Rigidity of superdense coding

Ashwin Nayak [*] Henry Yuen [†]

Abstract

The famous superdense coding protocol of Bennett and Wiesner demonstrates that it is possible to communicate two bits of classical information by sending only one qubit and using a shared EPR pair. Our first result is that an arbitrary protocol for achieving this task (where there are no assumptions on the sender's encoding operations or the dimension of the shared entangled state) is locally equivalent to the canonical Bennett-Wiesner protocol. In other words, the superdense coding task is *rigid*. In particular, we show that the sender and receiver only use additional entanglement (beyond the EPR pair) as a source of classical randomness.

We also investigate several questions about higher-dimensional superdense coding, where the goal is to communicate one of d^2 possible messages by sending a d-dimensional quantum state, for general dimensions d. Unlike the $d = 2$ case (i.e. sending a single qubit), there can be inequivalent superdense coding protocols for higher d. We present concrete constructions of inequivalent protocols, based on constructions of inequivalent orthogonal unitary bases for all $d > 2$. Finally, we analyze the performance of superdense coding protocols where the encoding operators are independently sampled from the Haar measure on the unitary group. Our analysis involves bounding the distinguishability of random maximally entangled states, which may be of independent interest.

Contents

[*]Department of Combinatorics and Optimization, and Institute for Quantum Computing, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada. Email: ashwin.nayak@uwaterloo.ca .

[†]Department of Computer Science, Columbia University, New York, USA. E-mail: henry.yuen@columbia.edu .

1 Introduction

In quantum information theory, rigidity is a phenomenon where optimal performance in an information processing task requires using a protocol satisfying extremely stringent constraints — in some cases, there is essentially a *unique* optimal protocol. The primary examples of rigidity come from nonlocal games (also known as *Bell tests* in the physics literature). In this setting two spatially separated parties Alice and Bob play a game with a third-party called the referee. In order to maximize their chances of winning, before the game starts Alice and Bob choose an entangled state to share as well as local measurements to perform on the state. For example, in the famous CHSH game the optimal winning probability is $\cos^2(\pi/8)$, and a canonical strategy that achieves this uses a (rotated) EPR pair and single-qubit Pauli measurements. The CHSH game is *rigid* in the sense that *any* optimal strategy for the CHSH game is identical to this canonical strategy, up to local changes of basis.

The study of rigidity in quantum information processing arguably started with the work of Mayers and Yao [MY98, MY04], who initiated the concept of *device-independent cryptography*. The idea behind this subject is that a classical user can verify that untrusted quantum hardware is behaving as intended — say, generating random keys or performing a quantum computation – simply by verifying that the hardware is employing a (near)-optimal strategy in a rigid nonlocal game. Since the work of Mayers and Yao, nonlocal game rigidity has been an extremely fruitful concept in quantum cryptography (see, e.g., Refs. [VV19] and [CGJV19]), complexity theory (see Ref. [JNV+20] and the references therein), and quantum information more generally [ŠB20]. This motivates the following question: what other tasks in quantum information also exhibit rigidity phenomena?

To our knowledge, the only other work on rigidity phenomena outside of nonlocal games is that reported in Refs. [TKV+18, FK19] on the rigidity of *quantum random access codes (QRACs)*. The authors study "$2^d \to 1$" QRACs, which encode 2 classical dits $x, y \in [d]$ into a d-dimensional system, such that either x or y may be retrieved by performing a suitable measurement. These works show that $2^d \to 1$ QRACs are rigid, and in fact certify measurements based on mutually unbiased bases (MUBs).

In this paper we investigate the rigidity properties of superdense coding, which plays a fundamental role in quantum Shannon theory (see, e.g., Ref. [Wil13, Chapter 6]). The superdense coding *task* is to communicate one of four possible messages while only transmitting one quantum bit across a channel. The superdense coding *protocol*, first proposed by Bennett and Wiesner [BW92], achieves this task in the following way: Alice and Bob share one qubit each of an EPR pair (i.e., the maximally entangled state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$) in advance, and to transmit a message $i \in \{1, 2, 3, 4\}$,

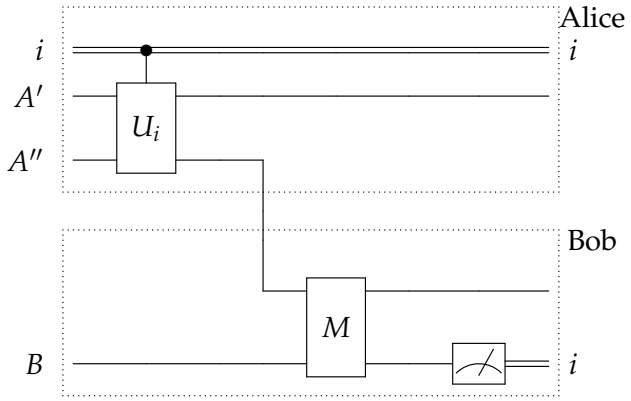

Figure 1: A general superdense coding protocol. This quantum circuit is modified from [Cha20].

Alice applies a one of four Pauli operators to her half of the EPR pair and sends her qubit. Bob then performs a Bell measurement on the qubit received from Alice and his qubit to determine i.

1.1 Rigidity for superdense coding of two classical bits

The first result in our paper is to show that superdense coding is rigid: *any* protocol that accomplishes this task is "locally equivalent" to the Bennett-Wiesner protocol. We model arbitrary protocols for superdense coding in the following manner: Alice and Bob share a density matrix τ on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where we assume without loss of generality that $\mathcal{H}_A$ factors into $\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$ where $\mathcal{H}_{A''}$ is isomorphic to $\mathbb{C}^2$. Given an input $i \in \{1,2,3,4\}$, Alice applies a unitary operator U_i (called an *encoding operator*) to her share of τ (with support in the space $\mathcal{H}_A$), sends the qubit A'' to Bob, and Bob then performs an optimal distinguishing measurement on the Hilbert space $\mathcal{H}_{A''} \otimes \mathcal{H}_B$ to determine what the input i was. See Figure 1 for an illustration of a general superdense coding protocol.

A priori it appears daunting to characterize the structure of an arbitrary superdense coding protocol. For one, the dimension of the spaces $\mathcal{H}_{A'}$ and $\mathcal{H}_B$ are unbounded, and the state τ is uncharacterized. Furthermore, the encoding unitary operators U_i can be extremely complicated, potentially performing complex entangling operations between the space $\mathcal{H}_{A'}$ and $\mathcal{H}_{A''}$ (the qubit to be sent over to Bob). However, the property of being a superdense coding protocol is extremely constraining. Theorem 1.1 gives a precise characterization of how an arbitrary superdense protocol is *locally equivalent* to the canonical Bennett-Wiesner protocol. In the statement of the theorem, "$=_{\tau'}$" denotes equality of two unitary operators with respect to the state τ'; in other words, $C =_\rho D$ means that $C\rho C^* = D\rho D^*$.

Theorem 1.1 (Rigidity for superdense coding)**.** *Let $(\tau, (U_i))$ denote a superdense coding protocol. Then there exist*

1. *Unitary operators V acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$ and $(C_i)_{i\in[4]}$ acting on $\mathcal{H}_{A'}$,*
2. *An isometry W mapping $\mathcal{H}_B$ to a Hilbert space $\mathcal{H}_{B'} \otimes \mathcal{H}_{B''}$ where $\mathcal{H}_{B''}$ is isomorphic to $\mathbb{C}^2$,*
3. *A density matrix ρ on $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$,*
4. *A set of pairwise orthogonal projectors $\{P_r\}$ that sum to the identity on $\mathcal{H}_{A'}$, and*
5. *A collection of 2×2 unitary operators $\{S_r\}$,*

such that, letting $\tau' := (V \otimes W)\tau(V \otimes W)^*$, *we have*

$$\tau' = \rho^{A'B'} \otimes |\mathrm{EPR}\rangle\langle\mathrm{EPR}|^{A''B''}$$

and for $i \in \{1,2,3,4\}$,

$$(C_i^* \otimes \mathbb{1})U_i V^* =_{\tau'} \sum_r P_r \otimes S_r \sigma_i S_r^*$$

where $\sigma_1 := \mathbb{1}$, $\sigma_2 := \mathrm{Z}$, $\sigma_3 := \mathrm{X}$, *and* $\sigma_4 := \mathrm{Y}$ *are the one-qubit Pauli matrices.*

Theorem 1.1 can be interpreted as expressing rigidity for superdense coding in the following way: given an arbitrary protocol $(\tau, (U_i))$ for superdense coding, there exists local isometries V, W where if Alice applies V and Bob applies W to their share of τ, then an EPR pair is extracted with an auxiliary state ρ remaining. By pre-applying V^* to Alice's unitary operators U_i, we discover that $U_i V^*$ has a very regular form: operationally, it can be interpreted as performing some projective measurement $\{P_r\}$ on Alice's part of the auxiliary state ρ to obtain some outcome r in a set $\mathcal{R}$, and then based on r, applying a rotated version of the standard Bennett-Wiesner superdense coding protocol to Alice's part of the EPR pair. Finally, after sending her EPR qubit, Alice then applies some unitary operator C_i on her remaining qubits (which does not affect Bob's measurement in any way). This considerably strengthens and extends the characterization of "tight" superdense coding protocols due to Vollbrecht and Werner [VW00, Lemma 3] (see also Ref. [Wer01]); they studied protocols in which the shared entangled state τ is a state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ (or on $\mathbb{C}^d \otimes \mathbb{C}^d$ in the case of d-dimensional superdense coding protocols; see Section 1.2). This difference would be significant in a cryptographic setting; in the context of quantum key distribution, this is the difference between the *device-independent* and *semi-device-independent* settings.

The proof of Theorem 1.1 is given in Section 3. It proceeds via a number of reductions: first, using an information-theoretic argument, we show every superdense coding protocol $(\tau, (U_i))$ is locally equivalent to one that uses an EPR pair in the state τ. Given this, we then show that each of the encoding operators U_i can be individually block-diagonalized with respect to the EPR pair. Finally, we show that the blocks across the different encoding operators U_i can be "matched up" in a way that they correspond to the Pauli matrices. Each of these steps requires carefully deducing the structure imposed on the state and the encoding operators by the correctness of the protocol.

1.2 Rigidity for higher dimensional superdense coding?

We then consider the generalization of superdense coding to communicating more than 2 classical bits. Specifically, we consider protocols for communicating one of d^2 possible messages by sending a d-dimensional quantum system over the channel — we call these d-dimensional superdense coding protocols. A canonical protocol for d-dimensional superdense coding is as follows: the players share a d-dimensional maximally entangled state $|\phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{e=0}^{d-1} |e\rangle|e\rangle$, and given message $i \in [d^2]$, Alice applies a unitary operator E_i to her share of $|\phi_d\rangle$, and sends it over to Bob. The family of unitary operators $\{E_i\}$ can be any orthogonal unitary basis for the space of $d \times d$ matrices. (The orthogonality property means that $\mathrm{Tr}(E_i^* E_j) = 0$ if and only if $i \neq j$.) An example of such a basis is the set of Heisenberg-Weyl operators. In dimension d, these are a set of $d \times d$ matrices $\{P_{i,j} : 0 \leq i, j < d\}$ defined as follows. Let $\omega_d := \exp\left(\frac{2\pi \mathrm{i}}{d}\right)$ be a primitive dth root of unity. For $i, j \in \{0, 1, 2, \ldots, d-1\}$, let $P_{i,j} = \mathrm{X}_d^i \mathrm{Z}_d^j$ where $\mathrm{X}_d := \sum_{k=0}^{d-1} |k+1 \pmod d\rangle\langle k|$ is the "shift" operator, and $\mathrm{Z}_d := \sum_{k=0}^{d-1} \omega_d^k |k\rangle\langle k|$ is the "clock" operator. Does the rigidity phenomenon also extend to dimensions d larger than 2?

The second result of this paper is that d-dimensional superdense coding for $d \geq 3$ is not rigid in the same sense as Theorem 1.1: there are d-dimensional superdense coding protocols which

are not locally equivalent to each other. This is because in dimensions three and higher there are *inequivalent* orthogonal unitary bases. (In contrast, all orthogonal unitary bases in dimension two are equivalent to the Pauli matrices.) Here, equivalence between two unitary bases $\{E_i\}$ and $\{F_i\}$ means there exist unitary operators U, V such that for all i, we have $F_i = \alpha_i U E_i V$ for some choice of complex phase α_i.

Theorem 1.2 (Existence of inequivalent orthogonal unitary bases for all $d \geq 3$)**.** *For every dimension $d \geq 3$, there are orthogonal unitary bases that are not equivalent to each other.*

The uniqueness of orthogonal unitary bases was first studied by Vollbrecht and Werner [VW00], and the existence of non-equivalent orthogonal unitary bases for all dimensions greater than 2 was observed in follow-up work by Werner [Wer01]. (We elaborate on prior work on the topic in Section 4.2.) Werner described, without proof, how non-equivalent bases may be constructed. We present explicit constructions of such bases in Section 4. The construction for $d \geq 4$ is based on the observation that the shift operator X_d corresponds to a perfect matching in $K_{d,d}$, the complete bipartite graph. Moreover, its powers $\{X_d^i : 0 \leq i < d\}$ correspond to a partition of the edge set of $K_{d,d}$ into d *disjoint* perfect matchings. By replacing this partition with another carefully chosen such partition, we obtain an orthogonal unitary basis that is not equivalent to the clock and shift construction. The proof of non-equivalence involves comparing the spectra of the operators in the two bases, taking into account the complex phase and unitary operators that witness a potential equivalence map. For $d = 3$, we follow a construction described by Werner [Wer01]. We prove non-equivalence to the clock and shift basis by showing that the resulting basis is not a commutative projective group (again accounting for a potential equivalence map).

In a previous version of this paper, we conjectured that rigidity for higher dimensional superdense coding holds *up to* choosing orthogonal unitary bases [NY20, Conjecture 1.3]. That is, every d-dimensional superdense coding protocol is locally equivalent (in the sense of Theorem 1.1) to one where Alice and Bob share an entangled state $\rho^{A'B'} \otimes |\phi_d\rangle\langle\phi_d|^{A''B''}$ for some density matrix $\rho^{A'B'}$, and Alice's encoding operators are of the form

$$U_i = \sum_r P_r \otimes E_{r,i}$$

where $\{P_r\}$ is a set of pairwise orthogonal projectors that sum to the identity on $\mathcal{H}_{A'}$ and for every r, the set $\{E_{r,i}\}_{i \in [d^2]}$ is an orthogonal unitary basis for the space of $d \times d$ complex matrices. This would be a natural extension of the statement of Theorem 1.1 to the case of general $d \geq 2$ where the registers $A'B'$ are treated as a source of "shared randomness" to help Alice and Bob synchronize their choice of orthogonal unitary basis.

We can show that when the shared entangled state between Alice and Bob is a pure state in $\mathbb{C}^d \otimes \mathbb{C}^d$, then this conjecture holds (see Section 4.1): up to local unitary operators, the shared state is necessarily the maximally entangled state $|\phi_d\rangle$ of local dimension d, and the encoding operators $\{U_i\}$ necessarily form an orthogonal unitary basis. However, this conjecture is false for protocols where Alice sends only a *part* of her entangled state. In work subsequent to ours, Farkas, Kaniewski, and Nayak [FKN22] show that there exist infinitely many superdense coding protocols that are not locally equivalent to a protocol of the form described in the conjecture. In particular, in these counterexample protocols Alice may perform a complicated entangling operation between her message and the rest of her state, rather than just treating the ancilla system as a source of shared randomness.

1.3 Superdense coding protocols with error

Finally, we consider *probabilistic* protocols for d-dimensional superdense coding, where Bob's decoding only needs to succeed with high probability. In particular, we say that $(\tau, (U_i))$ is a (d, ε)-superdense coding protocol if Bob is able to decode Alice's message i with probability at least $1 - \varepsilon$, for all i. We focus on the case where Alice and Bob share an entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$ (i.e., have local dimension d). As mentioned previously, in the exact case $\varepsilon = 0$, their shared state is necessarily the maximally entangled state and Alice's encoding unitary operators form an orthogonal unitary basis. We conjecture that even in the probabilistic setting, this characterization of d-dimensional superdense coding protocols is *robust*, in the following sense.

Conjecture 1.3. There exist functions $\delta_1, \delta_2 : [0,1] \to [0,1]$ where $\delta_1(\varepsilon)$ and $\delta_2(\varepsilon)$ monotonically decrease to 0 as $\varepsilon \to 0$, such that the following holds. For all (d, ε)-superdense coding protocols $(\tau, (U_i))$ such that τ is a density matrix on $\mathbb{C}^d \otimes C^d$ and U_i are unitary operators in $\mathsf{U}(\mathbb{C}^d)$, we have

$$\langle \phi_d | \tau | \phi_d \rangle \quad \geq \quad 1 - \delta_1(\varepsilon) \ ,$$

and there exists an orthogonal unitary basis $\{E_i\}_{i \in [d^2]}$ for the space of $d \times d$ complex matrices such that for all $i \in [d^2]$,

$$\|U_i - E_i\|_{\text{nhs}} \quad \leq \quad \delta_2(\varepsilon) \ ,$$

where $\|X\|_{\text{nhs}} := \sqrt{\frac{1}{d} \operatorname{Tr}(XX^\dagger)}$ denotes the normalized Hilbert-Schmidt norm on the space of $d \times d$ matrices.

We note that the choice of the normalized Hilbert-Schmidt norm in the statement of Conjecture 1.3 is somewhat arbitrary; one can also consider other formulations of the conjecture with other norms (such as the spectral norm, etc.).

The last part of our paper analyzes a possible challenge to Conjecture 1.3 proposed by Aram Harrow. Consider the following probabilistic construction for a potential (d, ε)-superdense coding protocol: independently sample d^2 matrices $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{d^2}$ from the Haar measure on $\mathsf{U}(\mathbb{C}^d)$, the group of $d \times d$ complex unitary matrices. Let $\tau := |\phi_d\rangle\langle\phi_d|$ denote the d-dimensional maximally entangled state. How well does the protocol $(\tau, (\boldsymbol{U}_i))$ accomplish superdense coding?

In classical and quantum communication, many tasks can be performed near-optimally via probabilistically constructed protocols. See, e.g., the text by Wilde [Wil13] for examples from Shannon theory. A simple example from communication complexity is the task of *quantum fingerprinting* [BCWdW01], which enables checking whether two n-bit strings x and y are equal by only comparing two $\mathrm{O}(\log n)$-qubit *fingerprints* of the strings. It can be shown that picking random $\mathrm{O}(\log n)$-qubit states for each n-bit string x yields a good quantum fingerprinting protocol.

Let $\boldsymbol{\Pi}_d$ denote the random superdense protocol specified by $(\phi_d, (\boldsymbol{U}_i))$. Note that the error $\boldsymbol{\varepsilon}$ of $\boldsymbol{\Pi}_d$, when averaged over the choice of random unitaries $(\boldsymbol{U}_i)$, is some function of d. We first argue that the conjecture implies that the error of a random superdense coding protocol, when averaged over the choice of $(\boldsymbol{U}_i)$, cannot be too small:

Proposition 1.4. *Suppose Conjecture 1.3 were true. Let $\delta_2(\varepsilon)$ be the function from Conjecture 1.3. Then the random superdense coding protocol $\boldsymbol{\Pi}_d$ specified by $(\phi_d, (\boldsymbol{U}_i))$ must have error $\boldsymbol{\varepsilon}$ satisfying*

$$\mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \delta_2(\boldsymbol{\varepsilon})^2 \quad \geq \quad (2d)^{-2} \ .$$

Put another way, it cannot be that both Conjecture 1.3 is true and also the random superdense protocol has error vanishing so quickly such that $\delta_2(\varepsilon)$ is smaller than $(2d)^{-2}$, on average. Due to

the concentration of measure phenomenon for Haar-random states and unitary operators (as expressed by, e.g., the Lévy-like property in Theorem 5.5), it is plausible *a priori* that the average error ε, and therefore also $\delta_2(\varepsilon)$, scale as $o(d^{-2})$. Thus, the random superdense protocol is potentially a counterexample to Conjecture 1.3.

We show that this probabilistic construction does *not* yield a good superdense coding protocol: with overwhelmingly high probability over the choice of random unitary operators (U_i), the protocol has a nonzero probability of error that is independent of d. Thus, Conjecture 1.3 is not ruled out by the random protocol construction.

Theorem 1.5 (Performance of a random superdense coding protocol)**.** *The random superdense coding protocol $\mathbf{\Pi}_d$ specified by $(\phi_d, (\boldsymbol{U}_i))$ where $\boldsymbol{U}_i \in \mathsf{U}(\mathbb{C}^d)$ are Haar-random unitary operators has error at least $1 - \frac{8}{3\pi} \approx 0.15$ as $d \to \infty$, with high probability over the choice of $(\boldsymbol{U}_i)$.*

We prove Theorem 1.5 by showing that the *distinguishability* of the ensemble of random states $\{(\boldsymbol{U}_i \otimes \mathbb{1})|\phi_d\rangle\}_{i\in[d^2]}$ is bounded away from 1 (with high probability). The generalized Holevo-Curlander bounds [Kho79, Cur79, ON99, Tys09b] relate the distinguishability of an ensemble $\{(p_i, \rho_i)\}$ to the quantity

$$\operatorname{Tr}\Big(\sum_i p_i^2 \rho_i^2\Big)^{1/2} \quad . \tag{1.1}$$

Our analysis of this quantity is largely inspired by work due to Montanaro [Mon07] on the distinguishability of random pure quantum states. However, extending his approach to the ensemble of interest to us—one consisting of random *maximally entangled* states—involves significant technical difficulties. The approach involves relating the distinguishability of an ensemble of states to the spectrum of the ensemble average. In the case of Haar-random pure states, the ensemble average is well approximated by the ensemble average of *unnormalized* complex gaussian vectors with suitably chosen variance. The spectrum of such matrices in the asymptotic limit is given by the Marčenko-Pastur Theorem from random matrix theory. In our case, the entries of the random vectors in the ensemble are *not* independent. We instead bound the generalized Holevo-Curlander quantity in Equation (1.1) by employing a recent generalization of the Marčenko-Pastur Theorem due to Yaskov [Yas16]. (The theorem was proven for ensembles of random real vectors. We verify that its proof also extends to complex random vectors with analogous properties.) In the process, we show that random maximally entangled states satisfy a *pseudo-isotropy* condition that suffices for the theorem to hold.

A subtlety in the use of the Marčenko-Pastur law is that we would like to deduce the convergence of a sequence of means to the mean of the limiting distribution from the convergence of a sequence of distributions. This does not necessarily hold in general. In order to prove such a relation between the two forms of convergence, we show that random maximally entangled states are *sub-gaussian*. This allows us to draw on a generalization of the Bai-Yin Theorem, which bounds the norm of matrices whose columns are given by i.i.d. sub-gaussian vectors. We thus show that the norm of the ensemble average has an exponentially decaying tail, which in turn guarantees the form of convergence we seek.

We believe the techniques used in our analysis are of independent interest. In fact, the subtlety mentioned above was overlooked by Montanaro; the ideas we develop may also be used to close a gap in his analysis (see Section 5.4 for the details).

1.4 Further remarks and open questions

In this paper we have initiated the study of rigidity phenomena in superdense coding protocols. Given its importance in quantum Shannon theory, our study may shed new light on protocols

based on superdense coding. The power of entanglement as a resource in distributed quantum computation, in particular in two-party communication complexity, remains a mystery. The rigidity theorem we establish (Theorem 1.1) gives a complete picture for a simple but fundamental task. The property shown in the analysis of random superdense coding protocols (Theorem 1.5) may also be interpreted as placing a limit on how closely a sequence of random unitary operators approximate an orthogonal basis. This may be of relevance to the theory of error-correction, where unitary error bases play a central role.

We list several open questions that arise from this work:

1. Is Conjecture 1.3 true? Does a robust version of Theorem 1.1 hold?
2. Do d-dimensional superdense coding protocols, in which the shared state between Alice and Bob may have local dimension larger than d, also exhibit some non-trivial form of rigidity?
3. Does rigidity also hold for quantum teleportation, a task that is "dual" to superdense coding? Can this be derived in a black-box way from the rigidity of superdense coding?
4. Are there any connections between the QRAC rigidity results of [TKV+18, FK19] and our results on the rigidity of superdense coding?
5. What other quantum information processing tasks have the rigidity property?

We believe the investigation of these questions will lead to significant new insights into the nature of quantum information, with wide ranging ramifications.

Acknowledgements. We thank the anonymous journal reviewers for their thorough feedback on an earlier version of the paper. We thank Adam Bouland, Chinmay Nirkhe, and Zeph Landau for stimulating discussions at the beginning of this project. H.Y. would like to especially thank Adam Bouland for his integral role in formulating the questions explored in this paper. We would like to thank Pavel Yaskov for the correspondence on his work on the Marčenko–Pastur theorem. We thank Jędrzej Kaniewski and Mate Farkas for their pointers to the literature on rigidity of QRACs. A.N. would like to thank Kanstantsin Pashkovich and Vern Paulsen for helpful discussions on bases of orthogonal unitary operators, and is grateful to the Berkeley CS Theory Group for their hospitality during a visit in Fall 2017, when this work was initiated. A.N.'s research is supported in part by a Discovery Grant from NSERC Canada. H.Y. is supported by an NSERC Discovery Grant, a Google Quantum Research Award, and AFOSR Grant No. FA9550-21-1-0040. This research was partly conducted at the Kavli Institute for Theoretical Physics during the Quantum Physics of Information program in 2017 (and thus this research was supported in part by the National Science Foundation under Grant No. NSF PHY17-48958).

2 Properties of superdense coding

2.1 Quantum information basics

We refer the reader to texts such as [NC11, Wat18, Wil13] for the basics of quantum information, and mention some notational conventions here.

We write $\mathbb{1}$ to denote the identity operator on a Hilbert space. We use superscripts on quantum states, e.g., $|\psi\rangle^{AB}$ or ρ^{AB}, to denote the registers in which they are stored. Similarly, we use subscripts on operators to indicate the registers on which they act, unless this is clear from the

context. Given a bipartite density matrix ρ^{AB}, we write ρ^A to denote its reduction to register A (i.e., the partial trace over B).

Given operators A, B and a density matrix ρ acting on a Hilbert space $\mathcal{H}$, we write $A =_\rho B$ to denote $A\rho A^* = B\rho B^*$. In other words, the operators A and B have the same action on the state ρ.

We write $|\mathrm{EPR}\rangle$ to denote the maximally entangled state $\frac{1}{\sqrt{2}}\Big(|00\rangle + |11\rangle\Big)$ on two qubits. We write $|\phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle|i\rangle$ to denote the d-dimensional maximally entangled state, or simply $|\phi\rangle$ if the dimension d is clear from context. We recall the single qubit Pauli matrices:

$$\mathbb{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathrm{X} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \mathrm{Y} := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \mathrm{Z} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

2.2 Basic properties of superdense coding

Here we give a formal definition of a general superdense coding protocol, and prove some basic properties about them.

Definition 2.1 (Superdense coding protocol)**.** *Let d be a positive integer. Let $\mathcal{H}_A := \mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$ and $\mathcal{H}_B$ be finite dimensional Hilbert spaces where $\mathcal{H}_{A''}$ is isomorphic to $\mathbb{C}^d$. Let τ denote a density matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$ and let $(U_i)_{i\in[d^2]}$ denote a sequence of d^2 unitary operators (called* encoding operators*) acting on $\mathcal{H}_A$. We say that $(\tau, (U_i))$ is a (d, ε)-*superdense coding protocol *if there exists a POVM $\{M_i\}_{i\in[d^2]}$ acting on $\mathcal{H}_{A''} \otimes \mathcal{H}_B$ such that*

$$\mathrm{Tr}(M_i \rho_i) \geq 1 - \varepsilon \qquad \forall\, i \in [d^2] \tag{2.1}$$

*where ρ_i denotes the reduced density matrix of $(U_i \otimes \mathbb{1})\tau(U_i \otimes \mathbb{1})^*$ on registers $A''B$. When $\varepsilon = 0$ we simply call $(\tau, (U_i))$ a d-*dimensional superdense coding protocol*.*

Lemma 2.2 (Orthogonality conditions I)**.** *Let $(\tau, (U_i))$ be a d-dimensional superdense coding protocol. Then letting ρ_i denote the reduced density matrix of $(U_i \otimes \mathbb{1})\tau(U_i \otimes \mathbb{1})^*$ on registers $A''B$, we have that*

$$\mathrm{Tr}(\rho_i \rho_j) = 0 \qquad \forall i \neq j \in [d^2] .$$

Proof. Let $\{M_i\}$ denote a POVM satisfying Equation (2.1) for $\varepsilon = 0$. Then for all $i \in [d^2]$, we have $\rho_i \leq M_i$ according to the positive semidefnite ordering. This is because if we write $\rho_i = \sum_k p_{ik} |\psi_{ik}\rangle\langle\psi_{ik}|$ for some probabilities $\{p_{ik}\}$, then $\mathrm{Tr}(\rho_i M_i) = 1$ implies that for all k, $\langle\psi_{ik}|M_i|\psi_{ik}\rangle = 1$, which implies that $|\psi_{ik}\rangle$ is an eigenvector of M_i with eigenvalue 1. This implies that $M_i = \sum_k |\psi_{ik}\rangle\langle\psi_{ik}| + M_i'$ for some positive semidefinite operator M_i', and this is at least ρ_i in the positive semidefinite ordering.

This means then that for all $j \neq i$,

$$0 \leq \mathrm{Tr}(\rho_i M_j) \leq \mathrm{Tr}(\rho_i (\mathbb{1} - M_i)) = \mathrm{Tr}(\rho_i) - \mathrm{Tr}(\rho_i M_i) = 0$$

where the first inequality is due to the positivity of ρ_i and M_j, the second inequality is due to the fact that $\sum_i M_i = \mathbb{1}$, and the last equality is due to the fact that $\mathrm{Tr}(\rho_i) = \mathrm{Tr}(\rho_i M_i) = 1$. Therefore $\mathrm{Tr}(\rho_i M_j) = 0$. Thus we have

$$0 \leq \mathrm{Tr}(\rho_i \rho_j) \leq \mathrm{Tr}(\rho_i M_j) = 0 \ ,$$

so $\mathrm{Tr}(\rho_i \rho_j) = 0$. □

Lemma 2.3 (Orthogonality conditions II)**.** *Let $(\tau, (U_i))$ be a d-dimensional superdense coding protocol. Then for all $i \neq j \in [d^2]$,*

$$\mathrm{Tr}_{A''}\left(U_i \tau^A U_j^*\right) = 0$$

where τ^A denotes the reduced density matrix of τ on register A and $\mathrm{Tr}_{A''}(\cdot)$ denotes the partial trace over register A''.

Proof. Let $|\tau\rangle$ denote a purification of τ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_R$, where $\mathcal{H}_R$ is the purifying space. Clearly, the protocol where Bob also has access to the purifying space $\mathcal{H}_R$ is also a d-dimensional superdense coding protocol.

Let $\{|1\rangle, \ldots, |\dim A'\rangle\}$ denote an orthonormal basis for $\mathcal{H}_{A'}$. Let $|\rho_{ik}\rangle$ be the (sub-normalized) pure state on registers $A''BR$ given by

$$|\rho_{ik}\rangle := (\langle k|^{A'} \otimes \mathbb{1})(U_i \otimes \mathbb{1})|\tau\rangle \ .$$

Intuitively, $|\rho_{ik}\rangle$ represents the residual state of the protocol on registers $A''B$ when Alice applies unitary operator U_i, and then measures the A' subsystem in the standard basis to obtain outcome $|k\rangle$.

Note that if we let ρ_i denote the state $(U_i \otimes \mathbb{1})|\tau\rangle\langle\tau|(U_i \otimes \mathbb{1})^*$ reduced to the registers $A''BR$, we have the identity

$$\rho_i = \sum_k |\rho_{ik}\rangle\langle\rho_{ik}| \ ,$$

because we can think of ρ_i as the result of first measuring the A' register in the standard basis, and discarding the outcome. Applying Lemma 2.2 to the purified protocol $(|\tau\rangle\langle\tau|, (U_i))$, we have for $i \neq j$,

$$0 = \mathrm{Tr}(\rho_i \rho_j) = \sum_{k,k'} \mathrm{Tr}(|\rho_{ik}\rangle\langle\rho_{ik}| \cdot |\rho_{jk'}\rangle\langle\rho_{jk'}|) = \sum_{k,k'} |\langle\rho_{jk'}|\rho_{ik}\rangle|^2$$

and therefore $|\langle\rho_{jk'}|\rho_{ik}\rangle|^2 = 0$ for all k, k'. This implies that $\langle\rho_{jk'}|\rho_{ik}\rangle = 0$ for all k, k', which can be rewritten as

$$\langle\tau|(U_j \otimes \mathbb{1})^*(|k'\rangle\langle k|^{A'} \otimes \mathbb{1})(U_i \otimes \mathbb{1})|\tau\rangle = \mathrm{Tr}\left((\langle k|^{A'} \otimes \mathbb{1})(U_i \otimes \mathbb{1})|\tau\rangle\langle\tau|(U_j^* \otimes \mathbb{1})(|k'\rangle^{A'} \otimes \mathbb{1})\right) = 0.$$

This is equivalent to the statement that

$$\langle k|^{A'} \ \mathrm{Tr}_{A''}\left(U_i \tau^A U_j^*\right) |k'\rangle^{A'} = 0.$$

Since this holds for all k, k', the matrix $\mathrm{Tr}_{A''}\left(U_i \tau^A U_j^*\right)$ is identically zero, which completes the proof of the Lemma. □

Next we define what it means for superdense protocols to be locally equivalent.

Definition 2.4. *Let $\mathcal{H}_A := \mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$ be a Hilbert space where $\mathcal{H}_{A''}$ is isomorphic to $\mathbb{C}^d$. Let τ, τ' be density matrices on $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $(U_i), (U_i')$ be unitary operators acting on $\mathcal{H}_A$. We say that $(\tau, (U_i))$ and $(\tau', (U_i'))$ are* locally equivalent *if there exists*

1. *A unitary operator V acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$,*
2. *A set of unitary operators $(C_i)_{i \in [d^2]}$ acting on $\mathcal{H}_{A'}$*

such that

1. $\tau' = (V \otimes \mathbb{1})\tau(V \otimes \mathbb{1})^*$, *and*

2. $U_i' = (C_i \otimes \mathbb{1})U_i V^*$.

Lemma 2.5 (Local unitary freedom of superdense coding protocols)**.** *The following properties hold for local equivalence.*

1. *If $(\tau, (U_i))$ and $(\tau', (U_i'))$ are locally equivalent, then $(\tau, (U_i))$ is a (d, ε)-superdense coding protocol if and only if $(\tau', (U_i'))$ is.*

2. *Local equivalence is transitive.*

Proof. For any fixed i, after Alice applies her encoding unitary operator, the reduced density matrix on registers $A''B$ is the same whether the protocol $(\tau, (U_i))$ or $(\tau', (U_i'))$ is used. Thus Bob's ability to distinguish between the different messages is exactly the same. This establishes Item 1.

If $(\tau, (U_i))$ and $(\tau', (U_i'))$ are locally equivalent, then $\tau' = (V \otimes \mathbb{1})\tau(V^* \otimes \mathbb{1})$, and $U_i' = (C_i \otimes \mathbb{1})U_i V^*$. If $(\tau', (U_i'))$ and $(\tau'', (U_i''))$ are locally equivalent, then $\tau'' = (V' \otimes \mathbb{1})\tau'((V')^* \otimes \mathbb{1})$, and $U_i'' = (C_i' \otimes \mathbb{1})U_i'(V')^*$. Thus we have

$$\begin{aligned} \tau'' &= (V'V \otimes \mathbb{1})\tau(V^*(V')^* \otimes \mathbb{1}) \ , \qquad \text{and} \\ U_i'' &= (C_i' C_i \otimes \mathbb{1})U_i V^*(V')^* \ , \end{aligned}$$

which implies that $(\tau, (U_i))$ is locally equivalent to $(\tau'', (U_i''))$. This establishes Item 2. □

2.3 Nice form protocols

In this section, we define *nice form* protocols and then show that every superdense coding protocol is locally equivalent to one that has a nice form.

Definition 2.6. *A d-dimensional superdense coding protocol $(\tau, (U_i))$ has a* nice form *if*

1. $U_1 = \mathbb{1}$,

2. *There exists an isometry $W : \mathcal{H}_B \to \mathcal{H}_{B'} \otimes \mathcal{H}_{B''}$ where $\mathcal{H}_{B''}$ is isomorphic to $\mathbb{C}^d$ such that*

$$(\mathbb{1} \otimes W)\tau(\mathbb{1} \otimes W)^* = \rho^{A'B'} \otimes |\phi_d\rangle\langle\phi_d|^{A''B''}$$

   for some density matrix ρ on $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$.

3. *For all $i \in [d^2]$, we have that*

$$U_i \operatorname{Tr}_B(\tau) U_i^* = U_i \left(\rho^{A'} \otimes \frac{\mathbb{1}}{d}\right) U_i^* = \rho^{A'} \otimes \frac{\mathbb{1}}{d}$$

   where $\rho^{A'}$ denotes the reduced density matrix of τ on $\mathcal{H}_{A'}$.

4. *Let the spectral decomposition of $\rho^{A'}$ be $\sum_k \lambda_k \Pi_k$ where $\lambda_k > 0$ for all k, with λ_k distinct. Then for all k and $i \neq j$, we have*

$$\operatorname{Tr}_{A''}((\Pi_k \otimes \mathbb{1})\, U_i U_j^* \,(\Pi_k \otimes \mathbb{1})) = 0.$$

Item 2 says that up to a local unitary operation, the two parties share a maximally entangled state (in addition to other entanglement), and Item 3 turns out to be a consequence of this. Item 4 is equivalent to saying that the encoding of distinct messages $i \neq j$ are orthogonal to each other. The proof of Lemma 2.7 below may give the reader further intuition into these properties.

In the proof of Lemma 2.7, we make use of an information-theoretic argument that involves quantities such as von Neumann entropy $\mathrm{H}(A)$, conditional entropy $\mathrm{H}(A|B)$, and mutual information $\mathrm{I}(A : B)$. For a comprehensive reference on these quantities and their basic properties, we recommend Wilde's textbook [Wil13]. It is an interesting question whether Lemma 2.7 can be proved *without* making use of these information-theoretic quantities.

Lemma 2.7. *All superdense coding protocols $(\tau, (U_i))$ are locally equivalent to a superdense coding protocol $(\tau', (U_i'))$ that has a nice form.*

Proof. We define a unitary operator V acting on $\mathcal{H}_{A'}$ and unitary operators $(C_i : i \in [d^2])$ acting on $\mathcal{H}_{A'}$ such that, letting $\tau' := (V \otimes \mathbb{1})\tau(V \otimes \mathbb{1})^*$ and $U_i' := (C_i \otimes \mathbb{1})U_i V^*$, the pair $(\tau', (U_i'))$ is a superdense coding protocol and has a nice form.

Let $V := U_1$ and let $C_1 := \mathbb{1}$. This already yields Item 1 of Definition 2.6.

Let $|\tau\rangle^{ABR}$ be a purification of τ where $\mathcal{H}_R$ is a reference system of dimension $\dim(\mathcal{H}_A \otimes \mathcal{H}_B)$. Consider the cq-state

$$\xi := \frac{1}{d^2} \sum_i |i\rangle\langle i|^X \otimes (U_i \otimes \mathbb{1})|\tau\rangle\langle\tau|^{ABR}(U_i^* \otimes \mathbb{1}) \ ,$$

where the Hilbert space of register X is $\mathcal{H}_X$. By Lemma 2.5, the protocol $(V\tau V^*, (U_i V^*))$ is a superdense coding protocol. Therefore the information contained in registers $A''B$ about X in the state ξ is

$$\mathrm{I}(X : A''B)_\xi = 2\log_2 d \ .$$

Intuitively, this is because Bob can perfectly recover the value of $i \in [d^2]$, i.e., $2\log_2 d$ bits of information, from the registers $A''B$ of ξ. On the other hand, we have that

$$\mathrm{I}(X : B)_\xi = 0$$

because without the qubit register A'', Bob has no information about X (the state of the register B is the same for all i). Therefore we get

$$\mathrm{I}(X : A''|B)_\xi = \mathrm{I}(X : A''B)_\xi - \mathrm{I}(X : B)_\xi = 2\log_2 d \ .$$

Using the entropy characterization of conditional mutual information, we get

$$2\log_2 d = \mathrm{I}(X : A''|B)_\xi = \mathrm{H}(A''|B)_\xi - \mathrm{H}(A''|XB)_\xi \ .$$

Since $\mathrm{H}(A''|B)_\xi \leq \log_2 d$ and $\mathrm{H}(A''|XB)_\xi \geq -\log_2 d$ (because the dimension of register A'' is d), we get that $\mathrm{H}(A''|B)_\xi = \log_2 d$ and $\mathrm{H}(A''|XB)_\xi = -\log_2 d$.

Since X is a classical register, we can write $\mathrm{H}(A''|XB)_\xi$ as

$$-\log_2 d = \mathrm{H}(A''|XB)_\xi = \mathop{\mathbb{E}}_i \mathrm{H}(A''|B, X = i)$$

where $\mathrm{H}(A''|B, X = i)$ is defined as $\mathrm{H}(A''|B)_{\xi_i}$ with $|\xi_i\rangle := (U_i \otimes \mathbb{1})|\tau\rangle$. Since $\mathrm{H}(A''|B, X = i) \geq -\log_2 d$, we have $\mathrm{H}(A''|B)_{\xi_i} = -\log_2 d$ for all i, and in particular for $i = 1$.

Then $\mathrm{H}(A''|B)_{\xi_i} = -\mathrm{H}(A''|RA')_{\xi_i}$ (because ξ_i is pure), so $\mathrm{H}(A''|RA')_{\xi_i} = \log_2 d$. On one hand, we have that $\mathrm{I}(A'' : RA')_{\xi_i} = \mathrm{H}(A'')_{\xi_i} - \mathrm{H}(A''|RA')_{\xi_i}$, and on the other hand, mutual information

is always nonnegative. Thus $\mathrm{H}(A'')_{\xi_i} = \log_2 d$, and the reduced density matrix of ξ_i on the A'' register is maximally mixed. Furthermore we have $\mathrm{I}(A'' : RA')_{\xi_i} = 0$, so ξ_i has no correlations between registers A'' and RA':

$$\mathrm{Tr}_B(\xi_i) = \rho_i^{RA'} \otimes \frac{\mathbb{1}}{d} \tag{2.2}$$

where ρ_i is some density matrix on the RA' registers.

Fix $i = 1$, and let $\rho^{RA'}$ denote $\rho_1^{RA'}$. Let $\mathcal{H}_{B'}$ be a Hilbert space with dimension $\dim(\mathcal{H}_R \otimes \mathcal{H}_{A'})$ and let $\mathcal{H}_{B''}$ be isomorphic to $\mathbb{C}^d$. Let $|\rho\rangle^{RA'B'}$ denote a purification of $\rho^{RA'}$. Notice that $|\rho\rangle^{RA'B'} \otimes |\phi_d\rangle^{A''B''}$ is a purification of the state in Equation (2.2). Using Uhlmann's Theorem [Uhl76] (also known as the Schrödinger-HJW Theorem [Sch35, HJW93]), there exists an isometry W on $\mathcal{H}_B$ such that

$$(\mathbb{1} \otimes W)|\xi_1\rangle^{RA'A''B} = |\rho\rangle^{RA'B'} \otimes |\phi_d\rangle^{A''B''} .$$

Since $|\xi_1\rangle = (V \otimes \mathbb{1})|\tau\rangle$, we have that

$$\rho^{A'B'} \otimes |\phi_d\rangle\langle\phi_d|^{A''B''} = \mathrm{Tr}_R((\mathbb{1} \otimes W)|\xi_1\rangle\langle\xi_1|(\mathbb{1} \otimes W)^*) = (V \otimes W)\tau(V \otimes W)^* .$$

Since $\tau' = (V \otimes \mathbb{1})\tau(V \otimes \mathbb{1})^*$ we obtain Item 2 of Definition 2.6 for the protocol $(\tau', (U_i V^*))$.

In what follows we let ζ denote $\rho^{A'}$ (which also equals $\tau'^{A'}$). We now establish Item 3 of Definition 2.6. Let $\sum_k \lambda_k \Pi_k$ be the spectral decomposition of ζ where the $\{\lambda_k\}$ are distinct and nonzero, and Π_k is the orthogonal projector onto the eigenspace of ζ corresponding to eigenvalue λ_k. Since by Equation (2.2) we have

$$U_i V^* \left(\zeta \otimes \frac{\mathbb{1}}{d}\right) (U_i V^*)^* = \mathrm{Tr}_{BR}(\xi_i) = \zeta_i \otimes \frac{\mathbb{1}}{d} \tag{2.3}$$

for some density matrix ζ_i, the states ζ_i and ζ have the same eigenvalues with the same multiplicities. That is, there are an orthogonal set of projectors $\{\Pi_k^{(i)}\}_k$ such that

$$\zeta_i = \sum_k \lambda_k \Pi_k^{(i)} ,$$

where $\dim(\Pi_k^{(i)}) = \dim(\Pi_k)$ for all $i \in [d^2]$. It follows that for all i,

$$U_i V^* \left(\Pi_k \otimes \frac{\mathbb{1}}{d}\right) (U_i V^*)^* = \Pi_k^{(i)} \otimes \frac{\mathbb{1}}{d} .$$

For $i \in [d^2]$ let C_i be a unitary operator on $\mathcal{H}_{A'}$ such that $C_i \Pi_k^{(i)} C_i^* = \Pi_k$ for all k. Since $\Pi_k^{(1)} = \Pi_k$, our choice of $C_1 = \mathbb{1}$ suffices. Let $U_i' = (C_i \otimes \mathbb{1}) U_i V^*$. By Lemma 2.5, we have $(\tau', (U_i'))$ is a superdense coding protocol. Furthermore, Equation (2.3) implies that

$$U_i' \Big(\sum_k \lambda_k \Pi_k \otimes \frac{\mathbb{1}}{d} \Big) (U_i')^* = \sum_k \lambda_k \Pi_k \otimes \frac{\mathbb{1}}{d} ,$$

which implies Item 3 of Definition 2.6.

Lemma 2.5 implies that $(\tau', (U_i'))$ is also a superdense coding protocol, so by Lemma 2.3 we have that for all $i \neq j$,

$$\mathrm{Tr}_{A''} \left(U_i' \left(\zeta \otimes \frac{\mathbb{1}}{d} \right) (U_j')^* \right) = 0 .$$

Since Π_k commutes with the (U'_i) for all k, we have

$$\begin{aligned} 0 &= \mathrm{Tr}_{A''}\left(U'_i\left(\rho\otimes\frac{\mathbb{1}}{d}\right)(U'_j)^*\right) \\ &= \sum_k \lambda_k \,\mathrm{Tr}_{A''}\left(U'_i\left(\Pi_k\otimes\frac{\mathbb{1}}{d}\right)(U'_j)^*\right) \\ &= \frac{1}{d}\sum_k \lambda_k \,\mathrm{Tr}_{A''}\left((\Pi_k\otimes\mathbb{1})\,U'_i(U'_j)^*\right) \quad . \end{aligned}$$

Since the λ_k's are positive, we have

$$\mathrm{Tr}_{A''}\left((\Pi_k\otimes\mathbb{1})U'_i(U'_j)^*(\Pi_k\otimes\mathbb{1})\right) = 0$$

for all k, and $i \neq j$. This implies Item 4 of Definition 2.6.

□

3 Rigidity for two-dimensional superdense coding

In this section we prove Theorem 1.1, that is, rigidity for 2-dimensional superdense coding protocols (coding 2 bits into one qubit, with no error). For the remainder of this section we drop the qualification "2-dimensional" for brevity.

The proof involves a number of steps. First, we invoke Lemma 2.7, which states that every superdense coding protocol is locally equivalent to one that has a nice form. Then, we argue that up to local equivalence, in every nice form superdense coding protocol, the encoding operators (U_i) can be block-diagonalized. That is, we can write $U_i = \sum_\ell Q_{i\ell}\otimes R_{i\ell}$ where the $\{Q_{i\ell}\}$ are a set of orthogonal projectors acting on $\mathcal{H}_{A'}$ summing to the identity, and the $\{R_{i\ell}\}$ are a set of Hermitian unitary operators acting on $\mathcal{H}_{A''}$. Next, we argue that (again up to local equivalence) across the different i's, the projectors $\{Q_{i\ell}\}$ can be "matched up", and the corresponding operators $R_{i\ell}$ are all pairwise orthogonal. This implies that in fact $\{R_{1\ell}, R_{2\ell}, R_{3\ell}, R_{4\ell}\}$ are unitarily equivalent to the standard Pauli matrices $\{\mathbb{1}, X, Y, Z\}$. Using the property that local equivalence is a transitive relation, this concludes the argument.

Lemma 2.7 is proven in Section 2.3. We now proceed to prove the remaining steps in detail.

3.1 Block-diagonalizing nice form protocols

In this section, we analyze the structure of the encoding operators in a nice form superdense coding protocol in dimension two. We show that they apply a 2×2 unitary operator on register A'', controlled by the state in register A'.

Theorem 3.1. *Let $(\tau, (U_i))$ be a nice-form protocol. Then there exists a locally-equivalent protocol $(\tau', (U'_i))$ that has a nice form and for $i \in \{2,3,4\}$ we have*

$$U'_i =_{\tau'} \sum_\ell Q_{i\ell}\otimes R_{i\ell}\ .$$

for some orthogonal projectors $\{Q_{i\ell}\}_\ell$ on $\mathcal{H}_{A'}$ that sum to $\mathbb{1}$, and 2×2 traceless, Hermitian unitary matrices $\{R_{i\ell}\}_\ell$ on $\mathcal{H}_{A''}$.

Proof. Fix an $i \in \{2,3,4\}$. Since $(\tau, (U_i))$ has a nice form (see Definition 2.6), this means that $\tau^A = \rho^{A'} \otimes \frac{\mathbb{1}}{2}$ for some density matrix ρ, and $\mathrm{Tr}_{A''}((\Pi_k \otimes \mathbb{1})U_i U_1^* (\Pi_k \otimes \mathbb{1})) = \mathrm{Tr}_{A''}((\Pi_k \otimes \mathbb{1})U_i(\Pi_k \otimes \mathbb{1})) = 0$, where Π_k is a non-zero eigenspace of ρ.

Fix a k. By property 3 of a nice-form protocol, U_i commutes with $\Pi_k \otimes \mathbb{1}$, and we can write

$$U_i \quad = \quad (\Pi_k \otimes \mathbb{1})U_i(\Pi_k \otimes \mathbb{1}) + (\mathbb{1} - \Pi_k \otimes \mathbb{1})U_i(\mathbb{1} - \Pi_k \otimes \mathbb{1}) \ .$$

Let $\hat{U}_{ik} = (\Pi_k \otimes \mathbb{1})U_i(\Pi_k \otimes \mathbb{1})$, and note that $\hat{U}_{ik}$ is unitary on the image of $\Pi_k \otimes \mathbb{1}$, i.e.:

$$\hat{U}_{ik}\hat{U}_{ik}^* = \hat{U}_{ik}^*\hat{U}_{ik} = \Pi_k \otimes \mathbb{1} \ .$$

For notational convenience we drop the subscripts i and k until the very end. Let $\hat{U} := \hat{U}_{ik}$ and let $\Pi := \Pi_k$.

The condition that $\mathrm{Tr}_{A''}(\hat{U}) = 0$ implies that we can write $\hat{U}$ as

$$\hat{U} = \left(\begin{array}{c|c} F & G \\ \hline H & -F \end{array}\right) \ ,$$

where F, G, H are block matrices that act on the image of Π and the block partitions are with respect to the tensor factor $\mathcal{H}_{A''}$.

Let

$$F = D_F T_F \ , \qquad G = D_G T_G \ , \qquad H = D_H T_H$$

give the polar decompositions of F, G, H respectively where D_F, D_G, D_H are positive semidefinite and T_F, T_G, T_H are unitary on the image of Π.

Then

$$\hat{U} = \left(\begin{array}{c|c} D_F T_F & D_G T_G \\ \hline D_H T_H & -D_F T_F \end{array}\right) .$$

The relation $\hat{U}\hat{U}^* = \Pi \otimes \mathbb{1}$ implies that $D_F^2 = \Pi - D_G^2 = \Pi - D_H^2$. For notational brevity we write $D := D_F$ and $\widetilde{D} := D_G = D_H = \sqrt{\Pi - D^2}$. Note that D and $\widetilde{D}$ have support in the image of Π and are simultaneously diagonalizable. Write $K := T_F^* D T_F$ and $\widetilde{K} := T_F^* \widetilde{D} T_F$. Note that K and $\widetilde{K}$ are positive semidefinite, and also simultaneously diagonalizable. Write $W_G := T_F^* T_G$ and $W_H^* := T_F^* T_H$. Continuing our simplification, we see that

$$(T_F^* \otimes \mathbb{1})\hat{U} = \left(\begin{array}{c|c} K & \widetilde{K} W_G \\ \hline \widetilde{K} W_H^* & -K \end{array}\right) .$$

Now our goal is to find a unitary operator E acting on $\mathcal{H}_{A'}$ such that $(E T_F^* \otimes \mathbb{1})\hat{U}$ is Hermitian. This is equivalent to the conditions that $EK = KE^*$ and $E\widetilde{K}W_G = (E\widetilde{K}W_H^*)^*$. We construct such an operator E using some relations between the operators $K, \widetilde{K}, W_G, W_H$ that we derive below.

We use the unitarity relation $\hat{U}^*\hat{U} = \Pi \otimes \mathbb{1}$ to obtain the equations

$$K^2 + W_G^* \widetilde{K}^2 W_G = \Pi \ , \qquad \text{and} \tag{3.1}$$

$$K^2 + W_H \widetilde{K}^2 W_H^* = \Pi \ . \tag{3.2}$$

These equations, along with the definitions of K and $\widetilde{K}$, imply that the unitary operators W_G, W_H are block-diagonal with respect to the eigenspaces of K and $\widetilde{K}$ (and therefore commute with K

and $\widetilde{K}$). Another relation we get from unitarity is $K\widetilde{K}W_G = W_H\widetilde{K}K$, which via commutativity of $W_H, \widetilde{K}, K$ implies

$$K\widetilde{K}W_G = K\widetilde{K}W_H \,. \tag{3.3}$$

Let Π_+ be the orthogonal projector onto $\mathrm{supp}(K)$. Since W_G commutes with $K, \widetilde{K}$ (and so does W_H), we have that Π_+ commutes with $\widetilde{K}, W_G, W_H$. By Equation (3.3), we have $\Pi_+\widetilde{K}W_G = \Pi_+\widetilde{K}W_H = W_H\widetilde{K}\,\Pi_+$. Since $K^2 + \widetilde{K}^2 = \Pi$, we also have that $(\Pi - \Pi_+)\widetilde{K} = (\Pi - \Pi_+)$. Note that Equation (3.3), together with the fact that W_H, W_G are block-diagonal with respect to the eigenspaces of the product $K\widetilde{K}$, implies that W_G and W_H must be equal on the support of $K\widetilde{K}$ (equivalently, the support of Π_+).

We now construct the desired unitary E. Let Π_0 denote $\Pi - \Pi_+$ (i.e., the projection onto the kernel of K within the image of Π). Left-multiplying both sides of Equation (3.3) with K^+ (the pseudoinverse of K on the image of Π_+) and right-multiplying both sides by Π_+ we get

$$\Pi_+\widetilde{K}W_G\Pi_+ = \Pi_+\widetilde{K}W_H\Pi_+ \,. \tag{3.4}$$

Recall that $\Pi_0\widetilde{K} = \Pi_0$; combined with the fact that W_G and W_H are both block-diagonal with respect to the eigenspaces of K and $\widetilde{K}$, we have

$$\widetilde{K}W_G = \Pi_+\widetilde{K}W_G\Pi_+ + \Pi_0 W_G \Pi_0 \qquad \text{and} \qquad \widetilde{K}W_H^* = \Pi_+\widetilde{K}W_H^*\Pi_+ + \Pi_0 W_H^* \Pi_0 \,. \tag{3.5}$$

Furthermore, $V_G := \Pi_0 W_G \Pi_0$ and $V_H := \Pi_0 W_H^* \Pi_0$ are unitary on Π_0. Let $M := V_H^* V_G$, and consider its spectral decomposition $M = \sum_j e^{\mathrm{i}\theta_j}|v_j\rangle\langle v_j|$ where $\{|v_j\rangle\}$ is an orthonormal basis for the support of Π_0. Let $M^{1/2} := \sum_j e^{\mathrm{i}\theta_j/2}|v_j\rangle\langle v_j|$ denote the principal square root of M. Define $E_0 := M^{1/2}V_G^*$. Observe that E_0 is supported only on Π_0 and satisfies

$$E_0 V_G = M^{1/2} = (M^{1/2}M^*)^* = (E_0 V_H)^*. \tag{3.6}$$

Consider the operator $E := \Pi_+ + E_0$ that is unitary on the support of Π, and acts non-trivially only on the support of Π_0. Combining Equations (3.4), (3.5) and (3.6) we get

$$\begin{aligned} E\widetilde{K}W_G &= \Pi_+\widetilde{K}W_G\Pi_+ + E_0 V_G \\ &= \Pi_+\widetilde{K}W_H\Pi_+ + (E_0 V_H)^* \\ &= (\Pi_+ W_H^*\widetilde{K}\Pi_+ + E_0 V_H)^* \\ &= (\Pi_+\widetilde{K}W_H^*\Pi_+ + E_0 V_H)^* \\ &= (E\widetilde{K}W_H^*)^* \,. \end{aligned}$$

where the second line follows from Equation (3.4) and Equation (3.6), the fourth line follows because $\widetilde{K}$ commutes with W_H^*, and the last line follows because of Equation (3.5). We also have that E commutes with $\widetilde{K}$: this is because Π_+ commutes with $\widetilde{K}$ and also E_0 acts nontrivially only on the eigenspace of $\widetilde{K}$ with eigenvalue 1.

Let $L := E\widetilde{K}W_G$. Putting everything together, we have

$$(ET_F^* \otimes \mathbb{1})\hat{U} = \left(\begin{array}{c|c} EK & E\widetilde{K}W_G \\ \hline E\widetilde{K}W_H^* & -EK \end{array}\right) = \left(\begin{array}{c|c} K & L \\ \hline L^* & -K \end{array}\right) . \tag{3.7}$$

where we used the fact that $EK = K$.

Let $K = \sum_r \alpha_r P_r$ and $\widetilde{K} = \sum_r \sqrt{1-\alpha_r^2}\, P_r$ be spectral decompositions of K and $\widetilde{K}$ where the reals α_r's are nonnegative and distinct, and the operators P_r are orthogonal projectors summing to Π. The operator $\widetilde{K}$ has such a spectral decomposition because $K^2 + \widetilde{K}^2 = \Pi$. Next, since the unitary operators E and W_G commute with $\widetilde{K}$, they are block-diagonal with respect to the projectors $\{P_r\}$. Thus

$$P_r L P_r = P_r E \widetilde{K} W_G P_r = P_r \sqrt{\widetilde{K}} E W_G \sqrt{\widetilde{K}} P_r = \sqrt{1-\alpha_r^2}\; P_r E W_G P_r \,.$$

The operator $P_r E W_G P_r$ is unitary on P_r and we can express it as $\sum_s \beta_{rs} Q_{rs}$ where the β_{rs}'s are complex numbers on the unit circle and $\{Q_{rs}\}_s$ are orthogonal projectors that sum to P_r. Thus we can write $K = \sum_{r,s} \alpha_r\, Q_{rs}$ and $L = \sum_{r,s} \sqrt{1-\alpha_r^2}\, \beta_{rs}\, Q_{rs}$, and $(E T_F^* \otimes \mathbb{1})\hat{U}$ can be written as

$$(E T_F^* \otimes \mathbb{1})\hat{U} = \sum_{r,s} Q_{rs} \otimes R_{rs} \tag{3.8}$$

where R_{rs} is the 2×2 matrix

$$\begin{pmatrix} \alpha_r & \sqrt{1-\alpha_r^2} \cdot \beta_{rs} \\ \sqrt{1-\alpha_r^2} \cdot \beta_{rs}^* & -\alpha_r \end{pmatrix} \,.$$

Notice that R_{rs} has determinant -1 and is traceless, therefore its eigenvalues are $\{+1, -1\}$.

Re-introducing the indices $i \in \{2,3,4\}$ and k, we have deduced that for every block of U_i corresponding to the eigenspace Π_k, there exists a map S_{ik} that is unitary on the image of Π_k such that

$$(S_{ik} \otimes \mathbb{1})\hat{U}_{ik} = \sum_{r,s} Q_{ikrs} \otimes R_{ikrs}$$

where the $(R_{ikrs})_{r,s}$ are 2×2 Hermitian unitary operators with trace 0. Define the unitary operator S_i on $\mathcal{H}_{A'}$ as $S_i := (\mathbb{1} - \sum_k \Pi_k) + \sum_k S_{ik}$. If we sum over k, we get

$$(S_i \otimes \mathbb{1})\hat{U}_i = \sum_{\ell} Q_{i\ell} \otimes R_{i\ell} \;\;,$$

where $\hat{U}_i := \sum_k \hat{U}_{ik}$ and we have re-indexed the sum over k, r, s to be a sum over indices ℓ. Let $U_i' := (S_i \otimes \mathbb{1}) U_i$ for all $i \in \{2,3,4\}$. Then, letting $P := \sum_k \Pi_k$ denote the projector onto the support of ρ, we have

$$U_i' \tau (U_i')^* = (S_i U_i P)\, \tau\, (S_i U_i P)^* = (S_i \hat{U}_i P)\, \tau\, (S_i \hat{U}_i P)^* = S_i \hat{U}_i\, \tau\, \hat{U}_i^* S_i^*$$

where have suppressed the tensoring with identity that extends all the operators to the same space, and used the property that $(P \otimes \mathbb{1})\tau(P \otimes \mathbb{1}) = \tau$, $\hat{U}_i P = U_i P$, and $\hat{U}_i P = \hat{U}_i$. Thus, the unitary operators U_i' satisfy the conclusions of the theorem statement. Let $\tau' := \tau$, so that $(\tau', (U_i'))$ is a superdense coding protocol by Lemma 2.5.

Furthermore, since $(\tau, (U_i))$ has a nice form, it can be verified that $(\tau', (U_i'))$ also has a nice form. First, since $S_1 = \mathbb{1}$, we have that $U_1' = \mathbb{1}$ (and hence Item 1 of Definition 2.6 is satisfied). Item 2 of Definition 2.6 is satisfied since $\tau' = \tau$. Third, since U_i commutes with $\rho \otimes \frac{\mathbb{1}}{2}$ and S_i is block-diagonal with respect to the eigenspaces of ρ, it follows that U_i' also commutes with $\rho \otimes \frac{\mathbb{1}}{2}$ (so Item 3 of Definition 2.6 is satisfied). Finally, we have

$$\begin{aligned}
\mathrm{Tr}_{A''}((\Pi_k \otimes \mathbb{1}) U_i' (U_j')^* (\Pi_k \otimes \mathbb{1})) &= \mathrm{Tr}_{A''}((S_{ik} \otimes \mathbb{1})\hat{U}_{ik} \hat{U}_{jk}^* (S_{jk}^* \otimes \mathbb{1})) \\
&= S_{ik} \left[\mathrm{Tr}_{A''}(\hat{U}_{ik} \hat{U}_{jk}^*)\right] S_{jk}^* \\
&= S_{ik} \left[\mathrm{Tr}_{A''}((\Pi_k \otimes \mathbb{1}) U_i U_j^* (\Pi_k \otimes \mathbb{1}))\right] S_{jk}^* \\
&= 0.
\end{aligned}$$

Thus, Item 4 of Definition 2.6 is satisfied. This completes the proof of the Theorem. □

3.2 Matching the blocks of the encoding operators

In the previous section we saw how, up to local equivalence of protocols, we can express the encoding operator U_i in a two-dimensional superdense coding protocol as a block-diagonal matrix with 2×2 Hermitian unitary operators on the diagonal. In this section we relate the decompositions to each other. Ultimately, the conclusion is that the blocks "line up", so that the operators in the same diagonal block of the four encoding operators are the four single qubit Pauli operators.

Theorem 3.2. *Let $(\tau, (U_i))$ be a superdense coding protocol that has a nice form where for $i \in \{2,3,4\}$ we have*

$$U_i =_\tau \sum_k Q_{ik} \otimes R_{ik} \ ,$$

for some orthogonal projectors $\{Q_{ik}\}_k$ on $\mathcal{H}_{A'}$ that sum to $\mathbb{1}$, and 2×2 traceless, Hermitian unitary matrices $\{R_{ik}\}_k$ on $\mathcal{H}_{A''}$. Then $(\tau, (U_i))$ is locally equivalent to a superdense coding protocol $(\tau', (U_i'))$ that has a nice form and satisfies the following: there exist orthogonal projectors $\{K_r\}$ on $\mathcal{H}_{A'}$ that sum to the identity and 2×2 traceless, Hermitian unitary operators $\{R_{ir}\}$ such that for all $i \in \{2,3,4\}$,

$$U_i' =_{\tau'} \sum_r K_r \otimes R_{ir} \ .$$

Furthermore, for all $r, i \neq j$ we have

$$\mathrm{Tr}(R_{ir} R_{jr}) = 0 \ .$$

Proof. The first step is to "coarse-grain" the projectors $\{Q_{ik}\}$ so that the associated operators R_{ik} are all inequivalent in the following sense. For each i, we say that k and k' are *i-equivalent* if $R_{ik} = \pm R_{ik'}$. For every i, this forms an equivalence relation on the k's. Let $p_i(k)$ denote the least k' such that k' and k are i-equivalent. Define $s_{ik} \in \{\pm 1\}$ to be such that $R_{ik} = s_{ik} R_{ip_i(k)}$.

For every $i \in \{2,3,4\}$, for every k, define the unitary operator $S_i = \sum_k s_{ik} Q_{ik}$ which acts on $\mathcal{H}_{A'}$ (and set $S_1 = \mathbb{1}$). Then if we define $U_i' = (S_i \otimes \mathbb{1}) U_i$ and $\tau' = \tau$, by Lemma 2.5 we get that the pair $(\tau', (U_i'))$ is a superdense coding protocol, and furthermore the operators (U_i') admit a block-diagonalization where for all i, the associated 2×2 unitary operators R_{ik} are all inequivalent. It is also straightforward to check that $(\tau', (U_i'))$ has a nice form.

Next, from Lemma 2.7 we have that for $i \neq j$

$$0 = \mathrm{Tr}_{A''}(U_i' (U_j')^*) = \sum_{k,\ell} Q_{ik} Q_{j\ell} \cdot \mathrm{Tr}(R_{ik} R_{j\ell}). \tag{3.9}$$

By left-multiplying the above expression by Q_{ik} for some k and right-multiplying by $Q_{j\ell}$ for some ℓ yields $Q_{ik} Q_{j\ell} \cdot \mathrm{Tr}(R_{ik} R_{j\ell}) = 0$. Therefore, if $Q_{ik} Q_{j\ell} \neq 0$, it follows that $\mathrm{Tr}(R_{ik} R_{j\ell}) = 0$.

Given three sets of projectors $\{Q_{2k}\}$, $\{Q_{3\ell}\}$, and $\{Q_{4m}\}$ we can define the following tripartite graph G, which we call the *overlap graph*. Associate a vertex with every projector Q_{ik} for $i \in \{2,3,4\}$. Include an edge between Q_{ik} and $Q_{j\ell}$ if and only if $Q_{ik} Q_{j\ell} \neq 0$. A *triangle* $T = (k, \ell, m)$ in the graph G corresponds to a triple of projectors $Q_{2k}, Q_{3\ell}, Q_{4m}$ such that the pairwise products are all nonzero. Given encoding operators as in Equation (3.9), we use triangles to match their blocks.

Lemma 3.3 (Reduction Lemma)**.** *Let $\{Q_{2k}\}$, $\{Q_{3\ell}\}$, and $\{Q_{4m}\}$ be sets of orthogonal projectors with the following properties:*

1. $\sum_k Q_{2k} = \sum_\ell Q_{3\ell} = \sum_m Q_{4m}$
2. *For $i \neq j$, for all k, ℓ, $Q_{ik} Q_{j\ell} \neq 0$ implies that $\mathrm{Tr}(R_{ik} R_{j\ell}) = 0$.*

3. *For all $i \in \{2,3,4\}$, the $\{R_{ik}\}_k$ are inequivalent.*

Then there exists a triangle $T = (k, \ell, m)$ and a unit vector $|v\rangle \in \mathcal{H}_{A'}$ such that

$$Q_{2k}|v\rangle = Q_{3\ell}|v\rangle = Q_{4m}|v\rangle = |v\rangle.$$

We shall assume for now that the Reduction Lemma holds. We show how this gives us an iterative decomposition procedure to construct the orthogonal projectors $\{K_r\}$ satisfying the conclusions of the Theorem.

The sets $Q_2^{(0)} := \{Q_{2k}\}$, $Q_3^{(0)} := \{Q_{3\ell}\}$, and $Q_4^{(0)} := \{Q_{4m}\}$ satisfy the required conditions of the Reduction Lemma with $\sum_k Q_{2k} = \sum_\ell Q_{3\ell} = \sum_m Q_{4m} = \mathbb{1}$. Thus there exists a triangle $T_0 = (k_0, \ell_0, m_0)$ and a vector $|v_0\rangle$ that is a common eigenvector of $Q_{2k_0}, Q_{3\ell_0}, Q_{4m_0}$. Thus we can write

$$Q_{2k_0} = |v_0\rangle\langle v_0| + Q'_{2k_0} \qquad Q_{3\ell_0} = |v_0\rangle\langle v_0| + Q'_{3\ell_0} \qquad Q_{4m_0} = |v_0\rangle\langle v_0| + Q'_{4m_0}$$

where Q'_{2k_0}, $Q'_{3\ell_0}$, and Q'_{4m_0} are orthogonal projectors with rank one smaller.

Define the sets $Q_2^{(1)}, Q_3^{(1)}, Q_4^{(1)}$ to be the sets $Q_2^{(0)}, Q_3^{(0)}, Q_4^{(0)}$ with the projectors $Q_{2k_0}, Q_{3\ell_0}, Q_{4m_0}$ replaced by $Q'_{2k_0}, Q'_{3\ell_0}, Q'_{4m_0}$.

Observe that $Q_2^{(1)}, Q_3^{(1)}, Q_4^{(1)}$ satisfies the required conditions of the Reduction Lemma, with

$$\sum_{F \in Q_i^{(1)}} F \quad = \quad \mathbb{1} - |v_0\rangle\langle v_0| \ ,$$

for all $i \in \{2,3,4\}$. Applying the Reduction Lemma again, we find another triangle T_1 and a common eigenvector $|v_1\rangle$ of the triangle. We continue this process of reducing the rank of at least one operator each in the sets $Q_2^{(r)}, Q_3^{(r)}, Q_4^{(r)}$ and finding common eigenvectors $|v_r\rangle$ until we have fully expressed

$$U'_i =_\tau \sum_r K_r \otimes R_{ir} \ ,$$

where $K_r := |v_r\rangle\langle v_r|$, and for every r, the pairwise inner products satisfy $\mathrm{Tr}(R_{2r}R_{3r}) = \mathrm{Tr}(R_{2r}R_{4r}) = \mathrm{Tr}(R_{3r}R_{4r}) = 0$. This concludes the proof of the Theorem. □

Before proving the Reduction Lemma we establish the following lemma, which claims that up to conjugation by the same unitary operator, the only collection of 2×2 traceless, Hermitian, unitary, mutually orthogonal matrices are the single-qubit Pauli matrices.

Lemma 3.4. *Let R_2, R_3, R_4 be 2×2 unitary matrices that are traceless, Hermitian, and satisfy $\mathrm{Tr}(R_iR_j) = 0$ for all $i \neq j$. Then there exists a 2×2 unitary operator S such that*

$$R_2 = SZS^* \qquad R_3 = SXS^* \qquad R_4 = SYS^*.$$

Proof. We find a sequence of unitary operators S_1, S_2, S_3 such that $S := S_1^* S_2^* S_3^*$ satisfies the conclusions of the lemma. Because R_2 is unitary, Hermitian and traceless, we can unitarily diagonalize it as $R_2 = |a\rangle\langle a| - |b\rangle\langle b|$. Define S_1 as the unitary operator with

$$S_1|a\rangle = |0\rangle \qquad S_1|b\rangle = |1\rangle.$$

Let $R'_i := S_1 R_i S_1^*$ for $i \in \{2,3,4\}$. These are all traceless, Hermitian, pairwise orthogonal unitary matrices, and furthermore $R'_2 = \mathrm{Z}$. Suppose

$$R'_3 = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \ .$$

Since $\mathrm{Tr}(R_2' R_3') = 0$ and R_3' is traceless, we have that $r = u = 0$, and since R_3' is Hermitian and unitary, we have $s = t^* = \mathrm{e}^{\mathrm{i}\theta}$ for some $\theta \in [0, 2\pi)$. Let

$$S_2 := \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\theta/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta/2} \end{pmatrix} ,$$

and $R_i'' := S_2 R_i' S_2^*$ for $i \in \{2, 3, 4\}$. Again, the operators R_i'' remain traceless, Hermitian unitary, and pairwise orthogonal, and furthermore $R_2'' = \mathrm{Z}$ and $R_3'' = \mathrm{X}$. Suppose

$$R_4'' = \begin{pmatrix} w & x \\ y & z \end{pmatrix} .$$

From $\mathrm{Tr}(R_2'' R_4'') = 0$, Hermiticity, unitarity, and tracelessness of R_4'' we have again that $w = z = 0$ and $x = y^* = \mathrm{e}^{\mathrm{i}\phi}$ for some $\phi \in [0, 2\pi)$. From $\mathrm{Tr}(R_3'' R_4'') = 0$ we have that $x = -y$, which means that $x = \pm\mathrm{i}$. If $x = -\mathrm{i}$, then set $S_3 := \mathbb{1}$. Otherwise, set $S_3 := -\mathrm{i}\,\mathrm{Z}$. Let $R_i''' := S_3 R_i'' S_3^*$ for $i \in \{2, 3, 4\}$. We have that $R_2''' = \mathrm{Z}$, $R_3''' = \mathrm{X}$, $R_4''' = \mathrm{Y}$. Thus, letting $S = S_1^* S_2^* S_3^*$, we obtain the desired conclusion of the lemma. □

We now turn to proving the Reduction Lemma.

Proof of Lemma 3.3. Define $\Pi = \sum_k Q_{2k}$.

No two triangles in G share an edge. Suppose we have two triangles corresponding to projectors $(Q_{2k}, Q_{3\ell}, Q_{4m})$ and $(Q_{2k}, Q_{3\ell}, Q_{4m'})$ for some k, l, m, m'. We then have the equations

$$\mathrm{Tr}(R_{2k} R_{3\ell}) = \mathrm{Tr}(R_{2k} R_{4m}) = \mathrm{Tr}(R_{3\ell} R_{4m}) = \mathrm{Tr}(R_{2k} R_{4m'}) = \mathrm{Tr}(R_{3\ell} R_{4m'}) = 0.$$

By Lemma 3.4, this implies that there exists a unitary operator S such that

$$R_{2k} = S\mathrm{Z}S^* \qquad R_{3\ell} = S\mathrm{X}S^* \qquad R_{4m} = S\mathrm{Y}S^* .$$

Therefore we obtain that

$$\mathrm{Tr}(\mathrm{Z} S^* R_{4m'} S) = \mathrm{Tr}(\mathrm{X} S^* R_{4m'} S) = 0 .$$

If

$$S^* R_{4m'} S = \begin{pmatrix} a & b \\ b^* & d \end{pmatrix} ,$$

the above equation implies that $a = d = 0$ and $b = -b^*$, or equivalently that $b = \pm\mathrm{i}$. Thus $S^* R_{4m'} S = \pm\mathrm{Y} = \pm S^* R_{4m} S$, or in other words $R_{4m} = \pm R_{4m'}$, contradicting the assumption that R_{4m} and $R_{4m'}$ are inequivalent.

Every vertex is in a triangle. Consider Q_{2k} for some k. There exists an index ℓ such that $Q_{2k} Q_{3\ell} \neq 0$, because the operators $\{Q_{3\ell}\}$ form a resolution of Π. Since $\{Q_{4m}\}$ also forms an orthogonal resolution of Π, we have that

$$0 \neq Q_{2k} Q_{3\ell} = Q_{2k} \left(\sum_m Q_{4m} \right) Q_{3\ell} .$$

This implies that there exists an index m such that $Q_{2k} Q_{4m} \neq 0$ and $Q_{4m} Q_{3\ell} \neq 0$.

Finding a common eigenvector of a triangle. Fix a triangle $T = (k, \ell, m)$. For notational simplicity we shall write $C := Q_{2k}$, $D := Q_{3\ell}$, and $E := Q_{4m}$. First, observe that if $C(\Pi - D)E \neq 0$, then there exists some index ℓ' such that $CQ_{3\ell'}E \neq 0$. This implies that (k, ℓ', m) forms a triangle in G. But this cannot happen as this triangle would share the edge (k, m) with T. Therefore $C(\Pi - D)E = 0$, i.e., $CE = CDE$.

By symmetry, we also get that $CD = CED$ and $ED = ECD$. Thus we have

$$0 \neq CE = CDE = CEDE = CECDE = CDECDE.$$

Since CDE is a product of three projectors, its spectral norm is at most 1. Let $|v\rangle$ be a unit vector realizing the spectral norm of CDE, i.e. such that $\|CDE|v\rangle\| = \|CDE\| > 0$. Then

$$\|CDE\| = \|CDE|v\rangle\| = \|CDECDE|v\rangle\| \leq \|CDECDE\| \leq \|CDE\|^2.$$

The inequality $\|CDE\| \leq \|CDE\|^2$ implies that $\|CDE\| = 1$ (since it is not zero and is at most one). Therefore $|v\rangle$ is a vector such that $C|v\rangle = D|v\rangle = E|v\rangle = |v\rangle$. □

We now put everything in this section together to prove Theorem 1.1, which we restate here for convenience.

Theorem 1.1 (Rigidity for superdense coding)**.** *Let $(\tau, (U_i))$ denote a superdense coding protocol. Then there exist*

1. *Unitary operators V acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_{A''}$ and $(C_i)_{i\in[4]}$ acting on $\mathcal{H}_{A'}$,*
2. *An isometry W mapping $\mathcal{H}_B$ to a Hilbert space $\mathcal{H}_{B'} \otimes \mathcal{H}_{B''}$ where $\mathcal{H}_{B''}$ is isomorphic to $\mathbb{C}^2$,*
3. *A density matrix ρ on $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$,*
4. *A set of pairwise orthogonal projectors $\{P_r\}$ that sum to the identity on $\mathcal{H}_{A'}$, and*
5. *A collection of 2×2 unitary operators $\{S_r\}$,*

such that, letting $\tau' := (V \otimes W)\tau(V \otimes W)^*$, we have

$$\tau' = \rho^{A'B'} \otimes |\mathrm{EPR}\rangle\langle\mathrm{EPR}|^{A''B''}$$

and for $i \in \{1, 2, 3, 4\}$,

$$(C_i^* \otimes \mathbb{1})U_i V^* =_{\tau'} \sum_r P_r \otimes S_r \sigma_i S_r^*$$

where $\sigma_1 := \mathbb{1}$, $\sigma_2 := \mathrm{Z}$, $\sigma_3 := \mathrm{X}$, and $\sigma_4 := \mathrm{Y}$ are the one-qubit Pauli matrices.

Proof. Putting together Lemma 2.7, Theorem 3.1, and Theorem 3.2, we get that all superdense coding protocols $(\tau, (U_i))$ are locally equivalent to one that has a nice form and satisfies the conclusions of Theorem 3.2. Finally, we apply Lemma 3.4 to the conclusions of Theorem 3.2 to obtain the conclusions of Theorem 1.1. □

4 Superdense coding and orthogonal unitary bases

In this section, we prove that there are multiple non-equivalent superdense coding protocols for transmitting d^2 messages for $d \geq 3$, even when no ancilla is used in the encoding process, and there is no error in decoding. This implies that rigidity of superdense coding protocols for $d \geq 3$ may only hold in a relaxed form: as we see in this section, rigidity may hold only up to the choice of an orthogonal unitary basis for the space of linear operators $\mathsf{L}(\mathbb{C}^d)$.

4.1 The connection with unitary bases

We draw a connection between superdense coding and bases for the vector space of $d \times d$ complex matrices. Although this connection may be inferred from Lemma 2.7, we give a simple and direct derivation here.

For any integer $d > 1$, consider a protocol for superdense coding of d^2 classical strings using a shared entangled state with local dimension d, and a single d-dimensional message. Assume that the protocol does not use any ancilla in the encoding process, and that there is no decoding error. Such a protocol necessarily has a simple form, as we describe below.

First, we argue that the initial shared state is maximally entangled. Bob's state after the message has support in a d^2-dimensional space. Since there are d^2 strings, and these are decoded without error, the corresponding states are orthogonal and pure. So the mixed state of the entire encoded state, corresponding to a uniformly random string, is completely mixed. However, the marginal of this state on the register initially held by Bob is be the same as the marginal for any fixed string. Thus Bob's share of the initial state is also the d-dimensional completely mixed state. This implies that the initial shared state is maximally entangled.

Any maximally entangled state with local dimension d is of the form $(U \otimes V)|\phi_d\rangle$, where U, V are unitary operators in $\mathsf{U}(\mathbb{C}^d)$, and $|\phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle|k\rangle$. Therefore, without loss of generality, we may assume that Alice and Bob initially share the state $|\phi_d\rangle$. When the dimension d is clear from the context, we omit it from the subscript.

Second, since the encoding of any message is pure, Alice's local operations satisfy the following properties. On input $i \in [d^2]$, Alice applies a unitary operator $U_i \in \mathsf{U}(\mathbb{C}^d)$ to her share of the state $|\phi\rangle$, and sends the share to the Bob. Since Bob can decode the input i with probability 1, the states $(U_i \otimes \mathbb{1})|\phi\rangle$ are all orthogonal, i.e., for all distinct $i, j \in [d^2]$, we have

$$\langle\phi|(U_i^* U_j \otimes \mathbb{1})|\phi\rangle \quad = \quad 0 \ .$$

This condition is equivalent to the property that the operators U_i are mutually orthogonal with respect to the Hilbert-Schmidt inner product:

$$\mathrm{Tr}(U_i^* U_j) \quad = \quad 0 \ , \qquad \text{for all } i, j \in [d^2],\ i \neq j \ . \tag{4.1}$$

Thus, the operators form an orthogonal unitary basis for the space of linear operators on $\mathbb{C}^d$.

It is straightforward to verify that any such basis for $\mathsf{L}(\mathbb{C}^d)$ leads to an errorless superdense coding protocol for d^2 classical messages. Thus, the study of rigidity of superdense coding protocols as above is equivalent to the study of orthogonal unitary bases.

A well-known example of an orthogonal unitary basis in dimension d is generated by the "clock" and "shift" operators. The elements of this basis are also known as the generalized Pauli operators or the Heisenberg-Weyl operators. Let $\omega_d := \exp\left(\frac{2\pi \mathrm{i}}{d}\right)$ be a primitive dth root of unity. For $i, j \in \{0, 1, \ldots, d-1\}$, the (i, j)th operator P_{ij} in the basis is defined as $P_{ij} := \mathsf{X}_d^i \mathsf{Z}_d^j$, where $\mathsf{X}_d := \sum_{k=0}^{d-1} |k + 1 \pmod d\rangle\langle k|$ is the shift (or Pauli X) operator, and $\mathsf{Z}_d := \sum_{k=0}^{d-1} \omega_d^k |k\rangle\langle k|$ is the clock (or Pauli Z) operator.

4.2 Uniqueness of orthogonal unitary bases

Given an orthogonal unitary basis B for $\mathsf{L}(\mathbb{C}^d)$, we may derive other such bases by conjugating elements of B by a pair of unitary operators, and mutliplying each basis element by a potentially different complex number of unit modulus. Since this is a rather straightforward method to derive new bases, we consider the new basis to be equivalent to B.

Definition 4.1. *Let* $B_1 := \{U_i : i \in [d^2]\}$ *be an orthogonal unitary basis for* $\mathsf{L}(\mathbb{C}^d)$. *We say that an orthogonal unitary basis* B_2 is equivalent to B_1 *if there exist unit complex numbers* $\alpha_i \in \mathsf{U}(\mathbb{C})$ *and a pair of unitary operators* $V, W \in \mathsf{U}(\mathbb{C}^d)$ *such that*

$$B_2 \quad = \quad \{\alpha_i V U_i W : \ i \in [d^2]\} \ . \tag{4.2}$$

We may verify that this defines an equivalence relation.

Another way to construct an orthogonal unitary basis is by taking tensor products of bases in lower dimensions. Suppose d is composite, with $d = d_1 d_2$ and $1 < d_1, d_2 < d$, and $\{U_i : i \in [d_1^2]\}$ and $\{V_j : j \in [d_2^2]\}$ are orthogonal unitary bases for $\mathsf{L}(\mathbb{C}^{d_1})$ and $\mathsf{L}(\mathbb{C}^{d_2})$, respectively. Then

$$\{U_i \otimes V_j : i \in [d_1^2];\ j \in [d_2^2]\}$$

is an orthogonal unitary basis for $\mathsf{L}(\mathbb{C}^d)$. This hints at the possibility that are bases that are not equivalent to each other under operations as in Eq. (4.2). The following proposition confirms this for dimensions which are powers of two.

Proposition 4.2. *Suppose* $d = 2^k$ *for an integer* $k > 1$. *Let* B_1 *be the basis for* $\mathsf{L}(\mathbb{C}^d)$ *obtained by taking tensor products of k two-dimensional Pauli X and Z operators, i.e.,*

$$B_1 \quad := \quad \left\{ \bigotimes_{i=1}^{k} P_i : P_i \in \{\mathbb{1}, \mathrm{X}_2, \mathrm{Z}_2, \mathrm{X}_2\mathrm{Z}_2\} \right\} \ .$$

The basis B_1 *is* not *equivalent to* B_2, *the d-dimensional clock and shift basis.*

Proof. The intuition behind the statement is that tensor products of the two-dimensional Pauli operators in B_1 all have at most two distinct eigenvalues (either 1, or ± 1, or $\pm \mathrm{i}$), whereas some of the operators in the clock and shift basis have d distinct complex eigenvalues. Due to the freedom available in generating equivalent bases, we need additional arguments to formalize this intuition.

Suppose that B_1 and B_2 *are* equivalent and consider unitary operators $V, W \in \mathsf{U}(\mathbb{C}^d)$ which show their equivalence. Consider the operator $P \in B_1$ that is mapped to the identity in B_2 under the equivalence. Let α be a complex number of unit modulus such that $\alpha V P W = \mathbb{1}$. Then $V = \alpha^* W^* P^*$.

Suppose the operator $Q \in B_1$ is mapped to the clock operator $\mathrm{Z}_d \in B_2$, and that $\mathrm{Z}_d = \beta V Q W$ for some complex number β. Then $\mathrm{Z}_d = \beta\alpha^* W^* P^* Q W$. The operator on the right hand side has at most two eigenvalues (either $\beta\alpha^*$, or $\pm\beta\alpha^*$, or $\pm\mathrm{i}\beta\alpha^*$) as P^*Q has eigenvalues 1, or ± 1, or $\pm\mathrm{i}$. However, the clock operator Z_d has d distinct eigenvalues, the dth complex roots of unity. Since $d \geq 4$, we get a contradiction, and we conclude that B_1 and B_2 are not equivalent. □

It is then natural to wonder if there is a unique orthogonal unitary basis in *prime* dimensions, up to the equivalence defined above. In Section 4.4 we show that even this does not hold, by giving an explicit construction of a basis in any dimension $d \geq 5$ that is not equivalent to the clock and shift basis.

After our discovery of non-equivalent bases, we learned that the question of uniqueness has been studied before by Vollbrecht and Werner [VW00]. They prove the uniqueness of the basis consisting of the Pauli operators in dimension two, and state that the problem of characterizing orthogonal unitary bases in dimensions larger than two is open. They also give a construction of "shift-and-multiply" bases from a collection of d complex Hadamard matrices of dimension $d \times d$ and a $d \times d$ Latin square. This construction and non-equivalent bases are discussed in more detail

by Werner in subsequent work [Wer01], although the notion of equivalence there does not include multiplication by phases (complex numbers of unit modulus). Werner states without proof that the existence of non-equivalent bases in dimension at least five follows from the existence of non-equivalent Hadamard matrices or non-equivalent Latin squares, even when the dimension is prime. In dimension three, Werner describes how we may construct non-equivalent bases, but does not explicitly present them. We present a concrete instance of this construction in Proposition 4.9. Altogether, we have the following result.

Theorem 4.3. *For every dimension $d \geq 3$, there are orthogonal unitary bases that are not equivalent to the clock and shift basis.*

The theorem implies that for any $d \geq 3$, there are non-equivalent superdense coding protocols for transmitting d^2 messages, even when no ancilla is used in the encoding process, and there is no error in decoding.

Orthogonal unitary bases have also been studied in the context of quantum error-correction under the name "unitary error bases" (see, e.g., Ref. [MV16] and the references therein). In addition to the shift-and-multiply construction, several other methods such as the "Hadamard method" and the "algebraic method" have been proposed for their construction. The "quantum shift-and-multiply" method due to Musto and Vicary [MV16] simultaneously generalizes the shift-and-multiply and Hadamard methods. Musto and Vicary give examples of orthogonal unitary bases resulting from this method that are not equivalent to those derived from any of the other methods mentioned above. However, they give explicit examples only in dimension 4. As far as we can tell, earlier explicit constructions, for example those due to Klappenecker and Rötteler [KR03], were also for a few small dimensions.

4.3 Some useful properties

Here we present two properties that are used in an explicit construction leading to Theorem 4.3. The following property of the eigenvalues of the clock and shift operators helps in proving non-equivalence to another basis. Recall that $\omega_d := \exp\left(\frac{2\pi \mathrm{i}}{d}\right)$ is a primitive dth root of unity.

Lemma 4.4. *Let $d > 1$ be an integer, and let $a, b \in \{0, 1, \ldots, d-1\}$. The eigenvalues of the operator $\mathrm{X}_d^a \mathrm{Z}_d^b$ are all of the form*

$$\omega_d^l \cdot \exp\left(\frac{ab(d-1)\pi \mathrm{i}}{d}\right) ,$$

for some $l \in \{0, 1, \ldots, d-1\}$.

Proof. Since $\mathrm{X}_d \mathrm{Z}_d = \omega_d^* \mathrm{Z}_d \mathrm{X}_d$, we have $\left(\mathrm{X}_d^a \mathrm{Z}_d^b\right)^d = \omega_d^{abd(d-1)/2} \mathrm{X}_d^{ad} \mathrm{Z}_d^{bd} = \omega_d^{abd(d-1)/2} \mathbb{1}$. So the eigenvalues of $\mathrm{X}_d^a \mathrm{Z}_d^b$ are dth complex roots of $\omega_d^{abd(d-1)/2}$, and the lemma follows. □

We also use the following simple number-theoretic property in the construction of new orthogonal unitary bases.

Lemma 4.5. *Any integer $d \geq 5$ has at most $d-2$ positive integer divisors.*

Proof. If d is prime, then it has exactly two positive integer divisors, $1, d$, and the lemma holds.

Suppose d is composite and has prime factorization $p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, where k and $a_1, a_2, \ldots, a_k$ are positive integers, and $p_1, p_2, \ldots, p_k$ are distinct prime numbers arranged in increasing order. The number of positive integer divisors of d equals $(a_1+1)(a_2+1)\cdots(a_k+1)$.

Since d is composite, either $k = 1$ and $a_1 \geq 2$, or $k \geq 2$.

Suppose $k = 1$. If $p_1 \geq 3$, the lemma follows since $n+1 \leq q^n - 2$ for all positive integers $n \geq 2$, for any $q \geq 3$. If $p_1 = 2$, we have $a_1 \geq 3$ since $d = p_1^{a_1} \geq 5$. Since $n+1 \leq 2^n - 2$ for all integers $n \geq 3$, the lemma again follows.

Now suppose $k \geq 2$. Since $n+1 \leq q^n$ and $n+1 \leq r^n - 1$ for all $n \geq 1$ whenever $q \geq 2$ and $r \geq 3$, the number of divisors of d is bounded as

$$\begin{aligned}(a_1+1)(a_2+1)\cdots(a_k+1) &\leq p_1^{a_1}(p_2^{a_2}-1)\,p_3^{a_3}\cdots p_k^{a_k} \\ &\leq p_1^{a_1}p_2^{a_2}p_3^{a_3}\cdots p_k^{a_k} - p_1^{a_1} \\ &\leq d-2\ ,\end{aligned}$$

as claimed. □

4.4 Explicit constructions

We now proceed to describe an explicit construction for all dimensions $d \geq 5$. The construction we give has the same form as the shift-and-multiply construction due to Vollbrecht and Werner [VW00]. In particular, the bases we present correspond to the construction with the Fourier transform over the cyclic group of order d as the Hadamard matrix, and certain Latin squares that are not equivalent to the one generated by X_d.

We construct bases that are not equivalent to the clock and shift basis by introducing a modification. In particular, we replace the operators generated by X_d by another sequence. Note that the operators X_d^i correspond to permutations on $\mathbb{Z}/d\mathbb{Z}$, i.e., they permute the standard basis of $\mathbb{C}^d$. Conversely any permutation P on $\mathbb{Z}/d\mathbb{Z}$ corresponds to the operator $\sum_{a\in\mathbb{Z}/d\mathbb{Z}} |P(a)\rangle\langle a|$ which permutes standard basis elements of $\mathbb{C}^d$. So this is a bijection.

It is also helpful to view a permutation P on $\mathbb{Z}/d\mathbb{Z}$ as a perfect matching in the complete bipartite graph $K_{d,d}$, with vertex a in one part being matched with the vertex $P(a)$ in the other. This mapping defines a bijection between permutations and perfect matchings. The construction we give relies on these three equivalent views of a permutation, and uses the same letter to refer to the corresponding matching and the linear operator on $\mathbb{C}^d$.

We start with the following observation.

Lemma 4.6. *Let $P_0, P_1, \ldots, P_{d-1}$ be a sequence of d* disjoint *matchings in the graph $K_{d,d}$. Then the matrices*

$$\left\{P_i Z_d^j : 0 \leq i, j < d\right\}$$

form an orthogonal unitary basis.

Proof. Since P_i permutes the standard basis vectors of $\mathbb{C}^d$, it is a unitary operator. Therefore $P_i Z_d^j$ is also unitary.

Now consider the inner product of $P_i Z_d^j$ and $P_k Z_d^l$ for two pairs (i,j) and (k,l). We have

$$\mathrm{Tr}(Z_d^{-j} P_i^{-1} P_k Z_d^l) \quad = \quad \sum_{m=0}^{d-1} \omega_d^{(l-j)m} \langle P_i(m) | P_k(m)\rangle \ .$$

If $i \neq k$, the inner product is 0 since for any m, the disjoint matchings P_i and P_k match the vertex m to distinct vertices. If $i = k$, but $l \neq j$, the inner product is again 0 as ω_d is a dth root of unity. □

We show that to derive non-equivalent bases, it suffices to have two of the matchings satisfy simple properties.

Lemma 4.7. *Let $d \geq 2$, and $P_0, P_1, \ldots, P_{d-1}$ be a sequence of d* disjoint *matchings in $K_{d,d}$ such that*

1. *P_0 is the identity permutation, i.e., matches vertex i in one part to vertex i in the other part; and*
2. *the permutation P_1 has a cycle of length k such that k does not divide d.*

Then the basis $B := \left\{ P_i Z_d^j : 0 \leq i, j < d \right\}$ is not equivalent to the clock and shift basis.

Proof. The intuition here is the following. The operator corresponding to the permutation P_1 has the k distinct kth roots of unity as eigenvalues. In particular, the eigenvalues include 1 and ω_k. On the other hand, the eigenvalues of any operator in the clock and shift basis are of the form $\gamma\omega_d^l$ for some integer l, and a fixed unit complex number $\gamma \in \mathsf{U}(\mathbb{C})$ depending only on the operator. Since k does not divide d, the operator P_1 does not belong to the clock and shift basis.

Formally, suppose that B is equivalent to the clock and shift basis, and the equivalence is given by unitary operators U, W. Suppose that the identity operator $P_0 \in B$ is mapped to $X_d^i Z_d^j$ and P_1 is mapped to $X_d^k Z_d^l$ under this equivalence. That is, $X_d^i Z_d^j = \alpha V P_0 W = \alpha V W$ and $X_d^k Z_d^l = \beta V P_1 W$ for some $\alpha, \beta \in \mathsf{U}(\mathbb{C})$.

From the equation for P_0 we have $V = \alpha^* X_d^i Z_d^j W^*$, so that $X_d^k Z_d^l = \beta\alpha^* X_d^i Z_d^j W^* P_1 W$. Equivalently, we have

$$\alpha\beta^*\omega_d^m X_d^{k-i} Z_d^{l-j} \quad = \quad W^* P_1 W \ , \tag{4.3}$$

where $m = -(k-i)j$. By Lemma 4.4, there is a fixed $\gamma \in \mathsf{U}(\mathbb{C})$ such that the eigenvalues of the operator on the left hand side of Eq. (4.3) are of the form $\gamma\omega_d^l$ for some integer l. On the other hand, the operator on the right hand side of the equation is similar to P_1. Since P_1 has eigenvalues 1 and ω_k, we have $1 = \gamma\omega_d^m$ and $\omega_k = \gamma\omega_d^n$ for some integers m, n. Eliminating γ, we get $\omega_k = \omega_d^{n-m}$. This implies that

$$\frac{2\pi\mathrm{i}}{k} \quad = \quad \frac{2\pi\mathrm{i}(n-m)}{d} + 2\pi\mathrm{i}p \ ,$$

for some integer p, or equivalently that $d = (n - m + pd)k$. This is a contradiction, as k does not divide d. □

Finally, we prove that matchings as in the hypothesis of Lemma 4.7 exist.

Lemma 4.8. *For any integer $d \geq 5$, there is a sequence of d* disjoint *matchings $P_0, P_1, \ldots, P_{d-1}$ in $K_{d,d}$ such that*

1. *P_0 is the identity permutation, i.e., matches vertex i in one part with vertex i in the other part; and*
2. *the permutation P_1 has a cycle of length k such that k does not divide d.*

Proof. By Lemma 4.5, for any $d \geq 5$, there is an integer $k \in [2, d-2]$ that does not divide d.

Let P_0 be the identity permutation, and let $P_1 := (0, 1, \ldots, k-1)(k, k+1, \ldots, d-1)$ be a permutation consisting of two cycles of length k and $d-k$, respectively. The perfect matchings corresponding to P_0 and P_1 are disjoint, as P_0 maps each element to itself while P_1 cyclically shifts every element within each of its two cycles (both of which are of length at least two).

Consider the graph G obtained by deleting the edges in the matchings P_0 and P_1 from $K_{d,d}$. The graph G is a $(d-2)$-regular bipartite graph. Thus, by the Hall theorem [Hal35], G can be decomposed into $(d-2)$ disjoint perfect matchings. □

Lemma 4.8 and Lemma 4.7 together imply that for any dimension $d \geq 5$, there are multiple non-equivalent orthogonal unitary bases. The same property holds for $d = 4$ due to Proposition 4.2, and for $d = 3$ due to Proposition 4.9 below. This proves Theorem 4.3.

Proposition 4.9. *There are orthogonal unitary bases for* $\mathsf{L}(\mathbb{C}^3)$ *that are not equivalent to the clock and shift basis.*

Proof. Denote the clock and shift basis by B. Note that B is a commutative projective group under operator composition, i.e., it is closed under taking products of the operators and the operators commute, all up to some phase (a unit complex number) that may depend on the operators. We construct a basis B' such that the equivalence of B and B' implies that B' also is a commutative projective group. However, the basis B' has elements that do not commute even up to a phase, which is a contradiction.

We construct B' following an idea due to Werner [Wer01]; see the discussion after Proposition 9 in the paper. Let $M := \beta|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|$, where $\beta \in \mathsf{U}(\mathbb{C})$ is a unit complex number such that $\beta \neq 1$. Let $B' := \{U_{ij} : 0 \leq i, j \leq 2\}$, where for any $i \in \{0,1,2\}$

$$U_{ij} \quad := \quad \begin{cases} \mathrm{X}_3^j \, \mathrm{Z}_3^i & j \in \{0,2\} \quad , \quad \text{and} \\ \mathrm{X}_3 \, \mathrm{Z}_3^i M & j = 1 \ . \end{cases}$$

We may verify that this is an orthogonal unitary basis for any choice of $\beta \in \mathsf{U}(\mathbb{C})$.

Suppose the basis B' is equivalent to B, and the equivalence is given by the operators $V, W \in \mathsf{U}(\mathbb{C}^3)$ and unit complex numbers $\alpha_{ij} \in \mathsf{U}(\mathbb{C})$. Consider the element $\mathrm{X}_3^a \, \mathrm{Z}_3^b$ of B that corresponds to the operator $U_{00} \in B'$. We have $\mathrm{X}_3^a \, \mathrm{Z}_3^b = \alpha_{00} V U_{00} W = \alpha_{00} V W$. Then $W = \alpha_{00}^* V^* \mathrm{X}_3^a \mathrm{Z}_3^b$, and

$$B \quad = \quad \left\{ \alpha_{ij} \, \alpha_{00}^* V U_{ij} V^* \mathrm{X}_3^a \, \mathrm{Z}_3^b \ : \ 0 \leq i, j \leq 2 \right\} \ .$$

Since B is closed under right multiplication by $(\mathrm{X}_3^a \, \mathrm{Z}_3^b)^*$ up to phases, the set of operators

$$\left\{ V U_{ij} M^* V^* : 0 \leq i, j \leq 2 \right\}$$

is also a commutative projective group, as is the basis B'.

We show next that not all operators in the set B' commute, even up to a phase. Consider the operators U_{01} and U_{02}. These operators commute up to a phase if and only if there is a unit complex number γ such that

$$\begin{array}{rrcl} & U_{01} U_{02} & = & \gamma \, U_{02} U_{01} \\ \iff & \mathrm{X}_3 M \, \mathrm{X}_3^2 & = & \gamma \, \mathrm{X}_3^2 \, \mathrm{X}_3 M \\ \iff & |0\rangle\langle 0| + \beta |1\rangle\langle 1| + |2\rangle\langle 2| & = & \gamma \left(\beta |0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2| \right) \ . \end{array}$$

This implies that $\gamma = \beta = 1$. As we chose $\beta \neq 1$, this is a contradiction, and B and B' are not equivalent. □

5 Random superdense coding protocols

In this section we study a random protocol for approximate superdense coding. Its analysis draws heavily on results in high-dimensional probability. We present these results in Section 5.1 and develop some properties of random entangled vectors in Section 5.2, before proceeding to the analysis in Section 5.3. Finally, in Section 5.4, we address a subtle issue that we encounter in the analysis.

5.1 Background from random matrix theory

In this section, we present some useful results from random matrix theory.

Definition 5.1 (Isotropic vector). *We say a random vector $|\boldsymbol{\xi}\rangle \in \mathbb{C}^n$ is* isotropic *if* $\mathbb{E}\,|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| = \mathbb{1}$.

Random variables which have tails that decay as fast as the normal distribution play an important role in high dimensional probability. Let S^{n-1} denote the set of unit vectors in $\mathbb{C}^n$.

Definition 5.2 (Sub-gaussian random variables and vectors). *A random variable $\boldsymbol{x} \in \mathbb{C}$ is* sub-gaussian *if there exists a parameter $\kappa > 0$ such that*

$$\Pr(|\boldsymbol{x}| \geq t) \quad \leq \quad 2\exp\big(-t^2/\kappa^2\big)$$

for all $t \geq 0$. The sub-gaussian norm *of $\boldsymbol{x}$, denoted by $\|\boldsymbol{x}\|_{\psi_2}$, is defined as*

$$\|\boldsymbol{x}\|_{\psi_2} \quad := \quad \inf\Big\{t > 0 : \mathbb{E}\exp(|\boldsymbol{x}|^2/t^2) \leq 2\Big\} \quad .$$

A random vector $|\boldsymbol{v}\rangle \in \mathbb{C}^n$ is sub-gaussian *if for all unit vectors $|u\rangle \in S^{n-1}$, the inner product $\langle u|\boldsymbol{v}\rangle$ is sub-gaussian. The sub-gaussian norm of $|\boldsymbol{v}\rangle$ is defined as*

$$\|\boldsymbol{v}\|_{\psi_2} \quad := \quad \sup_{u \in S^{n-1}} \|\langle u|\boldsymbol{v}\rangle\|_{\psi_2} \quad .$$

Sub-gaussian norm can be characterized in multiple ways. The following lemma describes two of them; see [Ver18, Proposition 2.5.2, Section 2.5.1].

Lemma 5.3. *There are positive universal constants c_1, c_2 such that for any random variable $\boldsymbol{x} \in \mathbb{C}$,*

1. *If for some parameter $\kappa_1 > 0$,*

$$\Pr(|\boldsymbol{x}| \geq t) \quad \leq \quad 2\exp\big(-t^2/\kappa_1^2\big)$$

for all $t \geq 0$, then $\boldsymbol{x}$ is sub-gaussian and $\|\boldsymbol{x}\|_{\psi_2} \leq c_1\kappa_1$.

2. *If $\boldsymbol{x}$ is sub-gaussian with $\|\boldsymbol{x}\|_{\psi_2} \leq \kappa_2$ for some parameter $\kappa_2 > 0$, then*

$$\Pr(|\boldsymbol{x}| \geq t) \quad \leq \quad 2\exp\big(-t^2/c_2\kappa_2^2\big)$$

for all $t \geq 0$.

The following theorem gives a sharp bound on the largest singular value of a class of random matrices; see the text by Verhsynin [Ver18] for this and related results. Vershynin states the result for real matrices, but the proof extends to complex matrices in a straightforward manner.

Theorem 5.4 ([Ver18], Theorem 4.6.1). *Let $\boldsymbol{A} := \sum_{i=1}^m |i\rangle\langle\boldsymbol{x}_i|$ be a complex $m \times n$ matrix whose rows $|\boldsymbol{x}_i\rangle$ are independent, mean zero, sub-gaussian isotropic vectors in $\mathbb{C}^n$. Then there is a universal constant $c > 0$ such that for all $t \geq 0$, we have*

$$\|\boldsymbol{A}\| \quad \leq \quad \sqrt{m} + c\kappa^2(\sqrt{n} + t)$$

with probability at least $1 - 2\exp(-t^2)$, where $\kappa := \max_i \|\boldsymbol{x}_i\|_{\psi_2}$.

Let $\|\cdot\|_2$ denote the Hilbert-Schmidt norm on $\mathsf{L}(\mathbb{C}^d)$:

$$\|A\|_2 \quad := \quad \sqrt{\mathrm{Tr}(A^*A)} \ .$$

This norm induces the following ℓ_2-sum metric on $\left(\mathsf{U}(\mathbb{C}^d)\right)^m$:

$$\|(U_1, U_2, \ldots, U_m) - (V_1, V_2, \ldots, V_m)\|_2 \quad := \quad \left(\sum_{i=1}^m \|U_i - V_i\|_2^2\right)^{1/2} \ .$$

Let $f : \left(\mathsf{U}(\mathbb{C}^d)\right)^m \to \mathbb{R}$ be a continuous function. We say f is κ-Lipschitz with respect to the ℓ_2-sum of Hilbert-Schmidt metrics if for all $(U_i), (V_i) \in \left(\mathsf{U}(\mathbb{C}^d)\right)^m$, we have

$$|f(U_1, U_2, \ldots, U_m) - f(V_1, V_2, \ldots, V_m)| \quad \leq \quad \kappa \, \|(U_1, U_2, \ldots, U_m) - (V_1, V_2, \ldots, V_m)\|_2 \ .$$

Let $\boldsymbol{U}_i \in \mathsf{U}(\mathbb{C}^d)$, $1 \leq i \leq m$ be i.i.d. Haar-random unitary operators. If κ is sufficiently smaller than the dimension d, with high probability, the random variable $f(\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_m)$ is close to its expectation. This concentration of measure property is formalized by the following theorem, which is a special case of Theorem 5.17 in the book on random matrix theory by Meckes [Mec19].

Theorem 5.5 ([Mec19], Theorem 5.17, page 159)**.** *Let $\boldsymbol{U}_i \in \mathsf{U}(\mathbb{C}^d)$, $i \in [m]$, be i.i.d. random unitary operators chosen according to the Haar measure. Suppose the function $f : \left(\mathsf{U}(\mathbb{C}^d)\right)^m \to \mathbb{R}$ is κ-Lipschitz with respect to the ℓ_2-sum of Hilbert-Schmidt metrics, with $\kappa > 0$. Then for every positive real number t, we have*

$$\Pr(f(\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_m) \geq \varphi + t) \quad \leq \quad \exp\left(-\frac{(d-2)t^2}{24\kappa^2}\right) \ ,$$

where $\varphi := \mathbb{E}\, f(\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_m)$.

The Marčenko–Pastur theorem characterises the spectrum of a wide class of random matrices in the limit of large dimension. We rely on a version of the theorem due to Yaskov [Yas16] that applies to matrices whose entries need not all be independent. While Yaskov states the result for *real* matrices, the proof extends to *complex* matrices with straightforward modifications. We sketch the observations and the modifications which enable this extension after the statement of the theorem.

The columns of the random matrices we consider satisfy a certain asymptotic isotropy condition.

Definition 5.6. *Let $m(n)$ be a sequence of positive integers such that $m \to \infty$ as $n \to \infty$. Let $(|\boldsymbol{x}_m\rangle)$ be a sequence of random vectors with $|\boldsymbol{x}_m\rangle \in \mathbb{C}^m$. We say that the sequence $(|\boldsymbol{x}_m\rangle)$ is pseudo-isotropic if for all sequences of complex matrices (A_m) with $A_m \in \mathbb{C}^{m\times m}$ and with uniformly bounded spectral norm (i.e., $\|A_m\| \leq \kappa$ for all m for a universal constant κ),*

$$\frac{1}{m}\left(\langle \boldsymbol{x}_m | A_m | \boldsymbol{x}_m \rangle - \mathrm{Tr}(A_m)\right) \xrightarrow{\mathrm{P}} 0$$

as $m \to \infty$.

Define the *empirical spectral distribution* (ESD) of an $m \times m$ positive semi-definite matrix A as $\frac{1}{m}\sum_{i=1}^m \delta(x - \lambda_i)$, where $(\lambda_i : i \in [m])$ are the eigenvalues of A, and δ is the Dirac-delta function. This is the probability density function of a uniformly random eigenvalue of A.

Theorem 5.7 (Marčenko-Pastur law [Yas16]). *Fix an $r > 0$, and let m, n be integers with $n, m \geq 1$ and m a function of n such that $m/n \to r$ as $n \to \infty$. For each m, let $|\boldsymbol{x}_m\rangle$ in $\mathbb{C}^m$ be a random vector such that the sequence of vectors $(|\boldsymbol{x}_m\rangle)$ is pseudo-isotropic. Let $(\boldsymbol{M}_{n,m})$ be a sequence of $m \times n$ random matrices whose columns are i.i.d. copies of the random vector $|\boldsymbol{x}_m\rangle$, and let $\boldsymbol{\mu}_{n,m}$ be the ESD of the matrix $\frac{1}{n}\boldsymbol{M}_{n,m}\boldsymbol{M}_{n,m}^*$. Then, as $n \to \infty$, the ESD $\boldsymbol{\mu}_{n,m}$ converges weakly to the density p_r almost surely, where*

$$\mathrm{p}_r(x) \quad := \quad \max\{0, 1 - 1/r\}\,\delta(x) + \frac{\sqrt{(x-a)(b-x)}}{2\pi r x}\,\mathbf{1}(a \leq x \leq b) \ ,$$

with $a := (1-\sqrt{r}\,)^2$, $b := (1+\sqrt{r}\,)^2$.

In other words, as $n \to \infty$, with probability 1, the cumulative distribution function of a uniformly random eigenvalue of the matrix $\frac{1}{n}\boldsymbol{M}_{n,m}\boldsymbol{M}_{n,m}^*$ converges point-wise to that given by the probability density function p_r.

Theorem 5.7 follows from the proof of Theorem 2.1 in Ref. [Yas16] by noting the following points. The eigenvalues of the matrix $\frac{1}{n}\boldsymbol{M}_{n,m}\boldsymbol{M}_{n,m}^*$ are all real, and therefore the Stieltjes continuity theorem applies to $\boldsymbol{\mu}_{n,m}$. Further, the Sherman-Morrison formula also extends to the sum $A + |u\rangle\langle v|$, where A is an invertible $m \times m$ complex matrix, and $|u\rangle, |v\rangle$ are in $\mathbb{C}^m$: the matrix $A + |u\rangle\langle v|$ is invertible if and only if $1 + \langle v|A^{-1}|u\rangle \neq 0$, and if the latter condition holds,

$$(A + |u\rangle\langle v|)^{-1} \quad = \quad A^{-1} - \frac{A^{-1}|u\rangle\langle v|A^{-1}}{1 + \langle v|A^{-1}|u\rangle} \ .$$

We can prove that the Stieltjes transform $\boldsymbol{s}_n(z)$ of $\boldsymbol{\mu}_{n,m}$ tends to its expectation $\mathbb{E}\,\boldsymbol{s}_n(z)$ almost surely as $n \to \infty$, following Step 1 in the proof of Theorem 1.1 in Ref. [BZ08]. The rest of the proof in Ref. [Yas16] now extends to the case of interest to us by replacing all instances of the transpose of a real vector by the conjugate transpose of the corresponding complex vector.

5.2 Pseudo-isotropy of random maximally entangled vectors

In this section, we develop properties of linear operators with certain symmetries, and use these to prove that a sequence of random maximally entangled vectors is pseudo-isotropic. This property is later used in the analysis of a random superdense coding protocol.

We consider operators on $\mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$, and label the four tensor factors with A, B, C, D, respectively. As is the convention in quantum information, we use superscripts to indicate the tensor factors on which an operator acts. Let $\mathrm{F} := \sum_{i,j=1}^d |i,j\rangle\langle j,i|$ be the *swap* operator on $\mathbb{C}^d \otimes \mathbb{C}^d$; it permutes the two tensor factors.

Lemma 5.8. *Let $W \in \mathsf{L}(\mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d)$. Suppose W commutes with $\mathbb{1}^{AB} \otimes U^C \otimes U^D$ for all unitary operators $U \in \mathsf{U}(\mathbb{C}^d)$, as well as with $U^A \otimes U^B \otimes \mathbb{1}^{CD}$. Then W is a linear combination of operators of the form $P^{AB} \otimes Q^{CD}$, where $P, Q \in \{\mathbb{1}, \mathrm{F}\}$.*

Proof. Let (E_i) be a basis for the vector space $\mathsf{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$. We may express W as $W = \sum_{i=1}^{d^2} E_i \otimes W_i$ for some operators $W_i \in \mathsf{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$. Since W commutes with $\mathbb{1}^{AB} \otimes U^C \otimes U^D$, we have

$$\sum_{i=1}^{d^2} E_i \otimes W_i \quad = \quad \sum_{i=1}^{d^2} E_i \otimes (U \otimes U) W_i (U^* \otimes U^*) \ ,$$

for all $U \in \mathsf{U}(\mathbb{C}^d)$. Since the operators E_i form a basis, we conclude that

$$W_i \quad = \quad (U \otimes U) W_i (U^* \otimes U^*) \ ,$$

i.e., the operator W_i commutes with $U \otimes U$ for every i. As a consequence of the von Neumann double commutant theorem [Wat18, Theorem 7.15, Section 7.1], each operator W_i may be written as a linear combination of $\{\mathbb{1}, \mathrm{F}\}$. So

$$W \quad = \quad \sum_{i=1}^{d^2} E_i \otimes (\alpha_i \mathbb{1} + \beta_i \mathrm{F}) \ ,$$

for some complex numbers α_i, β_i. Rearranging the sum, we get that

$$W \quad = \quad G \otimes \mathbb{1} + H \otimes \mathrm{F} \ ,$$

for some operators $G, H \in \mathsf{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$. Since W commutes with $U^A \otimes U^B \otimes \mathbb{1}^{CD}$ as well, and $\mathbb{1}$ and F are linearly independent, by [Wat18, Theorem 7.15, Section 7.1] we similarly get that G and H are also linear combinations of $\{\mathbb{1}, \mathrm{F}\}$. The lemma follows. □

Consider the random vector $|\psi\rangle$ defined as $|\psi\rangle := (\boldsymbol{U} \otimes \mathbb{1})|\phi_d\rangle$, where $\boldsymbol{U} \in \mathsf{U}(\mathbb{C}^d)$ is a Haar-random unitary operator and $|\phi_d\rangle$ is the maximally entangled state $\frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k\rangle|k\rangle$ with local dimension d. We would like to compute a closed form expression for the operator M on $\mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ defined as:

$$M \quad := \quad \mathbb{E}\, |\psi\rangle\langle\psi|^{\otimes 2} \ . \tag{5.1}$$

We use the symmetries of M in order to do so.

Lemma 5.9. *Let M be the operator defined in Eq. (5.1). Then*

$$M \quad = \quad \beta \left[\mathbb{1} + \mathrm{F}^{AC} \otimes \mathrm{F}^{BD}\right] + \gamma \left[\mathbb{1}^{AC} \otimes \mathrm{F}^{BD} + \mathrm{F}^{AC} \otimes \mathbb{1}^{BD}\right] \ ,$$

where $\beta := d^{-2}(d^2-1)^{-1}$ and $\gamma := -d^{-3}(d^2-1)^{-1}$.

Proof. Since

$$M \quad = \quad \mathbb{E}\ (\boldsymbol{U}^A \otimes \mathbb{1}^B)|\phi_d\rangle\langle\phi_d|(\boldsymbol{U}^{*A} \otimes \mathbb{1}^B) \otimes (\boldsymbol{U}^C \otimes \mathbb{1}^D)|\phi_d\rangle\langle\phi_d|(\boldsymbol{U}^{*C} \otimes \mathbb{1}^D) \ ,$$

and $\boldsymbol{U}$ is Haar random, the operator M commutes with $V^A \otimes V^C \otimes \mathbb{1}^{BD}$ for all $V \in \mathsf{U}(\mathbb{C}^d)$. Further, since $(\mathbb{1} \otimes V)|\phi_d\rangle = (V^\top \otimes \mathbb{1})|\phi_d\rangle$, the operator M also commutes with $\mathbb{1}^{AC} \otimes V^B \otimes V^D$. By Lemma 5.8, we have

$$M \quad = \quad \alpha\mathbb{1} + \beta(\mathrm{F}^{AC} \otimes \mathrm{F}^{BD}) + \gamma(\mathbb{1}^{AC} \otimes \mathrm{F}^{BD}) + \delta(\mathrm{F}^{AC} \otimes \mathbb{1}^{BD}) \ . \tag{5.2}$$

Consider the following linear functionals:

1. $X \mapsto \mathrm{Tr}(X)$
2. $X \mapsto \mathrm{Tr}((\mathrm{F}^{AC} \otimes \mathrm{F}^{BD})X)$
3. $X \mapsto \mathrm{Tr}((\mathrm{F}^{AC} \otimes \mathbb{1}^{BD})X)$
4. $X \mapsto \mathrm{Tr}((\mathbb{1}^{AC} \otimes \mathrm{F}^{BD})X)$

We apply these functionals to both sides of Eq. (5.2). We calculate the value of the functional on the left hand side directly from the definition of M, i.e., Eq. (5.1), and on the right hand side from Eq. (5.2). We thus obtain the following linear equations:

1. $1 = \alpha d^4 + \beta d^2 + \gamma d^3 + \delta d^3$,
2. $1 = \alpha d^2 + \beta d^4 + \gamma d^3 + \delta d^3$,
3. $1/d = \alpha d^3 + \beta d^3 + \gamma d^2 + \delta d^4$, and
4. $1/d = \alpha d^3 + \beta d^3 + \gamma d^4 + \delta d^2$, respectively.

Solving for $\alpha, \beta, \gamma, \delta$, we get the unique solution

$$\alpha = \beta = \frac{1}{d^2(d^2-1)} \quad , \quad \text{and} \quad \gamma = \delta = \frac{-1}{d^3(d^2-1)} \quad .$$

□

Let $n := d^2$. Consider the random vector $|\boldsymbol{\xi}_n\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ defined as $|\boldsymbol{\xi}_n\rangle := d|\psi\rangle = d(\boldsymbol{U} \otimes \mathbb{1})|\phi_d\rangle$. We prove that the sequence of these vectors is pseudo-isotropic.

Lemma 5.10. *The sequence of vectors $(|\boldsymbol{\xi}_n\rangle)$ is pseudo-isotropic.*

Proof. Let $(A_n \in \mathsf{L}(\mathbb{C}^d \otimes \mathbb{C}^d) : n \geq 1)$ be a sequence of complex matrices with spectral norm $\|A_n\|$ bounded by a constant κ, for each n. We use the Chebyshev Inequality to show that

$$\frac{1}{n}\left(\langle \boldsymbol{\xi}_n | A_n | \boldsymbol{\xi}_n \rangle - \mathrm{Tr}(A_n)\right) \xrightarrow{\mathrm{P}} 0 \tag{5.3}$$

as $n \to \infty$. Let $\boldsymbol{x}_n$ be the complex random variable defined as $\boldsymbol{x}_n := \langle \boldsymbol{\xi}_n | A_n | \boldsymbol{\xi}_n \rangle$. We may verify that $\mathbb{E}\,|\boldsymbol{\xi}_n\rangle\langle\boldsymbol{\xi}_n| = \mathbb{1}$, so that $\mathbb{E}\,\boldsymbol{x}_n = \mathbb{E}\,\mathrm{Tr}(|\boldsymbol{\xi}_n\rangle\langle\boldsymbol{\xi}_n| A_n) = \mathrm{Tr}(A_n)$. Eq. (5.3) is equivalent to showing that for every $\epsilon > 0$, $\Pr(|\boldsymbol{x}_n - \mathbb{E}\,\boldsymbol{x}_n| > \epsilon n) \to 0$ as $n \to \infty$. By the Chebyshev Inequality,

$$\Pr(|\boldsymbol{x}_n - \mathbb{E}\,\boldsymbol{x}_n| > \epsilon n) \quad \leq \quad \frac{1}{\epsilon^2 n^2}\,\mathbb{E}\,|\boldsymbol{x}_n - \mathbb{E}\,\boldsymbol{x}_n|^2 \quad .$$

So it suffices to show that the variance of $\boldsymbol{x}_n$ is $\mathrm{o}(n^2)$.

The variance $\mathbb{E}\,|\boldsymbol{x}_n - \mathbb{E}\,\boldsymbol{x}_n|^2 = \mathbb{E}\,|\boldsymbol{x}_n|^2 - |\mathbb{E}\,\boldsymbol{x}_n|^2 = \mathbb{E}\,|\boldsymbol{x}_n|^2 - |\mathrm{Tr}(A_n)|^2$. To calculate the second moment of $\boldsymbol{x}_n$, we rewrite it as follows.

$$\begin{aligned}
\mathbb{E}\,|\boldsymbol{x}_n|^2 &= \mathbb{E}\ \langle \boldsymbol{\xi}_n | A_n | \boldsymbol{\xi}_n \rangle \langle \boldsymbol{\xi}_n | A_n^* | \boldsymbol{\xi}_n \rangle \\
&= \mathbb{E}\ \mathrm{Tr}\left[(|\boldsymbol{\xi}_n\rangle\langle\boldsymbol{\xi}_n| \otimes |\boldsymbol{\xi}_n\rangle\langle\boldsymbol{\xi}_n|)(A_n \otimes A_n^*)\right] \\
&= n^2\,\mathrm{Tr}\left[M(A_n \otimes A_n^*)\right] \quad ,
\end{aligned}$$

where M is the matrix defined in Eq. (5.1). By Lemma 5.9, and the Hölder Inequality (namely, $|\mathrm{Tr}(AB)| \leq \|A\|_{\mathrm{tr}}\,\|B\|$),

$$\begin{aligned}
\mathbb{E}\,|\boldsymbol{x}_n|^2 &= n^2\Big[\beta\big(\,|\mathrm{Tr}(A_n)|^2 + \mathrm{Tr}(A_n^* A_n)\big) \\
&\quad + \gamma\big(\,\mathrm{Tr}\left((\mathrm{F}^{AC} \otimes \mathbb{1}^{BD})(A_n \otimes A_n^*)\right) + \mathrm{Tr}\left((\mathbb{1}^{AC} \otimes \mathrm{F}^{BD})(A_n \otimes A_n^*)\right)\big)\Big] \\
&\leq n^2\Big[\beta\big(\,|\mathrm{Tr}(A_n)|^2 + \kappa^2 n\big) + 2\,|\gamma|\,\kappa^2 n^2\Big] \quad ,
\end{aligned}$$

where $\beta = 1/n(n-1)$ and $\gamma = -1/n^{3/2}(n-1)$. Thus the variance is bounded as

$$\mathbb{E}\,|\boldsymbol{x}_n|^2 - |\mathrm{Tr}(A_n)|^2 \quad \leq \quad \frac{1}{n-1}\,|\mathrm{Tr}(A_n)|^2 + \frac{\kappa^2 n^2}{n-1} + \frac{2\kappa^2 n^{5/2}}{n-1} \quad ,$$

which is $\mathrm{o}(n^2)$ as $|\mathrm{Tr}(A_n)| \leq \kappa n$. This proves that the sequence $(|\boldsymbol{\xi}_n\rangle)$ is pseudo-isotropic. □

5.3 Analysis of a random protocol

Consider the following random protocol $\mathbf{\Pi}_d$. Let d be an integer ≥ 2, and $n := d^2$. Alice and Bob agree on a choice of n independently chosen Haar-random unitary operators $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n \in \mathsf{U}(\mathbb{C}^d)$. They also share the maximally entangled state $|\phi_d\rangle := \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k\rangle|k\rangle$ with local dimension d. When Alice gets message $i \in [n]$, she applies $\boldsymbol{U}_i$ to her half of $|\phi_d\rangle$, and sends it over to Bob. Bob now holds the state $|\psi_i\rangle := (\boldsymbol{U}_i \otimes \mathbb{1})|\phi_d\rangle$. He performs an optimal measurement to identify i, given that the state is drawn from the ensemble $\boldsymbol{\mathcal{E}}_d := \big(|\psi_j\rangle : j \in [n]\big)$.

Aram Harrow (personal communication) suggested the protocol $\mathbf{\Pi}_d$ as a candidate for an approximate (d, ϵ)-superdense coding protocol with vanishing error ϵ in the limit of large dimension. If this random construction of superdense coding protocols did indeed have error that vanishes rapidly as a function of d, then this could potentially refute Conjecture 1.3. This is formalized by the following proposition (which was stated in Section 1.3, and is reproduced here for convenience).

Proposition 1.4. *Suppose Conjecture 1.3 were true. Let $\delta_2(\varepsilon)$ be the function from Conjecture 1.3. Then the random superdense coding protocol $\mathbf{\Pi}_d$ specified by $(\phi_d, (\boldsymbol{U}_i))$ must have error ε satisfying*

$$\mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \delta_2(\varepsilon)^2 \quad \geq \quad (2d)^{-2} \ .$$

Proof. Let $\mathbf{\Pi}_d$ be the random protocol and let $(\boldsymbol{U}_i)$ be the ensemble of random unitaries specified in the Proposition statement. Suppose for contradiction that Conjecture 1.3 were true and the error ε of the protocol $\mathbf{\Pi}_d$ satisfied

$$\mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \delta_2(\varepsilon)^2 \quad < \quad (2d)^{-2} \ . \tag{5.4}$$

First we argue that

$$\mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j} \neq \boldsymbol{k}} |\operatorname{Tr}(\boldsymbol{U}_{\boldsymbol{j}} \boldsymbol{U}_{\boldsymbol{k}}^*)|^2 \quad = \quad 1 \ , \tag{5.5}$$

where the first expectation is over the ensemble of random unitary operators $(\boldsymbol{U}_i)$, and the second expectation is over a uniformly random pair of distinct indices $\boldsymbol{j}, \boldsymbol{k} \in [d]$, $\boldsymbol{j} \neq \boldsymbol{k}$. To prove this, note that for all $j \neq k$ $\mathbb{E}_{(\boldsymbol{U}_i)} |\operatorname{Tr}(\boldsymbol{U}_j \boldsymbol{U}_k^*)|^2 = \mathbb{E}_{(\boldsymbol{U}_i)} |\operatorname{Tr}(\boldsymbol{U}_1 \boldsymbol{U}_2^*)|^2$ because $\boldsymbol{U}_i$ are independent, identically distributed Haar-random unitaries operators. Furthermore, by the rotation invariance of the Haar measure, $\boldsymbol{U}_1 \boldsymbol{U}_2^*$ is also distributed according to the Haar measure. So the above quantity is equal to $\mathbb{E}_{\boldsymbol{U}} |\operatorname{Tr}(\boldsymbol{U})|^2$ for Haar-random $\boldsymbol{U}$. So the LHS of Equation (5.5) equals

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{U}} |\operatorname{Tr}(\boldsymbol{U})|^2 \quad &= \quad \mathop{\mathbb{E}}_{\boldsymbol{U}} \left| \sum_{j=1}^{d} \langle j|\boldsymbol{U}|j\rangle \right|^2 \\
&= \quad \mathop{\mathbb{E}}_{\boldsymbol{U}} \sum_{j,k=1}^{d} \langle j|\boldsymbol{U}|j\rangle\langle k|\boldsymbol{U}^*|k\rangle \\
&= \quad \sum_{j,k} \langle j| \Big(\mathop{\mathbb{E}}_{\boldsymbol{U}} \boldsymbol{U}|j\rangle\langle k|\boldsymbol{U}^* \Big) |k\rangle \ .
\end{aligned}$$

Since $\mathbb{E}_{\boldsymbol{U}} \boldsymbol{U}|j\rangle\langle j|\boldsymbol{U}^* = \mathbb{1}/d$ and $\mathbb{E}_{\boldsymbol{U}} \boldsymbol{U}|j\rangle\langle k|\boldsymbol{U}^* = 0$ when $j \neq k$, we have $\mathbb{E} |\operatorname{Tr}(\boldsymbol{U})|^2 = 1$, which establishes Equation (5.5).

On the other hand, the rigidity condition promised by Conjecture 1.3 implies that every collection of $d \times d$ unitary operators (U_i) yields a superdense protocol with some error ε, and in turn there exists an orthogonal unitary basis (E_i) such that $\|U_i - E_i\|_{\text{nhs}} \leq \delta_2(\varepsilon)$ for all $i \in [d^2]$.

Note that ε and (E_i) depend on (U_i), and let $\boldsymbol{\varepsilon}$ and $(\boldsymbol{E}_i)$ be the error of the protocol $\boldsymbol{\Pi}_d$ and the corresponding orthogonal unitary basis, respectively. Then

$$\begin{aligned}
\mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j}\neq\boldsymbol{k}} |\operatorname{Tr}(\boldsymbol{U_j}\boldsymbol{U_k}^*)|^2 &= \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j}\neq\boldsymbol{k}} |\operatorname{Tr}(\boldsymbol{U_j}\boldsymbol{U_k}^*) - \operatorname{Tr}(\boldsymbol{E_j}\boldsymbol{E_k}^*)|^2 \\
&\leq \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j}\neq\boldsymbol{k}} \Big(|\operatorname{Tr}((\boldsymbol{U_j}-\boldsymbol{E_j})\boldsymbol{U_k}^*)| + |\operatorname{Tr}(\boldsymbol{E_j}(\boldsymbol{U_k}^*-\boldsymbol{E_k}^*))| \Big)^2 \\
&\leq 2 \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j}\neq\boldsymbol{k}} |\operatorname{Tr}((\boldsymbol{U_j}-\boldsymbol{E_j})\boldsymbol{U_k}^*)|^2 + |\operatorname{Tr}(\boldsymbol{E_j}(\boldsymbol{U_k}^*-\boldsymbol{E_k}^*))|^2 \qquad (5.6)
\end{aligned}$$

where the first equality is due to the orthogonality condition $\operatorname{Tr}(\boldsymbol{E}_j\boldsymbol{E}_k^*) = 0$ whenever $j \neq k$, the second line is due to the triangle inequality, and the third line is due to the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ for real numbers a, b. By the Cauchy-Schwarz inequality for the Hilbert-Schmidt inner product, we have $|\operatorname{Tr}((\boldsymbol{U}_j - \boldsymbol{E}_j)\boldsymbol{U}_k^*)|^2 \leq \|\boldsymbol{U}_j - \boldsymbol{E}_j\|_2^2 \cdot \|\boldsymbol{U}_k\|_2^2$ and $|\operatorname{Tr}(\boldsymbol{E}_j(\boldsymbol{U}_k^* - \boldsymbol{E}_k^*))|^2 \leq \|\boldsymbol{E}_j\|_2^2 \cdot \|\boldsymbol{U}_k - \boldsymbol{E}_k\|_2^2$, where $\|X\|_2 = \sqrt{\operatorname{Tr}(XX^*)}$ denotes the (unnormalized) Hilbert-Schmidt norm. Since $\|A\|_2^2 = d$ for all $d \times d$ unitary matrices A, we can upper bound the RHS of Equation (5.6) as

$$\begin{aligned}
&\leq 2d \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \mathop{\mathbb{E}}_{\boldsymbol{j}\neq\boldsymbol{k}} \left(\|\boldsymbol{U_j}-\boldsymbol{E_j}\|_2^2 + \|\boldsymbol{U_k}-\boldsymbol{E_k}\|_2^2\right) \\
&= 4 \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \sum_j \|\boldsymbol{U}_j - \boldsymbol{E}_j\|_{\text{nhs}}^2 \\
&\leq 4d^2 \mathop{\mathbb{E}}_{(\boldsymbol{U}_i)} \delta_2(\boldsymbol{\varepsilon})^2 \\
&< 1 \ ,
\end{aligned}$$

where the last inequality follows from the assumption in Equation (5.4). However, this contradicts Equation (5.5). Thus, either the conjecture does not hold, or the random superdense coding protocol $\boldsymbol{\Pi}_d$ has error satisfying $\mathbb{E}\,\delta_2(\boldsymbol{\varepsilon}) \geq (2d)^{-2}$. □

In the rest of this section, we prove that for sufficiently large dimension, with high probability, the protocol $\boldsymbol{\Pi}_d$ has positive constant error. This indicates that random maximally entangled quantum states are not very reliable for transmitting classical information, and proves Theorem 1.5. Thus, the random protocol $\boldsymbol{\Pi}_d$ does not rule out a robust rigidity theorem for superdense coding.

To analyze the decoding error of the protocol, we study the *distinguishability* of the ensemble $\boldsymbol{\mathcal{E}}_d$. This is the probability that, if the pure state $|\psi_i\rangle$ is selected uniformly at random from the ensemble, an optimal measurement correctly identifies the state.

Definition 5.11. *Let $\mathcal{F} := \big((p_i, \rho_i) : \rho_i \in \mathsf{D}(\mathbb{C}^k), i \in [m]\big)$ be an ensemble of states in which state ρ_i occurs with probability p_i. We define the* distinguishability *of $\mathcal{F}$ as*

$$\gamma(\mathcal{F}) \quad := \quad \max_{POVM\ M} \sum_{i=1}^m p_i \operatorname{Tr}(M_i\rho_i) \ ,$$

where the maximization is over all measurements (i.e., POVMs) M with elements $M_1, \ldots, M_m$.

We can estimate the distinguishability of an ensemble of states via the generalized Holevo-Curlander bounds [Kho79, Cur79, ON99, Tys09b].

Theorem 5.12 (generalized Holevo-Curlander bounds [ON99, Tys09b])**.** *Let $\mathcal{F} := \big((p_i, \rho_i) : \rho_i \in \mathsf{D}(\mathbb{C}^k), i \in [m]\big)$ be an ensemble of m quantum states. Then the distinguishability of $\mathcal{F}$ satisfies*

$$\big(\alpha_{Holevo}(\mathcal{F})\big)^2 \quad \leq \quad \gamma(\mathcal{F}) \quad \leq \quad \alpha_{Holevo}(\mathcal{F}) \ ,$$

where

$$\alpha_{Holevo}(\mathcal{F}) \quad := \quad \operatorname{Tr} \sqrt{\sum_{i=1}^{m} p_i^2 \rho_i^2} \ .$$

We only need the upper bound on distinguishability above for a uniform ensemble of pure states. This bound was given by Curlander [Cur79] in the case of linearly independent states. It was generalized to the case of equiprobable, possibly mixed states by Ogawa and Nagaoka [ON99, Lemma 1]. The proof they gave also extends with minor modifications to non-uniform ensembles. The two bounds in Theorem 5.12 were proven — re-proven independently in the case of the upper bound [Tys09a] — by Tyson[Tys09b, Theorem 10]. Tyson later gave another proof of the bounds which also generalizes to error-recovery [Tys10, Section III].

We show that the expectation of the quantity $\alpha_{\text{Holevo}}(\boldsymbol{\mathcal{E}}_d)$ for the ensemble of random maximally entangled states is at most a constant strictly less than 1, for sufficiently large dimension d. This implies that the distinguishability $\gamma(\boldsymbol{\mathcal{E}}_d)$ is also strictly less than 1 in expectation, and that any measurement Bob makes has a non-zero constant probability of failure, on average.

Theorem 5.13. *The distinguishability $\gamma(\boldsymbol{\mathcal{E}}_d)$ of the random superdense coding protocol $\boldsymbol{\Pi}_d$ tends to $\frac{8}{3\pi} \approx 0.85$ as $d \to \infty$.*

Proof. Define the matrix $\boldsymbol{Q}$ as $\boldsymbol{Q} := \sum_{i=1}^{n} |\psi_i\rangle\langle\psi_i|$, and $\boldsymbol{\Lambda}_d$ as a uniformly random eigenvalue of $\boldsymbol{Q}$. Then $\mathbb{E}\, \alpha_{\text{Holevo}}(\boldsymbol{\mathcal{E}}_d)$ is the expectation of the random variable $\sqrt{\boldsymbol{\Lambda}_d}$, and we aim to bound this from above.

Define the $n \times n$ matrix $\boldsymbol{R} := d \sum_{i=1}^{n} |\psi_i\rangle\langle i|$ so that $\boldsymbol{Q} = \frac{1}{n}\boldsymbol{R}\boldsymbol{R}^*$. Consider the random vector $|\boldsymbol{\xi}_n\rangle$ defined as $|\boldsymbol{\xi}_n\rangle := d(\boldsymbol{U} \otimes \mathbb{1})|\phi_d\rangle$, where $\boldsymbol{U} \in \mathsf{U}(\mathbb{C}^d)$ is a Haar-random unitary operator. I.e., $|\boldsymbol{\xi}_n\rangle$ is a scaled random maximally entangled state. Lemma 5.10 shows that the sequence $(|\boldsymbol{\xi}_n\rangle)$ is pseudo-isotropic. Since the columns of the matrix $\boldsymbol{R}$ are i.i.d. copies of the random vector $|\boldsymbol{\xi}_n\rangle$, the random matrix $\boldsymbol{Q}$ is of the form described in Theorem 5.7. Thus the limiting distribution of the uniformly random eigenvalue $\boldsymbol{\Lambda}_d$ of $\boldsymbol{Q}$ follows the Marčenko-Pastur law with density p_1 (i.e., with parameter $r = 1$):

$$\mathrm{p}_1(x) \quad := \quad \begin{cases} \frac{1}{2\pi}\sqrt{\frac{4-x}{x}} & \text{if } 0 \le x \le 4 \ , \text{ and} \\ 0 & \text{otherwise} \ . \end{cases} \tag{5.7}$$

We would like to use the density p_1 to estimate the limit of the expectation of $\sqrt{\boldsymbol{\Lambda}_d}$. A subtle issue is that weak convergence (i.e., convergence in distribution of the random variables) does not necessarily imply that the limit of the expectation values $\mathbb{E}\sqrt{\boldsymbol{\Lambda}_d}$ equals the expectation of the limiting random variable. A simple example for which this does not hold is described in Section 5.4. Nonetheless, in Theorem 5.16 in Section 5.4, we show that the sequence of random variables $\boldsymbol{\Lambda}_d$ satisfies the stronger property we need. Namely, $\mathbb{E}\sqrt{\boldsymbol{\Lambda}_d}$ converges to $\mathbb{E}\sqrt{\boldsymbol{\Lambda}}$ as $d \to \infty$, where $\boldsymbol{\Lambda}$ is a random variable with density p_1. We may thus bound the distinguishability of the ensemble $\boldsymbol{\mathcal{E}}_d$ as $d \to \infty$ as follows.

$$\begin{aligned} \lim_{d\to\infty} \mathbb{E}\, \gamma(\boldsymbol{\mathcal{E}}_d) \quad &\le \quad \lim_{d\to\infty} \mathbb{E}\, \alpha_{\text{Holevo}}(\boldsymbol{\mathcal{E}}_d) \\ &= \quad \lim_{d\to\infty} \mathbb{E}\sqrt{\boldsymbol{\Lambda}_d} \quad = \quad \mathbb{E}\sqrt{\boldsymbol{\Lambda}} \\ &= \quad \int_{-\infty}^{\infty} \sqrt{x}\ \mathrm{p}_1(x)\, \mathrm{d}x \\ &= \quad \frac{1}{2\pi}\int_0^4 \sqrt{4-x}\, \mathrm{d}x \end{aligned}$$

$$= \quad \frac{8}{3\pi} \quad \approx \quad 0.85 \ .$$

□

We can strengthen this result to show that the distinguishability of $\boldsymbol{\mathcal{E}}_d$ is tightly concentrated around the mean. So all but an exponentially small fraction of the superdense coding protocols using $|\phi_d\rangle$ succeed with probability smaller than ≈ 0.85.

Theorem 5.14. *The distinguishability $\gamma(\boldsymbol{\mathcal{E}}_d)$ of the random superdense coding protocol $\boldsymbol{\Pi}_d$ satisfies*

$$\Pr\Big(\gamma(\boldsymbol{\mathcal{E}}_d) \geq \mathbb{E}\,\gamma(\boldsymbol{\mathcal{E}}_d) + t\Big) \leq \exp\left(-\frac{d^3(d-2)t^2}{96}\right) \ .$$

Proof. Define the function $f : \big(\mathsf{U}(\mathbb{C}^d)\big)^n \to \mathbb{R}$ as

$$f(U_1, \ldots, U_n) \quad \coloneqq \quad \sup_{\text{POVM } M} \frac{1}{n} \sum_{i=1}^{n} \mathrm{Tr}(M_i \psi_i) \ ,$$

where $|\psi_i\rangle = (U_i \otimes \mathbb{1})|\phi_d\rangle$, and we denote $|\psi\rangle\langle\psi|$ by ψ. So $f(\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n) = \gamma(\boldsymbol{\mathcal{E}}_d)$, the distinguishability of the ensemble $\boldsymbol{\mathcal{E}}_d$. We bounded the expected distinguishability by $\approx .85 < 1$ in Theorem 5.13. We show that $\gamma(\boldsymbol{\mathcal{E}}_d)$ is tightly concentrated around its expectation using Theorem 5.5. To do so, we compute a bound on the Lipschitz constant of f.

Fix unitary operators $U_1, \ldots, U_n, U_1', \ldots, U_n' \in \mathsf{U}(\mathbb{C}^d)$ and let $|\psi_i\rangle = (U_i \otimes \mathbb{1})|\phi_d\rangle$ and $|\psi_i'\rangle = (U_i' \otimes \mathbb{1})|\phi_d\rangle$. Since the space of n-dimensional POVMs with n outcomes is compact, for any sequence (U_i) the supremum in the definition of f is attained at some POVM M. Let M, M' correspond to the POVMs achieving $f(U_1, \ldots, U_n)$ and $f(U_1', \ldots, U_n')$, respectively, and let α, α' denote these quantities. Assume without loss of generality that $\alpha' \leq \alpha$.

We have that

$$|\alpha - \alpha'| \quad = \quad \alpha - \alpha' \quad = \quad \frac{1}{d^2}\Big(\sum_i \mathrm{Tr}(M_i \psi_i) - \mathrm{Tr}(M_i' \psi_i')\Big) \quad \leq \quad \frac{1}{d^2} \sum_i \mathrm{Tr}(M_i(\psi_i - \psi_i')) \ ,$$

as the POVM M may not be an optimal distinguishing measurement for the ensemble (ψ_i'). We bound this by

$$\begin{aligned}
&\frac{1}{d^2} \sum_i \Big| \mathrm{Tr}(M_i(\psi_i - \psi_i'))\Big| \\
&\quad \leq \quad \frac{1}{d^2} \sum_i \|M_i\| \cdot \|\psi_i - \psi_i'\|_1 && \text{(Hölder inequality)} \\
&\quad \leq \quad \frac{1}{d^2} \sum_i \|\psi_i - \psi_i'\|_1 && (M \text{ is a POVM}) \\
&\quad \leq \quad \frac{1}{d^2} \sum_i 2 \| |\psi_i\rangle - |\psi_i'\rangle \| \\
&\quad = \quad \frac{2}{d^2} \sum_i \sqrt{\langle \phi | (U_i - U_i')^* (U_i - U_i') | \phi \rangle} \\
&\quad \leq \quad 2 \sqrt{\frac{1}{d^2} \sum_i \frac{1}{d} \|U_i - U_i'\|_2^2} && \text{(Jensen inequality)}
\end{aligned}$$

where in the fourth line we used the property that for any two pure states $|\varphi\rangle$ and $|\theta\rangle$, the trace distance between $|\varphi\rangle\langle\varphi|$ and $|\theta\rangle\langle\theta|$ is at most $2\||\varphi\rangle - |\theta\rangle\|$. In the last line, we used the identity $\langle\phi_d|(A \otimes \mathbb{1})|\phi_d\rangle$ for any $d \times d$ matrix is equal to $\mathrm{Tr}(A)/d$.

Thus $|\alpha - \alpha'| \leq 2d^{-3/2}\sqrt{\sum_i \|U_i - U_i'\|_2^2}$, which implies that f is $2d^{-3/2}$-Lipschitz. Applying Theorem 5.5 we obtain

$$\Pr\left(\gamma(\boldsymbol{\mathcal{E}}_d) \geq \mathbb{E}\,\gamma(\boldsymbol{\mathcal{E}}_d) + t\right) \quad \leq \quad \exp\left(-\frac{d^3(d-2)t^2}{96}\right) \ .$$

□

5.4 A subtle issue

Recall from the proof of Theorem 5.13 in Section 5.3 that $\boldsymbol{\Lambda}_d$ denotes a uniformly random eigenvalue of the matrix $\boldsymbol{Q}$. The generalized Marčenko-Pastur Law (Theorem 5.7) tells us that $\boldsymbol{\Lambda}_d$ converges in distribution to a random variable $\boldsymbol{\Lambda}$ with density p_1 given in Eq. (5.7) as $d \to \infty$. We used this limiting distribution to estimate the limit of the mean of $\sqrt{\boldsymbol{\Lambda}_d}$ in Theorem 5.13. We pointed out the subtle issue that convergence in distribution does not necessarily imply that the limit of means equals the mean of the limiting random variable. A simple example which illustrates this issue is the following. For any positive integer k, let the random variable $\boldsymbol{x}_k$ take value k with probability $1/k$, and value 0 with the remaining probability. Then $\boldsymbol{x}_k$ converges in distribution to the constant 0, whereas $\mathbb{E}\,\boldsymbol{x}_k = 1$ for all k.

The example above highlights the reason underlying this phenomenon: while the probability of an interval on the line may go to zero in the limit, the rate of convergence may not be fast enough to dampen the contribution to the mean from that interval. We show that the probability that the random variable $\boldsymbol{\Lambda}_d$ deviates from zero decays exponentially. This helps us conclude the convergence of the mean $\mathbb{E}\,\sqrt{\boldsymbol{\Lambda}_d}$ to $\mathbb{E}\,\sqrt{\boldsymbol{\Lambda}}$.

A similar property was assumed to hold by Montanaro [Mon07] in his work on the distinguishability of random quantum states. Let $\boldsymbol{S} := \sum_{i=1}^k |\zeta_i\rangle\langle i|$, where $|\zeta_i\rangle$ are i.i.d. random vectors in $\mathbb{C}^d$ with i.i.d. complex gaussian entries with mean 0 and variance 1. Montanaro approximates $\mathbb{E}\,\mathrm{Tr}(\boldsymbol{S}\boldsymbol{S}^*)^{1/2}$ using the Marčenko-Pastur Law. He justifies this using estimates on the rate of convergence of the expected distribution of a uniformly random eigenvalue of $(1/k)\boldsymbol{S}\boldsymbol{S}^*$ to the limiting distribution given by the Marčenko-Pastur Law (see the discussion after Lemma 5 in Ref. [Mon07]). The rate of convergence is measured in terms of the Kolmogorov distance between the two distributions. (The Kolmogorov distance between the cumulative distribution functions F_1 and F_2 of real random variables is defined as $\sup_{x\in\mathbb{R}} |F_1(x) - F_2(x)|$.) The Kolmogorov distance was shown to be $\mathrm{O}(k^{-5/48})$ by Bai [Bai93]. However, vanishing Kolmogorov distance does not necessarily imply the convergence of the mean to the mean of the limiting distribution. For example, the Kolmogorov distance of the distribution of the random variable $\boldsymbol{x}_k$ defined above from the constant 0 is $1/k$. The approach we take in this section can also be used to fill the gap in Montanaro's work. In fact, the analogue of Lemma 5.15 we need for this purpose follows directly, as the columns of $\boldsymbol{S}$ are gaussian. We leave the details to the reader, and return to the analysis of the random variable $\boldsymbol{\Lambda}_d$.

In order to show that $\boldsymbol{\Lambda}_d$ has an exponentially decaying tail, it suffices to show that the spectral norm of the matrix $\boldsymbol{Q}$—i.e., its largest eigenvalue—has this property. So we proceed by deriving a tail bound for the spectral norm of $\boldsymbol{Q}$.

Lemma 5.15. *Let $d \geq 3$. There are positive universal constants c_1, c_2 such that for all $t \geq c_1$,*

$$\Pr\left(\|\boldsymbol{Q}\| > t\right) \quad \leq \quad 2\exp(-tn/c_2) \ . \tag{5.8}$$

Proof. Recall that $n := d^2$, $\boldsymbol{Q} := \sum_{i=1}^n |\psi_i\rangle\langle\psi_i|$, and that $\boldsymbol{R} := d\sum_{i=1}^n |\psi_i\rangle\langle i|$. We have $\|\boldsymbol{Q}\| = n^{-1}\|\boldsymbol{R}\boldsymbol{R}^*\| = n^{-1}\|\boldsymbol{R}\|^2$, so

$$\Pr\left(\|\boldsymbol{Q}\| > t\right) \quad = \quad \Pr\left(\|\boldsymbol{R}\| > \sqrt{nt}\right) \ .$$

It thus suffices to give a suitable tail bound for $\|\boldsymbol{R}\|$.

The vectors $d|\psi_i\rangle$ are i.i.d. copies of the random vector $|\xi_n\rangle$ defined as $|\xi_n\rangle := d(\boldsymbol{U}\otimes\mathbb{1})|\phi_d\rangle$, where $\boldsymbol{U}$ is a Haar-random unitary operator on $\mathbb{C}^d$. So the vectors $d|\psi_i\rangle$ have zero mean. We can verify that $\mathbb{E}\,|\xi_n\rangle\langle\xi_n| = \mathbb{1}$, so the vectors $d|\psi_i\rangle$ are isotropic. We prove below that the vector $|\xi_n\rangle$ is sub-gaussian, with sub-gaussian norm at most a universal constant κ. So the matrix $\boldsymbol{R}^* = d\sum_{i=1}^n |i\rangle\langle\psi_i|$ satisfies the conditions of Theorem 5.4, and we have that for some positive universal constant c_3,

$$\|\boldsymbol{R}\| = \|\boldsymbol{R}^*\| \quad > \quad \sqrt{n} + c_3\kappa^2(\sqrt{n} + t_1)$$

with probability at most $2\exp(-t_1^2)$, for all $t_1 \geq 0$. Let t_1 be such that the right hand side above equals $\sqrt{nt}$, i.e.,

$$t_1 \quad := \quad \frac{\sqrt{n}}{c_3\kappa^2}\left(\sqrt{t} - (1 + c_3\kappa^2)\right) \ .$$

Let $c_1 := 4(1 + c_3\kappa^2)^2$. Whenever $t \geq c_1$, we see that $t_1 \geq \sqrt{nt}/2c_3\kappa^2 \geq 0$. So

$$\Pr\left(\|\boldsymbol{R}\| > \sqrt{nt}\right) \quad \leq \quad 2\exp\big(-nt/4c_3^2\kappa^4\big) \ ,$$

and the theorem holds with $c_2 := 4c_3^2\kappa^4$.

It remains to prove that the sub-gaussian norm κ of $|\xi_n\rangle$ is at most some universal constant. By Lemma 5.3, it suffices to show that for any unit vector $|u\rangle \in \mathbb{C}^n$, the random variable $\boldsymbol{x} := \langle u|\xi_n\rangle$ has sub-gaussian tails: for a positive universal constant κ_1,

$$\Pr(|\boldsymbol{x}| \geq t) \quad \leq \quad 2\exp\big(-t^2/\kappa_1^2\big) \ , \tag{5.9}$$

for all $t \geq 0$. We establish this by appealing to Theorem 5.5.

Since $|\xi_n\rangle$ is isotropic,

$$\mathbb{E}\,|\boldsymbol{x}|^2 \quad = \quad \langle u|\left(\mathbb{E}\,|\xi_n\rangle\langle\xi_n|\right)|u\rangle \quad = \quad 1 \ .$$

So $\mathbb{E}\,|\boldsymbol{x}| \leq \left(\mathbb{E}\,|\boldsymbol{x}|^2\right)^{1/2} \leq 1$.

Define the function $f : \mathsf{U}(\mathbb{C}^d) \to \mathbb{C}$ as $f(U) := d\,|\langle u|(U\otimes\mathbb{1})|\phi_d\rangle|$. Then $|\boldsymbol{x}| = f(\boldsymbol{U})$. To show that f is Lipschitz, consider $U, V \in \mathsf{U}(\mathbb{C}^d)$. Since $|u\rangle$ is a unit vector and $\langle\phi_d|(W\otimes\mathbb{1})|\phi_d\rangle = \frac{1}{d}\operatorname{Tr}(W)$ we have

$$\begin{aligned}
|f(U) - f(V)| \quad &\leq \quad d\,|\langle u|((U-V)\otimes\mathbb{1})|\phi_d\rangle| \\
&\leq \quad d\,\|((U-V)\otimes\mathbb{1})|\phi_d\rangle\| \\
&= \quad d\left(\langle\phi_d|((U-V)^*(U-V)\otimes\mathbb{1})|\phi_d\rangle\right)^{1/2} \\
&= \quad \sqrt{d}\,\|U-V\|_2 \ ,
\end{aligned}$$

where $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm on $\mathsf{L}(\mathbb{C}^d)$. So the function f is $\sqrt{d}$-Lipschitz with respect to Hilbert-Schmidt metric. By Theorem 5.5,

$$\Pr(|\boldsymbol{x}| \geq 1 + t_1) \quad \leq \quad \exp\left(-\frac{(d-2)t_1^2}{24d}\right)$$

for every $t_1 > 0$. For $d \geq 3$, we have $d - 2 \geq d/3$. So the right hand side is at most $\exp(-t_1^2/72)$, and we have

$$\Pr(|\boldsymbol{x}| \geq t) \quad \leq \quad \exp\left(-\frac{(t-1)^2}{3 \cdot 24}\right) \quad \leq \quad \exp\left(-\frac{t^2}{12 \cdot 24}\right) \ ,$$

for all $t \geq 2$ (as $t - 1 \geq t/2$). Note that Eq. (5.9) holds trivially for $t \in [0, 2]$ for any choice of positive constant κ_1 such that $2\exp(-4/\kappa_1^2) \geq 1$. Thus, taking $\kappa_1 := 24$, we see that Eq. (5.9) holds for all $t \geq 0$ whenever $d \geq 3$. □

Lemma 5.15 implies that for a large enough constant α, the contribution to the mean $\mathbb{E}\sqrt{\boldsymbol{\Lambda}_d}$ outside an interval $[0, \alpha]$ goes to 0 as $d \to \infty$. Within this interval, the contribution to the mean tends to that for $\mathbb{E}\sqrt{\boldsymbol{\Lambda}}$. This helps us derive the limiting value of the mean.

Theorem 5.16. $\lim_{d\to\infty} \mathbb{E}\sqrt{\boldsymbol{\Lambda}_d} \quad = \quad \mathbb{E}\sqrt{\boldsymbol{\Lambda}}$.

Proof. We formalize the intuition given above by appealing to a weaker property implied by convergence in distribution, namely that the expectation of any *bounded* continuous function f of the random variable $\boldsymbol{\Lambda}_d$ converges to $\mathbb{E} f(\boldsymbol{\Lambda})$.

Fix $\alpha \geq \max\{c_1, 4\}$, where c_1 is the constant in the statement of Lemma 5.15 and consider the function f_α defined as follows:

$$f_\alpha(x) \quad := \quad \begin{cases} 0 & x \leq 0 \\ \sqrt{x} & 0 < x \leq \alpha \\ \sqrt{\alpha} & \alpha < x \ . \end{cases}$$

Since f_α is continuous and bounded, and $\boldsymbol{\Lambda} \in [0, 4]$,

$$\lim_{d\to\infty} \mathbb{E} f_\alpha(\boldsymbol{\Lambda}_d) \quad = \quad \mathbb{E} f_\alpha(\boldsymbol{\Lambda}) \quad = \quad \mathbb{E}\sqrt{\boldsymbol{\Lambda}} \ .$$

On the other hand, $\boldsymbol{\Lambda}_d \geq 0$ and $f_\alpha(x) \leq \sqrt{x}$ for all $x \geq 0$. So $\mathbb{E} f_\alpha(\boldsymbol{\Lambda}_d) \leq \mathbb{E}\sqrt{\boldsymbol{\Lambda}_d}$, and

$$\mathbb{E}\sqrt{\boldsymbol{\Lambda}} \quad = \quad \lim_{d\to\infty} \mathbb{E} f_\alpha(\boldsymbol{\Lambda}_d) \quad \leq \quad \lim_{d\to\infty} \mathbb{E}\sqrt{\boldsymbol{\Lambda}_d} \ .$$

We prove the reverse inequality using Lemma 5.15. Let $p(x)$ be the probability density function of $\boldsymbol{\Lambda}_d$. By the definition of f_α,

$$\mathbb{E}\sqrt{\boldsymbol{\Lambda}_d} \quad \leq \quad \mathbb{E} f_\alpha(\boldsymbol{\Lambda}_d) + \int_{x \geq \alpha} \sqrt{x} p(x) \, \mathrm{d}x \ . \tag{5.10}$$

Let $g(\alpha, d)$ denote the second term on the right hand side of Eq. (5.10) above. This is the contribution to $\mathbb{E}\sqrt{\boldsymbol{\Lambda}_d}$ outside of the interval $[0, \alpha]$. Using $\alpha \geq 4$, Fubini's Theorem, $\boldsymbol{\Lambda}_d \leq \|\boldsymbol{Q}\|$, and Lemma 5.15, we have

$$\begin{aligned} g(\alpha, d) \quad &\leq \quad \int_{x \geq \alpha} x p(x) \, \mathrm{d}x \\ &= \quad \int_{x \geq \alpha} \int_{y \in [0,x]} p(x) \, \mathrm{d}y \, \mathrm{d}x \\ &= \quad \int_{y \geq 0} \int_{x \geq \max\{\alpha, y\}} p(x) \, \mathrm{d}x \, \mathrm{d}y \\ &= \quad \int_{y \in [0,\alpha]} \int_{x \geq \alpha} p(x) \, \mathrm{d}x \, \mathrm{d}y + \int_{y \geq \alpha} \int_{x \geq y} p(x) \, \mathrm{d}x \, \mathrm{d}y \end{aligned}$$

$$
\begin{aligned}
&= \int_{y\in[0,\alpha]} \Pr(\mathbf{\Lambda}_d \geq \alpha)\,\mathrm{d}y + \int_{y\geq\alpha} \Pr(\mathbf{\Lambda}_d \geq y)\,\mathrm{d}y \\
&\leq 2\alpha\exp(-\alpha n/c_2) + 2\int_{y\geq\alpha} \exp(-yn/c_2)\,\mathrm{d}y \\
&= 2(\alpha + c_2/n)\exp(-\alpha n/c_2)\ ,
\end{aligned}
$$

where c_2 is the universal constant in the statement of Lemma 5.15. Since $n = d^2$, $g(\alpha, d)$ vanishes as d goes to ∞. By Eq. (5.10),

$$
\begin{aligned}
\lim_{d\to\infty} \mathbb{E}\sqrt{\mathbf{\Lambda}_d} &\leq \lim_{d\to\infty} \mathbb{E}\, f_\alpha(\mathbf{\Lambda}_d) + \lim_{d\to\infty} g(\alpha, d) \\
&= \mathbb{E}\, f_\alpha(\mathbf{\Lambda}) \quad = \quad \mathbb{E}\sqrt{\mathbf{\Lambda}}\ .
\end{aligned}
$$

This proves the theorem. □

References

[Bai93] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. Part II. sample covariance matrices. *Annals of Probability*, 21(2):649–672, April 1993.

[BCWdW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Physical review letters*, 69(20):2881, 1992.

[BZ08] Zhidong Bai and Wang Zhou. Large sample covariance matrices without independence structures in columns. *Statistica Sinica*, 18(2):425–442, 2008.

[CGJV19] Andrea Coladangelo, Alex B. Grilo, Stacey Jeffery, and Thomas Vidick. Verifier-on-a-leash: New schemes for verifiable delegated quantum computation, with quasilinear resources. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2019*, volume 11478 of *Lecture Notes in Computer Science*, pages 247–277, Cham, 2019. Springer International Publishing.

[Cha20] Michael Charezma. Quantum circuit diagrams. https://warwick.ac.uk/fac/sci/physics/research/cfsa/people/pastmembers/charemzam/pastprojects, 2006 (accessed October 14, 2020).

[Cur79] Paul Joseph Curlander. *Quantum Limitations on Communication Systems*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 1979.

[FK19] Máté Farkas and Jędrzej Kaniewski. Self-testing mutually unbiased bases in the prepare-and-measure scenario. *Physical Review A*, 99(3):032316, 2019.

[FKN22] Máté Farkas, Jędrzej Kaniewski, and Ashwin Nayak. Mutually unbiased measurements, Hadamard matrices, and Superdense Coding. Technical Report arXiv:2204.11886 [quant-ph], ArXiv.org Preprint Archive, https://www.arxiv.org/, April 2022.

[Hal35] Philip Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.

[HJW93] Lane P. Hughston, Richard Jozsa, and William K. Wootters. A complete classification of quantum ensembles having a given density matrix. *Physics Letters A*, 183(1):14–18, 1993.

[JNV+20] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. $\mathsf{MIP}^* = \mathsf{RE}$. Technical Report arXiv:2001.04383 [quant-ph], ArXiv.org Preprint Archive, `https://www.arxiv.org/`, January 2020.

[Kho79] Alexander S. Kholevo. On asymptotically optimal hypothesis testing in quantum statistics. *Theory of Probability & Its Applications*, 23(2):411–415, 1979.

[KR03] Andreas Klappenecker and Martin Rötteler. Unitary error bases: Constructions, equivalence, and applications. In Marc Fossorier, Tom Høholdt, and Alain Poli, editors, *Proceedings of the 15th International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes (AAECC)*, volume 2643 of *Lecture Notes in Computer Science*, pages 139–149. Springer, Berlin / Heidelberg, Germany, May12–16, 2003.

[Mec19] Elizabeth S. Meckes. *The Random Matrix Theory of the Classical Compact Groups*, volume 218 of *Cambridge Tracts in Mathematics*. Cambridge University Press, July 2019.

[Mon07] Ashley Montanaro. On the distinguishability of random quantum states. *Communications in Mathematical Physics*, 273(3):619–636, August 2007.

[MV16] Benjamin Musto and Jamie Vicary. Quantum Latin squares and unitary error bases. *Quantum Information and Computation*, 16(15-16):1318–1332, November 2016.

[MY98] Dominic Mayers and Andrew Yao. Quantum cryptography with imperfect apparatus. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280)*, pages 503–509. IEEE, 1998.

[MY04] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *Quantum Information & Computation*, 4(4):273–286, July 2004.

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, NY, USA, 2011. 10th Anniversary Edition.

[NY20] Ashwin Nayak and Henry Yuen. Rigidity of superdense coding. Technical Report arXiv:2012.01672v1 [quant-ph], arXiv Pre-print server, `https://arxiv.org/abs/2012.01672`, December 2020.

[ON99] Tomohiro Ogawa and Hiroshi Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory*, 45(7):2486–2489, 1999.

[ŠB20] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review. *Quantum*, 4:337, 2020.

[Sch35] Erwin Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):555–563, 1935.

[TKV+18] Armin Tavakoli, Jędrzej Kaniewski, Tamás Vértesi, Denis Rosset, and Nicolas Brunner. Self-testing quantum states and measurements in the prepare-and-measure scenario. *Physical Review A*, 98(6):062307, 2018.

[Tys09a] Jon Tyson. Erratum: "Minimum-error quantum distinguishability bounds from matrix monotone functions: A comment on 'Two-sided estimates of minimum-error distinguishability of mixed quantum states via generalized Holevo-Curlander bounds'" [J. Math. Phys. 50, 062102 (2009)]. *Journal of Mathematical Physics*, 50(10):109902, 2009.

[Tys09b] Jon Tyson. Two-sided estimates of minimum-error distinguishability of mixed quantum states via generalized Holevo-Curlander bounds. *Journal of Mathematical Physics*, 50(3):032106, 2009.

[Tys10] Jon Tyson. Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on maximum overlap using directional iterates. *Journal of Mathematical Physics*, 51(9):092204, 2010.

[Uhl76] Armin Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, UK, 2018.

[VV19] Umesh Vazirani and Thomas Vidick. Fully device independent quantum key distribution. *Communications of the ACM*, 62(4):133–133, 2019.

[VW00] Karl Gerd H. Vollbrecht and Reinhard F. Werner. Why two qubits are special. *Journal of Mathematical Physics*, 41(10):6772–6782, 2000.

[Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, May 2018.

[Wer01] Reinhard F. Werner. All teleportation and dense coding schemes. *Journal of Physics A: Mathematical and General*, 34(35):7081–7094, August 2001.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, Cambridge, UK, 2013.

[Yas16] Pavel Yaskov. A short proof of the Marchenko–Pastur theorem. Une courte démonstration du théorème de Marchenko–Pastur. *Comptes Rendus Mathématique*, 354(3):319–322, March 2016.