2}. \quad (304)$$

Therefore,

$$v_{\text{wc}} = \frac{\sqrt{3} + \sqrt{2}}{\sqrt{3} + s}. \quad (305)$$

whenever

$$a(\sqrt{6} - 1) \leq B + \sqrt{3} - \sqrt{2}. \quad (306)$$

H.4 Case II: Unbalanced branch, $R_1 > R_2 + R_3$

Suppose the balanced condition in Eq. (306) no longer holds,

$$a(\sqrt{6} - 1) > B + \sqrt{3} - \sqrt{2}, \quad (307)$$

then, the closest point lies in the unbalanced branch, where

$$T(\vec{R}) = \sqrt{R_1^2 + (R_2 + R_3)^2}. \quad (308)$$

Defining

$$u := R_1, \quad L := R_2 + R_3, \quad (309)$$

gives

$$T(\vec{R}) = \sqrt{u^2 + L^2}. \quad (310)$$

Hence the condition $T(\vec{R}) = t$ becomes

$$u^2 + L^2 = t^2. \quad (311)$$

Now we compute the distance to the target. For fixed u and L , the closest choice of R_2, R_3 to b, c , subject to $L = R_2 + R_3$ is

$$R_2 = b + \frac{L - B}{2}, \quad R_3 = c + \frac{L - B}{2}. \quad (312)$$

Therefore,

$$(R_2 - b)^2 + (R_3 - c)^2 = 2 \left(\frac{L - B}{2} \right)^2 = \frac{(L - B)^2}{2}. \quad (313)$$

Hence the distance condition is given by

$$(u - a)^2 + \frac{(L - B)^2}{2} = (t - 1)^2. \quad (314)$$

Thus, at the optimal point, the unbalanced problem is two-dimensional:

$$\begin{cases} (u - a)^2 + \frac{(L - B)^2}{2} = (t - 1)^2 \\ u^2 + L^2 = t^2, \end{cases} \quad (315)$$

For a fixed t , the boundary given by the upper equation describes a circle in the (u, L) -plane, while the lower equation describes an elliptic boundary centered at (a, B) . Since the problem is to find the smallest t for which the circle and ellipse intersect, at the smallest such t , the two boundaries should just touch, i.e., tangent each other.

H.5 Tangency condition in the unbalanced branch

At the optimum, the distance curve is tangent to the $T = t$ curve. Thus, we minimize

$$D^2(u, L) = (u - a)^2 + \frac{(L - B)^2}{2} \quad (316)$$

subject to

$$u^2 + L^2 = t^2. \quad (317)$$

Introducing the Lagrange multiplier μ :

$$\mathcal{L} = (u - a)^2 + \frac{(L - B)^2}{2} + \mu(u^2 + L^2 - t^2). \quad (318)$$

Differentiating with respect to u gives:

$$\frac{\partial \mathcal{L}}{\partial u} = 2(u - a) + 2\mu u = 0 \iff u - a + \mu u = 0 \iff a = (1 + \mu)u. \quad (319)$$

Further differentiating with respect to L :

$$\frac{\partial \mathcal{L}}{\partial L} = (L - B) + 2\mu L = 0 \iff B = (1 + 2\mu)L. \quad (320)$$

Therefore,

$$a = (1 + \mu)u, \quad B = (1 + 2\mu)L. \quad (321)$$

Moreover, the distance is

$$t - 1 = \sqrt{(u - a)^2 + \frac{(L - B)^2}{2}}. \quad (322)$$

Using Eq. (321), we get

$$u - a = -\mu u, \quad L - B = -2\mu L. \quad (323)$$

Hence,

$$(t - 1)^2 \mu^2 u^2 + \frac{4\mu^2 L^2}{2} = \mu^2 (u^2 + 2L^2). \quad (324)$$

This gives,

$$t - 1 = \mu \sqrt{u^2 + 2L^2}. \quad (325)$$

H.6 Parametrization by λ

Setting

$$\lambda := \frac{u}{L} \iff u = \lambda L. \quad (326)$$

Since the unbalanced branch has $u \geq L$, $\lambda \geq 1$. Further defining

$$A := \sqrt{\lambda^2 + 2}, \quad C := \sqrt{\lambda^2 + 1}. \quad (327)$$

Then,

$$\begin{aligned} t &= \sqrt{u^2 + L^2} = L\sqrt{\lambda^2 + 1} = LC, \\ \sqrt{u^2 + 2L^2} &= L\sqrt{\lambda^2 + 2} = LA. \end{aligned} \quad (328)$$

Moreover, the distance condition becomes

$$t - 1 = \mu LA, \quad (329)$$

and since $t = LC$,

$$LC - 1 = \mu LA \iff L = \frac{1}{C - \mu A}. \quad (330)$$

Using the tangency equations (315), we get

$$a = (1 + \mu)u = (1 + \mu)\lambda L, \quad B = (1 + 2\mu)L. \quad (331)$$

Moreover,

$$a\lambda + B = [(1 + \mu)\lambda^2 + (1 + 2\mu)] L [\lambda^2 + 1 + \mu(\lambda^2 + 2)] L, \quad (332)$$

but since

$$\lambda^2 + 1 = C^2, \quad \lambda^2 + 2 = A^2, \quad (333)$$

we have

$$a\lambda + B = (C^2 + \mu A^2)L \iff a\lambda + B = \frac{C^2 + \mu A^2}{C - \mu A}. \quad (334)$$

where in the last equality

$$L = \frac{1}{C - \mu A}, \quad (335)$$

has been used. Further letting $X := a\lambda + B$, gives

$$X(C - \mu A) = C^2 + \mu A^2, \quad (336)$$

which leads to

$$\mu = \frac{C(X - C)}{A(X + A)} = \frac{C(a\lambda + B - C)}{A(a\lambda + B + A)}. \quad (337)$$

where $X = a\lambda + B$ has been substituted in the last equality. Since $t = LC$ and $L = \frac{1}{C - \mu A}$, we have

$$t = \frac{C}{C - \mu A}. \quad (338)$$

Substituting the expression for μ Eq. (337),

$$t = \frac{C}{C - A \frac{C(a\lambda + B - C)}{A(a\lambda + B + A)}} = \frac{C}{C - \frac{C(a\lambda + B - C)}{a\lambda + B + A}} = \frac{1}{1 - \frac{a\lambda + B - C}{a\lambda + B + A}}. \quad (339)$$

Combining the denominator,

$$1 - \frac{a\lambda + B - C}{a\lambda + B + A} = \frac{a\lambda + B + A - (a\lambda + B - C)}{a\lambda + B + A} = \frac{A + C}{a\lambda + B + A}. \quad (340)$$

Therefore,

$$t(\lambda) = \frac{A + a\lambda + B}{A + C}, \quad \lambda \geq 1. \quad (341)$$

Equivalently, from Eq. (327), we get

$$t(\lambda) = \frac{\sqrt{\lambda^2 + 2} + a\lambda + B}{\sqrt{\lambda^2 + 2} + \sqrt{\lambda^2 + 1}}, \quad \lambda \geq 1. \quad (342)$$

The worst tangent determines the inverse visibility, hence

$$t_{\text{wc}} = \max_{\lambda \geq 1} t(\lambda). \quad (343)$$

Hence,

$$v_{\text{wc}} = \frac{1}{t_{\text{wc}}}. \quad (344)$$

H.7 Stationary equation for λ_*

Defining

$$N = A + a\lambda + B, \quad D = A + C. \quad (345)$$

Then, from Eq. (341), we have

$$t(\lambda) = \frac{N}{D}, \quad (346)$$

with the derivatives

$$N' = \frac{\lambda}{A} + a, \quad D' = \frac{\lambda}{A} + \frac{\lambda}{C}. \quad (347)$$

The stationary point satisfies

$$N'D - ND' = 0. \quad (348)$$

Hence,

$$\left(\frac{\lambda}{A} + a\right)(A + C) - (A + a\lambda + B)\left(\frac{\lambda}{A} + \frac{\lambda}{C}\right) = 0. \quad (349)$$

Multiplying by AC , simplifying, and using $A + C > 0$, gives

$$a(AC - \lambda^2) = \lambda(B + A - C). \quad (350)$$

If the derivative at $\lambda = 1$ is non-positive, the maximum is at $\lambda = 1$, which corresponds to the balanced branch. If the derivative at $\lambda = 1$ is positive, the maximum occurs at the unique interior solution $\lambda_* > 1$.

H.8 Algebraic root form

To express the root algebraically, define $p := A - C$. Then

$$A - C = p, \quad A + C = \frac{1}{p}. \quad (351)$$

The reason for the second form above is the identity $A^2 - C^2 = 1$, which explicitly can be written as

$$\left(\sqrt{\lambda^2 + 2} - \sqrt{\lambda^2 + 1}\right)\left(\sqrt{\lambda^2 + 2} + \sqrt{\lambda^2 + 1}\right) = 1. \quad (352)$$

Hence, if

$$p = \sqrt{\lambda^2 + 2} - \sqrt{\lambda^2 + 1}, \quad (353)$$

then

$$\sqrt{\lambda^2 + 2} + \sqrt{\lambda^2 + 1} = \frac{1}{p}. \quad (354)$$

Hence, from Eq. (351)

$$A = \frac{1}{2}\left(p + \frac{1}{p}\right), \quad C = \frac{1}{2}\left(\frac{1}{p} - p\right). \quad (355)$$

Moreover,

$$\lambda = \frac{\sqrt{1 - 6p^2 + p^4}}{2p}. \quad (356)$$

Hence, stationary equation becomes

$$ap(3 - p^2) = (B + p)\sqrt{1 - 6p^2 + p^4}. \quad (357)$$

Squaring gives

$$(B + p)^2(p^4 - 6p^2 + 1) - a^2p^2(3 - p^2)^2 = 0. \quad (358)$$

The relevant root lies in

$$0 < p < \sqrt{3} - \sqrt{2}, \quad (359)$$

so

$$v_{\text{wc}} = \frac{2}{1 + p_*^2 + 2Bp_* + a\sqrt{1 - 6p_*^2 + p_*^4}}. \quad (360)$$

H.9 Final formula

Defining

$$x = \cos\left(\frac{\psi}{2}\right), \quad y = \cos\left(\frac{\theta}{2}\right), \quad z = \sqrt{-\frac{\cos\psi + \cos\theta}{2}}. \quad (361)$$

Furthermore, defining

$$s := x + y + z, \quad a = \max\{x, y, z\}, \quad B = s - a, \quad (362)$$

leads to

$$v_{\text{wc}}(\theta, \psi) = \begin{cases} \frac{\sqrt{3} + \sqrt{2}}{\sqrt{3} + s}, & a(\sqrt{6} - 1) \leq B + \sqrt{3} - \sqrt{2}, \\ \frac{2}{1 + p_*^2 + 2Bp_* + a\sqrt{1 - 6p_*^2 + p_*^4}}, & a(\sqrt{6} - 1) > B + \sqrt{3} - \sqrt{2}. \end{cases} \quad (363)$$

Here p_* is the unique root in

$$0 < p_* < \sqrt{3} - \sqrt{2} \quad (364)$$

of

$$(B + p)^2(p^4 - 6p^2 + 1) - a^2p^2(3 - p^2)^2 = 0. \quad (365)$$

Hence, the worst-case noise visibility can be derived using the same T -function as in the white-noise case. The difference is that worst-case noise is not radial. Therefore, after introducing T , one must solve a tangency problem, which gives the one-parameter expression $t(\lambda)$, and then the algebraic root equation for p_* .

I Proof of redundancy of classical post-processing of outcomes in PI-simulation

Proof. ¹⁰ Here it is shown that classical post-processing in the PI simulation can be neglected w.l.o.g.

Upon considering a deterministic post-processing rule, which takes the form $p(a|a', \lambda) = \delta_{a, f(a', \lambda)}$, where f is some function of a' and λ , Eq. (4.1) from the main text can be rewritten as

$$\mathcal{I}_a(\rho) = \sum_{\lambda} q_{\lambda} \underbrace{\sum_{a'} \Lambda_{a', \lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \delta_{a, f(a', \lambda)}}_{\Gamma_{a|\lambda}^f} = \sum_{\lambda} q_{\lambda} \Gamma_{a|\lambda}^f, \quad (366)$$

where the every $\Gamma_{a|\lambda}^f$ is a CP map whose sum is TP, $\text{Tr}\left(\sum_a \Gamma_{a|\lambda}^f(\rho)\right) = 1$. Hence, the set $\{\Gamma_{a|\lambda}^f\}_a$ is a quantum instrument for all choices of f and λ .

¹⁰This proof is based on [7].

Now, defining PVMs $\tilde{E}_{a|\lambda}^f := \sum_{a'} E_{a'|\lambda} \delta_{a,f(a',\lambda)}$, where $\sum_a \tilde{E}_{a|\lambda}^f = \mathbb{1}$. To complete the proof that $\tilde{E}_{x|\lambda}^f$ is a PVM, we can show that it also fulfills orthogonality:

$$\begin{aligned}
\tilde{E}_{a|\lambda}^f \tilde{E}_{b|\lambda}^f &= \sum_{a',b'} E_{a'|\lambda} E_{b'|\lambda} \delta_{a,f(a',\lambda)} \delta_{b,f(b',\lambda)} \\
&= \sum_{a'b'} \delta_{a',b'} \delta_{a,f(a',\lambda)} \delta_{b,f(b',\lambda)} E_{a'|\lambda} \\
&= \sum_{a'} \delta_{a,f(a',\lambda)} \delta_{b,f(a',\lambda)} E_{a'|\lambda} \\
&= \sum_{a'} \delta_{a,b} \delta_{a,f(a',\lambda)} E_{a'|\lambda} \\
&= \delta_{a,b} \tilde{E}_{a|\lambda}^f,
\end{aligned} \tag{367}$$

where in the second equality, $E_{a'|\lambda} E_{b'|\lambda} = \delta_{a',b'} E_{a'|\lambda}$, has been used. It is further possible to express $\Gamma_{a|\lambda}^f$ as

$$T_{a|\lambda}^f(\rho) \stackrel{!}{=} \tilde{\Lambda}_\lambda \left[\tilde{E}_{a|\lambda}^f \rho \tilde{E}_{a|\lambda}^f \right] = \tilde{\Lambda}_\lambda \left[\underbrace{\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} E_{l|\lambda} \rho E_{l'|\lambda}}_X \right]. \tag{368}$$

Comparing to

$$\Gamma_{a|\lambda}^f = \sum_{a'} \Lambda_{a',\lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \delta_{a,f(a',\lambda)}, \tag{369}$$

we can select the CPTP maps as

$$\tilde{\Lambda}_\lambda[X] := \sum_{a'} \Lambda_{a',\lambda} [E_{a'|\lambda} X E_{a'|\lambda}]. \tag{370}$$

Inserting into Eq. (368) gives

$$\begin{aligned}
\tilde{\Lambda}_\lambda \left[\tilde{E}_{a|\lambda}^f \rho \tilde{E}_{a|\lambda}^f \right] &= \tilde{\Lambda}_\lambda \left[\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} E_{l|\lambda} \rho E_{l'|\lambda} \right] \\
&= \sum_{a'} \Lambda_{a',\lambda} \left[E_{a'|\lambda} \left(\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} E_{l|\lambda} \rho E_{l'|\lambda} \right) E_{a'|\lambda} \right] \\
&= \sum_{a'} \Lambda_{a',\lambda} \left[\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} \underbrace{E_{a'|\lambda} E_{l|\lambda}}_{=\delta_{a',l} E_{a'|\lambda}} \rho \underbrace{E_{l'|\lambda} E_{a'|\lambda}}_{=\delta_{l',a'} E_{a'|\lambda}} \right] \\
&= \sum_{a'} \Lambda_{a',\lambda} \left[\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} \delta_{a',l} \delta_{l',a'} E_{a'|\lambda} \rho E_{a'|\lambda} \right] \\
&= \sum_{a'} \Lambda_{a',\lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \delta_{a,f(a',\lambda)} \\
&:= \Gamma_{a|\lambda}^f(\rho),
\end{aligned} \tag{371}$$

where in the first equality, it is used that:

$$\tilde{\Lambda}_\lambda[X] = \sum_{a'} \Lambda_{a',\lambda} \left[E_{a'|\lambda} \left(\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} E_{l|\lambda} \rho E_{l'|\lambda} \right) \right] E_{a'|\lambda}. \quad (372)$$

In the second last equality $\delta_{a',l} \delta_{l',a'}$ enforces $l = a'$ and $l' = a'$, collapsing the double sum:

$$\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} \delta_{a',l} \delta_{l',a'} = \delta_{a,f(a',\lambda)} \delta_{a,f(a',\lambda)} = \delta_{a,f(a',\lambda)}. \quad (373)$$

Moreover, in the penultimate equality,

$$\sum_{l,l'} \delta_{a,f(l,\lambda)} \delta_{a,f(l',\lambda)} \delta_{a',l} \delta_{l',a'} = \delta_{a,f(a',\lambda)} \delta_{a,f(a',\lambda)} = \delta_{a,f(a',\lambda)} \quad (374)$$

was used. The result is exactly on the form in Eq. (366). Therefore, a generic PI with deterministic classical post-processing is given by:

$$\mathcal{I}_a(\rho) = \sum_{\lambda} q_{\lambda} \sum_{a'} \Lambda_{a',\lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \delta_{a,f(a',\lambda)} = \sum_{\lambda} q_{\lambda} \tilde{\Lambda}_{\lambda} [\tilde{E}_{a|\lambda}^f \rho \tilde{E}_{a|\lambda}^f], \quad (375)$$

where $\tilde{\Lambda}_{a,\lambda}$ is a CPTP channel for all a and λ and $\{\tilde{E}_{a|\lambda}^f\}_a$ is a PVM.

Extending the argument directly to stochastic post-processing rules finishes the proof. Any general post-processing strategy $p(a|a', \lambda)$ can be written as a convex combination over deterministic strategies $p(a|a', \lambda) = \sum_f p_f \delta_{a,f(a',\lambda)}$, where p_f is the probability of selecting the deterministic post-processing function f . Hence, general post-processing changes Eq. (4.1) to

$$\begin{aligned} \mathcal{I}_a(\rho) &= \sum_{\lambda} q_{\lambda} \sum_{a'} p(a|a', \lambda) \Lambda_{a',\lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \\ &= \sum_{\lambda} q_{\lambda} \sum_f p_f \sum_{a'} \Lambda_{a',\lambda} [E_{a'|\lambda} \rho E_{a'|\lambda}] \delta_{a,f(a',\lambda)} \\ &= \sum_{\lambda} q_{\lambda} \sum_f p_f \tilde{\Lambda}_{\lambda} [\tilde{E}_{a|\lambda}^f \rho \tilde{E}_{a|\lambda}^f]. \end{aligned} \quad (376)$$

Re-defining $\lambda \rightarrow (\lambda, f)$ leads to a new distribution $q_{\lambda} \rightarrow q_{\lambda} p_f$. Hence, a PI was obtained without any classical post-processing. The quantum post-processing channel no longer needs to depend on the classical outcome, which can be absorbed into the channel itself. $\square$

J Derivation of Eq. 4.3

$$\eta_a = \left(\left(\sum_{\lambda} q_{\lambda} \Lambda_{\lambda} [E_{a|\lambda} \rho E_{a|\lambda}] \right) \otimes \mathbb{1} \right) [\phi^+] \quad (377)$$

$$= \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} \left((\Lambda_{\chi, \vec{r}} \circ \underbrace{[E_{a|\chi, \vec{r}} \rho E_{a|\chi, \vec{r}}]}_{\mathcal{P}_{a|\chi, \vec{r}}(\rho)}) \otimes \mathbb{1} \right) [\phi^+] \quad (378)$$

$$= \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} (\Lambda_{\chi, \vec{r}} \otimes \mathbb{1}) \left[\underbrace{(\mathcal{P}_{a|\chi, \vec{r}} \otimes \mathbb{1})[\phi^+]}_{(E_{a|\chi, \vec{r}} \otimes \mathbb{1})\phi^+ (E_{a|\chi, \vec{r}} \otimes \mathbb{1}) =: \nu_{a|\chi, \vec{r}}} \right] \quad (379)$$

$$= \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} (\Lambda_{\chi, \vec{r}} \otimes \mathbb{1}) [(E_{a|\chi, \vec{r}} \otimes \mathbb{1})\phi^+ (E_{a|\chi, \vec{r}} \otimes \mathbb{1})[\phi^+]] \quad (380)$$

$$= \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} (\Lambda_{\chi, \vec{r}} \otimes \mathbb{1}) [\nu_{a|\chi, \vec{r}}], \quad (381)$$

where linearity and the decomposition $\lambda = (\chi, \vec{r})$ is used in the second equality, and $(\Lambda \circ \mathcal{P}_{a|\chi, \vec{r}}) \otimes \mathbb{1} = (\Lambda \otimes \mathbb{1}) \circ (\mathcal{P}_{a|\chi, \vec{r}} \otimes \mathbb{1})$ is used in the third equality. In the penultimate equality, $\mathcal{P}_{a|\chi, \vec{r}}(\rho) = E_{a|\chi, \vec{r}} \rho E_{a|\chi, \vec{r}}$ can be viewed as a Kraus map.

K Proof of Theorem 1

Theorem 1. (*Instrument simulation*). For any N -outcome PI with $\{\eta_a\}_a$ being its Choi representation, there exists a decomposition:

$$\eta_a = \sum_{\vec{r}} \sigma_{a|\vec{r}}, \quad \text{where } \sigma_{a|\vec{r}} \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathcal{H}_A) \quad (382)$$

$$\text{Tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{r_a}{d}, \quad \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{\mathbb{1}}{d}, \quad SN(\sigma_{a|\vec{r}}) \leq r_a,$$

where $\vec{r}$ runs over all rank-vectors and SN represents the Schmidt number.

Proof. ¹¹ The post-measurement state for an outcome a of an N -outcome PI $\mathcal{I}$ operating from a Hilbert space of dimension $d = \dim(\mathcal{H}_A)$ is defined by

$$\mathcal{I}_a(\rho) = \sum_{\lambda} q_{\lambda} \Lambda_{\lambda} [E_{a|\lambda} \rho E_{a|\lambda}], \quad (383)$$

where $\{E_{a|\lambda}\}_a$ are PVMs, $\{\Lambda_{\lambda}\}$ are CPTP maps, and $\{q_{\lambda}\}$ is a probability distribution. Each N -outcome PI, $E = \{E_a\}_{a=1}^N$, is associated with a rank vector of non-negative integers $\vec{r} = (r_1, \dots, r_N)$, where $r_a = \text{rank}(E_a)$. Hence, all N -tuples satisfying $\sum_a r_a = d$ are valid rank vectors. Therefore, we can write $\lambda = (\chi, \vec{r})$, where χ is a random variable for selecting PMs of rank vector $\vec{r}$. As a result, Eq. (383) is equivalent to:

$$\mathcal{I}_a(\rho) = \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} \Lambda_{\chi, \vec{r}} [E_{a|\chi, \vec{r}} \rho E_{a|\chi, \vec{r}}] \equiv \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} \mathcal{S}_{a|\chi, \vec{r}}[\rho], \quad (384)$$

¹¹This proof is based on the proof in [7].

where $\mathcal{S}$ is a CP trace-non-increasing map. Now, using channel-state duality to represent these maps as bipartite operators. Furthermore, by associating the Choi operators η_a and $\tilde{\sigma}_{a|\chi,\vec{r}}$ to the instruments $\mathcal{I}_a$ and $\mathcal{S}_{a|\chi,\vec{r}}$, respectively, we have:

$$\eta_a = \sum_{\chi,\vec{r}} q_{\chi,\vec{r}} (\mathcal{S}_{a|\chi,\vec{r}} \otimes \mathbb{1})[\phi^+] = \sum_{\chi,\vec{r}} q_{\chi,\vec{r}} \tilde{\sigma}_{a|\chi,\vec{r}} = \sum_{\vec{r}} \sigma_{a|\vec{r}}, \quad (385)$$

where $\sigma_{a|\vec{r}} := \sum_{\chi} q_{\chi,\vec{r}} \tilde{\sigma}_{a|\chi,\vec{r}}$. Note that $\sigma_{a|\vec{r}} \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathcal{H}_A)$ due to complete positivity of the involved maps. This corresponds to the first constraint in Eq. (1).

Next, we consider the first constraint

$$\begin{aligned} \text{Tr}(\sigma_{a|\vec{r}}) &= \sum_{\chi} q_{\chi,\vec{r}} \text{Tr}(\tilde{\sigma}_{a|\chi,\vec{r}}) \\ &= \sum_{\chi} q_{\chi,\vec{r}} \text{Tr}(\mathcal{S}_{a|\chi,\vec{r}} \otimes \mathbb{1}[\phi^+]) \\ &= \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \sum_{i,j} \text{Tr}(\Lambda_{\chi,\vec{r}}[E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}] \otimes |i\rangle \langle j|) \\ &= \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \text{Tr} \left(\Lambda_{\chi,\vec{r}}[E_{a|\chi,\vec{r}}] \left(\sum_i |i\rangle \langle i| \right) E_{a|\chi,\vec{r}} \right) \\ &= \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \text{Tr}(\Lambda_{\chi,\vec{r}}[E_{a|\chi,\vec{r}}]) \\ &= \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \text{Tr}(E_{a|\chi,\vec{r}}) \\ &= q_{\vec{r}} \frac{r_a}{d}, \end{aligned} \quad (386)$$

where in the second equality the Choi form of the $\tilde{\sigma}_{a|\chi,\vec{r}}$ is used, while in the third equality, $\phi^+ = |\phi^+\rangle \langle \phi^+| = \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} |i\rangle \langle i| \langle j| \langle j|$ is substituted. Moreover, the fifth equality uses the completeness relation ($\sum_i |i\rangle \langle i| = \mathbb{1}$), cyclicity of trace, and projectivity of $\{E_{a|\chi,\vec{r}}\}_a$. In the second last step, trace-preservation of $\Lambda_{\chi,\vec{r}}$ is used, while it is used that $\text{rank}(E_{a|\chi,\vec{r}}) = \text{Tr}(E_{a|\chi,\vec{r}}) = r_a$ in the last equality, where the marginal probability distribution $q_{\vec{r}} := \sum_{\chi} q_{\chi,\vec{r}}$ was introduced. This corresponds to the first constraint in the second row of Eq. (1).

For the second constraint of the second row in Eq. (1), consider

$$\begin{aligned} \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) &= \sum_{a,\chi} q_{\chi,\vec{r}} \text{Tr}_{A'}(\tilde{\sigma}_{a|\chi,\vec{r}}) \\ &= \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \sum_{i,j} \sum_a \text{Tr}(\Lambda_{\chi,\vec{r}}[E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}] \otimes |i\rangle \langle j|), \end{aligned} \quad (387)$$

where

$$\begin{aligned} \sum_a \text{Tr}(\Lambda_{\chi,\vec{r}}[E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}]) &= \sum_a \text{Tr}(E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}) \\ &= \text{Tr} \left(|i\rangle \langle j| \sum_a E_{a|\chi,\vec{r}} \right) = \delta_{i,j}. \end{aligned} \quad (388)$$

In the first equality, trace-preservation of $\Lambda_{\chi,\vec{r}}$ is used, and in the second equality, the projectivity of $E_{a|\chi,\vec{r}}$ is used. In the last equality, the completeness of the measurements is used. Finally, we have

$$\sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = \frac{1}{d} \sum_{\chi} q_{\chi,\vec{r}} \sum_{i,j} \delta_{i,j} |i\rangle \langle j| = q_{\vec{r}} \frac{1}{d}. \quad (389)$$

Recalling that $\nu_{a|\chi,\vec{r}} := (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) \phi^+ (E_{a|\chi,\vec{r}} \otimes \mathbb{1})$ and using $\phi^+ = |\phi^+\rangle \langle \phi^+| = \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} (|i\rangle_{A'} \langle j|_{A'}) \otimes (|i\rangle_A \langle j|_A)$ together with the identity $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, we obtain the following:

$$\begin{aligned} \nu_{a|\chi,\vec{r}} &= (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) \left(\frac{1}{d} \sum_{i,j} (|i\rangle \langle j|) \otimes (|i\rangle \langle j|) \right) (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) \\ &= \frac{1}{d} \sum_{i,j} (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) [(|i\rangle \langle j|) \otimes (|i\rangle \langle j|)] (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) \\ &= \frac{1}{d} \sum_{i,j} [(E_{a|\chi,\vec{r}} |i\rangle \langle j|) \otimes (|i\rangle \langle j|)] (E_{a|\chi,\vec{r}} \otimes \mathbb{1}) \\ &= \frac{1}{d} \sum_{i,j} (E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}) \otimes (|i\rangle \langle j| \mathbb{1}) \\ &= \frac{1}{d} \sum_{i,j} (E_{a|\chi,\vec{r}} |i\rangle \langle j| E_{a|\chi,\vec{r}}) \otimes (|i\rangle \langle j|), \end{aligned} \quad (390)$$

where $(E_{a|\chi,\vec{r}} |i\rangle \langle j|) \otimes (\mathbb{1} |i\rangle \langle j|) = (E_{a|\chi,\vec{r}} |i\rangle \langle j|) \otimes (|i\rangle \langle j|)$ was used in the second equality. Since $E_{a|\chi,\vec{r}}$ is a projector of rank r_a , it projects the system A onto an r_a -dimensional subspace $\mathbb{C}^{r_a} \subseteq \mathbb{C}^d$. That means that $\nu_{a|\chi,\vec{r}} \in (A') \otimes (\mathbb{C}^{r_a} \subseteq A)$. Hence, the smaller local dimension of the bipartite support is at most r_a . Hence, the local projection of ϕ^+ onto $E_{a|\chi,\vec{r}}$ confines it to a r_a -dimensional subspace.

The last constraint is motivated by the following. If the pure bipartite state $|\psi\rangle_{AB}$ is defined on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ of local dimensions d_A and d_B , respectively [59]. Then, it has a Schmidt decomposition $|\psi\rangle = \sum_{i=1}^r \sqrt{\lambda_i} |\alpha_i\rangle \otimes |\beta_i\rangle$, where $r = \min(d_A, d_B)$. The Schmidt number is the extension to mixed states ρ and is the smallest integer k such that ρ can be written as a convex mixture of pure states, each having Schmidt rank $\leq k$ [59, 60]:

$$SN(\rho) = \min\{k : \rho = \sum_l p_l |\psi_l\rangle \langle \psi_l|, SR(|\psi_l\rangle) \leq k \quad \forall l\}. \quad (391)$$

From this, we get two criteria: (i) $SN(\rho) = 1$ if and only if ρ is separable, and (ii) $SN(\rho) \leq \min(d_A, d_B)$ always. The first condition follows directly from the definition of the Schmidt number itself; $SN(\rho)$ is the smallest k such that:

$$\rho = \sum_l p_l |\psi_l\rangle \langle \psi_l|, \quad p_l \geq 0, \sum_l p_l = 1, \quad (392)$$

with every $|\psi_l\rangle$ having Schmidt rank $\leq k$. Putting $k = 1$, a pure state has Schmidt rank one if and only if it is a product state:

$$SR(|\psi\rangle) = 1 \iff |\psi\rangle = |a\rangle \otimes |b\rangle. \quad (393)$$

So, $SN(\rho) \leq 1$ means ρ can be written as a mixture of product pure states:

$$\rho = \sum_l p_l (|a_l\rangle \langle a_l| \otimes |b_l\rangle \langle b_l|), \quad (394)$$

which is the definition of a separable state [36]. Moreover, $SN(\rho) \geq 1$ always, since $SR \neq 0$ for non-zero pure states $\ker \psi$. Hence,

$$SN(\rho) = 1 \iff \rho \text{ separable}. \quad (395)$$

This means that the sub-normalised state $(E_{a|\chi, \vec{r}} \otimes \mathbb{1})\phi^+(E_{a|\chi, \vec{r}} \otimes \mathbb{1})$ lives on a Hilbert space isomorphic to $\mathcal{L}_+(\mathbb{C}^{r_a} \otimes \mathbb{C}^d)$.

We are allowed to stochastically apply CPTP maps $\Lambda_{\chi, \vec{r}}$ on one of the systems. However, these are local operations without post-selection. Such operations cannot increase the Schmidt number [60]. Hence, it is concluded that $SN(\sigma_{a|\vec{r}}) \leq r_a$. $\square$

L Proof of Theorem 2

Theorem 2. (*Qubit instrument simulation*) For instruments with qubit input, Theorem 1 is both necessary and sufficient.

Proof. ¹² For an N -outcome qubit instrument with Choi operators η_a , Theorem 1 is necessary for the instrument to be a PI. Here, it is proven that it is also sufficient.

Assume there exist $\sigma_{a|\vec{r}}$ fulfilling the conditions in Eq. (1). Qubit PVMs have two distinct types of rank vectors: $\vec{s} = (2, 0, \dots, 0)$ and $\vec{t} = (1, 1, 0, \dots, 0)$, up to permutations. The permutation of $\vec{s}$ with 2 in the i -th element position is denoted $\vec{s}_i$, while the permutation of $\vec{t}$ with 1 as the j -th and k -th ($j \neq k$) elements as $\vec{t}_{jk}$. The outcome $a = i$ for rank vector $\vec{s}_i$ corresponds to the rank-2 projector $E_{i|\vec{s}_i} = \mathbb{1}$, while the outcomes $a \neq i$ correspond to null matrices. Similarly, for the rank vector $\vec{t}_{jk}$, outcomes $a = j, k$ correspond to orthogonal rank-1 projectors $E_{a|\chi, \vec{t}_{jk}} = \varphi_{a|\chi, \vec{t}_{jk}}$ satisfying $\varphi_{j|\chi, \vec{t}_{jk}} = \mathbb{1} - \varphi_{k|\chi, \vec{t}_{jk}}$. Meanwhile, $a \neq j, k$ are associated with null matrices.

Recalling from the main text (Eq. (4.3)), we have

$$\eta_a = \sum_{\chi, \vec{r}} q_{\chi, \vec{r}} (\Lambda_{\chi, \vec{r}} \otimes \mathbb{1}) [\nu_{a|\chi, \vec{r}}], \quad (396)$$

where $\nu_{a|\chi, \vec{r}} = (E_{a|\chi, \vec{r}} \otimes \mathbb{1})\phi^+(E_{a|\chi, \vec{r}} \otimes \mathbb{1})$. Combining this with the above rank-decomposition, we get for a PI that the Choi operator becomes:

$$\eta_a = \sum_{i=1}^N q_{\vec{s}_i} (\Lambda_{\vec{s}_i} \otimes \mathbb{1}) [E_{a|\vec{s}_i} \otimes \mathbb{1} \phi^+ E_{a|\vec{s}_i} \otimes \mathbb{1}] \quad (397)$$

$$+ \sum_{j < k} \sum_{\chi} q_{\chi, \vec{t}_{jk}} (\Lambda_{\chi, \vec{t}_{jk}} \otimes \mathbb{1}) [E_{a|\chi, \vec{t}_{jk}} \otimes \mathbb{1} \phi^+ E_{a|\chi, \vec{t}_{jk}} \otimes \mathbb{1}] \quad (398)$$

$$= q_{\vec{s}_a} (\Lambda_{\vec{s}_a} \otimes \mathbb{1}) [\phi^+] + \sum_{j < k} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \frac{1}{2} \Lambda_{\chi, \vec{t}_{jk}} [\varphi_{a|\chi, \vec{t}_{jk}}] \otimes \varphi_{a|\chi, \vec{t}_{jk}}^T, \quad (399)$$

¹²This proof is based on [7].

where the identity¹³ $(E \otimes \mathbb{1})\phi^+(E \otimes \mathbb{1}) = (E \otimes E^T)/2$ was used. The first term in the last equality corresponds to a generic form of a sub-normalized Choi state, which corresponds to the constraints in Eq. (1) for $\sigma_{a|\vec{s}_a}$ ¹⁴. It comes from the fact that $E_{a|\vec{s}_i} = \mathbb{1}$ iff $a = i$ and 0 otherwise, leading to $(\mathbb{1} \otimes \mathbb{1})\phi^+(\mathbb{1} \otimes \mathbb{1}) = \phi^+$. Moreover, since $\vec{t}_{jk} = (0, \dots, 1, \dots, 1, \dots, 0)$, $j < k$ is to avoid count same rank vector twice, could have been $k < j$ instead. Equivalently we could have written $\frac{1}{2} \sum_{j \neq k}$.

Considering the second term in the last equality, it can be noted that $\Lambda_{\chi, \vec{t}_{jk}}$ is an arbitrary CPTP map for PIs. Hence, whenever $a \in \{j, k\}$, we can generate any desired state $\xi_{a, \chi, \vec{t}_{jk}} := \Lambda_{\chi, \vec{t}_{jk}}[\varphi_{a|\chi, \vec{t}_{jk}}]$. Consequently, the second term becomes

$$\sum_{j < k} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \frac{1}{2} \xi_{a, \chi, \vec{t}_{jk}} \otimes \varphi_{a|\chi, \vec{t}_{jk}}^T = \sum_{j < k} \begin{cases} T_{a, jk} & a \in \{j, k\} \\ 0 & \text{otherwise} \end{cases}, \quad (400)$$

where $T_{a, jk} = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \xi_{a, \chi, \vec{t}_{jk}} \otimes \varphi_{a|\chi, \vec{t}_{jk}}^T$ is a separable operator with $\text{Tr}(T_{a, jk}) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} = \frac{q_{\vec{t}_{jk}}}{2}$. Each $T_{a, jk}$ given by

$$T_{a, jk} := \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \xi_{a, \chi, \vec{t}_{jk}} \otimes \varphi_{a|\chi, \vec{t}_{jk}}^T \quad (401)$$

is the generic form of a separable ($SN = 1$) trace-half (up to sub-normalization $q_{\vec{t}_{jk}}$) operator. This is a positive product operator, and $T_{a|jk}$ is a sum of them, which follows the definition of a separable form. This is also a generic form since any separable (subnormalized) operator can be written as a mixture of pure product terms. Following Eq. (400), $T_{a, jk}$ is 0 for $a \notin \{j, k\}$ since $E_{a|\chi, \vec{t}_{jk}} = 0$ for that case. Fixing the instrument outcome label a . Hence, the conditions in Theorem (1) imply projective simulability for qubit instruments.

To see why $\text{Tr}(T_{a, jk}) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} = \frac{q_{\vec{t}_{jk}}}{2}$, consider the following. We have already defined

$$T_{a, jk} := \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \xi_{a, \chi, \vec{t}_{jk}} \otimes \varphi_{a|\chi, \vec{t}_{jk}}^T, \quad (402)$$

with $\xi_{a, \chi, \vec{t}_{jk}} := \Lambda_{\chi, \vec{t}_{jk}}[\varphi_{a|\chi, \vec{t}_{jk}}]$. Taking trace and using $\text{Tr}(A \otimes B) = \text{Tr}(A) \text{Tr}(B)$:

$$\text{Tr}(T_{a, jk}) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \text{Tr}(\xi_{a|\chi, \vec{t}_{jk}}) \text{Tr}(\varphi_{a|\chi, \vec{t}_{jk}}^T). \quad (403)$$

Since $\varphi_{a|\chi, \vec{t}_{jk}}$ is a rank-1 projector, so $\text{Tr}(\varphi) = 1$ and transpose does not change trace:

$$\text{Tr}(\varphi_{a|\chi, \vec{t}_{jk}}^T) = \text{Tr}(\varphi_{a|\chi, \vec{t}_{jk}}) = 1. \quad (404)$$

Moreover, $\Lambda_{\chi, \vec{t}_{jk}}$ is trace-preserving, hence, it preserves trace of its input:

$$\text{Tr}(\xi_{a, \chi, \vec{t}_{jk}}) = \text{Tr}(\Lambda_{\chi, \vec{t}_{jk}}[\varphi_{a|\chi, \vec{t}_{jk}}]) = \text{Tr}(\varphi_{a|\chi, \vec{t}_{jk}}) = 1. \quad (405)$$

¹³This is proven at the end of the section.

¹⁴Note that $SN \leq 2$ always holds for all $\sigma_{a|\vec{r}} \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathbb{C}^2)$.

Hence, each summand contributes just $q_{\chi, \vec{t}_{jk}}$, giving:

$$\text{Tr}(T_{a,jk}) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}}. \quad (406)$$

We can define the total weight of the rank-vector as:

$$q_{\vec{t}_{jk}} := \sum_{\chi} q_{\chi, \vec{t}_{jk}}, \quad (407)$$

so

$$\text{Tr}(T_{a,jk}) = \frac{q_{\vec{t}_{jk}}}{2}. \quad (408)$$

This is exactly the trace condition from Theorem 1 with $d = 2$ and $r_a = 1$. For the partial-trace normalization. To see this, consider the following. For fixed j, k , the two rank-one projectors satisfy

$$\varphi_{j|\chi, \vec{t}_{jk}} + \varphi_{k|\chi, \vec{t}_{jk}} = 1, \quad (409)$$

because they are the two outcomes of a qubit PVM. Therefore,

$$\sum_{a=j,k} \text{Tr}_{A'}(T_{a,jk}) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \left(\varphi_{j|\chi, \vec{t}_{jk}}^T + \varphi_{k|\chi, \vec{t}_{jk}}^T \right) = \frac{1}{2} \sum_{\chi} q_{\chi, \vec{t}_{jk}} \mathbb{1} = q_{\vec{t}_{jk}} \frac{1}{2}. \quad (410)$$

This corresponds exactly to the second normalization condition in Theorem 1, specifically for $d = 2$. Hence, $\sigma_{a|\vec{t}_{jk}}$ can be identified with $T_{a,jk}$. $\square$

Proof. We end the section by showing where the identity $(E \otimes \mathbb{1})\phi^+(E \otimes \mathbb{1}) = (E \otimes E^T)/2$ comes from. Considering qubits, $d = 2$, we have, in general, for any operator A

$$(A \otimes \mathbb{1})|\phi^+\rangle = (\mathbb{1} \otimes A^T)|\phi^+\rangle, \quad (411)$$

and

$$\phi^+ = \frac{1}{d} \sum_{ij}^{d-1} |i\rangle \langle j| \otimes |i\rangle \langle j|. \quad (412)$$

Apply $(E \otimes \mathbb{1})$ on the left:

$$(E \otimes \mathbb{1})\phi^+ = \frac{1}{d} \sum_{i,j} (E|i\rangle \langle j|) \otimes (|i\rangle \langle j|), \quad (413)$$

and on the right:

$$(E \otimes \mathbb{1})\phi^+(E \otimes \mathbb{1}) = \frac{1}{d} \sum_{i,j} (E|i\rangle \langle j| E) \otimes (|i\rangle \langle j|). \quad (414)$$

The key identity

$$(\mathbb{1} \otimes E^T)|i\rangle \otimes |j\rangle = |i\rangle \otimes \sum_k (E^T)_{kj} |k\rangle = |i\rangle \otimes \sum_k E_{jk} |k\rangle, \quad (415)$$

comes from

$$(1 \otimes E^T)(|i\rangle \otimes |j\rangle) = |i\rangle \otimes (E^T |j\rangle). \quad (416)$$

Expand $E^T |j\rangle$ in the computational basis

$$A |j\rangle = \sum_k |k\rangle \langle k| A |j\rangle = \sum_k A_{kj} |k\rangle, \quad (417)$$

where completeness was inserted and $A_{kj} = \langle k| A |j\rangle$ are the matrix elements in that basis. Applying this to $A = E^T$ gives

$$E^T |j\rangle = \sum_k (E^T)_{kj} |k\rangle. \quad (418)$$

Now replace the transpose matrix elements by the original ones. By definition of transpose is $(E^T)_{kj} = E_{jk}$. Therefore,

$$(\mathbb{1} \otimes E^T) |i\rangle \otimes |j\rangle = |i\rangle \otimes \sum_k E_{jk} |k\rangle. \quad (419)$$

From this identity, we can use the transpose trick

$$(E \otimes \mathbb{1}) |\phi^+\rangle = (\mathbb{1} \otimes E^T) |\phi^+\rangle. \quad (420)$$

This is seen by

$$(E \otimes \mathbb{1}) |\phi^+\rangle = \frac{1}{\sqrt{d}} \sum_i (E |i\rangle) \otimes |i\rangle = \frac{1}{\sqrt{d}} \sum_{i,k} E_{ki} |k\rangle \otimes |i\rangle \quad (421)$$

$$= \frac{1}{\sqrt{d}} \sum_{i,k} |i\rangle \otimes E_{ik} |k\rangle = (\mathbb{1} \otimes E^T) |\phi^+\rangle. \quad (422)$$

Using Eq. (420) on both sides of $\phi^+ = |\phi^+\rangle \langle \phi^+|$ gives

$$(E \otimes \mathbb{1}) \phi^+ (E \otimes \mathbb{1}) = (\mathbb{1} \otimes E^T) \phi^+ (\mathbb{1} \otimes E^T)^\dagger. \quad (423)$$

If E is Hermitian (true for a projector), then $(E^T)^\dagger = E^T$. Hence,

$$(E \otimes \mathbb{1}) \phi^+ (E \otimes \mathbb{1}) = (\mathbb{1} \otimes E^T) \phi^+ (\mathbb{1} \otimes E^T). \quad (424)$$

Evaluating $(\mathbb{1} \otimes E^T) \phi^+ (\mathbb{1} \otimes E^T)$ yields

$$(\mathbb{1} \otimes E^T) \phi^+ (\mathbb{1} \otimes E^T) = \frac{1}{d} \sum_{i,j} |i\rangle \langle j| \otimes (E^T |i\rangle \langle j| E^T). \quad (425)$$

If E is a projector, then $E^2 = E$, likewise $(E^T)^2 = E^T$. In that case, the support is restricted and we can rewrite this compactly as

$$(E \otimes \mathbb{1}) \phi^+ (E \otimes \mathbb{1}) = \frac{1}{d} (E \otimes E^T), \quad (426)$$

where ϕ^+ is the unnormalized ϕ^+ and where trace cyclicity was used. With the normalized convention, the identity is

$$(E \otimes \mathbb{1}) \phi^+ (E \otimes \mathbb{1}) = \frac{1}{d} (E \otimes E^T) \quad (427)$$

for rank-1 projectors E . For qubits, $d = 2$. □

M Motivation for using the positive partial transpose criterion for qubit-qubit instruments

We showed in the previous section that the separability across the Choi bipartition shows why Theorem 1 becomes exact for qubit instruments. For qubit-qubit instruments, the separability condition can be written as the sufficient PPT constraint $\sigma_{a|\vec{r}}^{T_A}$.

To see this, consider the following. For $d = 2$, the ranks r_a can only take on values 0, 1 or 2. Hence, there are three possible constraints for the Schmidt number: $SN \leq 0, 1, 2$. Since valid rank-vectors satisfy $\sum_a r_a = d = 2$, there are two types of rank vectors (up to permutation): $\vec{s} = (2, 0, \dots, 0)$ and $\vec{t} = (1, 1, 0, \dots, 0)$. For $\vec{s}$, the only non-zero projector has rank 2 (identity). Hence, the the Schmidt number constraint in Theorem 1 becomes $SN(\sigma_{a|\vec{s}}) \leq 2$, which is trivial because any bipartite operator on $\mathbb{C}^2 \otimes \mathbb{C}^2$ has $SN \leq 2$. For $\vec{t}$, the two relevant outcomes correspond to rank-one projectors. For those outcomes $r_a = 1$ and Theorem 1 becomes:

$$SN(\sigma_{a|\vec{t}}) \leq 1, \quad (428)$$

and $SN \leq 1$ means exactly separability (no entanglement across the Choi bipartition). The PPT criterion is a sufficient condition for separability on a qubit-qubit system. Therefore, for qubit-qubit systems, we can replace the Schmidt-number constraint ($SN \leq 1$) by:

$$\sigma_{a|\vec{r}}^{T_A} \succeq 0. \quad (429)$$

For qubit-qubit instruments, Theorem 1 is therefore exact, while for other cases, we have to resort to an SDP relaxation [7].

N Unsharp observables

Consider an instrument realizing an unsharp Pauli observable with an eigenbasis $\{|0\rangle, |1\rangle\}$. The Kraus operators of the Lüders instrument corresponding to an unsharp Pauli- Z observable are given by [7]:

$$K_a = \sqrt{\frac{1 - (-1)^a \gamma}{2}} |0\rangle \langle 0| + \sqrt{\frac{1 + (-1)^a \gamma}{2}} |1\rangle \langle 1|, \quad a \in \{1, 2\}, \quad \gamma \in [0, 1], \quad (430)$$

where γ is the sharpness parameter. The corresponding Choi representation becomes $\eta_a = (K_a \otimes \mathbb{1})\phi^+(K_a^\dagger \otimes \mathbb{1})$. In line with our objective, an instrument corresponding to a mixture of the instrument described with a noisy instrument $\{\eta_a^{\text{noise}}\}_a$:

$$\eta_a^v \equiv v\eta_a + (1 - v)\eta_a^{\text{noise}}, \quad v \in [0, 1]. \quad (431)$$

The critical visibility for PI-simulability is equivalent to the largest v such that $\{\eta_a^v\} \in \mathcal{P}$. Using Theorem 1 and 2, the primal SDP is given by [7]:

$$\begin{aligned}
& \max_{v, \{\sigma_{a|\vec{r}}\}} v \\
& \text{s.t.} \quad \eta_a^v = \sigma_{a|(1,1)} + \sigma_{a|(2,0)} + \sigma_{a|(0,2)}, \\
& \quad \text{Tr}(\sigma_{1|(1,1)}) = \text{Tr}(\sigma_{2|(1,1)}) = \frac{q_{(1,1)}}{2}, \quad \text{Tr}(\sigma_{1|(2,0)}) = q_{(2,0)}, \quad \text{Tr}(\sigma_{2|(0,2)}) = q_{(0,2)}, \\
& \quad \sum_a \text{Tr}_{A'}(\sigma_{a|(1,1)}) = q_{(1,1)} \frac{1}{2}, \quad \text{Tr}_{A'}(\sigma_{1|(2,0)}) = q_{(2,0)} \frac{1}{2}, \quad \text{Tr}_{A'}(\sigma_{2|(0,2)}) = q_{(0,2)} \frac{1}{2}, \\
& \quad \sigma_{1|(1,1)}^{T_A} \succeq 0, \quad \sigma_{2|(1,1)}^{T_A} \succeq 0, \quad \sigma_{2|(2,0)} = \sigma_{1|(0,2)} = 0, \\
& \quad \sigma_{1|(1,1)}, \sigma_{2|(1,1)}, \sigma_{1|(2,0)}, \sigma_{2|(0,2)} \in \mathcal{L}_+(\mathbb{C}_{A'}^2 \otimes \mathbb{C}_A^2).
\end{aligned} \tag{432}$$

O Dual SDP for fixed noise

We start from the primal SDP given by

$$\begin{aligned}
(\mathbf{P}) \quad & \max_{v, \{\sigma_{a|\vec{r}_i}\}} v \\
& \text{s.t.} \quad \sum_{\vec{r}} \sigma_{a|\vec{r}} = v\eta_a + (1-v)\eta_a^{\text{noise}}, \quad \forall a \\
& \quad \text{Tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{r_a}{d}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \\
& \quad \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{1}{d}, \quad \forall \vec{r} \\
& \quad \sigma_{a|\vec{r}}^{T_{A'}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1 \\
& \quad \sigma_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2},
\end{aligned} \tag{433}$$

where three sector sets are defined in the following way: $\mathcal{R}_0 := \{(a, \vec{r}) : r_a = 0\}$, $\mathcal{R}_1 = \{(a, \vec{r}) : r_a = 1\}$, $\mathcal{R}_2 := \{(a, \vec{r}) : r_a = 2\}$, and $\mathcal{R}_{1,2} := \mathcal{R}_1 \cup \mathcal{R}_2$.

Introducing a Lagrangian multiplier for each constraint:

$$\eta_a^v = \sum_{\vec{r}} \sigma_{a|\vec{r}} \rightarrow \Gamma_a, \quad \forall a \tag{434}$$

$$\text{Tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{r_a}{d} \rightarrow y_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \tag{435}$$

$$\sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{1}{d} \rightarrow Y_{\vec{r}}, \quad \forall \vec{r} \tag{436}$$

$$\sigma_{a|\vec{r}} \succeq 0 \rightarrow Z_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \tag{437}$$

$$\sigma_{a|\vec{r}}^{T_{A'}} \succeq 0 \rightarrow R_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \tag{438}$$

$$\tag{439}$$

gives the following Lagrangian of the problem:

$$\begin{aligned}
\mathcal{L} = & v + \sum_a \text{tr} \left[\Gamma_a \left(v\eta_a + (1-v)\eta_a^{\text{noise}} - \sum_{\vec{r}} \sigma_{a|\vec{r}} \right) \right] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_{1,2}} \text{tr} [Z_{a|\vec{r}} \sigma_{a|\vec{r}}] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_1} \text{tr} [R_{a|\vec{r}} \sigma_{a|\vec{r}}^{T_{A'}}] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_{1,2}} y_{a|\vec{r}} \left(\text{tr}(\sigma_{a|\vec{r}}) - q_{\vec{r}} \frac{r_a}{d} \right) \\
& + \sum_{\vec{r}} \text{tr} \left[Y_{\vec{r}} \left(\sum_a \text{tr}_{A'}(\sigma_{a|\vec{r}}) - q_{\vec{r}} \frac{\mathbb{1}_A}{d} \right) \right].
\end{aligned} \tag{440}$$

Rewriting the Lagrangian using standard adjoint identities $\text{Tr}[R\sigma^{T_{A'}}] = \text{Tr}[R^{T_{A'}}\sigma]$, and $\text{Tr}[Y \text{Tr}_{A'}(\sigma)] = \text{Tr}[(\mathbb{1}_{A'} \otimes Y)\sigma]$ gives

$$\begin{aligned}
\mathcal{L} = & \sum_a \text{tr}(\Gamma_a \eta_a^{\text{noise}}) \\
& + v \left[1 + \sum_a \text{tr}(\Gamma_a(\eta_a - \eta_a^{\text{noise}})) \right] \\
& + \sum_{\vec{r}} q_{\vec{r}} \left[-\frac{1}{d} \text{tr}(Y_{\vec{r}}) - \frac{1}{d} \sum_{a:(a,\vec{r}) \in \mathcal{R}_{1,2}} r_a y_{a|\vec{r}} \right] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_2} \text{tr} \left[\sigma_{a|\vec{r}} (-\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}}) \right] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_1} \text{tr} \left[\sigma_{a|\vec{r}} (-\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}}) \right].
\end{aligned} \tag{441}$$

Now, making the Lagrangian independent of the primal variables yields:

$$v : \quad 1 + \sum_a \text{Tr}(\Gamma_a(\eta_a - \eta_a^{\text{noise}})) = 0 \tag{442}$$

$$q_{\vec{r}} : \quad -\frac{1}{d} \text{Tr}(Y_{\vec{r}}) - \frac{1}{d} \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \tag{443}$$

and for the coefficients for $\sigma_{a|\vec{r}}$:

$$-\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \tag{444}$$

$$-\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2. \tag{445}$$

Combined with:

$$Z_{a|\vec{r}} \succeq 0, \quad (a, \vec{r}) \in \mathcal{R}_{1,2}, \quad R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \tag{446}$$

the Lagrangian reduces to:

$$\mathcal{L} = \sum_a \text{Tr}(\Gamma_a \eta_a^{\text{noise}}). \quad (447)$$

The full dual becomes:

$$\begin{aligned} (\text{D}) \quad & \min \quad \sum_a \text{tr}(\Gamma_a \eta_a^{\text{noise}}) \\ \text{s.t.} \quad & 1 + \sum_a \text{tr}(\Gamma_a (\eta_a - \eta_a^{\text{noise}})) = 0, \\ & \text{tr}(Y_{\vec{r}}) + \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\ & -\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\ & -\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\ & Z_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \\ & R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1. \end{aligned} \quad (448)$$

By eliminating the slack variables,

$$Z_{a|\vec{r}} = \Gamma_a - R_{a|\vec{r}}^{T_{A'}} - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad (449)$$

$$Z_{a|\vec{r}} = \Gamma_a - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad (450)$$

and similarly, rewriting the objective function gives the final reduced dual:

$$\begin{aligned} (\text{D}) \quad & \min_{\{\Gamma_a\}, \{Y_{\vec{r}}\}, \{y_{a|\vec{r}}\}, \{R_{a|\vec{r}}\}_{\mathcal{R}_1}} \quad 1 + \sum_a \text{tr}(\Gamma_a \eta_a) \\ \text{s.t.} \quad & 1 + \sum_a \text{tr}(\Gamma_a (\eta_a - \eta_a^{\text{noise}})) = 0, \\ & \text{tr}(Y_{\vec{r}}) + \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\ & \Gamma_a - R_{a|\vec{r}}^{T_{A'}} - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\ & \Gamma_a - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\ & R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1. \end{aligned} \quad (451)$$

The feasible solutions $\{\Gamma_a\}$ of the dual problem above provide a witness to detect $\mathcal{I} \notin \mathcal{P}_{\mathcal{I}}$. Hence, we can define:

$$W(\eta) = \sum_a \text{Tr}(\Gamma_a \eta_a). \quad (452)$$

P Dual SDP for worst-case noise

Considering the noise contribution to be a variable: $\tilde{\eta}_a^{\text{noise}} := (1 - v)\eta_a^{\text{noise}}$. This choice avoids a variable multiplied by a variable. This is subnormalized and PSD, $\tilde{\eta}_a^{\text{noise}} \succeq 0$, since $\{\eta_a^{\text{noise}}\}_a$ is an instrument, which also means that $\sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) = \frac{1-v}{d} \mathbb{1}_A$. Hence,

the primal SDP becomes

$$\begin{aligned}
(\mathbf{P}) \quad & \max_{v, \{\sigma_{a|\vec{r}}\}, \{\tilde{\eta}_a^{\text{noise}}\}, \{q_{\vec{r}}\}} v \\
\text{s.t.} \quad & \sum_{\vec{r}} \sigma_{a|\vec{r}} = q_{\vec{r}} \frac{r_a}{d}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \\
& \sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{1_A}{d}, \quad \forall \vec{r}, \\
& \sigma_{a|\vec{r}}^{T_{A'}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& \sigma_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \\
& \tilde{\eta}_a^{\text{noise}} \succeq 0, \quad \forall a, \\
& \sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) = \frac{1-v}{d} \mathbb{1}_A.
\end{aligned} \tag{453}$$

We introduce the following multipliers:

$$v\eta_a + \tilde{\eta}_a^{\text{noise}} = \sum_{\vec{r}} \sigma_{a|\vec{r}} \rightarrow \Gamma_a, \quad \forall a \tag{454}$$

$$\text{Tr}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{r_a}{d} \rightarrow y_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \tag{455}$$

$$\sum_a \text{Tr}_{A'}(\sigma_{a|\vec{r}}) = q_{\vec{r}} \frac{\mathbb{1}}{d} \rightarrow Y_{\vec{r}}, \quad \forall \vec{r} \tag{456}$$

$$\sigma_{a|\vec{r}} \succeq 0 \rightarrow Z_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \tag{457}$$

$$\sigma_{a|\vec{r}}^{T_{A'}} \succeq 0 \rightarrow R_{a|\vec{r}}, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \tag{458}$$

$$\tilde{\eta}_a^{\text{noise}} \succeq 0 \rightarrow N_a, \quad \forall a, \tag{459}$$

$$\sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) = \frac{1-v}{d} \mathbb{1} \rightarrow \beta, \quad \forall \beta \in \mathcal{L}_+(\mathcal{H}_A). \tag{460}$$

Then,

$$\begin{aligned}
\mathcal{L} = & v + \sum_a \text{tr} \left[\Gamma_a \left(v\eta_a + \tilde{\eta}_a^{\text{noise}} - \sum_{\vec{r}} \sigma_{a|\vec{r}} \right) \right] \\
& + \sum_{(a, \vec{r}) \in \mathcal{R}_{1,2}} \text{tr} [Z_{a|\vec{r}} \sigma_{a|\vec{r}}] \\
& + \sum_{(a, \vec{r}) \in \mathcal{R}_1} \text{tr} [R_{a|\vec{r}} \sigma_{a|\vec{r}}^{T_{A'}}] \\
& + \sum_{(a, \vec{r}) \in \mathcal{R}_{1,2}} y_{a|\vec{r}} \left(\text{tr}(\sigma_{a|\vec{r}}) - q_{\vec{r}} \frac{r_a}{d} \right) \\
& + \sum_{\vec{r}} \text{tr} \left[Y_{\vec{r}} \left(\sum_a \text{tr}_{A'}(\sigma_{a|\vec{r}}) - q_{\vec{r}} \frac{\mathbb{1}_A}{d} \right) \right] \\
& + \sum_a \text{Tr}(N_a \tilde{\eta}_a^{\text{noise}}) + \text{Tr} \left[\beta \left(\sum_a \text{Tr}_{A'}(\tilde{\eta}_a^{\text{noise}}) - \frac{1-v}{d} \mathbb{1}_A \right) \right].
\end{aligned} \tag{461}$$

Using the identities $\text{Tr}[R\sigma^{T_{A'}}] = \text{Tr}[R^{T_{A'}}\sigma]$, and $\text{Tr}[Y \text{Tr}_{A'}(\sigma)] = \text{Tr}[(\mathbb{1}_{A'} \otimes Y)\sigma]$, the Lagrangian becomes

$$\begin{aligned}
\mathcal{L} = & -\frac{1}{d} \text{Tr}(\beta) + v \left[1 + \sum_a \text{Tr}(\Gamma_a \eta_a) + \frac{1}{d} \text{Tr}(\beta) \right] \\
& + v \left[1 + \sum_a \text{tr}(\Gamma_a(\eta_a - \eta_a^{\text{noise}})) \right] \\
& + \sum_{\vec{r}} q_{\vec{r}} \left[-\frac{1}{d} \text{tr}(Y_{\vec{r}}) - \frac{1}{d} \sum_{a:(a,\vec{r}) \in \mathcal{R}_{1,2}} r_a y_{a|\vec{r}} \right] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_2} \text{tr} \left[\sigma_{a|\vec{r}} (-\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}}) \right] \\
& + \sum_{(a,\vec{r}) \in \mathcal{R}_1} \text{tr} \left[\sigma_{a|\vec{r}} (-\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}}) \right] \\
& + \sum_a \text{Tr} [N_a (\Gamma_a + N_a + \mathbb{1}_{A'} \otimes \beta)] .
\end{aligned} \tag{462}$$

To get independence of the primal variables:

$$v : \quad 1 + \sum_a \text{Tr}(\Gamma_a \eta_a) + \frac{1}{d} \text{Tr}(\beta) = 0, \tag{463}$$

$$q_{\vec{r}} : \quad -\frac{1}{d} \text{Tr}(Y_{\vec{r}}) - \frac{1}{d} \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2} \tag{464}$$

$$\sigma_{a|\vec{r}} : \quad -\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \tag{465}$$

$$\sigma_{a|\vec{r}} : \quad -\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2 \tag{466}$$

$$N_a : \quad \Gamma_a + N_a + \mathbb{1}_{A'} \otimes \beta = 0, \quad \forall a. \tag{467}$$

Hence, the Lagrangian reduces to

$$\mathcal{L} = -\frac{1}{d} \text{Tr}(\beta). \tag{468}$$

Then, the full dual SDP is given by

$$\begin{aligned}
(\mathbf{D}) \quad & \min_{\Gamma, Y, y, Z, R, W, \beta} \quad -\frac{1}{d} \text{tr}(\beta) \\
\text{s.t.} \quad & 1 + \sum_a \text{tr}(\Gamma_a \eta_a) + \frac{1}{d} \text{tr}(\beta) = 0, \\
& \text{tr}(Y_{\vec{r}}) + \sum_{a:(a,\vec{r}) \in \mathcal{R}_{1,2}} r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\
& -\Gamma_a + Z_{a|\vec{r}} + R_{a|\vec{r}}^{T_{A'}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& -\Gamma_a + Z_{a|\vec{r}} + y_{a|\vec{r}} \mathbb{1}_{A'A} + \mathbb{1}_{A'} \otimes Y_{\vec{r}} = 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\
& \Gamma_a + W_a + \mathbb{1}_{A'} \otimes \beta = 0, \quad \forall a, \\
& Z_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_{1,2}, \\
& R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& W_a \succeq 0, \quad \forall a.
\end{aligned} \tag{469}$$

Eliminating slack variables $Z_{a|\vec{r}}$ and N_a as before and rewriting the objective function gives

$$\begin{aligned}
(\mathbf{D}) \quad & \min_{\{\Gamma_a\}, \{Y_{\vec{r}}\}, \{y_{a|\vec{r}}\}, \{R_{a|\vec{r}}\}_{\mathcal{R}_1}} 1 + \sum_a \text{tr}(\Gamma_a \eta_a) \\
\text{s.t.} \quad & 1 + \sum_a \text{tr}(\Gamma_a(\eta_a)) + \frac{1}{d} \text{Tr}(\beta) = 0, \\
& \text{tr}(Y_{\vec{r}}) + \sum_a r_a y_{a|\vec{r}} = 0, \quad \forall \vec{r}, \\
& \Gamma_a - R_{a|\vec{r}}^{T_{A'}} - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1, \\
& \Gamma_a - y_{a|\vec{r}} \mathbb{1}_{A'A} - \mathbb{1}_{A'} \otimes Y_{\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_2, \\
& -\Gamma_a - \mathbb{1}_{A'} \otimes \beta \succeq 0, \quad \forall a, \\
& R_{a|\vec{r}} \succeq 0, \quad \forall (a, \vec{r}) \in \mathcal{R}_1.
\end{aligned} \tag{470}$$

Q Derivation of $v_{\text{wc}}^{(3)}(\zeta)$

In the spirit of the method used to obtain the three-outcome POVM worst-case noise visibility, the worst-case noise visibility for the three-outcome instrument can be derived by using symmetry to reduce the SDP equality to a small set of scalar equations, then solve the remaining optimization analytically. The idea is to express the noisy decomposition equation in the SDP using the Bloch representation of a POVM $E_a = \lambda_a^E (\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma})$. Then, by imposing the symmetry of the weights and Bloch vectors alongside variable substitutions, the problem reduces to a quadratic equation for v . However, for a Lüders instrument, the POVM Bloch vectors become insufficient and we therefore make use of the associated Choi operators of the instrument [7], since a QI takes into account both the outcome probabilities and the post-measurement state. Hence, the Choi matrices give rise to the scalar equations in this derivation.

Since the POVM extracted from the optimal instrument represented by η_{opt} corresponds to an equal-weight trine POVM, the analytical visibility curve is restricted to the equal-weight trine POVM family parameterized by a sharpness/shrinking factor ζ . No other parameters are needed as all equal-weight trine POVMs will always lie in a plane due to symmetry.

Q.1 Preliminaries

Recall that a qubit POVM element can be written as

$$E_a = \lambda_a^E (\mathbb{1} + \vec{n}_a^E \cdot \vec{\sigma}) \tag{471}$$

with $|\vec{n}_a| \leq 1$, and $\sum_a \lambda_a = 1$, and $\sum_a \lambda_a^E \vec{n}_a^E = 0$. For an equal-weight trine POVM, these constraints lead to $\lambda_a^E = \frac{1}{3}$ and

$$\sum_a \vec{n}_a = 0, \quad \vec{n}_a \cdot \vec{n}_b = -\frac{1}{2} \quad a \neq b. \tag{472}$$

We can further conveniently fix the three Bloch vector directions to be

$$\vec{n}_1 = (0, 0, 1), \quad \vec{n}_2 = \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2} \right), \quad \vec{n}_3 = \left(-\frac{\sqrt{3}}{2}, 0, -\frac{1}{2} \right). \quad (473)$$

Since the geometry is fixed up to a global rotation for the equal-weight trine, the only parameter in this class of POVMs is the scalar variable $\zeta \in [0, 1]$. Hence, the complete equal-weight trine POVM family considered is

$$E_a(\zeta) = \frac{1}{3}(\mathbb{1} + \zeta \tilde{\vec{n}}_a^E \cdot \vec{\sigma}), \quad 0 \leq \zeta \leq 1, \quad (474)$$

where $\zeta \tilde{\vec{n}}_a^E := \vec{n}_a^E$, comparing to Eq. (471), and $\zeta = |\vec{n}_a^E|$. It can be noted that a value of $\zeta = 0$ gives a random-output POVM while $0 < \zeta < 1$ produces an unsharp trine and $\zeta = 1$ generates a sharp rank-one trine. Any other equal-weight trine is obtained by a global Bloch rotation, corresponding to a unitary conjugation

$$E_a \mapsto U E_a U^\dagger. \quad (475)$$

The SDP value is invariant under this change of basis, hence $v_{\text{opt}} = v_{\text{opt}}(\zeta)$.

Moreover, the Lüders instrument corresponding to $E_a(\zeta)$ has the associated Kraus operators

$$K_a(\zeta) = \sqrt{E_a(\zeta)}, \quad (476)$$

and the Choi operator for outcome a is (see Eq. (2.23) from the main text):

$$\eta_a(\zeta) = (K_a \otimes \mathbb{1}) |\phi^+\rangle \langle \phi^+| (K_a^\dagger \otimes \mathbb{1}). \quad (477)$$

In order to write E_a on Lüders form, it is useful to write it on diagonal form. The diagonalization of a trine effect can be done in the following way. Fix outcome a in the eigenbasis of $\vec{n}_a^E \cdot \vec{\sigma}$ ($\{|+_a\rangle, |-_a\rangle\}$) and consider the identity

$$(\zeta \vec{n}_a^E \cdot \vec{\sigma})^2 = \zeta^2 |\vec{n}_a^E|^2 \mathbb{1} = \mathbb{1}. \quad (478)$$

From this, the eigenvalue of $\vec{n}_a^E \cdot \vec{\sigma}$, denoted λ , fulfills

$$\lambda^2 = \zeta^2 \iff \lambda = \pm \zeta. \quad (479)$$

Hence, in the basis $\{|+_a\rangle, |-_a\rangle\}$, the matrix, on spectral form is given by

$$\vec{n}_a \cdot \vec{\sigma} = \begin{bmatrix} \zeta & 0 \\ 0 & -\zeta \end{bmatrix}. \quad (480)$$

Therefore, the matrix form of the POVM effect given by Eq. (474) is given by

$$E_a(\zeta) = \frac{1}{3} \begin{bmatrix} 1 + \zeta & 0 \\ 0 & 1 - \zeta \end{bmatrix} \quad (481)$$

in this basis. Hence, the Kraus operator becomes

$$K_a(\zeta) = \sqrt{E_a(\zeta)} = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{1 + \zeta} & 0 \\ 0 & \sqrt{1 - \zeta} \end{bmatrix}. \quad (482)$$

Applying it to one side of $|\phi^+\rangle$ ¹⁵, i.e.,

$$(K_a \otimes \mathbb{1}) |\phi^+\rangle = \frac{1}{\sqrt{6}} \left(\sqrt{1+\zeta} |+_a+_a\rangle + \sqrt{1-\zeta} |-_a-_a\rangle \right), \quad (483)$$

gives the Choi operator matrix form

$$\eta_a(\zeta) = \frac{1}{6} \begin{bmatrix} 1+\zeta & 0 & 0 & \sqrt{1-\zeta^2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sqrt{1-\zeta^2} & 0 & 0 & 1-\zeta \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 1+\zeta & 0 & 0 & \delta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \delta & 0 & 0 & 1-\zeta \end{bmatrix} \quad (484)$$

in the basis $\{|+_a+_a\rangle, |+_a-_a\rangle, |-_a+_a\rangle, |-_a-_a\rangle\}$ where $\delta := \sqrt{1-\zeta^2}$. The determinant of this matrix is zero ($\det \eta_a(\zeta) = (1+\zeta)(1-\zeta) - \delta^2 = 1 - \zeta^2 - (1 - \zeta^2) = 0$). Therefore, each Lüders Choi operator is of rank one [42].

Q.2 Positivity condition

Recall that the noisy implementation for the worst-case noise is

$$\eta_a^{(v)} = v\eta_a(\zeta) + N_a \iff \eta_a^{(v)} - v\eta_a(\zeta) = N_a \succeq 0, \quad (485)$$

is projectively simulable, where $N_a \succeq 0$ is the arbitrary worst-case noise to be optimized. It should be noted that the full condition for positivity is

$$\eta_a^{(v)} - v\eta_a(\zeta) \succeq 0 \quad \forall a = 1, 2, 3. \quad (486)$$

However, the Bloch vectors of the POVM family considered are cyclic by applying rotations, hence POVM effects are cyclic. Thereby, the Lüders Kraus operators are cyclic too. Hence, all three constraints are the same up to unitary rotation:

$$\eta_a(\zeta) = U_a \eta_1(\zeta) U_a^\dagger, \quad (487)$$

and since unitary conjugation preserves positive semidefiniteness,

$$\eta_1^{(v)}(t) - v\eta_1(\zeta) \succeq 0 \iff \eta_a^{(v)} - v\eta_a^{(v)}(\zeta) \succeq 0, \quad \forall a. \quad (488)$$

The global visibility for a fixed t is

$$v_{\max}(\zeta, t) = \min_a v_a(\zeta, t). \quad (489)$$

However, by symmetry, we have

$$v_1(\zeta, t) = v_2(\zeta, t) = v_3(\zeta, t). \quad (490)$$

Hence, $v_{\max}(\zeta, t) = v_a(\zeta, t)$. Hence, we compute $v_{\max}$ from one outcome. While the matrices for the different outcomes would look different in a fixed basis (e.g. the computational basis) when they are transformed by unitaries, we can work in a local basis for outcome a , given by $\{|+_a\rangle, |-_a\rangle\}$, for which the matrices for the different outcomes would be equal.

¹⁵In the a -dependent eigenbasis, $|00\rangle \equiv |+_a+_a\rangle$, $|11\rangle = |-_a-_a\rangle$.

To continue, we note that for a three-outcome qubit POVM, there are two rank-vector types $\vec{s} = (2, 0, 0)$ and $\vec{t} = (1, 1, 0)$, up to permutations¹⁶ [7]. The former corresponds to the rank-two projector $\{\mathbb{1}, 0, 0\}$ up to relabeling, while the latter corresponds to a usual rank-one binary projective measurement with two rank-one outcomes and one zero outcome $\{P, I - P, 0\}$, up to relabeling. Putting this together, the full PI-simulable instrument is given by

$$\eta_a^{(v)} = \sum_i \sigma_{a|\vec{s}_i} + \sum_{j < k} \sigma_{a|\vec{t}_{jk}}. \quad (491)$$

Recall from Theorem 1 that each rank-vector block has a total weight of $q_{\vec{r}} = \text{Tr}(\sum_a \sigma_{a|\vec{r}})$. Since the full projective instrument is normalized ($\sum_{\vec{r}} q_{\vec{r}} = 1$). As noted by the numerical SDP, all weights corresponding to the same rank-vector type ($\vec{s}$ or $\vec{t}$) are equal but weights belonging to different types are not. Hence, we can define

$$t := \sum_i q_{s_i} \implies 1 - t = \sum_{j < k} q_{t_{jk}}, \quad (492)$$

where the implication follows from normalization. This leads to the following decomposition

$$\eta_a^{(v)}(t) = t \sum_i \tilde{\sigma}_{a|\vec{s}_i} + (1 - t) \sum_{j < k} \tilde{\sigma}_{a|\vec{t}_{jk}}, \quad 0 \leq t \leq 1, \quad (493)$$

where t is the weight of the identity-channel part, and $1 - t$ is the weight of the binary PVM part and $\tilde{\sigma}_{a|\vec{s}}, \tilde{\sigma}_{a|\vec{t}_{jk}}$ are the weighted projective operators $\frac{1}{t}\sigma_{a|\vec{s}_i}$ and $\frac{1}{1-t}\sigma_{a|\vec{t}_{jk}}$, respectively.

The first term in the relation above is an instrument corresponding to the identity operator, which has the Choi state

$$t \sum_i \tilde{\sigma}_{a|\vec{s}_i} = t \frac{1}{3} |\phi^+\rangle \langle \phi^+| = \frac{t}{6} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad (494)$$

which comes from Eq. (477) where $K_i = \mathbb{1}$.

The second term corresponds to rank-one PVMs. To compute this term, the key argument is that every projectively simulable instrument induces a projectively simulable POVM. We can construct a PM-simulable trine POVM directly from the PVMs. For instance, each POVM Bloch vector symmetrically separated can be simulated using two rank-one PVM Bloch vectors separated by 30° relative to the POVM Bloch vector, see Fig. 6a. Hence, pairs of PVM Bloch vectors produce effects of the exact projective decomposition

$$E_a^{\text{eff}} = \frac{1}{3}(\mathbb{1} + \mu \vec{n}_a^{E^{\text{eff}}} \cdot \vec{\sigma}), \quad a = 1, 2, 3, \quad (495)$$

where $1/3$ is the weight (which is equal for all outcomes) and μ is the Bloch radius of the effective trine POVM. This is an equal-weight trine POVM with Bloch directions $\vec{n}_a^{E^{\text{eff}}}$, as Fig. 6a indicates. The POVM defined in the equation above is the measurement part of the instrument defined by the second term in Eq. (493). Since this term is an instrument,

¹⁶The rank vector $\vec{t}$ should not be confused with the weight t .

E_a^{eff} for all a needs to be converted back to instrument form.

While rank-one PVM effects themselves have a Bloch radius of one, the value $\mu = \frac{\sqrt{3}}{2}$ is the Bloch radius of the trine POVM obtained by mixing the PVMs. For the equal-weight trine POVM simulation, we are mixing three binary PVMs. Take three binary PVMs with directions $\vec{n}_a^P$, $a = 1, 2, 3$, also forming a trine:

$$P_a^\pm = \frac{1}{2}(\mathbb{1} \pm \vec{n}_a^P \cdot \vec{\sigma}). \quad (496)$$

For the a -th outcome, we combine two neighboring PVM outcomes, e.g.

$$E_a^{\text{eff}} = \frac{1}{3}P_a^+ + \frac{1}{3}P_{a-2}^-. \quad (497)$$

The reason for the factor $1/3$ is that the trine has three symmetric outcomes. To preserve this symmetry, we mix three binary PVMs with equal probability $1/3$. Then

$$E_a^{\text{eff}} = \frac{1}{3} \left[\frac{1}{2}(\mathbb{1} + \vec{n}_a^{P^+} \cdot \vec{\sigma}) + \frac{1}{2}(\mathbb{1} - \vec{n}_{a-2}^{P^+} \cdot \vec{\sigma}) \right] = \frac{1}{3} \left[\mathbb{1} + \frac{\vec{n}_a^{P^+} - \vec{n}_{a-2}^{P^+}}{2} \cdot \vec{\sigma} \right]. \quad (498)$$

We can choose the projective directions $\vec{n}_a^{P^+}$ such that they are rotated 30° with respect to the target trine directions $\vec{n}_a^{E^{\text{eff}}}$, as Fig. 6 shows. As Fig. 6b shows, the chord between two neighboring $\vec{n}_a^{P^+}$ -vectors points exactly along $\vec{n}_a^E$. Since neighboring trine Bloch vectors are separated by 120° we get

$$\vec{n}_a^{P^+} \cdot \vec{n}_{a-1}^{P^+} = \cos(120^\circ) = -\frac{1}{2}, \quad (499)$$

so

$$\left\| \vec{n}_a^{P^+} - \vec{n}_{a-2}^{P^+} \right\| = \sqrt{|\vec{n}_a^{P^+}|^2 + |\vec{n}_{a-2}^{P^+}|^2 - 2\vec{n}_a^{P^+} \cdot \vec{n}_{a-2}^{P^+}} = \sqrt{1 + 1 - 2\left(-\frac{1}{2}\right)} = \sqrt{3}. \quad (500)$$

Hence,

$$\vec{n}_a^{P^+} - \vec{n}_{a-2}^{P^+} = \sqrt{3} \vec{n}_a^E. \quad (501)$$

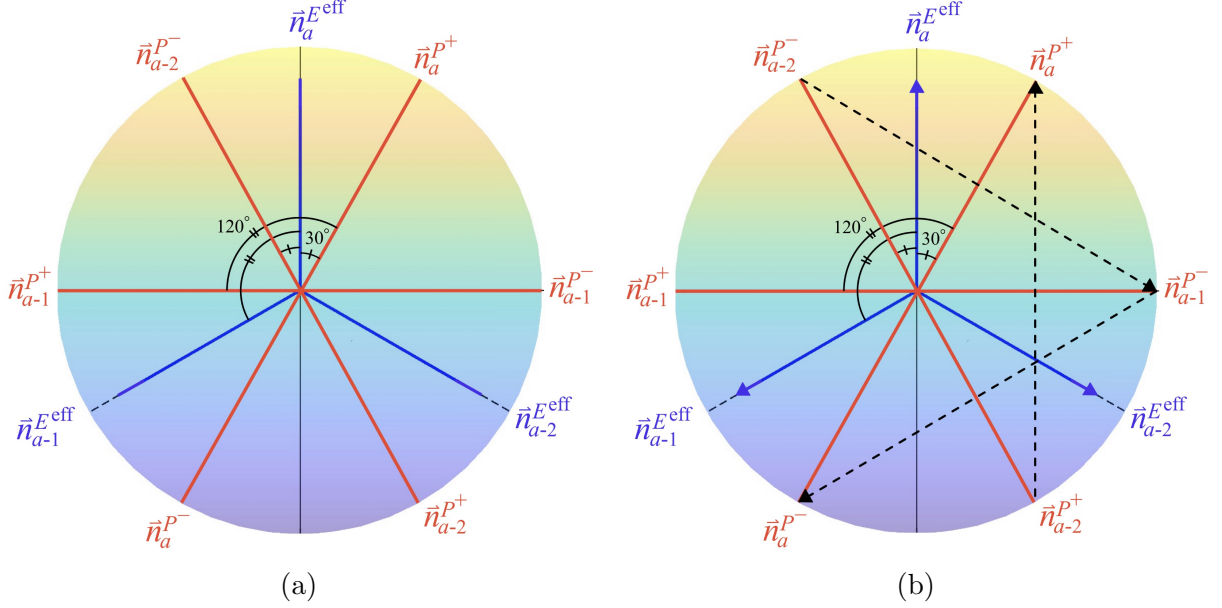


Figure 6: a) Projective decomposition of E_a^{eff} , $a = 1, 2, 3$, with Bloch vectors shown in blue by rank-one normalized binary PVMs with Bloch vectors shown in red. The $-$ sign denotes the antipodal vector of a PVM. Illustrated in the xz -plane of the Bloch sphere. b) Bloch vector of a PVM pair.

Therefore, we have $\mu = \sqrt{3}/2$:

$$E_a^{\text{eff}} = \frac{1}{3} \left(\mathbb{1} + \frac{\sqrt{3}}{2} \vec{n}_a^{\text{eff}} \cdot \vec{\sigma} \right) = \frac{1}{3} \begin{bmatrix} 1 + \mu & 0 \\ 0 & 1 - \mu \end{bmatrix}. \quad (502)$$

This value is the optimal, as seen by Fig. 6a ($\mu = \sqrt{3}/2 = \cos(30^\circ)$), projectively simulable shrinking factor of the measurement part. Since the projectively simulable set is the convex hull of binary PVMs, and this line is a supporting line of that convex hull, no projectively simulable trine POVM can have a larger Bloch radius. The largest value of μ is of relevance since we are looking for the largest visibility.

Next, we need to find the full instrument from the measurement part defined by Eq. (502) above. Recall that a Choi operator is given by

$$\eta_a = \sum_l (K_{a,l} \otimes \mathbb{1}) |\phi^+\rangle \langle \phi^+| (K_{a,l}^\dagger \otimes \mathbb{1}). \quad (503)$$

To get the instrument realization of E_a^{eff} we first note that, in the a -dependent eigenbasis $\{|+_a\rangle, |-_a\rangle\}$, we have

$$E_a^{\text{eff}} = \frac{1}{3} (\mathbb{1} + \mu \vec{n}_a \cdot \vec{\sigma}) = \frac{1}{3} \begin{bmatrix} 1 + \mu & 0 \\ 0 & 1 - \mu \end{bmatrix} \iff E_a^{\text{eff}} = \frac{1 + \mu}{3} |+_a\rangle \langle +_a| + \frac{1 - \mu}{3} |-_a\rangle \langle -_a|, \quad (504)$$

which is the spectral decomposition of E_a^{eff} . A valid instrument only needs to reproduce the effect. For an instrument branch with Kraus operators $\{K_{a,l}\}_l$, the associated POVM effect is

$$E_a = \sum_l K_{a,l}^\dagger K_{a,l}. \quad (505)$$

Comparing this to the spectral form above, we can choose

$$K_{a,+} = \sqrt{\frac{1+\mu}{3}} |+_a\rangle \langle +_a|, \quad K_{a,-} = \sqrt{\frac{1-\mu}{3}} |-_a\rangle \langle -_a|. \quad (506)$$

Then,

$$K_{a,+}^\dagger K_{a,+} = \frac{1+\mu}{3} |+_a\rangle \langle +_a|, \quad K_{a,-}^\dagger K_{a,-} = \frac{1-\mu}{3} |-_a\rangle \langle -_a|. \quad (507)$$

Hence,

$$K_{a,+}^\dagger K_{a,+} + K_{a,-}^\dagger K_{a,-} = E_a^{\text{eff}}. \quad (508)$$

So, these Kraus operators implement exactly the desired effective POVM effect. The Choi matrix then follows immediately. Using the maximally entangled state written in the a -dependent eigenbasis, $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|+_a+_a\rangle + |-_a-_a\rangle)$, we get

$$(K_{a,+} \otimes \mathbb{1})|\phi^+\rangle = \sqrt{\frac{1+\mu}{6}} |+_a+_a\rangle, \quad (K_{a,-} \otimes \mathbb{1})|\phi^+\rangle = \sqrt{\frac{1-\mu}{6}} |-_a-_a\rangle. \quad (509)$$

To see this, consider first the chosen Kraus operator from Eq. (506)

$$K_{a,+} = \sqrt{\frac{1+\mu}{3}} |+_a\rangle \langle +_a|. \quad (510)$$

Then

$$(K_{a,+} \otimes \mathbb{1})|\phi^+\rangle = (K_{a,+} \otimes I) \frac{1}{\sqrt{2}} (|+_a+_a\rangle + |-_a-_a\rangle). \quad (511)$$

Apply it term by term. First,

$$(K_{a,+} \otimes \mathbb{1})|+_a+_a\rangle = K_{a,+}|+_a\rangle \otimes |+_a\rangle = \sqrt{\frac{1+\mu}{3}} |+_a\rangle \otimes |+_a\rangle = \sqrt{\frac{1+\mu}{3}} |+_a+_a\rangle. \quad (512)$$

Second,

$$(K_{a,+} \otimes \mathbb{1})|-_a-_a\rangle = K_{a,+}|-_a\rangle \otimes |-_a\rangle = 0, \quad (513)$$

since $|+_a\rangle \langle +_a| |-_a\rangle = 0$. Therefore,

$$(K_{a,+} \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}} \sqrt{\frac{1+\mu}{3}} |+_a+_a\rangle = \sqrt{\frac{1+\mu}{6}} |+_a+_a\rangle. \quad (514)$$

Similarly, for the other chosen Kraus operator in Eq. (506)

$$K_{a,-} = \sqrt{\frac{1-\mu}{3}} |-_a\rangle \langle -_a|. \quad (515)$$

Then

$$(K_{a,-} \otimes \mathbb{1})|+_a+_a\rangle = 0, \quad (516)$$

since $K_{a,-}|+_a\rangle = 0$. Meanwhile,

$$(K_{a,-} \otimes \mathbb{1})|-_a-_a\rangle = K_{a,-}|-_a\rangle \otimes |-_a\rangle = \sqrt{\frac{1-\mu}{3}} |-_a-_a\rangle. \quad (517)$$

Thus,

$$(K_{a,-} \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}}\sqrt{\frac{1-\mu}{3}}|-_a-_a\rangle = \sqrt{\frac{1-\mu}{6}}|-_a-_a\rangle. \quad (518)$$

The corresponding Choi contributions are therefore

$$\frac{1 \pm \mu}{6} |+_a+_a\rangle \langle +_a+_a| + \frac{1 - \mu}{6} |-_a-_a\rangle \langle -_a-_a|. \quad (519)$$

Then, the Choi operator is

$$\eta_a = \frac{1+\mu}{6} |+_a+_a\rangle \langle +_a+_a| + \frac{1-\mu}{6} |-_a-_a\rangle \langle -_a-_a|. \quad (520)$$

Finally, the second term in Eq. (493) is given by:

$$\sum_{j < k} \tilde{\sigma}_{a|\vec{t}_{jk}} = \frac{1}{6} \begin{bmatrix} 1+\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\mu \end{bmatrix}. \quad (521)$$

Hence, with Eq. (494) and (521) substituted into Eq. (493) becomes

$$\eta_a^{(v)}(t) = \frac{1}{6} \begin{bmatrix} 1+\mu(1-t) & 0 & 0 & t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ t & 0 & 0 & 1-\mu(1-t) \end{bmatrix}, \quad (522)$$

where $\mu = \frac{\sqrt{3}}{2}$. The matrix above has two intuitive limits, when $t = 0$, we have

$$\eta_a^{(v)}(0) = \frac{1}{6} \begin{bmatrix} 1+\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\mu \end{bmatrix}, \quad (523)$$

which is fully measurement-like, and for $t = 1$

$$\eta_a^{(v)}(1) = \frac{1}{6} \begin{bmatrix} 1+\mu(1-t) & 0 & 0 & t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ t & 0 & 0 & 1-\mu(1-t) \end{bmatrix}, \quad (524)$$

which is the random-output identity-channel part. Hence, t controls the tradeoff between information extraction and coherence preservation.

Hence, with Eq. (522) and (484) positivity condition in Eq. (485) becomes

$$\begin{aligned} \eta_a^{(v)}(t) - v\eta_a(\zeta) &\succeq 0 \\ \iff \frac{1}{6} \begin{bmatrix} 1+\mu(1-t) & 0 & 0 & t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ t & 0 & 0 & 1-\mu(1-t) \end{bmatrix} &- v\frac{1}{6} \begin{bmatrix} 1+\zeta & 0 & 0 & \delta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \delta & 0 & 0 & 1-\zeta \end{bmatrix}. \end{aligned} \quad (525)$$

Q.3 Expression for critical visibility, $v_{\max}$

By defining

$$A(t) := \begin{bmatrix} 1 + \mu(1-t) & 0 & 0 & t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ t & 0 & 0 & 1 - \mu(1-t) \end{bmatrix}, \quad B(\zeta) := \begin{bmatrix} 1 + \zeta & 0 & 0 & \delta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \delta & 0 & 0 & 1 - \zeta \end{bmatrix}, \quad (526)$$

we get

$$A(t) - vB(\zeta) \succeq 0. \quad (527)$$

Since $B(\zeta)$ is rank-one ($\det B = 0$), it can be expressed as an outer product [42, 46]:

$$B(\zeta) = u_\zeta u_\zeta^T, \quad u_\zeta = \begin{bmatrix} \sqrt{1+\zeta} \\ 0 \\ 0 \\ \sqrt{1-\zeta} \end{bmatrix}. \quad (528)$$

As a result, the largest visibility is

$$v_{\max} = \frac{\det A}{u^T \operatorname{adj}(A) u} \quad (529)$$

for a positive matrix A and a rank-one matrix $vu_\zeta u_\zeta^T$. The reason for the formula above can be seen from the following. We seek the maximal v satisfying

$$A - vu_\zeta u_\zeta^T \succeq 0. \quad (530)$$

If $A \succ 0$, it is possible to factorize¹⁷ the LHS above:

$$A - vu u^T = A^{1/2}(\mathbb{I} - vA^{-1/2}uu^T A^{-1/2})A^{1/2}. \quad (531)$$

However, while the matrix $A(t)$ given by Eq. (526) is positive semidefinite, it has two zero eigenvalues. Hence, the determinant is zero, and the decomposition above no longer holds. However, we can instead work with the restriction operator $A(t)|_U : U \rightarrow U$ [42] on the invariant subspace $U = \operatorname{span}\{|+_a+_a\rangle, |-_a-_a\rangle\}$. It is given by

$$A(t)|_U = \begin{bmatrix} 1 + \mu(1-t) & t \\ t & 1 - \mu(1-t) \end{bmatrix}, \quad (532)$$

with $\mu = \sqrt{3}/2$ and $t \in [0, 1]$. This new reduced $A(t)$ -matrix¹⁸ can safely be assumed to be positive definite. Hence, from now on, we will work in this invariant subspace instead of $\{|+_a+_a\rangle, |+_a-_a\rangle, |-_a+_a\rangle, |-_a-_a\rangle\}$.

By then defining

$$w := A^{1/2}u, \quad (533)$$

¹⁷If $A \succ 0$, it is non-singular and so the inverse matrix $A^{-1/2}$ exists.

¹⁸We will stick to the notation of $A(t)$ or A even for the restriction of the operator A to U but whenever A is written, it means $A(t)|_U$.

it is possible to write

$$A - vu u^T \succeq 0 \iff \mathbb{1} - v w w^T \succeq 0. \quad (534)$$

Now, define

$$\hat{w} := \frac{w}{\|w\|}. \quad (535)$$

Then, we have

$$w w^T = \|w\|^2 \hat{w} \hat{w}^T \iff w w^T \hat{w} = \|w\|^2 \hat{w}, \quad (536)$$

where it has been used that

$$\hat{w}^T \hat{w} = \frac{w^T w}{\|w\|^2} = \frac{\|w\|^2}{\|w\|^2} = 1. \quad (537)$$

Moreover, Eq. (536) is equivalently given by

$$(1 - v w w^T) \hat{w} = (1 - v \|w\|^2) \hat{w}. \quad (538)$$

Hence, $1 - v \|w\|^2$ is an eigenvalue of $1 - v w w^T$. For every vector x orthogonal to $\hat{w}$, we have

$$w^T x = 0 \implies w w^T x = 0. \quad (539)$$

Consequently,

$$(1 - v w w^T) x = x, \quad (540)$$

i.e., all eigenvectors orthogonal to $\hat{w} \parallel w$ have eigenvalue one. In total, the matrix $1 - v w w^T$ has the distinct eigenvalues $1 - v \|w\|^2$ and 1 up to some multiplicity. Since the matrix is positive semidefinite, all eigenvalues must be larger than or equal to 0. Therefore

$$\mathbb{1} - v w w^T \succeq 0 \iff 1 - v \|w\|^2 \geq 0 \iff v \leq \frac{1}{\|w\|^2}. \quad (541)$$

Furthermore, using Eq. (533) with the expression for $v_{\max}$ above gives

$$\|w\|^2 = w^T w = u^T A^{-1} u \implies v_{\max} = \frac{1}{u^T A^{-1} u}. \quad (542)$$

For any 2×2 matrix, we have

$$A^{-1} = \frac{\text{adj}(A)}{\det A} \implies u^T A^{-1} u = \frac{u^T \text{adj}(A) u}{\det A}. \quad (543)$$

Finally,

$$v_{\max} = \frac{\det A}{u^T \text{adj}(A) u}. \quad \square \quad (544)$$

To compute $\det A(t)$, let $s := 1 - t$. Then,

$$A(t) = \begin{bmatrix} 1 + \mu s & t \\ t & 1 - \mu s \end{bmatrix}, \quad (545)$$

with

$$\begin{aligned} \det A(t) &= (1 + \mu s)(1 - \mu s) - t^2 = 1 - \mu^2 s^2 - t^2 = 1 - \frac{3}{4}(1 - t)^2 - t^2 \\ &= \frac{1}{4} + \frac{3}{2}t - \frac{7}{4}t^2 = \frac{(1 - t)(1 + 7t)}{4}. \end{aligned} \quad (546)$$

Next, we calculate $u_\zeta^T \text{adj}(A)u_\zeta$. Recalling that A is given by

$$A(t) = \begin{bmatrix} 1 + \mu(1-t) & t \\ t & 1 - \mu(1-t) \end{bmatrix}, \quad (547)$$

then

$$\begin{aligned} u_\zeta^T \text{adj}(A)u_\zeta &= \begin{bmatrix} \sqrt{1+\zeta} & \sqrt{1-\zeta} \end{bmatrix} \begin{bmatrix} 1 + \mu(1-t) & t \\ t & 1 - \mu(1-t) \end{bmatrix} \begin{bmatrix} \sqrt{1+\zeta} \\ \sqrt{1-\zeta} \end{bmatrix} \\ &= [1 - \mu(1-t)](1+\eta) + [1 + \mu(1-t)](1-\eta) - 2t\sqrt{1-\eta^2}. \end{aligned} \quad (548)$$

Expanding the first two terms in the last equality gives

$$(1 + \zeta) + (1 - \zeta) - \mu(1-t)(1 + \zeta) + \mu(1-t)(1 - \zeta) = 2 - 2\mu\zeta(1-t), \quad (549)$$

and therefore

$$u_\zeta^T \text{adj}(A)u_\zeta = 2 \left[1 - \mu\zeta(1-t) - t\sqrt{1-\zeta^2} \right], \quad (550)$$

where $\mu = \frac{\sqrt{3}}{2}$ was used. Now, the visibility can be written entirely in terms of ζ and t , by substituting the expressions for $\det A(t)$ and $u_\zeta^T \text{adj}(A)u_\zeta$:

$$\begin{aligned} v(\zeta, t) &= \frac{\det A(t)}{u_\zeta^T \text{adj}(A)u_\zeta} = \frac{\frac{(1-t)(1+7t)}{4}}{2 \left[1 - \frac{\sqrt{3}}{2}\zeta(1-t) - t\sqrt{1-\zeta^2} \right]} \\ &= \frac{(1-t)(1+7t)}{8 \left[1 - \frac{\sqrt{3}}{2}\zeta(1-t) - t\sqrt{1-\zeta^2} \right]}. \end{aligned} \quad (551)$$

The optimal visibility is

$$v_{\text{opt}}(\zeta) = \max_{0 \leq t \leq 1} v(\zeta, t), \quad (552)$$

hence transferring the problem to an optimization over t .

Q.4 Solution to the t -optimization

Defining¹⁹

$$A := \frac{\sqrt{3}}{2}\zeta, \quad B := \sqrt{1-\zeta^2}. \quad (553)$$

gives

$$v(\zeta, t) = \frac{(1-t)(1+7t)}{8[1 - A(1-t) - Bt]}. \quad (554)$$

Further define

$$\begin{aligned} P(t) &:= (1-t)(1+7t) = 1 + 6t - 7t^2 \implies P'(t) = 6 - 14t, \\ D(t) &:= 1 - A(1-t) - Bt = 1 - A + (A-B)t \implies D'(t) = A - B, \end{aligned} \quad (555)$$

¹⁹These are not to be confused with the matrices A and B in the previous section.

ignoring the constant factor 8 in the denominator. Furthermore, at an interior optimum, we have

$$\frac{d}{dt} \frac{P(t)}{D(t)} = 0 \implies P'(t)D(t) - P(t)D'(t) = 0. \quad (556)$$

Substituting Eq. (555) gives a one-dimensional optimization over t

$$\begin{aligned} (6 - 14t)[1 - A + (A - B)t] - (1 - 6t - 7t^2)(A - B) &= 0 \\ \iff 7(B - A)t^2 + 14(A - 1)t + 6 - 7A + B &= 0. \end{aligned} \quad (557)$$

Consequently, the optimal $t_*(\zeta)$ is one of the roots of this quadratic, or one of the endpoints, $t = 0, 1$. Putting this together, the full analytical formula is

$$v_{\text{opt}}(\zeta) = v(\zeta, t_*(\zeta)), \quad (558)$$

where $t_*(\zeta)$ is chosen from $t \in \{0, 1, t_+(\zeta), t_-(\zeta)\} \cap [0, 1]$ to maximize $v(\zeta, t)$. For $B \neq A$, the two roots are

$$\begin{aligned} t_{\pm}(\zeta) &= \frac{1 - A \pm \sqrt{(1 - A)^2 - \frac{1}{7}(B - A)(6 - 7A + B)}}{B - A} \\ &= \frac{14(1 - A) \pm \sqrt{196(A - 1)^2 - 28(B - A)(6 - 7A + B)}}{14(B - A)}. \end{aligned} \quad (559)$$

The discriminant of the quadratic equation (Eq. (557)) is given by $\Delta(\zeta) = 196(1 - A)^2 - 28(B - A)(6 - 7A + B)$. The physical root is the minus root, given by

$$t_* = \frac{14(1 - A) - \sqrt{\Delta(\zeta)}}{14(B - A)}, \quad 0 \leq t \leq 1, \quad (560)$$

as it has a finite limit near the special point $A = B$. To avoid 0/0-type cancellation when $A \approx B$, the above formula can be rationalized

$$t_* = \frac{14(1 - A) - \sqrt{\Delta}}{14(B - A)} \cdot \frac{14(1 - A) + \sqrt{\Delta}}{14(1 - A) + \sqrt{\Delta}} = \dots = \frac{2(6 - 7A(\zeta) + B(\zeta))}{14(1 - A(\zeta)) + \sqrt{\Delta(\zeta)}}. \quad (561)$$

From this, the visibility can be written as a piecewise function in the following way. Using the definitions in Eq. (553).

$$v_{\text{opt}}(\zeta) = \begin{cases} 1, & \zeta = 0, \\ \frac{(1 - t_*(\zeta))(1 + 7t_*(\zeta))}{8[1 - A(\zeta)(1 - t_*(\zeta)) - B(\zeta)t_*(\zeta)]}, & 0 < \zeta < \zeta_c, \\ \frac{1}{8(1 - A(\zeta))}, & \zeta_c \leq \zeta \leq 1, \end{cases} \quad (562)$$

where ζ_c is the transition point, which is where the interior optimizer $t_*(\zeta)$ hits the boundary $t_* = 0$:

$$\begin{aligned} t_*(\zeta) &= \frac{2(6 - 7A + B)}{14(1 - A) + \sqrt{28(1 - B)(7 + B - 8A)}} = 0 \\ \iff 6 - 7A + B &= 0 \iff \sqrt{1 - \zeta^2} = \frac{7\sqrt{3}}{2}\zeta - 6. \end{aligned} \quad (563)$$

where $A = \frac{\sqrt{3}}{2}\zeta$ and $B = \sqrt{1-\zeta^2}$ was used in the last equality. Squaring both sides, expanding and rearranging, leads to the quadratic equation

$$\frac{151}{4}\zeta^2 - 42\sqrt{3}\zeta + 35 = 0 \iff 151\zeta^2 - 168\sqrt{3}\zeta + 140 = 0. \quad (564)$$

Solving the quadratic above gives

$$\zeta = \frac{168\sqrt{3} \pm 4\sqrt{7}}{302}. \quad (565)$$

Only the positive root satisfies the condition $\delta = \sqrt{1-\zeta^2} \geq 0$, we get

$$\zeta_c = \frac{84\sqrt{3} + 2\sqrt{7}}{151} \approx 0.9986. \quad (566)$$

Hence, $t_* = 0$ for the last interval, $\zeta_c \leq \zeta \leq 1$ and $t_*(\zeta) > 0$ for $0 < \zeta < \zeta_c$. Since the numerator of v_{opt} vanishes at $t_* = 1$, $v(\zeta, 1) = 0$.

R Derivation of $v_{\text{wc}}^{(4)}(\zeta)$

Here, the same method as above is used, hence the parts of the procedure specific to the tetrahedral instrument are written below.

R.1 Preliminaries

The tetrahedral equal-weight POVM family considered here is given by

$$E_a(\zeta) = \frac{1}{4}(\mathbb{1} + \zeta \vec{n}_a^E \cdot \vec{\sigma}), \quad 0 \leq \zeta \leq 1. \quad (567)$$

Hence, the Kraus operators for the Lüders instrument become

$$K_a(\zeta) = \sqrt{E_a(\zeta)} = \frac{1}{2} \begin{bmatrix} \sqrt{1+\zeta} & 0 \\ 0 & \sqrt{1-\zeta} \end{bmatrix} \quad (568)$$

in the eigenbasis of $\vec{n}_a^{E^{\text{eff}}} \cdot \vec{\sigma}$. Then, the restricted Choi operator of the Lüders branch in the invariant subspace defined by the same basis is

$$\eta_a(\zeta) = \frac{1}{8} \begin{bmatrix} 1+\zeta & \sqrt{1-\zeta^2} \\ \sqrt{1-\zeta^2} & 1-\zeta \end{bmatrix}. \quad (569)$$

After following the exact same procedure as for the trine derivation, the PI candidate has the Choi block

$$\eta_a^{(v)}(t) = \frac{1}{8} \begin{bmatrix} 1 + \mu(1-t) & t \\ t & 1 - \mu(1-t) \end{bmatrix}, \quad (570)$$

where μ is different from the value in the section above.

R.2 The value of μ

The effective equal-weight tetrahedral POVM is given by

$$E_a^{\text{eff}} = \frac{1}{4}(\mathbb{1} + \mu \vec{n}_a^{E^{\text{eff}}} \cdot \vec{\sigma}), \quad a = 1, 2, 3, \quad (571)$$

where μ is the Bloch radius.

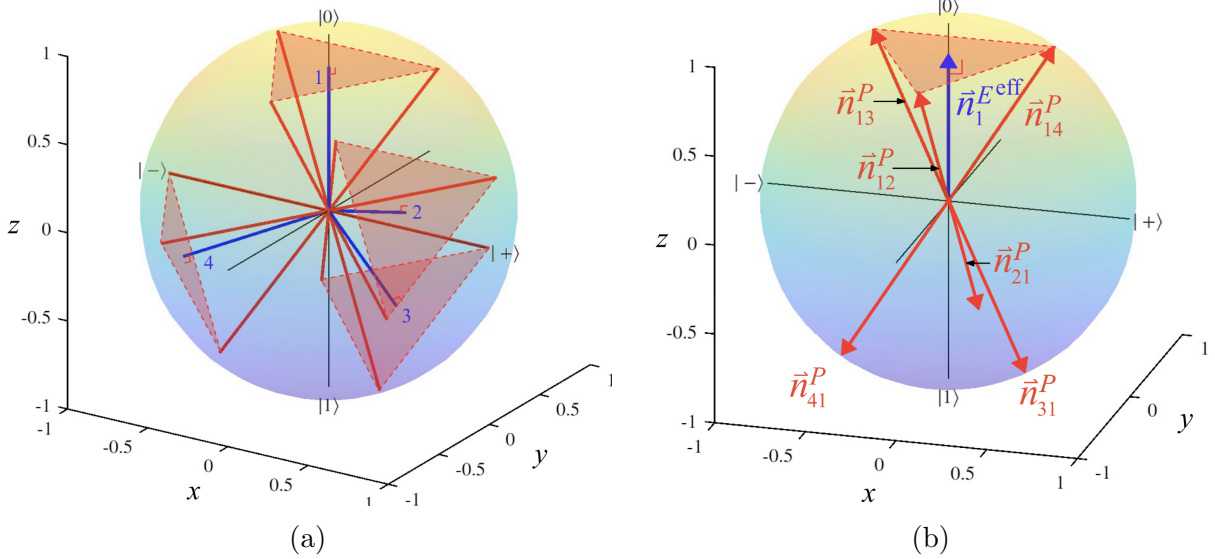


Figure 7: a) Projective decomposition of E_a^{eff} , $a = 1, 2, 3$, with Bloch vectors shown in blue by rank-one normalized binary PVMs with Bloch vectors shown in red. The $-$ sign denotes the antipodal vector of a PVM. Illustrated in the xz -plane of the Bloch sphere. b) Bloch vector of a PVM pair.

The binary rank-one PVM decomposition becomes:

$$E_a^{\text{eff}} = \frac{1}{6} \sum_{b \neq a} P_{ab} = \frac{1}{6} \sum_{b \neq a} \frac{1}{2} (\mathbb{1} + \vec{n}_{ab}^P \cdot \vec{\sigma}), \quad (572)$$

where $a, b \in \{1, 2, 3, 4\}$. The factor $1/6$ comes from mixing six binary PVMs equally to achieve symmetric tetrahedral outcomes as in Fig. 7. This is illustrated in Fig. 7a. For a binary qubit PVM assigned to outcomes a and b , the effects are

$$P_{ab} = \frac{1}{2}(\mathbb{1} + \vec{n}_{ab}^P \cdot \vec{\sigma}), \quad P_{ba} = \frac{1}{2}(\mathbb{1} + \vec{n}_{ba}^P \cdot \vec{\sigma}), \quad (573)$$

where $\vec{n}_{ba}^P = -\vec{n}_{ab}^P$ and the first index in the subscript denotes the outcome of the POVM to which the binary PVMs contribute to. Hence, P_{ab} contributes to E_a^{eff} while P_{ba} contributes to E_b^{eff} . To make this binary PVM best aligned with the two target tetrahedral directions $\vec{n}_a^{E^{\text{eff}}}$ and $\vec{n}_b^{E^{\text{eff}}}$, $\vec{n}_{ab}^P$ needs to point as much as possible toward $\vec{n}_a^{E^{\text{eff}}}$ and $\vec{n}_{ba}^P = -\vec{n}_{ab}^P$ point as much as possible toward $\vec{n}_b^{E^{\text{eff}}}$. This means that

$$\vec{n}_a^{E^{\text{eff}}} \cdot \vec{n}_{ab}^P + \vec{n}_b^{E^{\text{eff}}} \cdot (-\vec{n}_{ab}^P) = (\vec{n}_a^{E^{\text{eff}}} - \vec{n}_b^{E^{\text{eff}}}) \cdot \vec{n}_{ab}^P, \quad (574)$$

needs to be maximized over all unit vectors $\vec{n}_{ab}^P$. Using the scalar product formula, we get

$$(\vec{n}_a^{E^{\text{eff}}} - \vec{n}_b^{E^{\text{eff}}}) \cdot \vec{n}_{ab}^P = \left\| \vec{n}_a^{E^{\text{eff}}} - \vec{n}_b^{E^{\text{eff}}} \right\| \cos(\alpha), \quad (575)$$

where $\|\vec{n}_{ab}^P\|$ has been used. This is maximized when $\cos(\alpha) = 1$, i.e. when $\alpha = 0$. Consequently, $\vec{n}_{ab}^P \parallel \vec{n}_a - \vec{n}_b$. Therefore,

$$\vec{n}_{ab}^P = \frac{\vec{n}_a^{E^{\text{eff}}} - \vec{n}_b^{E^{\text{eff}}}}{\|\vec{n}_a^{E^{\text{eff}}} - \vec{n}_b^{E^{\text{eff}}}\|}, \quad (576)$$

while the other PVM outcome has the opposite direction $\vec{n}_{ba}^P = -\vec{n}_{ab}^P$. This is illustrated in Fig. 7b.

Following Fig. 7a, the POVM vectors are given by

$$\begin{aligned} \vec{n}_1^{E^{\text{eff}}} &= (0, 0, 1), \\ \vec{n}_2^{E^{\text{eff}}} &= \left(\frac{2\sqrt{2}}{3}, 0, -\frac{1}{3} \right), \\ \vec{n}_3^{E^{\text{eff}}} &= \left(-\frac{\sqrt{2}}{3}, \sqrt{\frac{2}{3}}, -\frac{1}{3} \right), \\ \vec{n}_4^{E^{\text{eff}}} &= \left(-\frac{\sqrt{2}}{3}, -\sqrt{\frac{2}{3}}, -\frac{1}{3} \right). \end{aligned} \quad (577)$$

which are valid tetrahedral directions since they satisfy:

$$\sum_{a=1}^4 \vec{n}_a^{E^{\text{eff}}} = 0, \quad \vec{n}_a^{E^{\text{eff}}} \cdot \vec{n}_b^{E^{\text{eff}}} = -\frac{1}{3}, \quad a \neq b. \quad (578)$$

Due to symmetry in the Bloch sphere, it is enough to compute the three red PVM directions contributing to outcome one. First,

$$\vec{n}_{12}^P = \frac{\vec{n}_1^{E^{\text{eff}}} - \vec{n}_2^{E^{\text{eff}}}}{\|\vec{n}_1^{E^{\text{eff}}} - \vec{n}_2^{E^{\text{eff}}}\|}, \quad (579)$$

where

$$\vec{n}_1^{E^{\text{eff}}} - \vec{n}_2^{E^{\text{eff}}} = \left(-\frac{2\sqrt{2}}{3}, 0, \frac{4}{3} \right), \quad \|\vec{n}_1^{E^{\text{eff}}} - \vec{n}_2^{E^{\text{eff}}}\| = \sqrt{\frac{8}{3}}, \quad (580)$$

we get

$$\vec{n}_{12}^P = \left(-\frac{1}{\sqrt{3}}, 0, \sqrt{\frac{2}{3}} \right). \quad (581)$$

Similarly,

$$\vec{n}_{13}^P = \frac{\vec{n}_1^{E^{\text{eff}}} - \vec{n}_3^{E^{\text{eff}}}}{\|\vec{n}_1^{E^{\text{eff}}} - \vec{n}_3^{E^{\text{eff}}}\|} = \left(\frac{1}{2\sqrt{3}}, -\frac{1}{2}, \sqrt{\frac{2}{3}} \right), \quad \vec{n}_{14}^P = \frac{\vec{n}_1^{E^{\text{eff}}} - \vec{n}_4^{E^{\text{eff}}}}{\|\vec{n}_1^{E^{\text{eff}}} - \vec{n}_4^{E^{\text{eff}}}\|} = \left(\frac{1}{2\sqrt{3}}, \frac{1}{2}, \sqrt{\frac{2}{3}} \right). \quad (582)$$

Their sum is

$$\vec{n}_{12}^P + \vec{n}_{13}^P + \vec{n}_{14}^P = \left(0, 0, 3\sqrt{\frac{2}{3}} \right). \quad (583)$$

Since $\vec{n}_1^{E^{\text{eff}}} = (0, 0, 1)$, the Bloch vector equation is given by

$$\vec{n}_{12}^P + \vec{n}_{13}^P + \vec{n}_{14}^P = \sqrt{6} \vec{n}_1^{E^{\text{eff}}}. \quad (584)$$

Hence, Eq. (572) becomes

$$E_a^{\text{eff}} = \frac{1}{6} \frac{1}{2} \left(3\mathbb{1} + \sqrt{6} \vec{n}_a^{E^{\text{eff}}} \right) = \frac{1}{4} \left(\mathbb{1} + \sqrt{\frac{2}{3}} \vec{n}_a^{E^{\text{eff}}} \cdot \vec{\sigma} \right). \quad (585)$$

Hence, the effective blue POVM vector in Fig. 7b has length $\mu = \sqrt{\frac{2}{3}}$, which is optimal.

R.3 Expression of $v(\zeta, t)$

Thus, for fixed t , the largest admissible visibility is

$$v(\zeta, t) = \frac{1 - \mu^2(1-t)^2 - t^2}{2 \left[1 - \mu\zeta(1-t) - t\sqrt{1-\zeta^2} \right]}. \quad (586)$$

Substituting $\mu^2 = 2/3$, this becomes

$$v(\zeta, t) = \frac{(1-t)(1+5t)}{6 \left[1 - \sqrt{\frac{2}{3}} \zeta(1-t) - t\sqrt{1-\zeta^2} \right]}. \quad (587)$$

Optimize over t gives

$$v_{\text{wc}}(\zeta) = \max_{0 \leq t \leq 1} v(\zeta, t). \quad (588)$$

Defining

$$A(\zeta) = \sqrt{\frac{2}{3}} \zeta, \quad B(\zeta) = \sqrt{1-\zeta^2}, \quad (589)$$

leads to

$$v(\zeta, t) = \frac{(1-t)(1+5t)}{6 \left[1 - A(1-t) - Bt \right]}. \quad (590)$$

The stationary condition is

$$5(B-A)t^2 + 10(A-1)t + 4 - 5A + B = 0. \quad (591)$$

The physical root is

$$t_{\star}(\zeta) = \frac{2(4 - 5A(\zeta) + B(\zeta))}{10(1 - A(\zeta)) + \sqrt{20(1 - B(\zeta))(5 + B(\zeta) - 6A(\zeta))}}. \quad (592)$$

The transition point is obtained from $t_{\star} = 0$, i.e.

$$4 - 5A + B = 0, \quad (593)$$

which gives

$$\zeta_c = \frac{20\sqrt{6} + \sqrt{15}}{53} \approx 0.9974. \quad (594)$$

Therefore the final expression is

$$v_{\text{wc}}(\zeta) = \begin{cases} 1, & \zeta = 0, \\ \frac{(1 - t_\star)(1 + 5t_\star)}{6 [1 - A(1 - t_\star) - Bt_\star]}, & 0 < \zeta < \zeta_c, \\ \frac{1}{6(1 - A)}, & \zeta_c \leq \zeta \leq 1. \end{cases} \quad (595)$$

Here

$$A = \sqrt{\frac{2}{3}} \zeta, \quad B = \sqrt{1 - \zeta^2}. \quad (596)$$

The minimum occurs at

$$A = B \quad \Longleftrightarrow \quad \sqrt{\frac{2}{3}} \zeta = \sqrt{1 - \zeta^2}. \quad (597)$$

Thus

$$\zeta_{\min} = \sqrt{\frac{3}{5}}. \quad (598)$$

At this point,

$$t_\star = \frac{2}{5}, \quad (599)$$

and the visibility is

$$v_{\min} = \frac{1 + \sqrt{2/5}}{2} \approx 0.8162. \quad (600)$$

S Computational cost

Since $\eta_a \in \mathcal{L}_+(\mathcal{H}_{A'} \otimes \mathcal{H}_A)$, the positive semidefinite matrix size is $D \times D$, where $D = d'd$. If both input and output dimensions are d , we have $D = d^2$, which is far larger than for POVMs, where the variables are of size $d \times d$.

The expensive part is the number rank vectors satisfying

$$\vec{r} = (r_1, \dots, r_N), \quad r_1 \geq 0, \quad \sum_{a=1}^N r_a = d. \quad (601)$$

The number of such vectors is given by the binomial coefficient

$$R(d, N) = \binom{d + N - 1}{N - 1}. \quad (602)$$

To see this, consider an example. We take $d = 5$ and $N = 3$. This is represented by

$$** \mid *** \mid \quad (603)$$

and corresponds to

$$(r_1, r_2, r_3) = (2, 3, 0). \quad (604)$$

There are d indistinguishable stars and $N - 1$ indistinguishable bars, which divide the stars into N groups. Altogether there are

$$d + N - 1 \quad (605)$$

objects. Therefore the number of combinations for grouping the stars into N groups is given by

$$R(d, N) = \binom{d + N - 1}{N - 1} = \frac{(d + N - 1)!}{d!(N - 1)!}. \quad (606)$$

Therefore, for each outcome a and each rank vector $\vec{r}$, the number of blocks is roughly

$$N \binom{d + N - 1}{N - 1}. \quad (607)$$

Each block is $D \times D$. For $d' = d$, this is $d^2 \times d^2$, hence, each Hermitian block is of the order $D^2 = d^4$ real degrees of freedom. Therefore,

$$\text{SDP size} \propto N \binom{d + N - 1}{N - 1} d^4. \quad (608)$$

From this, we can conclude that for a fixed N , we get a polynomial in d .

$$\binom{d + N - 1}{N - 1} \propto \frac{d^{N-1}}{(N - 1)!}, \quad (609)$$

meaning that the variable size scales like:

$$N \binom{d + N - 1}{N - 1} d^4 \propto \frac{N}{(N - 1)!} d^{N+3}. \quad (610)$$

However, if the number of outcomes scales with the dimension, $N \propto d$ or $N \propto d^2$, an exponential in d is obtained. To see this, consider the following:

$$R(d, d) = \binom{2d - 1}{d} \propto \frac{4^d}{\sqrt{d}}. \quad (611)$$

Then,

$$NR(d, N)d^4 \propto d \frac{4^d}{\sqrt{d}} d^4. \quad (612)$$

To get Eq. (611), we start from:

$$R(d, N) = \binom{d + N - 1}{N - 1}. \quad (613)$$

For basis-type qudit instruments, the number of outcomes is $N = d$. So,

$$R(d, d) = \binom{d + d - 1}{d - 1} = \binom{2d - 1}{d - 1}. \quad (614)$$

By symmetry of binomial coefficients, we also have

$$\begin{bmatrix} 2d-1 \\ d-1 \end{bmatrix} = \begin{bmatrix} 2d-1 \\ d \end{bmatrix} = R(d, d). \quad (615)$$

Comparing

$$\begin{bmatrix} 2d \\ d \end{bmatrix} = \frac{(2d)!}{d!d!} \quad (616)$$

to

$$\begin{bmatrix} 2d-1 \\ d \end{bmatrix} = \frac{(2d-1)!}{d!(d-1)!}, \quad (617)$$

gives

$$\begin{bmatrix} 2d \\ d \end{bmatrix} = \frac{(2d)!}{d!d!} = \frac{2d(2d-1)!}{d \cdot d!(d-1)!} = 2 \frac{(2d-1)!}{d!(d-1)!}. \quad (618)$$

Therefore,

$$\begin{bmatrix} 2d \\ d \end{bmatrix} = 2 \begin{bmatrix} 2d-1 \\ d \end{bmatrix}. \quad (619)$$

Using the standard Stirling approximation $n! \sim \sqrt{2\pi n}(n/e)^n$ gives

$$(2d)! \sim \sqrt{2\pi(2d)} \left(\frac{2d}{e}\right)^{2d} = \sqrt{4\pi d} \left(\frac{2d}{e}\right)^{2d}, \quad d!d! \sim \left[\sqrt{2\pi d} \left(\frac{d}{e}\right)^d \right]^2 = 2\pi d \left(\frac{d}{e}\right)^{2d}, \quad (620)$$

which means that

$$\begin{bmatrix} 2d \\ d \end{bmatrix} \sim \frac{\sqrt{4\pi d} \left(\frac{2d}{e}\right)^{2d}}{2\pi d \left(\frac{d}{e}\right)^{2d}} = \frac{4^d}{\sqrt{\pi d}}. \quad (621)$$

Hence,

$$R(d, d) = \begin{bmatrix} 2d-1 \\ d \end{bmatrix} \propto \frac{4^d}{2\sqrt{\pi d}}. \quad (622)$$

The important point is that the dominant term is 4^d , so the number of rank vectors grows exponentially in d when $N = d$. The factor $\frac{1}{2\sqrt{\pi d}}$ is only a polynomial correction. Furthermore, for the POVM SDP numerical cost size, the same procedure as above applies, but with d^2 instead of d^4 .