




# Asymptotic Performance of Port-Based Teleportation

Matthias Christandl<sup>1</sup>, Felix Leditzky<sup>2,3</sup>, Christian Majenz<sup>4,5,7</sup> ,  
Graeme Smith<sup>2,3,6</sup>, Florian Speelman<sup>4,7,10</sup>, Michael Walter<sup>4,5,8,9</sup>

- <sup>1</sup> QMATH, Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark
- <sup>2</sup> JILA, University of Colorado/NIST, Boulder, USA. E-mail: [felix.leditzky@jila.colorado.edu](mailto:felix.leditzky@jila.colorado.edu)
- <sup>3</sup> Center for Theory of Quantum Matter, University of Colorado Boulder, Boulder, CO, USA
- <sup>4</sup> QuSoft, Amsterdam, The Netherlands. E-mail: [michael.walter@gmail.com](mailto:michael.walter@gmail.com); [c.majenz@uva.nl](mailto:c.majenz@uva.nl)
- <sup>5</sup> Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands
- <sup>6</sup> Department of Physics, University of Colorado Boulder, Boulder, CO, USA.
- <sup>7</sup> CWI, Amsterdam, The Netherlands
- <sup>8</sup> Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands
- <sup>9</sup> Institute for Theoretical Physics, University of Amsterdam, Amsterdam, The Netherlands
- <sup>10</sup> Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

Received: 29 March 2019 / Accepted: 23 June 2020  
© The Author(s) 2020

**Abstract:** Quantum teleportation is one of the fundamental building blocks of quantum Shannon theory. While ordinary teleportation is simple and efficient, port-based teleportation (PBT) enables applications such as universal programmable quantum processors, instantaneous non-local quantum computation and attacks on position-based quantum cryptography. In this work, we determine the fundamental limit on the performance of PBT: for arbitrary fixed input dimension and a large number  $N$  of ports, the error of the optimal protocol is proportional to the inverse square of  $N$ . We prove this by deriving an achievability bound, obtained by relating the corresponding optimization problem to the lowest Dirichlet eigenvalue of the Laplacian on the ordered simplex. We also give an improved converse bound of matching order in the number of ports. In addition, we determine the leading-order asymptotics of PBT variants defined in terms of maximally entangled resource states. The proofs of these results rely on connecting recently-derived representation-theoretic formulas to random matrix theory. Along the way, we refine a convergence result for the fluctuations of the Schur–Weyl distribution by Johansson, which might be of independent interest.

## Contents

1.	Introduction	.....
1.1	Quantum teleportation protocols	.....
1.2	Summary of main results	.....
1.3	Structure of this paper	.....
2.	Preliminaries	.....
2.1	Notation and definitions	.....
2.2	Representation theory of the symmetric and unitary group	.....
3.	Port-Based Teleportation	.....
3.1	Deterministic PBT	.....
3.2	Probabilistic PBT	.....

3.3	Symmetries	.....
3.4	Representation-theoretic characterization	.....
4.	The Schur–Weyl Distribution	.....
5.	Probabilistic PBT	.....
6.	Deterministic PBT	.....
6.1	Asymptotics of the standard protocol	.....
6.2	Asymptotics of the optimal protocol	.....
7.	Converse Bound	.....
8.	Conclusion	.....
A.	Proof of Lemma 3.6	.....
B.	A Family of Explicit Protocols for Deterministic PBT	.....
C.	The Maximal Eigenvalue of a $2 \times 2$ GUE <sub>0</sub> Matrix	.....
D.	Technical Lemmas	.....

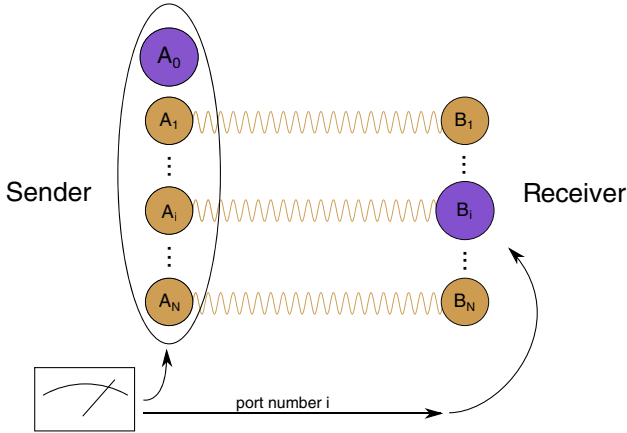
## 1. Introduction

*1.1. Quantum teleportation protocols.* Quantum teleportation [1] is a fundamental quantum information-processing task, and one of the hallmark features of quantum information theory: Two parties Alice and Bob may use a shared entangled quantum state together with classical communication to “teleport” an unknown quantum state from Alice to Bob. The original protocol in [1] consists of Alice measuring the unknown state together with her half of the shared entangled state and letting Bob know about the outcome of her measurement. Based on this information Bob can then manipulate his half of the shared state by applying a suitable correction operation, thus recovering the unknown state in his lab.

From an information-theoretic point of view, quantum teleportation implements a quantum channel between Alice and Bob. If the shared entangled state is a noiseless, maximally entangled state (a so-called EPR state, named after a famous paper by Einstein, Podolski, and Rosen [2]), then this quantum channel is in fact a perfect, noiseless channel. On the other hand, using a noisy entangled state as the shared resource in the teleportation protocol renders the effective quantum channel imperfect or noisy. A common way to measure the noise in a quantum channel is by means of the *entanglement fidelity*, which quantifies how well the channel preserves generic correlations with an inaccessible environment system (see Sect. 2.1 for a definition).

Port-based teleportation (PBT) [3,4] is a variant of the original quantum teleportation protocol [1], where the receiver’s correction operation consists of merely picking the right subsystem, called *port*, of their part of the entangled resource state. Figure 1 provides a schematic description of the protocol (see Sect. 3 for a more detailed explanation). While being far less efficient than the ordinary teleportation protocol, the simple correction operation allows the receiver to apply a quantum operation to the output of the protocol before receiving the classical message. This *simultaneous unitary covariance property* enables all known applications that require PBT instead of just ordinary quantum teleportation, including the construction of universal programmable quantum processors [3], quantum channel discrimination [5] and instantaneous non-local quantum computation (INQC) [6].

In the INQC protocol, which was devised by Beigi and König [6], two spatially separated parties share an input state and wish to perform a joint unitary on it. To do so, they are only allowed a single simultaneous round of communication. INQC provides a generic attack on any quantum position-verification scheme [7], a protocol in the field of



**Fig. 1.** Schematic representation of port-based teleportation (PBT). Like in ordinary teleportation, the sender applies a joint measurement to her input system  $A$  and her parts of the entangled resource,  $A_i$ ,  $i = 1, \dots, N$ , and sends the outcome to the receiver, who applies a correction operation. In PBT, however, this correction operation merely consists of choosing one of the subsystems  $B_i$ , the *ports*, of the entangled resource. A PBT protocol cannot implement a perfect quantum channel with a finite number of ports. There are different variants of PBT. The four commonly studied ones are characterized by whether failures are announced, or heralded (probabilistic PBT) or go unnoticed (deterministic PBT), and whether simplifying constraints on the resource state and the sender’s measurement are enforced

position-based cryptography [6, 8–10]. It is therefore of great interest for cryptography to characterize the resource requirements of INQC: it is still open whether a computationally secure quantum position-verification scheme exists, as all known generic attacks require an exponential amount of entanglement. Efficient protocols for INQC are only known for special cases [11–14]. The best lower bounds for the entanglement requirements of INQC are, however, linear in the input size [6, 15, 16], making the hardness of PBT, the corner stone of the best known protocol, the only indication for a possible hardness of INQC.

PBT comes in two variants, deterministic and probabilistic, the latter being distinguished from the former by the fact that the protocol implements a perfect quantum channel whenever it does not fail (errors are “heralded”). In addition, two classes of protocols have been considered in the literature, one using maximally entangled resource states and the other using more complex resources optimized for the protocol. The appeal of the former type of protocol is mostly due to the fact that maximally entangled states are a standard resource in quantum information processing and can be prepared efficiently. Using a protocol based on maximally entangled resources thus removes one parameter from the total complexity of the protocol, the complexity of preparing the resource, leaving the amount of resources as well as the complexity of the involved quantum measurement as the remaining two complexity contributions.

In their seminal work [3, 4], Ishizaka and Hiroshima completely characterize the problem of PBT for qubits. They calculate the performance of the standard<sup>1</sup> and optimized protocols for deterministic and the EPR and optimized protocols for probabilistic PBT, and prove the optimality of the ‘pretty good’ measurement in the standard deterministic case. They also show a lower bound on the performance of the standard protocol for

<sup>1</sup> The standard protocol uses a maximally entangled resource state and the so-called Pretty Good Measurement.

deterministic PBT, which was later reproven in [6]. Further properties of PBT were explored in [17], in particular with respect to recycling part of the resource state. Converse bounds for the probabilistic and deterministic versions of PBT have been proven in [18] and [19], respectively. In [20], exact formulas for the fidelity of the standard protocol for deterministic PBT with  $N = 3$  or 4 in arbitrary dimension are derived using a graphical algebra approach. Recently, exact formulas for arbitrary input dimension in terms of representation-theoretic data have been found for all four protocols, and the asymptotics of the optimized probabilistic case have been derived [21, 22].

Note that, in contrast to ordinary teleportation, a protocol obtained from executing several PBT protocols is not again a PBT protocol. This is due to the fact that the whole input system has to be teleported to the same output port for the protocol to have the mentioned simultaneous unitary covariance property. Therefore, the characterization of protocols for any dimension  $d$  is of particular interest. The mentioned representation-theoretic formulas derived in [21, 22] provide such a characterization. It is, however, not known how to evaluate these formulas efficiently for large input dimension.

*1.2. Summary of main results.* In this paper we provide several characterization results for port-based teleportation. As our main contributions, we characterize the leading-order asymptotic performance of fully optimized deterministic port-based teleportation (PBT), as well as the standard protocol for deterministic PBT and the EPR protocol for probabilistic PBT. In the following, we provide a detailed summary of our results. These results concern asymptotic characterizations of the entanglement fidelity of deterministic PBT, defined in Sect. 3.1, and the success probability of probabilistic PBT, defined in Sect. 3.2.

Our first, and most fundamental, result concerns deterministic PBT and characterizes the leading-order asymptotics of the optimal fidelity for a large number of ports.

**Theorem 1.1.** *For arbitrary but fixed local dimension  $d$ , the optimal entanglement fidelity for deterministic port-based teleportation behaves asymptotically as*

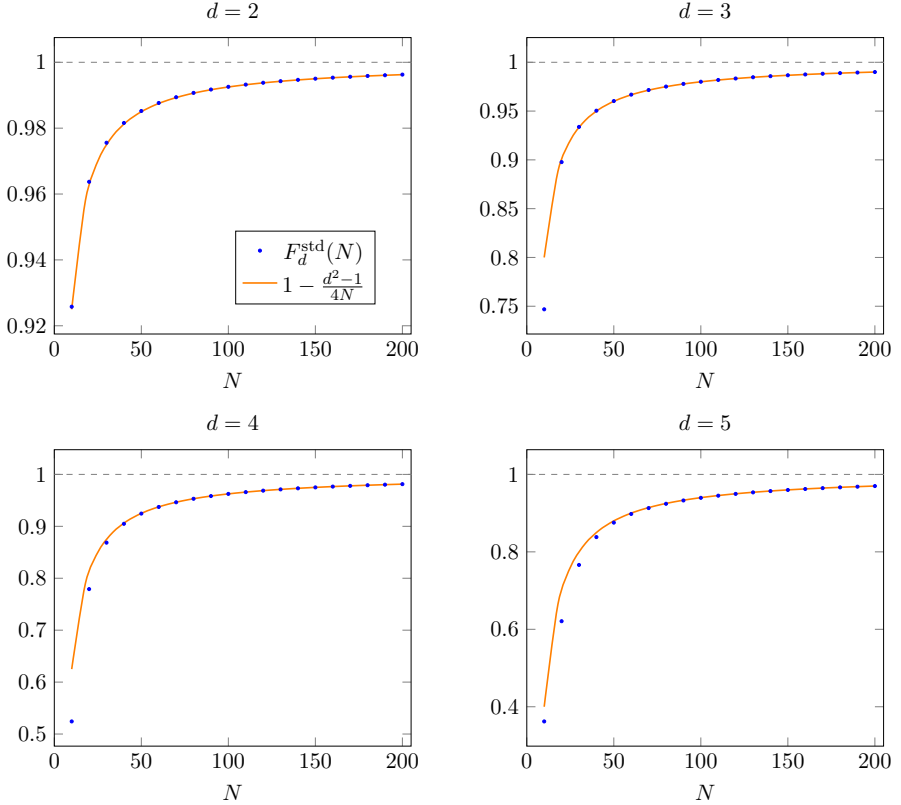
$$F_d^*(N) = 1 - \Theta(N^{-2}).$$

Theorem 1.1 is a direct consequence of Theorem 1.5 below. Prior to our work, it was only known that  $F_d^*(N) = 1 - \Omega(N^{-2})$  as a consequence of an explicit converse bound [19]. We prove that this asymptotic scaling is in fact achievable, and we also provide a converse with improved dependency on the local dimension, see Corollary 1.6.

For deterministic port-based teleportation using a maximally entangled resource and the pretty good measurement, a closed expression for the entanglement fidelity was derived in [21], but its asymptotics for fixed  $d > 2$  and large  $N$  remained undetermined. As our second result, we derive the asymptotics of deterministic port-based teleportation using a maximally entangled resource and the pretty good measurement, which we call the *standard protocol*.

**Theorem 1.2.** *For arbitrary but fixed  $d$  and any  $\delta > 0$ , the entanglement fidelity of the standard protocol of PBT is given by*

$$F_d^{\text{std}}(N) = 1 - \frac{d^2 - 1}{4N} + O(N^{-\frac{3}{2} + \delta}).$$



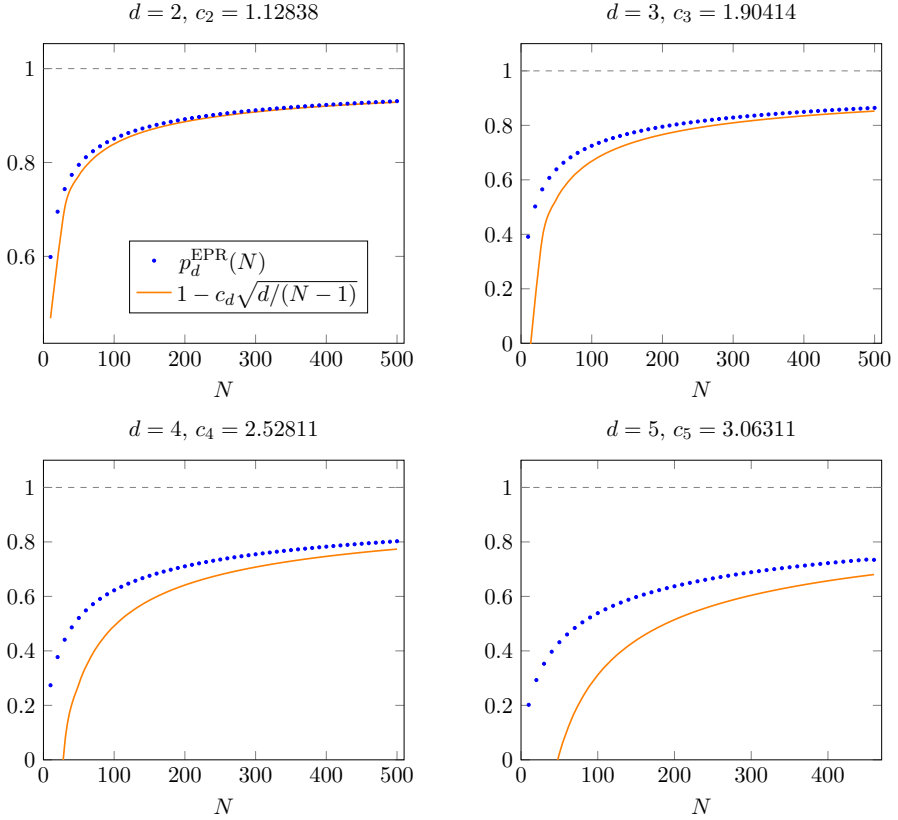
**Fig. 2.** Entanglement fidelity of the standard protocol for deterministic port-based teleportation in local dimension  $d = 2, 3, 4, 5$  using  $N$  ports [23]. We compare the exact formula (3.5) for  $F_d^{\text{std}}$  (blue dots) with the first-order asymptotics obtained from Theorem 1.2 (orange curve)

Previously, the asymptotic behavior given in the above theorem was only known for  $d = 2$  in terms of an exact formula for finite  $N$ ; for  $d > 2$ , it was merely known that  $F_d^{\text{std}}(N) = 1 - O(N^{-1})$  [4]. In Fig. 2 we compare the asymptotic formula of Theorem 1.2 to a numerical evaluation of the exact formula derived in [21] for  $d \leq 5$ .

For probabilistic port-based teleportation, Mozrymas et al. [22] obtained the following expression for the success probability  $p_d^*$  optimized over arbitrary entangled resources:

$$p_d^*(N) = 1 - \frac{d^2 - 1}{d^2 - 1 + N},$$

valid for all values of  $d$  and  $N$  (see the detailed discussion in Sect. 3). In the case of using  $N$  maximally entangled states as the entangled resource, an exact expression for the success probability in terms of representation-theoretic quantities was also derived in [21]. We state this expression in (3.9) in Sect. 3. However, its asymptotics for fixed  $d > 2$  and large  $N$  have remained undetermined to date. As our third result, we derive the following expression for the asymptotics of the success probability of the optimal protocol among the ones that use a maximally entangled resource, which we call the *EPR protocol*.



**Fig. 3.** Success probability of the EPR protocol for probabilistic port-based teleportation in local dimension  $d = 2, 3, 4, 5$  using  $N$  ports [23]. We compare the exact formula (3.9) for  $p_d^{\text{EPR}}(N)$  (blue dots) with the first-order asymptotic formula obtained from Theorem 1.3 (orange curve). The first-order coefficient  $c_d \equiv \mathbb{E}[\lambda_{\max}(\mathbf{G})]$  appearing in the formula in Theorem 1.3 was obtained by numerical integration from the eigenvalue distribution of  $\text{GUE}_d$

**Theorem 1.3.** *For probabilistic port-based teleportation in arbitrary but fixed dimension  $d$  with EPR pairs as resource states,*

$$p_d^{\text{EPR}}(N) = 1 - \sqrt{\frac{d}{N-1}} \mathbb{E}[\lambda_{\max}(\mathbf{G})] + o\left(N^{-1/2}\right),$$

where  $\mathbf{G} \sim \text{GUE}_d^0$ .

The famous Wigner semicircle law [24] provides an asymptotic expression for the expected maximal eigenvalue,  $\mathbb{E}[\lambda_{\max}(\mathbf{G})] \sim 2\sqrt{d}$  for  $d \rightarrow \infty$ . Additionally, there exist explicit upper and lower bounds for all  $d$ , see the discussion in Sect. 5.

To establish Theorems 1.2 and 1.3, we analyze the asymptotics of the Schur–Weyl distribution, which also features in other fundamental problems of quantum information theory including spectrum estimation, tomography, and the quantum marginal problem [25–35]. Our main technical contribution is a new convergence result for its fluctuations that strengthens a previous result by Johansson [36]. This result, which may be of independent interest, is stated as Theorem 4.1 in Sect. 4.

Theorem 1.1 is proved by giving an asymptotic lower bound for the optimal fidelity of deterministic PBT, as well as an upper bound that is valid for any number of ports and matches the lower bound asymptotically. For the lower bound, we again use an expression for the entanglement fidelity of the optimal deterministic PBT protocol derived in [22]. The asymptotics of this formula for fixed  $d$  and large  $N$  have remained undetermined so far. We prove an asymptotic lower bound for this entanglement fidelity in terms of the lowest Dirichlet eigenvalue of the Laplacian on the ordered  $(d - 1)$ -dimensional simplex.

**Theorem 1.4.** *The optimal fidelity for deterministic port-based teleportation is bounded from below by*

$$F_d^*(N) \geq 1 - \frac{\lambda_1(\text{OS}_{d-1})}{dN^2} - O(N^{-3}),$$

where

$$\text{OS}_{d-1} = \left\{ x \in \mathbb{R}^d \mid \sum_i x_i = 1, x_i \geq x_{i+1}, x_d \geq 0 \right\}$$

is the  $(d - 1)$ -dimensional simplex of ordered probability distributions with  $d$  outcomes and  $\lambda_1(\Omega)$  is the first eigenvalue of the Dirichlet Laplacian on a domain  $\Omega$ .

Using a bound from [37] for  $\lambda_1(\text{OS}_d)$ , we obtain the following explicit lower bound.

**Theorem 1.5.** *For the optimal fidelity of port-based teleportation with arbitrary but fixed input dimension  $d$  and  $N$  ports, the following bound holds,*

$$F_d^*(N) \geq 1 - \frac{d^5 + O(d^{9/2})}{4\sqrt{2}N^2} + O(N^{-3}).$$

As a complementary result, we give a strong upper bound for the entanglement fidelity of any deterministic port-based teleportation protocol. While valid for any finite number  $N$  of ports, its asymptotics for large  $N$  are given by  $1 - O(N^{-2})$ , matching Theorem 1.5.

**Corollary 1.6.** *For a general port-based teleportation scheme with input dimension  $d$  and  $N$  ports, the entanglement fidelity  $F_d^*$  and the diamond norm error  $\varepsilon_d^*$  can be bounded as*

$$F_d^*(N) \leq \begin{cases} \frac{\sqrt{N}}{d} & \text{if } N \leq \frac{d^2}{2} \\ 1 - \frac{d^2-1}{16N^2} & \text{otherwise} \end{cases} \quad \varepsilon_d^*(N) \geq \begin{cases} 2\left(1 - \frac{\sqrt{N}}{d}\right) & \text{if } N \leq \frac{d^2}{2} \\ 2\frac{d^2-1}{16N^2} & \text{otherwise.} \end{cases}$$

Previously, the best known upper bound on the fidelity [19] had the same dependence on  $N$ , but was increasing in  $d$ , thus failing to reflect the fact that the task becomes harder with increasing  $d$ . Interestingly, a lower bound from [38] on the program register size of a universal programmable quantum processor also yields a converse bound for PBT that is incomparable to the one from [19] and weaker than our bound.

Finally we provide a proof of the following ‘folklore’ fact that had been used in previous works on port-based teleportation. The unitary and permutation symmetries of port-based teleportation imply that the entangled resource state and Alice’s measurement can be chosen to have these symmetries as well. Apart from simplifying the optimization over resource states and POVMs, this implies that characterizing the entanglement fidelity is sufficient to give worst-case error guarantees. Importantly, this retrospectively

justifies the use of the entanglement fidelity  $F$  in the literature about deterministic port-based teleportation in the sense that any bound on  $F$  implies a bound on the diamond norm error without losing dimension factors. This is also used to show the diamond norm statement of Corollary 1.6.

**Proposition 1.7** (Proposition 3.4 and 3.3 and Corollary 3.5, informal). *There is an explicit transformation between port-based teleportation protocols that preserves any unitarily invariant distance measure on quantum channels, and maps an arbitrary port-based teleportation protocol with input dimension  $d$  and  $N$  ports to a protocol that*

- (i) *has a resource state and a POVM with  $U(d) \times S_N$  symmetry, and*
- (ii) *implements a unitarily covariant channel.*

*In particular, the transformation maps an arbitrary port-based teleportation protocol to one with the symmetries (i) and (ii) above, and entanglement fidelity no worse than the original protocol. Point (ii) implies that*

$$\varepsilon_d^* = 2(1 - F_d^*),$$

*where  $F_d^*$  and  $\varepsilon_d^*$  denote the optimal entanglement fidelity and optimal diamond norm error for deterministic port-based teleportation.*

*1.3. Structure of this paper.* In Sect. 2 we fix our notation and conventions and recall some basic facts about the representation theory of the symmetric and unitary groups. In Sect. 3 we define the task of port-based teleportation (PBT) in its two main variants, the probabilistic and deterministic setting. Moreover, we identify the inherent symmetries of PBT, and describe a representation-theoretic characterization of the task. In Sect. 4 we discuss the Schur–Weyl distribution and prove a convergence result that will be needed to establish our results for PBT with maximally entangled resources. Our first main result is proved in Sect. 5, where we discuss the probabilistic setting in arbitrary dimension using EPR pairs as ports, and determine the asymptotics of the success probability  $p_d^{\text{EPR}}$  (Theorem 1.3). Our second main result, derived in Sect. 6.1, concerns the deterministic setting in arbitrary dimension using EPR pairs, for which we compute the asymptotics of the optimal entanglement fidelity  $F_d^{\text{std}}$  (Theorem 1.2). Our third result, an asymptotic lower bound on the entanglement fidelity  $F_d^*$  of the optimal protocol in the deterministic setting (Theorem 1.5), is proved in Sect. 6.2. Finally, in Sect. 7 we derive a general non-asymptotic converse bound on deterministic port-based teleportation protocols using a non-signaling argument (Theorem 7.5). We also present a lower bound on the communication requirements for approximate quantum teleportation (Corollary 7.4). We make some concluding remarks in Sect. 8. The appendices contain technical proofs.

## 2. Preliminaries

*2.1. Notation and definitions.* We denote by  $A, B, \dots$  quantum systems with associated Hilbert spaces  $\mathcal{H}_A, \mathcal{H}_B, \dots$ , which we always take to be finite-dimensional, and we associate to a multipartite quantum system  $A_1 \dots A_n$  the Hilbert space  $\mathcal{H}_{A_1 \dots A_n} = \mathcal{H}_{A_1} \otimes \dots \otimes \mathcal{H}_{A_n}$ . When the  $A_i$  are identical, we also write  $A^n = A_1 \dots A_n$ . The set of linear operators on a Hilbert space  $\mathcal{H}$  is denoted by  $\mathcal{B}(\mathcal{H})$ . A *quantum state*  $\rho_A$  on quantum system  $A$  is a positive semidefinite linear operator  $\rho_A \in \mathcal{B}(\mathcal{H}_A)$  with unit trace, i.e.,  $\rho_A \geq 0$  and  $\text{tr}(\rho_A) = 1$ . We denote by  $I_A$  or  $1_A$  the identity operator on



$\mathcal{H}_A$ , and by  $\tau_A = I_A/|A|$  the corresponding *maximally mixed* quantum state, where  $|A| := \dim \mathcal{H}_A$ . A pure quantum state  $\psi_A$  is a quantum state of rank one. We can write  $\psi_A = |\psi\rangle\langle\psi|_A$  for a unit vector  $|\psi\rangle_A \in \mathcal{H}_A$ . For quantum systems  $A, A'$  of dimension  $\dim \mathcal{H}_A = \dim \mathcal{H}_{A'} = d$  with bases  $\{|i\rangle_A\}_{i=1}^d$  and  $\{|i\rangle_{A'}\}_{i=1}^d$ , the vector  $|\phi^+\rangle_{A'A} = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle_{A'} \otimes |i\rangle_A$  defines the *maximally entangled* state of Schmidt rank  $d$ . The *fidelity*  $F(\rho, \sigma)$  between two quantum states is defined by  $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$ , where  $\|X\|_1 = \text{tr}(\sqrt{X^\dagger X})$  denotes the *trace norm* of an operator. For two pure states  $|\psi\rangle$  and  $|\phi\rangle$ , the fidelity is equal to  $F(\psi, \phi) = |\langle\psi|\phi\rangle|^2$ . A *quantum channel* is a completely positive, trace-preserving linear map  $\Lambda: \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ . We also use the notation  $\Lambda: A \rightarrow B$  or  $\Lambda_{A \rightarrow B}$ , and we denote by  $\text{id}_A$  the identity channel on  $A$ . Given two quantum channels  $\Lambda_1, \Lambda_2: A \rightarrow B$ , the *entanglement fidelity*  $F(\Lambda_1, \Lambda_2)$  is defined as

$$F(\Lambda_1, \Lambda_2) := F((\text{id}_{A'} \otimes \Lambda_1)(\phi_{A'A}^+), (\text{id}_{A'} \otimes \Lambda_2)(\phi_{A'A}^+)),$$

and we abbreviate  $F(\Lambda) := F(\Lambda, \text{id})$ . The *diamond norm* of a linear map  $\Lambda: \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is defined by

$$\|\Lambda\|_\diamond := \sup_{\|X_{A'A}\|_1 \leq 1} \|(\text{id}_{A'} \otimes \Lambda)(X_{A'A})\|_1.$$

The induced distance on quantum channels is called the *diamond distance*. A *positive operator-valued measure (POVM)*  $E = \{E_x\}$  on a quantum system  $A$  is a collection of positive semidefinite operators  $E_x \geq 0$  satisfying  $\sum_x E_x = I_A$ .

We denote random variables by bold letters ( $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \dots$ ) and the valued they take by the non-bold versions ( $X, Y, Z, \dots$ ). We denote by  $\mathbf{X} \sim \mathbb{P}$  that  $\mathbf{X}$  is a random variable with probability distribution  $\mathbb{P}$ . We write  $\Pr(\dots)$  for the probability of an event and  $\mathbb{E}[\dots]$  for expectation values. The notation  $\mathbf{X}_n \xrightarrow{P} \mathbf{X}$  ( $n \rightarrow \infty$ ) denotes *convergence in probability* and  $\mathbf{X}_n \xrightarrow{D} \mathbf{X}$  ( $n \rightarrow \infty$ ) denotes *convergence in distribution*. The latter can be defined, e.g., by demanding that  $\mathbb{E}[f(\mathbf{X}_n)] \rightarrow \mathbb{E}[f(\mathbf{X})]$  ( $n \rightarrow \infty$ ) for every continuous, bounded function  $f$ . The *Gaussian unitary ensemble*  $\text{GUE}_d$  is the probability distribution on the set of Hermitian  $d \times d$ -matrices  $H$  with density  $Z_d^{-1} \exp(-\frac{1}{2} \text{tr} H^2)$ , where  $Z_d$  is the appropriate normalization constant. Alternatively, for  $\mathbf{X} \sim \text{GUE}_d$ , the entries  $\mathbf{X}_{ii}$  for  $1 \leq i \leq d$  are independently distributed as  $\mathbf{X}_{ii} \sim N(0, 1)$ , whereas the elements  $\mathbf{X}_{ij}$  for  $1 \leq i < j \leq d$  are independently distributed as  $\mathbf{X}_{ij} \sim N(0, \frac{1}{2}) + iN(0, \frac{1}{2})$ . Here,  $N(0, \sigma^2)$  denotes the centered normal distribution with variance  $\sigma^2$ . The *traceless Gaussian unitary ensemble*  $\text{GUE}_d^0$  can be defined as the distribution of the random variable  $\mathbf{Y} := \mathbf{X} - \frac{\text{tr} \mathbf{X}}{d} I$ , where  $\mathbf{X} \sim \text{GUE}_d$ .

For a complex number  $z \in \mathbb{C}$ , we denote by  $\Re(z)$  and  $\Im(z)$  its real and imaginary part, respectively. We denote by  $\mu \vdash n$  a partition  $(\mu_1, \dots, \mu_d)$  of  $n$  into  $d$  parts. That is,  $\mu \in \mathbb{Z}^d$  with  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_d \geq 0$  and  $\sum_i \mu_i = n$ . We also call  $\mu$  a *Young diagram* and visualize it as an arrangement of boxes, with  $\mu_i$  boxes in the  $i$ th row. For example,  $\mu = (3, 1)$  can be visualized as  $\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array}$ . We use the notation  $(i, j) \in \mu$  to mean that  $(i, j)$  is a box in the Young diagram  $\mu$ , that is,  $1 \leq i \leq d$  and  $1 \leq j \leq \mu_i$ . We denote by  $\text{GL}(\mathcal{H})$  the general linear group and by  $U(\mathcal{H})$  the unitary group acting on a Hilbert space  $\mathcal{H}$ . When  $\mathcal{H} = \mathbb{C}^d$ , we write  $\text{GL}(d)$  and  $U(d)$ . Furthermore, we denote by  $S_n$  the symmetric group on  $n$  symbols. A *representation*  $\varphi$  of a group  $G$  on a vector space  $\mathcal{H}$  is a map  $G \ni g \mapsto \varphi(g) \in \text{GL}(\mathcal{H})$  satisfying  $\varphi(gh) = \varphi(g)\varphi(h)$  for all  $g, h \in G$ . In this paper all representations are unitary, which means that  $\mathcal{H}$  is a Hilbert space and  $\varphi(g) \in U(\mathcal{H})$  for every  $g \in G$ . A representation is *irreducible* (or an *irrep*) if  $\mathcal{H}$  contains no nontrivial invariant subspace.

*2.2. Representation theory of the symmetric and unitary group.* Our results rely on the representation theory of the symmetric and unitary groups and Schur–Weyl duality (as well as their semiclassical asymptotics which we discuss in Sect. 4). In this section we introduce the relevant concepts and results (see e.g., [39,40]).

The irreducible representations of  $S_n$  are known as *Specht modules* and labeled by Young diagrams with  $n$  boxes. We denote the Specht module of  $S_n$  corresponding to a Young diagram  $\mu \vdash_d n$  by  $[\mu]$  ( $d$  is arbitrary). Its dimension is given by the *hook length formula* [39, pp. 53–54],

$$d_\mu = \frac{n!}{\prod_{(i,j) \in \mu} h_\mu(i,j)}, \quad (2.1)$$

where  $h_\mu(i,j)$  is the *hook length* of the hook with corner at the box  $(i,j)$ , i.e., the number of boxes below  $(i,j)$  plus the number of boxes to the right of  $(i,j)$  plus one (the box itself).

The polynomial irreducible representations of  $U(d)$  are known as *Weyl modules* and labeled by Young diagrams with no more than  $d$  rows. We denote the Weyl module of  $U(d)$  corresponding to a Young diagram  $\mu \vdash_d n$  by  $V_\mu^d$  ( $n$  is arbitrary). Its dimension can be computed using *Stanley’s hook length formula* [39, p. 55],

$$m_{d,\mu} = \prod_{(i,j) \in \mu} \frac{d + c(i,j)}{h_\mu(i,j)}, \quad (2.2)$$

where  $c(i,j) = j - i$  is the so-called *content* of the box  $(i,j)$ . This is an alternative to the *Weyl dimension formula*, which states that

$$m_{d,\mu} = \prod_{1 \leq i < j \leq d} \frac{\mu_i - \mu_j + j - i}{j - i}. \quad (2.3)$$

We stress that  $m_{d,\mu}$  depends on the dimension  $d$ .

Consider the representations of  $S_n$  and  $U(d)$  on  $(\mathbb{C}^d)^{\otimes n}$  given by permuting the tensor factors, and multiplication by  $U^{\otimes n}$ , respectively. Clearly the two actions commute. *Schur–Weyl duality* asserts that the decomposition of  $(\mathbb{C}^d)^{\otimes n}$  into irreps takes the form (see, e.g., [40])

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\mu \vdash_d n} [\mu] \otimes V_\mu^d. \quad (2.4)$$

### 3. Port-Based Teleportation

The original quantum teleportation protocol for qubits (henceforth referred to as ordinary teleportation protocol) is broadly described as follows [1]: Alice (the sender) and Bob (the receiver) share an EPR pair (a maximally entangled state on two qubits), and their goal is to transfer or ‘teleport’ another qubit in Alice’s possession to Bob by sending only classical information. Alice first performs a joint Bell measurement on the quantum system to be teleported and her share of the EPR pair, and communicates the classical measurement outcome to Bob using two bits of classical communication. Conditioned on this classical message, Bob then executes a correction operation consisting of one of the Pauli operators on his share of the EPR pair. After the correction operation, he

has successfully received Alice’s state. The ordinary teleportation protocol can readily be generalized to qudits, i.e.,  $d$ -dimensional quantum systems. Note that while the term ‘EPR pair’ is usually reserved for a maximally entangled state on two qubits ( $d = 2$ ), we use the term more freely for maximally entangled states of Schmidt rank  $d$  on two qudits, as defined in Sect. 2.

Port-based teleportation, introduced by Ishizaka and Hiroshima [3, 4], is a variant of quantum teleportation where Bob’s correction operation solely consists of picking one of a number of quantum subsystems upon receiving the classical message from Alice. In more detail, Alice and Bob initially share an entangled resource quantum state  $\psi_{A^N B^N}$ , where  $\mathcal{H}_{A_i} \cong \mathcal{H}_{B_i} \cong \mathbb{C}^d$  for  $i = 1, \dots, N$ . We may always assume that the resource state is pure, for we can give a purification to Alice and she can choose not to use it.<sup>2</sup> Bob’s quantum systems  $B_i$  are called *ports*. Just like in ordinary teleportation, the goal is for Alice to teleport a  $d$ -dimensional quantum system  $A_0$  to Bob. To achieve this, Alice performs a joint POVM  $\{(E_i)_{A_0 A^N}\}_{i=1}^N$  on the input and her part of the resource state and sends the outcome  $i$  to Bob. Based on the index  $i$  he receives, Bob selects the  $i$ th port, i.e. the system  $B_i$ , as being the output register (renaming it to  $B_0$ ), and discards the rest. That is, in contrast to ordinary teleportation, Bob’s decoding operation solely consists of selecting the correct port  $B_i$ . The quality of the teleportation protocol is measured by how well it simulates the identity channel from Alice’s input register  $A_0$  to Bob’s output register  $B_0$ .

Port-based teleportation is impossible to achieve perfectly with finite resources [3], a fact first deduced from the application to universal programmable quantum processors [41]. There are two ways to deal with this fact: either one can just accept an imperfect protocol, or one can insist on simulating a perfect identity channel, with the caveat that the protocol will fail from time to time. This leads to two variants of PBT, which are called *deterministic* and *probabilistic* PBT in the literature [3].<sup>3</sup>

**3.1. Deterministic PBT.** A protocol for deterministic PBT proceeds as described above, implementing an imperfect simulation of the identity channel whose merit is quantified by the entanglement fidelity  $F_d$  or the diamond norm error  $\varepsilon_d$ . We denote by  $F_d^*(N)$  and  $\varepsilon_d^*(N)$  the maximal entanglement fidelity and the minimal diamond norm error for deterministic PBT, respectively, where both the resource state and the POVM are optimized. We will often refer to this as the *fully optimized* case.

Let  $\psi_{A^N B^N}$  be the entangled resource state used for a PBT protocol. When using the entanglement fidelity as a figure of merit, it is shown in [4] that the problem of PBT for the fixed resource state  $\psi_{A^N B^N}$  is equivalent to the state discrimination problem given by the collection of states

$$\eta_{A^N B_0}^{(i)} = \text{id}_{B_i \rightarrow B_0} \text{tr}_{B_i^c} |\psi\rangle\langle\psi|_{A^N B^N}, \quad i = 1, \dots, N. \quad (3.1)$$

with uniform prior (here we trace over all  $B$  systems but  $B_i$ , which is relabeled to  $B_0$ ). More precisely, the success probability  $q$  for state discrimination with some fixed POVM  $\{E_i\}_{i=1}^N$  and the entanglement fidelity  $F_d$  of the PBT protocol with Alice’s POVM equal

<sup>2</sup> This can even be done while keeping the structure of the resource state if desired, by adding the purifying system to one of Alice’s systems  $A_i$ , or splitting it between all of them.

<sup>3</sup> Alternatively, one could call “deterministic PBT” just “PBT” and for “probabilistic PBT” use the term “heralded PBT”, which is borrowed from quantum optics terminology as used in, e.g., [42]. However, we will stick to the widely used terms.

to  $\{E_i\}_{i=1}^N$ , but acting on  $A^N A_0$ , are related by the equation  $q = \frac{d^2}{N} F_d$ . This link with state discrimination provides us with the machinery developed for state discrimination to optimize the POVM. In particular, it suggests the use of the pretty good measurement [43, 44].

As in ordinary teleportation, it is natural to consider PBT protocols where the resource state is fixed to be  $N$  maximally entangled states (or EPR pairs) of local dimension  $d$ . This is because EPR pairs are a standard resource in quantum information theory that can easily be produced in a laboratory. We will denote by  $F_d^{\text{EPR}}(N)$  the optimal entanglement fidelity for any protocol for deterministic PBT that uses maximally entangled resource states. A particular protocol is given by combining maximally entangled resource states with the pretty good measurement (PGM) POVM [43,44]. We call this the *standard protocol* for deterministic PBT and denote the corresponding entanglement fidelity by  $F_d^{\text{std}}(N)$ . For qubits ( $d = 2$ ), the pretty good measurement was shown to be optimal for maximally entangled resource states [4]:

$$F_2^{\text{std}}(N) = F_2^{\text{EPR}}(N) = 1 - \frac{3}{4N} + o(1/N). \quad (3.2)$$

According to [22], the PGM is optimal in this situation for  $d > 2$  as well.

In [3] it is shown that the entanglement fidelity  $F_d^{\text{std}}$  for the standard protocol is at least

$$F_d^{\text{std}}(N) \geq 1 - \frac{d^2 - 1}{N}. \quad (3.3)$$

Beigi and König [6] rederived the same bound with different techniques. In [19], a converse bound is provided in the fully optimized setting:

$$F_d^*(N) \leq 1 - \frac{1}{4(d-1)N^2} + O(N^{-3}). \quad (3.4)$$

Note that the dimension  $d$  is part of the denominator instead of the numerator as one might expect in the asymptotic setting. Thus, the bound lacks the right qualitative behavior for large values of  $d$ . A different, incomparable, bound can be obtained from a recent lower bound on the program register dimension of a universal programmable quantum processor obtained by Kubicki et al. [38],

$$\varepsilon_d^*(N) \geq 2 \left( 1 - c \frac{\log d}{d} \left( 2N + \frac{2}{3} \right) \right),$$

where  $c$  is a constant. By Proposition 1.7, this bound is equivalent to

$$F_d^*(N) \leq c \frac{\log d}{d} \left( 2N + \frac{2}{3} \right).$$

Earlier works on programmable quantum processors [45,46] also yield (weaker) converse bounds for PBT.

Interestingly, and of direct relevance to our work, *exact* formulas for the entanglement fidelity have been derived both for the standard protocol and in the fully optimized case. In [21], the authors showed that

$$F_d^{\text{std}}(N) = d^{-N-2} \sum_{\alpha \vdash_{d} N-1} \left( \sum_{\mu=\alpha+\square} \sqrt{d_{\mu} m_{d,\mu}} \right)^2. \quad (3.5)$$

Here, the inner sum is taken over all Young diagrams  $\mu$  that can be obtained by adding one box to a Young diagram  $\alpha \vdash_d N - 1$ , i.e., a Young diagram with  $N - 1$  boxes and at most  $d$  rows. Equation (3.5) generalizes the result of [4] for  $d = 2$ , whose asymptotic behavior is stated in Eq. (3.2).

In the fully optimized case, Mozrzykas et al. [22] obtained a formula similar to Eq. (3.5) in which the dimension  $d_\mu m_{d,\mu}$  of the  $\mu$ -isotypic component in the Schur–Weyl decomposition is weighted by a coefficient  $c_\mu$  that is optimized over all probability densities with respect to the Schur–Weyl distribution (defined in Sect. 4). More precisely,

$$F_d^*(N) = d^{-N-2} \max_{c_\mu} \sum_{\alpha \vdash_d N-1} \left( \sum_{\mu=\alpha+\square} \sqrt{c_\mu d_\mu m_{d,\mu}} \right)^2, \quad (3.6)$$

where the optimization is over all nonnegative coefficients  $\{c_\mu\}$  such that  $\sum_{\mu \vdash_d N} c_\mu \frac{d_\mu m_{d,\mu}}{d^N} = 1$ .

**3.2. Probabilistic PBT.** In the task of probabilistic PBT, Alice’s POVM has an additional outcome that indicates the failure of the protocol and occurs with probability  $1 - p_d$ . For all other outcomes, the protocol is required to simulate the identity channel perfectly. We call  $p_d$  the probability of success of the protocol. As before, we denote by  $p_d^*(N)$  the maximal probability of success for probabilistic PBT using  $N$  ports of local dimension  $d$ , where the resource state as well as the POVM are optimized.

Based on the no-signaling principle and a version of the no-cloning theorem, Pitalúa-García [18] showed that the success probability  $p_{2^n}^*(N)$  of teleporting an  $n$ -qubit input state using a general probabilistic PBT protocol is at most

$$p_{2^n}^*(N) \leq 1 - \frac{4^n - 1}{4^n - 1 + N}. \quad (3.7)$$

Subsequently, Mozrzykas et al. [22] showed for a general  $d$ -dimensional input state that the converse bound in (3.7) is also achievable, establishing that

$$p_d^*(N) = 1 - \frac{d^2 - 1}{d^2 - 1 + N}. \quad (3.8)$$

This fully resolves the problem of determining the optimal probability of success for probabilistic PBT in the fully optimized setting.

As discussed above, it is natural to also consider the scenario where the resource state is fixed to be  $N$  maximally entangled states of rank  $d$  and consider the optimal POVM given that resource state. We denote by  $p_d^{\text{EPR}}$  the corresponding probability of success. We use the superscript EPR to keep the analogy with the case of deterministic PBT, as the measurement is optimized for the given resource state and no simplified measurement like the PGM is used. In [4], it was shown for qubits ( $d = 2$ ) that

$$p_2^{\text{EPR}}(N) = 1 - \sqrt{\frac{8}{\pi N}} + o(1/\sqrt{N}).$$

For arbitrary input dimension  $d$ , Studziński et al. [21] proved the exact formula

$$p_d^{\text{EPR}}(N) = \frac{1}{d^N} \sum_{\alpha \vdash N-1} m_{d,\alpha}^2 \frac{d_{\mu^*}}{m_{d,\mu^*}}, \quad (3.9)$$

where  $\mu^*$  is the Young diagram obtained from  $\alpha$  by adding one box in such a way that

$$\gamma_\mu(\alpha) = N \frac{m_{d,\mu} d_\alpha}{m_{d,\alpha} d_\mu} \quad (3.10)$$

is maximized (as a function of  $\mu$ ).

Finally, we note that any protocol for probabilistic PBT with success probability  $p_d$  can be converted into a protocol for deterministic PBT by sending over a random port index to Bob whenever Alice's measurement outcome indicates an error. The entanglement fidelity of the resulting protocol can be bounded as  $F_d \geq p_d + \frac{1-p_d}{d^2}$ . When applied to the fully optimized protocol corresponding to Eq. (3.8), this yields a protocol for deterministic PBT with better entanglement fidelity than the standard protocol for deterministic PBT. It uses, however, an optimized resource state that might be difficult to produce, while the standard protocol uses  $N$  maximally entangled states.

**3.3. Symmetries.** The problem of port-based teleportation has several natural symmetries that can be exploited. Intuitively, we might expect a  $U(d)$ -symmetry and a permutation symmetry, since our figures of merit are unitarily invariant and insensitive to the choice of port that Bob has to select. For the resource state, we might expect an  $S_N$ -symmetry, while the POVM elements have a marked port, leaving a possible  $S_{N-1}$ -symmetry among the non-marked ports. This section is dedicated to making these intuitions precise.

The implications of the symmetries have been known for some time in the community and used in other works on port-based teleportation (e.g. in [22]). We provide a formal treatment here for the convenience of the interested reader as well as to highlight the fact that the unitary symmetry allows us to directly relate the entanglement fidelity (which a priori quantifies an average error) to the diamond norm error (a worst case figure of merit). This relation is proved in Corollary 3.5.

While the sub-structure of the resource state on Alice's side in terms of  $N$  subsystems is natural from a mathematical point of view, it does not correspond to an operational feature of the task of PBT. This is in contrast to the port sub-structure on Bob's side, in terms of which the port-based condition on the teleportation protocol is defined. In the following, it will be convenient to allow resource states for PBT to have an arbitrary sub-structure on Alice's side.

We begin with a lemma on purifications of quantum states with a given group symmetry (see [47, 48] and [49, Lemma 5.5]):

**Lemma 3.1.** *Let  $\rho_A$  be a quantum state invariant under a unitary representation  $\varphi$  of a group  $G$ , i.e.,  $[\rho_A, \varphi(g)] = 0$  for all  $g \in G$ . Then there exists a purification  $|\rho\rangle_{AA'}$  such that  $(\varphi(g) \otimes \varphi^*(g))|\rho\rangle_{AA'} = |\rho\rangle_{AA'}$  for all  $g \in G$ . Here,  $\varphi^*$  is the dual representation of  $\varphi$ , which can be written as  $\varphi^*(g) = \varphi(g)$ .*

Starting from an arbitrary port-based teleportation protocol, it is easy to construct a modified protocol that uses a resource state such that Bob's marginal is invariant under the natural action of  $S_N$  as well as the diagonal action of  $U(d)$ . In slight abuse of notation, we denote by  $\zeta_{B^N}$  the unitary representation of  $\zeta \in S_N$  that permutes the tensor factors of  $\mathcal{H}_B^{\otimes N}$ .

**Lemma 3.2.** *Let  $\rho_{A^N B^N}$  be the resource state of a protocol for deterministic PBT with input dimension  $d$ . Then there exists another protocol for deterministic PBT with resource*

state  $\rho'_{A^N B^N I J}$ , where  $I$  and  $J$  are additional registers held by Alice, such that  $\rho'_{B^N}$  is invariant under the above-mentioned group actions,

$$\begin{aligned} U_B^{\otimes N} \rho'_{B^N} \left( U_B^{\otimes N} \right)^\dagger &= \rho'_{B^N} \quad \text{for all } U_B \in U(d), \\ \zeta_{B^N} \rho'_{B^N} \zeta_{B^N}^\dagger &= \rho'_{B^N} \quad \text{for all } \zeta \in S_N, \end{aligned} \quad (3.11)$$

and such that the new protocol has diamond norm error and entanglement fidelity no worse than the original one.

In fact, Lemma 3.2 applies not only to the diamond norm distance and the entanglement fidelity, but any convex functions on quantum channels that is invariant under conjugation with a unitary channel.

*Proof of Lemma 3.2.* Define the resource state

$$\tilde{\rho}_{A^N B^N I} = \frac{1}{N!} \sum_{\zeta \in S_N} \zeta_{B^N} \rho_{A^N B^N} \zeta_{B^N}^\dagger \otimes |\zeta\rangle\langle\zeta|_I, \quad (3.12)$$

where  $\zeta_{B^N}$  is the action of  $S_N$  on  $\mathcal{H}_B^{\otimes N}$  that permutes the tensor factors, and  $I$  is a classical ‘flag’ register with orthonormal basis  $\{|\zeta\rangle\}_{\zeta \in S_N}$ . The following protocol achieves the same performance as the preexisting one: Alice and Bob start sharing  $\tilde{\rho}_{A^N B^N I}$  as an entangled resource, with Bob holding  $B^N$  as usual and Alice holding registers  $A^N I$ . Alice begins by reading the classical register  $I$ . Suppose that its content is a permutation  $\zeta$ . She then continues to execute the original protocol, except that she applies  $\zeta$  to the index she is supposed to send to Bob after her measurement, which obviously yields the same result as the original protocol.

A similar argument can be made for the case of  $U(d)$ . Let  $D \subset U(d)$ ,  $|D| < \infty$  be an exact unitary  $N$ -design, i.e., a subset of the full unitary group such that taking the expectation value of any polynomial  $P$  of degree at most  $N$  in both  $U$  and  $U^\dagger$  over the uniform distribution on  $D$  yields the same result as taking the expectation of  $P$  over the normalized Haar measure on  $U(d)$ . Such exact  $N$ -designs exist for all  $N$  ([50]; see [51] for a bound on the size of exact  $N$ -designs). We now define a further modified resource state  $\rho'_{A^N B^N I J}$  from  $\tilde{\rho}_{A^N B^N I}$  in analogy to (3.12):

$$\rho'_{A^N B^N I J} = \frac{1}{|D|} \sum_{U \in D} U_B^{\otimes N} \tilde{\rho}_{A^N B^N I} (U_B^\dagger)^{\otimes N} \otimes |U\rangle\langle U|_J,$$

where  $\{|U\rangle\}_{U \in D}$  is an orthonormal basis for the flag register  $J$ . Again, there exists a modified protocol, in which Bob holds the registers  $B^N$  as usual, but Alice holds registers  $A^N I J$ . Alice starts by reading the register  $J$  which records the unitary  $U \in D$  that has been applied to Bob’s side. She then proceeds with the rest of the protocol after applying  $U^\dagger$  to her input state. Note that  $\rho'_{B^N}$  clearly satisfies the symmetries in (3.11), and furthermore the new PBT protocol using  $\rho'_{A^N B^N I J}$  has the same performance as the original one using  $\rho_{A^N B^N}$ , concluding the proof.  $\square$

Denote by  $\text{Sym}^N(\mathcal{H})$  the symmetric subspace of a Hilbert space  $\mathcal{H}^{\otimes N}$ , defined by

$$\text{Sym}^N(\mathcal{H}) := \{|\psi\rangle \in \mathcal{H}^{\otimes N} : \pi|\psi\rangle = |\psi\rangle \text{ for all } \pi \in S_N\}.$$

Using the above two lemmas we arrive at the following result.

**Proposition 3.3.** *Let  $\rho_{A^N B^N}$  be the resource state of a PBT protocol with input dimension  $d$ . Then there exists another protocol with properties as in Lemma 3.2 except that it has a resource state  $|\psi\rangle\langle\psi|_{A^N B^N}$  with  $|\psi\rangle_{A^N B^N} \in \text{Sym}^N(\mathcal{H}_A \otimes \mathcal{H}_B)$  that is a purification of a symmetric Werner state, i.e., it is invariant under the action of  $U(d)$  on  $\mathcal{H}_A^{\otimes N} \otimes \mathcal{H}_B^{\otimes N}$  given by  $U^{\otimes N} \otimes \bar{U}^{\otimes N}$ .*

*Proof.* We begin by transforming the protocol according to Lemma 3.2, resulting in a protocol with resource state  $\rho'_{A^N B^N I J}$  such that Bob's part is invariant under the  $S_N$  and  $U(d)$  actions. By Lemma 3.1, there exists a purification  $|\psi\rangle_{A^N B^N} \in \text{Sym}^N(\mathbb{C}^d \otimes \mathbb{C}^d)$  of  $\rho'_{B^N}$  that is invariant under  $U^{\otimes N} \otimes \bar{U}^{\otimes N}$  (note that the  $S_n$ -representation is self-dual, so the representation  $\phi \otimes \phi^*$  referred to in Lemma 3.1 just permutes the pairs of systems  $A_i B_i$ ). But Uhlmann's Theorem ensures that there exists an isometry  $V_{A^N \rightarrow A^N I J E}$  for some Hilbert space  $\mathcal{H}_E$  such that  $V_{A^N \rightarrow A^N I J E} |\psi\rangle_{A^N B^N}$  is a purification of  $\rho'_{A^N B^N I J}$ . The following is a protocol using the resource state  $|\psi\rangle$ : Alice applies  $V$  and discards  $E$ . Then the transformed protocol from Lemma 3.2 is performed.  $\square$

Using the symmetries of the resource state, we can show that the POVM can be chosen to be symmetric as well. In the proposition below, we omit identity operators.

**Proposition 3.4.** *Let  $\{(E_i)_{A_0 A^N}\}_{i=1}^N$  be Alice's POVM for a PBT protocol with a resource state  $|\psi\rangle$  with the symmetries from Proposition 3.3. Then there exists another POVM  $\{(E'_i)_{A_0 A^N}\}_{i=1}^N$  such that the following properties hold:*

- (i)  $\zeta_{A^N} (E'_i)_{A_0 A^N} \zeta_{A^N}^\dagger = (E'_{\zeta(i)})_{A_0 A^N}$  for all  $\zeta \in S_N$ ;
- (ii)  $(U_{A_0} \otimes \bar{U}_A^{\otimes N}) (E'_i)_{A_0 A^N} (U_{A_0} \otimes \bar{U}_A^{\otimes N})^\dagger = (E'_i)_{A_0 A^N}$  for all  $U \in U(d)$ ;
- (iii) the channel  $\Lambda'$  implemented by the PBT protocol is unitarily covariant, i.e.,

$$\Lambda'_{A_0 \rightarrow B_0}(X) = U_{B_0} \Lambda'_{A_0 \rightarrow B_0}(U_{A_0}^\dagger X U_{A_0}) U_{B_0}^\dagger \text{ for all } U \in U(d);$$

- (iv) the resulting protocol has diamond norm distance (to the identity channel) and entanglement fidelity no worse than the original one.

*Proof.* Define an averaged POVM with elements

$$(E'_i)_{A_0 A^N} = \int_{U(\mathcal{H}_A)} dU \frac{1}{N!} \sum_{\zeta \in S_N} (U_{A_0} \otimes \bar{U}_A^{\otimes N} \zeta_{A^N}) (E_{\zeta^{-1}(i)})_{A_0 A^N} (U_{A_0}^\dagger \otimes \zeta_{A^N}^\dagger (U_A^T)^{\otimes N}),$$

which clearly has the symmetries (i) and (ii). The corresponding channel can be written as

$$\Lambda'_{A_0 \rightarrow B_0} = \int_{U(\mathcal{H}_A)} \frac{1}{N!} \sum_{\zeta \in S_N} \Lambda_{A_0 \rightarrow B_0}^{(U, \zeta)},$$

where

$$\begin{aligned} \Lambda_{A_0 \rightarrow B_0}^{(U, \zeta)}(X_{A_0}) &= \sum_{i=1}^N \text{tr}_{A_0 A^N B_i^c} [((U_{A_0} \otimes \bar{U}_A^{\otimes N} \zeta_{A^N}) (E_{\zeta^{-1}(i)})_{A_0 A^N} (U_{A_0}^\dagger \otimes \zeta_{A^N}^\dagger (U_A^T)^{\otimes N})) \\ &\quad (X_{A_0} \otimes |\psi\rangle\langle\psi|_{A^N B^N})] \end{aligned}$$



$$\begin{aligned}
 &= \sum_{i=1}^N \text{tr}_{A_0 A^N B_i^c} \left[ (E_{\zeta^{-1}(i)})_{A_0 A^N} (U_{A_0}^\dagger X_{A_0} U_{A_0} \right. \\
 &\quad \left. \otimes (\zeta_{A^N}^\dagger (U_A^T)^{\otimes N} \otimes I_{B^N}) |\psi\rangle\langle\psi|_{A^N B^N} (\overline{U}_A^{\otimes N} \zeta_{A^N} \otimes I_{B^N}) \right] \\
 &= \sum_{i=1}^N \text{tr}_{A_0 A^N B_i^c} \left[ (E_{\zeta^{-1}(i)})_{A_0 A^N} (U_{A_0}^\dagger X_{A_0} U_{A_0} \right. \\
 &\quad \left. \otimes (I_{A^N} \otimes U_B^{\otimes N} \zeta_{B^N}) |\psi\rangle\langle\psi|_{A^N B^N} (I_{A^N} \otimes \zeta_{B^N}^\dagger (U_B^\dagger)^{\otimes N}) \right] \\
 &= U_{B_0} \Lambda_{A_0 \rightarrow B_0} (U_{A_0}^\dagger X_{A_0} U_{A_0}) U_{B_0}^\dagger,
 \end{aligned}$$

where we suppressed  $\text{id}_{B_i \rightarrow B_0}$ . Here we used the cyclicity property

$$\text{tr}_A[(X_A Y_A \otimes I_B) Z_{AB}] = \text{tr}_A[(Y_A \otimes I_B) Z_{AB} (X_A \otimes I_B)]$$

of the partial trace and the symmetries of the resource state, and  $\Lambda_{A_0 \rightarrow B_0}$  denotes the channel corresponding to the original protocol. It follows at once that  $\Lambda'_{A_0 \rightarrow B_0}$  is covariant in the sense of (iii). Finally, since the identity channel is itself covariant, property (iv) follows from the concavity (convexity) and unitary covariance of the entanglement fidelity and the diamond norm distance, respectively.  $\square$

Similarly as mentioned below Lemma 3.2, the statement in Proposition 3.4(iv) can be generalized to any convex function on the set of quantum channels that is invariant under conjugation with unitary channels.

The unitary covariance allows us to apply a lemma from [5] (stated as Lemma D.3 in ‘‘Appendix D’’) to relate the optimal diamond norm error and entanglement fidelity of port-based teleportation. This shows that the achievability results Eqs. (3.5) to (3.4) for the entanglement fidelity of deterministic PBT, as well as the ones mentioned in the introduction, imply similar results for the diamond norm error without losing a dimension factor.

**Corollary 3.5.** *Let  $F_d^*$  and  $\varepsilon_d^*$  be the optimal entanglement fidelity and optimal diamond norm error for deterministic PBT with input dimension  $d$ . Then,  $\varepsilon_d^* = 2(1 - F_d^*)$ .*

Note that the same formula was proven for the standard protocol in [5].

**3.4. Representation-theoretic characterization.** The symmetries of PBT enable the use of representation-theoretic results, in particular Schur–Weyl duality. This was extensively done in [21, 22] in order to derive the formulas Eqs. (3.5)–(3.9). The main ingredient used in [21] to derive Eqs. (3.5) and (3.9) was the following technical lemma. For the reader’s convenience, we give an elementary proof in ‘‘Appendix A’’ using only Schur–Weyl duality and the classical Pieri rule. In the statement below,  $B_i^c$  denotes the quantum system consisting of all  $B$ -systems except the  $i$ th one.

**Lemma 3.6** [21]. *The eigenvalues of the operator*

$$T(N)_{AB^N} = \frac{1}{N} \left( \phi_{AB_1}^+ \otimes I_{B_1^c} + \dots + \phi_{AB_N}^+ \otimes I_{B_N^c} \right)$$

on  $(\mathbb{C}^d)^{\otimes(1+N)}$  are given by the numbers

$$\frac{1}{dN} \gamma_\mu(\alpha) = \frac{1}{d} \frac{d_\alpha m_{d,\mu}}{d_\mu m_{d,\alpha}},$$

where  $\alpha \vdash_d N - 1$ , the Young diagram  $\mu \vdash_d N$  is obtained from  $\alpha$  by adding a single box, and  $\gamma_\mu(\alpha)$  is defined in Eq. (3.10).

Note that the formula in Lemma 3.6 above gives *all* eigenvalues of  $T(N)_{AB^N}$ , i.e., including multiplicities.

The connection to deterministic PBT is made via the equivalence with state discrimination. In particular, when using a maximally entangled resource,  $T(N)$  is a rescaled version of the density operator corresponding to the ensemble of quantum states  $\eta_i$  from Eq. (3.1),

$$T(N) = \frac{d^{N-1}}{N} \sum_i \eta_i.$$

Using the hook length formulas Eqs. (2.1) and (2.2), we readily obtain the following simple expression for the ratio  $\gamma_\mu(\alpha)$  defined in Eq. (3.10):

**Lemma 3.7.** [52] *Let  $\mu = \alpha + e_i$ . Then,*

$$\gamma_\mu(\alpha) = \mu_i - i + d = \alpha_i - i + d + 1,$$

i.e.,

$$\frac{d_\alpha m_{d,\mu}}{d_\mu m_{d,\alpha}} = \frac{\alpha_i - i + d + 1}{N}.$$

*Proof.* Using Eqs. (2.1) and (2.2), we find

$$\begin{aligned} \gamma_\mu(\alpha) &= N \frac{m_{d,\mu} d_\alpha}{m_{d,\alpha} d_\mu} = N \prod_{(i,j) \in \mu} \frac{d + c(i,j)}{h_\mu(i,j)} \frac{\prod_{(i,j) \in \mu} h_\mu(i,j)}{N!} \frac{(N-1)!}{\prod_{(i,j) \in \alpha} h_\alpha(i,j)} \\ &\quad \times \prod_{(i,j) \in \alpha} \frac{h_\alpha(i,j)}{d + c(i,j)} \\ &= \prod_{(i,j) \in \mu} \frac{d + c(i,j)}{1} \prod_{(i,j) \in \alpha} \frac{1}{d + c(i,j)} = d + c(i, \mu_i) = d + \mu_i - i, \end{aligned}$$

which concludes the proof.  $\square$

*Remark 3.8.* It is clear that  $\gamma_\mu(\alpha)$  is maximized for  $\alpha = (N - 1, 0, \dots, 0)$  and  $i = 1$ . Therefore,

$$\|T(N)\|_\infty = \frac{N + d - 1}{dN}.$$

This result can be readily used to characterize the extendibility of isotropic states, providing an alternative proof of the result by Johnson and Viola [53].

#### 4. The Schur–Weyl Distribution

Our results rely on the asymptotics of the Schur–Weyl distribution, a probability distribution defined below in (4.1) in terms of the representation-theoretic quantities that appear in the Schur–Weyl duality (2.4). These asymptotics can be related to the random matrix ensemble  $\text{GUE}_d^0$ . In this section we explain this connection and provide a refinement of a convergence result (stated in (4.4)) by Johansson [36] that is tailored to our applications. While representation-theoretic techniques have been extensively used in previous analyses, the connection between the Schur–Weyl distribution and random matrix theory has, to the best of our knowledge, not been previously recognized in the context of PBT (see however [31] for applications in the the context of quantum state tomography).

Recalling the Schur–Weyl duality  $(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\alpha \vdash_d n} [\alpha] \otimes V_\alpha^d$ , we denote by  $P_\alpha$  the orthogonal projector onto the summand labeled by the Young diagram  $\alpha \vdash_d n$ . The collection of these projectors defines a projective measurement, and hence

$$p_{d,n}(\alpha) := \text{tr} \left( P_\alpha \tau_d^{\otimes n} \right) = \frac{d_\alpha m_{d,\alpha}}{d^n} \quad (4.1)$$

with  $\tau_d = \frac{1}{d} \mathbf{1}_{\mathbb{C}^d}$  defines a probability distribution on Young diagrams  $\alpha \vdash_d n$ , known as the *Schur–Weyl distribution*. Now suppose that  $\alpha^{(n)} \sim p_{d,n}$  for  $n \in \mathbb{N}$ . By spectrum estimation [25–27, 54, 55], it is known that

$$\frac{\alpha^{(n)}}{n} \xrightarrow{P} \left( \frac{1}{d}, \dots, \frac{1}{d} \right) \text{ as } n \rightarrow \infty. \quad (4.2)$$

This can be understood as a *law of large numbers*. Johansson [36] proved a corresponding *central limit theorem*: Let  $\mathbf{A}^{(n)}$  be the centered and renormalized random variable defined by

$$\mathbf{A}^{(n)} := \frac{\alpha^{(n)} - \left( \frac{n}{d}, \dots, \frac{n}{d} \right)}{\sqrt{n/d}}. \quad (4.3)$$

Then Johansson [36] proved that

$$\mathbf{A}^{(n)} \xrightarrow{D} \text{spec}(\mathbf{G}) \quad (4.4)$$

for  $n \rightarrow \infty$ , where  $\mathbf{G} \sim \text{GUE}_d^0$ . The result for the first row is by Tracy and Widom [56] (cf. [36, 57]; see [31] for further discussion).

In the following sections, we would like to use this convergence of random variables stated in Eqs. (4.2) and (4.4) to determine the asymptotics of Eqs. (3.9) and (3.5). To this end, we rewrite the latter as expectation values of some functions of Young diagrams drawn according to the Schur–Weyl distribution. However, in order to conclude that these expectation values converge to the corresponding expectation values of functions on the spectrum of  $\text{GUE}_d^0$ -matrices, we need a stronger sense of convergence than what is provided by the former results. Indeed, we need to establish convergence for functions that diverge polynomially as  $n \rightarrow \infty$  when  $A_j = \omega(1)$  or when  $A_j = O(n^{-1/2})$ .<sup>4</sup> The former are easily handled using the bounds from spectrum estimation [27], but for the latter a refined bound on  $p_{d,n}$  corresponding to small  $A$  is needed. To this end, we prove

<sup>4</sup> Here,  $f(n) = \omega(g(n))$  means that  $|f(n)/g(n)|$  diverges as  $n \rightarrow \infty$ .

the following result, which shows convergence of expectation values of a large class of functions that includes all polynomials in the variables  $\mathbf{A}_i$ .

In the following, we will need the cone of sum-free non-increasing vectors in  $\mathbb{R}^d$ ,

$$C^d = \left\{ x \in \mathbb{R}^d : \sum_i x_i = 0, x_i \geq x_{i+1} \right\},$$

and its interior  $\text{int}(C^d) = \{x \in C^d : x_i \neq 0 \text{ for } i = 1, \dots, d\}$ .

**Theorem 4.1.** *Let  $g : \text{int}(C^d) \rightarrow \mathbb{R}$  be a continuous function satisfying the following: There exist constants  $\eta_{ij}$  satisfying  $\eta_{ij} > -2 - \frac{1}{d-1}$  such that for*

$$\varphi_\eta(x) := \prod_{i < j} (x_i - x_j)^{\eta_{ij}}$$

there exists a polynomial  $q$  with

$$\frac{g(x)}{\varphi_\eta(x)} \leq q(\|x\|_1).$$

For every  $n$ , let  $\alpha^{(n)} \sim p_{d,n}$  be drawn from the Schur–Weyl distribution,  $\mathbf{A}^{(n)} := \sqrt{d/n}(\alpha^{(n)} - n/d)$  the corresponding centered and renormalized random variable, and  $\tilde{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} + \frac{d-i}{\sqrt{n}}$ . Then the family of random variables  $\left\{ g\left(\tilde{\mathbf{A}}^{(n)}\right) \right\}_{n \in \mathbb{N}}$  is uniformly integrable and

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ g\left(\tilde{\mathbf{A}}^{(n)}\right) \right] = \mathbb{E} [g(\mathbf{A})],$$

where  $\mathbf{A} = \text{spec}(\mathbf{G})$  and  $\mathbf{G} \sim \text{GUE}_d^0$ .

As a special case we recover the uniform integrability of the moments of  $\mathbf{A}$  (Corollary 4.5), which implies convergence in distribution in the case of an absolutely continuous limiting distribution. Therefore, Theorem 4.1 is a refinement of the result (4.4) by Johansson. The remainder of this section is dedicated to proving Theorem 4.1.

The starting point for what follows is Stirling’s approximation, which states that

$$\sqrt{2\pi} \sqrt{n} \left(\frac{n}{e}\right)^n \leq n! \leq e \sqrt{n} \left(\frac{n}{e}\right)^n \quad \text{for all } n \in \mathbb{N}.$$

It will be convenient to instead use the following variant,

$$\frac{\sqrt{2\pi}}{e} \sqrt{n+1} \left(\frac{n}{e}\right)^n \leq n! \leq e \sqrt{n} \left(\frac{n}{e}\right)^n, \quad (4.5)$$

where the upper bound is unchanged and the lower bound follows using  $n! = \frac{(n+1)!}{n+1}$ . The dimension  $d_\alpha$  is equal to the multinomial coefficient up to inverse polynomial factors [27]. Defining the normalized Young diagram  $\tilde{\alpha} = \frac{\alpha}{n}$  for  $\alpha \vdash n$ , the multinomial coefficient  $\binom{n}{\alpha}$  can be bounded from above using Eq. (4.5) as

$$\binom{n}{\alpha} = \frac{n!}{\alpha_1! \dots \alpha_d!} \leq C_d \sqrt{\frac{n}{\prod_{i=1}^d (\alpha_i + 1)}} \frac{n^n}{\alpha_1^{\alpha_1} \dots \alpha_d^{\alpha_d}},$$

where  $C_d := \frac{e^{d+1}}{(2\pi)^{d/2}}$ . Hence,

$$\begin{aligned}
 d^{-n} \binom{n}{\alpha} &\leq C_d \sqrt{\frac{n}{\prod_{i=1}^d (\alpha_i + 1)}} \exp(-nD(\bar{\alpha} \parallel \tau)) \\
 &\leq C_d \sqrt{\frac{n}{\prod_{i=1}^d (\alpha_i + 1)}} \exp\left(-\frac{n}{2} \|\bar{\alpha} - \tau\|_1^2\right) \\
 &= C_d n^{-\frac{d-1}{2}} \left[ \prod_{i=1}^d \left(\bar{\alpha}_i + \frac{1}{n}\right)^{-\frac{1}{2}} \right] \exp\left(-\frac{n}{2} \|\bar{\alpha} - \tau\|_1^2\right). \quad (4.6)
 \end{aligned}$$

Here,  $D(p \parallel q) := \sum_i p_i \log p_i/q_i$  is the Kullback-Leibler divergence defined in terms of the natural logarithm,  $\tau = (1/d, \dots, 1/d)$  is the uniform distribution, and we used Pinsker's inequality [58] in the second step.

We go on to derive an upper bound on the probability of Young diagrams that are close to the boundary of the set of Young diagrams under the Schur–Weyl distribution. More precisely, the following lemma can be used to bound the probability of Young diagrams that have two rows that differ by less than the generic  $O(\sqrt{n})$  in length.

**Lemma 4.2.** *Let  $d \in \mathbb{N}$  and  $c_1, \dots, c_{d-1} \geq 0$ ,  $\gamma_1, \dots, \gamma_{d-1} \geq 0$ . Let  $\alpha \vdash_d n$  be a Young diagram with (a)  $\alpha_i - \alpha_{i+1} \leq c_i n^{\gamma_i}$  for all  $i$ . Finally, set  $A := \sqrt{d/n}(\alpha - n/d)$ . Then,*

$$p_{d,n}(\alpha) \leq C n^{-\frac{d^2-1}{2} + 2 \sum_{i < j} \gamma_{ij}} \left[ \prod_{i=1}^d \left(1 + \sqrt{\frac{d}{n}} A_i + \frac{d}{n}\right)^{i-d-\frac{1}{2}} \right] \exp\left(-\frac{1}{2d} \|A\|_1^2\right),$$

where  $\gamma_{ij} := \max\{\gamma_i, \gamma_{i+1}, \dots, \gamma_{j-1}\}$  and  $C = C(c_1, \dots, c_{d-1}, d)$  is a suitable constant.

*Proof.* We need to bound  $p_{d,n}(\alpha) = m_{d,\alpha} d_\alpha / d^n$  and begin with  $m_{d,\alpha}$ . By assumption (a), there exist constants  $C_{ij} > 0$  (depending on  $c_i, \dots, c_{j-1}$  as well as on  $d$ ) such that the inequality  $\alpha_i - \alpha_j + j - i \leq C_{ij} n^{\gamma_{ij}}$  holds for all  $i < j$ . Using the Weyl dimension formula (2.3) and assumption (a), it follows that

$$m_{d,\alpha} = \prod_{i < j} \frac{\alpha_i - \alpha_j + j - i}{j - i} \leq C_1 n^{\sum_{i < j} \gamma_{ij}} \quad (4.7)$$

for a suitable constant  $C_1 = C_1(c_1, \dots, c_{d-1}, d) > 0$ . Next, consider  $d_\alpha$ . By comparing the hook-length formulas (2.1) and (2.2), we have

$$\begin{aligned}
 d_\alpha &= n! m_{d,\alpha} \left[ \prod_{(i,j) \in \alpha} (d + j - i) \right]^{-1} \\
 &= n! m_{d,\alpha} \left[ \prod_{i=1}^d \frac{(\alpha_i + d - i)!}{(d - i)!} \right]^{-1} \\
 &\leq n! m_{d,\alpha} \left[ \prod_{i=1}^d \frac{(\alpha_i + 1)^{d-i} \alpha_i!}{(d - i)!} \right]^{-1}
 \end{aligned}$$

$$\begin{aligned}
&= m_{d,\alpha} \left[ \prod_{i=1}^d \frac{(d-i)!}{(\alpha_i+1)^{d-i}} \right] \binom{n}{\alpha} \\
&= C_2 m_{d,\alpha} n^{-\frac{d(d-1)}{2}} \left[ \prod_{i=1}^d \left( \bar{\alpha}_i + \frac{1}{n} \right)^{i-d} \right] \binom{n}{\alpha}, \tag{4.8}
\end{aligned}$$

where  $C_2 = C_2(d) > 0$ , and  $\bar{\alpha}_i = \alpha_i/n$ . In the inequality, we used that  $\alpha_i + d - i \geq \alpha_i + 1$  for  $1 \leq i \leq d-1$ , and for  $i = d$ , the exponent of  $\alpha_i + 1$  on the right hand side is zero.

Combining Eqs. (4.7)–(4.6) and setting  $C_3 = C_1^2 C_2 C_d$ , we obtain

$$\begin{aligned}
p_{d,n}(\alpha) &= \frac{m_{d,\alpha} d_\alpha}{d^n} \leq C_2 m_{d,\alpha}^2 n^{-\frac{d(d-1)}{2}} \left[ \prod_{i=1}^d \left( \bar{\alpha}_i + \frac{1}{n} \right)^{i-d} \right] d^{-n} \binom{n}{\alpha} \\
&\leq C_3 n^{-\frac{d^2-1}{2} + 2\sum_{i<j} \gamma_{ij}} \left[ \prod_{i=1}^d \left( \bar{\alpha}_i + \frac{1}{n} \right)^{i-d-\frac{1}{2}} \right] \exp\left(-\frac{n}{2} \|\bar{\alpha} - \tau\|_1^2\right).
\end{aligned}$$

Substituting  $\bar{\alpha}_i = \frac{1}{d} + \frac{A_i}{\sqrt{nd}}$  we obtain the desired bound.  $\square$

In order to derive the asymptotics of entanglement fidelities for port-based teleportation, we need to compute limits of certain expectation values. As a first step, the following lemma ensures that the corresponding sequences of random variables are uniformly integrable. We recall that a family of random variables  $\{\mathbf{X}^{(n)}\}_{n \in \mathbb{N}}$  is called *uniformly integrable* if, for every  $\varepsilon > 0$ , there exists  $K < \infty$  such that  $\sup_n \mathbb{E} \left[ |\mathbf{X}^{(n)}| \cdot \mathbb{1}_{|\mathbf{X}^{(n)}| \geq K} \right] \leq \varepsilon$ .

**Lemma 4.3.** *Under the same conditions as for Theorem 4.1, the family of random variables  $\left\{ g \left( \tilde{\mathbf{A}}^{(n)} \right) \right\}_{n \in \mathbb{N}}$  is uniformly integrable.*

*Proof.* Let  $\mathbf{X}^{(n)} := g \left( \tilde{\mathbf{A}}^{(n)} \right)$ . The claimed uniform integrability follows if we can show that

$$\sup_n \mathbb{E} \left[ |\mathbf{X}^{(n)}| \right] < \infty \tag{4.9}$$

for every choice of the  $\eta_{ij}$ . Indeed, to show that  $\{\mathbf{X}^{(n)}\}$  is uniformly integrable it suffices to show that  $\sup_n \mathbb{E} \left[ |\mathbf{X}^{(n)}|^{1+\delta} \right] < \infty$  for some  $\delta > 0$  [59, Ex. 5.5.1]. If we choose  $\delta > 0$  such that  $\eta'_{ij} := (1+\delta)\eta_{ij} > -2 - \frac{1}{d-1}$  for all  $1 \leq i < j \leq d$ , then it is clear that Eq. (4.9) for  $\eta'_{ij}$  implies uniform integrability of the original family.

Moreover, we may also assume that  $h_\eta \equiv g/\varphi_\eta = 1$ , since the general case then follows from the fact that  $p_{d,n}(\alpha)$  decays exponentially in  $\|A\|_1$  (see Lemma 4.2). More precisely, for any polynomial  $r$  and any constant  $\theta_1 > 0$  there exist constants  $\theta_2, \theta_3 > 0$  such that

$$r(\|x\|_1) \exp(-\theta_1 \|x\|_1) \leq \theta_2 \exp(-\theta_3 \|x\|_1).$$

In particular, this holds for the polynomial  $q$  bounding  $h$  from above by assumption. When proving the statement  $\sup_n \mathbb{E} \left[ |\mathbf{X}^{(n)}| \right] < \infty$ , the argument above allows us to reduce the general case  $h_\eta = g/\varphi_\eta \neq 1$  to the case  $h_\eta = 1$ , or equivalently, to

$$g(x) = \varphi_\eta(x) = \prod_{i < j} (x_i - x_j)^{\eta_{ij}}.$$

Thus, it remains to be shown that

$$\sup_n \mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \right] < \infty, \quad (4.10)$$

where

$$f^{(n)}(A) := \varphi_\eta(\tilde{A}) = \prod_{i < j} \left( A_i - A_j + \frac{j-i}{\sqrt{n/d}} \right)^{\eta_{ij}}$$

for some constants  $\eta_{ij}$  satisfying the assumption of Theorem 4.1 that we fix for the rest of this proof. Define  $\Gamma_{ij} := A_i - A_j + \frac{j-i}{\sqrt{n/d}}$ . Then we have  $f^{(n)}(A) = \prod_{i < j} \Gamma_{ij}^{\eta_{ij}}$ , while the Weyl dimension formula (2.3) becomes

$$m_{d,\alpha} = \left( \frac{n}{d} \right)^{\frac{d(d-1)}{4}} \prod_{i < j} \frac{\Gamma_{ij}}{j-i}.$$

Hence, together with Eqs. (4.8) and (4.6) we can bound  $p_{d,n}(\alpha) = m_{d,\alpha} d_\alpha / d^n$  as

$$\begin{aligned} p_{d,n}(\alpha) &\leq C n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^2 \right) \left[ \prod_{i=1}^d \left( \bar{\alpha}_i + \frac{1}{n} \right)^{i-d-\frac{1}{2}} \right] \exp \left( \frac{n}{2} \|\bar{\alpha} - \tau\|_1^2 \right) \\ &= C n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^2 \right) \left[ \prod_{i=1}^d \left( 1 + \sqrt{\frac{d}{n}} A_i + \frac{d}{n} \right)^{i-d-\frac{1}{2}} \right] \exp \left( -\frac{1}{2d} \|A\|_1^2 \right), \end{aligned} \quad (4.11)$$

where  $C = C(d)$  is some constant, and we used  $\bar{\alpha}_i + \frac{1}{n} = \frac{1}{d} \left( \sqrt{\frac{d}{n}} A_i + \frac{d}{n} + 1 \right)$  and  $\tau = (1/d, \dots, 1/d)$  in the equality. Using  $f^{(n)}(A) = \prod_{i < j} \Gamma_{ij}^{\eta_{ij}}$ , this yields the bound

$$\begin{aligned} p_{d,n}(\alpha) f^{(n)}(A) &\leq C n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \left[ \prod_{i=1}^d \left( 1 + \sqrt{\frac{d}{n}} A_i + \frac{d}{n} \right)^{i-d-\frac{1}{2}} \right] \\ &\quad \exp \left( -\frac{1}{2d} \|A\|_1^2 \right). \end{aligned} \quad (4.12)$$

We now want to bound the expectation value in Eq. (4.10) and begin by splitting the sum over Young diagrams according to whether  $\exists i : |A_i| > n^\varepsilon$  for some  $\varepsilon \in (0, \frac{1}{2})$  to be determined later, or  $|A_i| \leq n^\varepsilon$  for all  $i$ . We denote the former event by  $\mathcal{E}$  and obtain

$$\mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \right] = \mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}} \right] + \mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}^c} \right]. \quad (4.13)$$

We treat the two expectation values in (4.13) separately and begin with the first one. If  $|A_i| > n^\varepsilon$  for some  $i$ , then  $\|A\|_1^2 \geq n^{2\varepsilon}$ , so it follows by Eq. (4.12) that

$$\begin{aligned}
& \mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}} \right] \\
&= \sum_{\substack{\alpha \vdash_d n \text{ s.t.} \\ \exists i: |A_i| > n^\varepsilon}} p_{d,n}(\alpha) f^{(n)}(A) \\
&\leq C \sum_{\substack{\alpha \vdash_d n \text{ s.t.} \\ \exists i: |A_i| > n^\varepsilon}} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \left[ \prod_{i=1}^d \left( 1 + \sqrt{\frac{d}{n}} A_i + \frac{d}{n} \right)^{i-d-\frac{1}{2}} \right] \\
&\quad \exp \left( -\frac{1}{2d} \|A\|_1^2 \right) \\
&\leq \text{poly}(n) \exp \left( -\frac{1}{2d} n^{2\varepsilon} \right).
\end{aligned}$$

Here,  $\text{poly}(n)$  denotes some polynomial in  $n$  and we also used that, for fixed  $d$ , the number of Young diagrams is polynomial in  $n$ . This shows that the first expectation value in (4.13) vanishes for  $n \rightarrow \infty$ .

For the second expectation value, note that  $|A_i| \leq n^\varepsilon = o(\sqrt{n})$  for all  $i$ , and hence there exists a constant  $K > 0$  such that we have

$$\prod_{i=1}^d \left( 1 + \sqrt{\frac{d}{n}} A_i + \frac{d}{n} \right)^{i-d-\frac{1}{2}} \leq K. \quad (4.14)$$

Using Eqs. (4.12) and (4.14), we can therefore bound

$$\begin{aligned}
\mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}^c} \right] &= \sum_{\substack{\alpha \vdash_d n \text{ s.t.} \\ \forall i: |A_i| \leq n^\varepsilon}} p_{d,n}(\alpha) f^{(n)}(A) \\
&\leq CK \sum_{A \in \mathcal{D}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \exp \left( -\frac{1}{2d} \|A\|_1^2 \right),
\end{aligned}$$

where we have introduced  $\mathcal{D}_n := \{A : \alpha \vdash_d n\}$ . The summands are nonnegative, even when evaluated on any point in the larger set  $\hat{\mathcal{D}}_n := \left\{ A \in \sqrt{\frac{d}{n}} (\mathbb{Z} - \frac{n}{d})^d : \sum_i A_i = 0, A_i \geq A_{i+1} \forall i \right\} \supset \mathcal{D}_n$ , so that we have the upper bound

$$\begin{aligned}
\mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}^c} \right] &\leq CK \sum_{A \in \mathcal{D}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \exp \left( -\frac{1}{2d} \|A\|_1^2 \right) \\
&\leq CK \sum_{A \in \hat{\mathcal{D}}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \exp \left( -\frac{1}{2d} \|A\|_1^2 \right). \quad (4.15)
\end{aligned}$$

Let  $x_i = A_i - A_{i+1}$ ,  $i = 1, \dots, d-1$ . Next, we will upper bound the exponential in Eq. (4.15). For this, define  $\tilde{x}_i = \max(\frac{1}{d-1}, x_i)$  and let  $S = \{i \in \{1, \dots, d-1\} \mid x_i \leq \frac{1}{d-1}\}$ .



Then, assuming  $S^c \neq \emptyset$ ,

$$\begin{aligned}
 \sum_{i=1}^{d-1} \tilde{x}_i &\leq \left( \sum_{i=1}^{d-1} \tilde{x}_i \right)^2 = \left( \sum_{i \in S} \tilde{x}_i + \sum_{i \in S^c} \tilde{x}_i \right)^2 = \left( \frac{|S|}{d-1} + \sum_{i \in S^c} x_i \right)^2 \\
 &= \left( \frac{|S|}{d-1} \right)^2 + 2 \frac{|S|}{d-1} \left( \sum_{i \in S^c} x_i \right) + \left( \sum_{i \in S^c} x_i \right)^2 \\
 &\leq \left( \frac{|S|}{d-1} \right)^2 + 2 \frac{|S|}{d-1} \frac{d-1}{|S^c|} \left( \sum_{i \in S^c} x_i \right)^2 + \left( \sum_{i \in S^c} x_i \right)^2 \\
 &= \left( \frac{|S|}{d-1} \right)^2 + \left( 1 + 2 \frac{|S|}{|S^c|} \right) \left( \sum_{i \in S^c} x_i \right)^2 \\
 &\leq 1 + (2d-1) \left( \sum_{i=1}^{d-1} x_i \right)^2
 \end{aligned}$$

since  $\sum_{i \in S^c} x_i \geq \frac{|S^c|}{d-1}$ . This bounds also holds when  $S^c = \emptyset$ . Hence,

$$\exp\left(-\frac{1}{2d} \|A\|_1^2\right) \leq \exp\left(-\frac{1}{2d} \left(\sum_{i=1}^{d-1} x_i\right)^2\right) \leq R \exp\left(-\gamma \sum_{i=1}^{d-1} \tilde{x}_i\right) = R \prod_{i=1}^{d-1} \exp(-\gamma \tilde{x}_i), \quad (4.16)$$

where  $\gamma := \frac{1}{2d(2d-1)}$  and  $R := e^\gamma$ . The first inequality follows from  $\sum_{i=1}^{d-1} x_i = A_1 - A_d = |A_1| + |A_d| \leq \|A\|_1$ . If we use Eq. (4.16) in Eq. (4.15) we obtain the upper bound

$$\mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}^c} \right] \leq C' \sum_{A \in \hat{\mathcal{D}}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \prod_{i=1}^{d-1} \exp(-\gamma \tilde{x}_i) \quad (4.17)$$

where  $C' := CKR$ .

Let us first assume that all  $\eta_{ij} \leq -2$ , so that  $2 + \eta_{ij} \in (-\frac{1}{d-1}, 0]$ . Since

$$\Gamma_{ij} = \left( \sum_{l=i}^{j-1} x_l \right) + \frac{j-i}{\sqrt{\frac{n}{d}}} = \sum_{l=i}^{j-1} \left( x_l + \frac{1}{\sqrt{\frac{n}{d}}} \right) \geq x_i + \frac{1}{\sqrt{\frac{n}{d}}}$$

and  $\eta_{ij} + 2 \leq 0$ , we have that

$$\Gamma_{ij}^{2+\eta_{ij}} \leq \left( x_i + \frac{1}{\sqrt{\frac{n}{d}}} \right)^{2+\eta_{ij}}, \quad (4.18)$$

as power functions with non-positive exponent are non-increasing. We can then upper-bound Eq. (4.17) as follows,

$$\begin{aligned}
\mathbb{E} \left[ f^{(n)}(\mathbf{A}^{(n)}) \mathbb{1}_{\mathcal{E}^c} \right] &\leq C' \sum_{A \in \hat{\mathcal{D}}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \Gamma_{ij}^{2+\eta_{ij}} \right) \prod_{i=1}^{d-1} \exp(-\gamma \tilde{x}_i) \\
&\leq C' \sum_{A \in \hat{\mathcal{D}}_n} n^{-\frac{d-1}{2}} \left( \prod_{i < j} \left( x_i + \frac{1}{\sqrt{\frac{n}{d}}} \right)^{2+\eta_{ij}} \right) \prod_{i=1}^{d-1} \exp(-\gamma \tilde{x}_i), \\
&= C' \sum_{A \in \hat{\mathcal{D}}_n} n^{-\frac{d-1}{2}} \left( \prod_{i=1}^{d-1} \left( x_i + \frac{1}{\sqrt{\frac{n}{d}}} \right)^{\sum_{j=i+1}^d (2+\eta_{ij})} \right) \prod_{i=1}^{d-1} \exp(-\gamma \tilde{x}_i), \\
&= C' \prod_{i=1}^{d-1} \left( n^{-\frac{1}{2}} \sum_{x_i \in \sqrt{\frac{d}{n}} \mathbb{N}} \left( x_i + \frac{1}{\sqrt{\frac{n}{d}}} \right)^{\sum_{j=i+1}^d (2+\eta_{ij})} \exp(-\gamma \tilde{x}_i) \right)
\end{aligned}$$

where the first inequality is Eq. (4.17) and in the second inequality we used Eq. (4.18). Since  $\eta_{ij} > -2 - \frac{1}{d-1}$  by assumption, it follows that  $\sum_{j=i+1}^d (2 + \eta_{ij}) > -\frac{d-i}{d-1} \geq -1$ . Thus, each term in the product is a Riemann sum for an improper Riemann integral, as in Lemma D.4, which then shows that the expression converges for  $n \rightarrow \infty$ .

The case where some  $\eta_{ij} > -2$  is treated by observing that

$$\prod_{\substack{i < j: \\ \eta_{ij} > -2}} \Gamma_{ij}^{2+\eta_{ij}} \exp\left(-\frac{1}{2d} \|A\|_1^2\right) \leq c_1 \exp\left(-\frac{c_2}{2d} \|A\|_1^2\right)$$

for suitable constants  $c_1, c_2 > 0$ . We can use this bound in Eq. (4.17) to replace each  $\eta_{ij} > -2$  by  $\eta_{ij} = -2$ , at the expense of modifying the constants  $C'$  and  $\gamma$ , and then proceed as we did before. This concludes the proof of Eq. (4.10).  $\square$

The uniform integrability result of Lemma 4.3 implies that the corresponding expectation values converge. To determine their limit in terms of the expectation value of a function of the spectrum of a  $\text{GUE}_d^0$ -matrix, however, we need to show that we can take the limit of the dependencies on  $n$  of the function and the random variable  $\mathbf{A}^{(n)}$  separately. This is proved in the following lemma, where we denote the interior of a set  $E$  by  $\text{int}(E)$ .

**Lemma 4.4.** *Let  $\{\mathbf{A}^{(n)}\}_{n \in \mathbb{N}}$  and  $\mathbf{A}$  be random variables on a Borel measure space  $E$  such that  $\mathbf{A}^{(n)} \xrightarrow{D} \mathbf{A}$  for  $n \rightarrow \infty$  and  $\mathbf{A}$  is absolutely continuous. Let  $f : \text{int}(E) \rightarrow \mathbb{R}$ . Let further  $f_n : E \rightarrow \mathbb{R}$ ,  $n \in \mathbb{N}$ , be a sequence of continuous bounded functions such that  $f_n \rightarrow f$  pointwise on  $\text{int}(E)$  and, for any compact  $S \subset \text{int}(E)$ ,  $\{f_n|_S\}_{n \in \mathbb{N}}$  is uniformly equicontinuous and  $f_n|_S \rightarrow f|_S$  uniformly. Then for any such compact  $S \subset \text{int}(E)$ , the expectation value  $\mathbb{E}[f(\mathbf{A}) \mathbb{1}_S(\mathbf{A})]$  exists and*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ f_n(\mathbf{A}^{(n)}) \mathbb{1}_S(\mathbf{A}^{(n)}) \right] = \mathbb{E} [f(\mathbf{A}) \mathbb{1}_S(\mathbf{A})].$$

*Proof.* For  $n, m \in \mathbb{N} \cup \{\infty\}$ , define

$$b_{nm}(S) = \mathbb{E} \left[ f_n(\mathbf{A}^{(m)}) \mathbb{1}_S(\mathbf{A}^{(m)}) \right]$$

with  $f_\infty := f$ ,  $\mathbf{A}^{(\infty)} := \mathbf{A}$  and  $S \subset \text{int}(E)$  compact. These expectation values readily exist as  $f_n$  is bounded for all  $n$ , and the uniform convergence of  $f_n|_S$  implies that  $f|_S$  is continuous and bounded as well. The uniform convergence  $f_n|_S \rightarrow f|_S$  implies that  $f_n|_S$  is uniformly bounded, so by Lebesgue's theorem of dominated convergence  $b_{\infty m}(S)$  exists for all  $m \in \mathbb{N}$  and

$$\lim_{n \rightarrow \infty} b_{nm}(S) = b_{\infty m}(S) \quad \forall m \in \mathbb{N} \cup \{\infty\}. \quad (4.19)$$

This convergence is even uniform in  $m$  which follows directly from the uniform convergence of  $f_n|_S$ . The sequence  $\{\mathbf{A}^{(n)}\}_{n \in \mathbb{N}}$  of random variables converges in distribution to the absolutely continuous  $\mathbf{A}$ , so the expectation value of any continuous bounded function converges. Therefore,

$$\lim_{m \rightarrow \infty} b_{nm}(S) = b_{n\infty}(S) \quad \forall n \in \mathbb{N} \cup \{\infty\}. \quad (4.20)$$

An inspection of the proof of Theorem 1, Chapter VIII in [60] reveals the following: The fact that the uniform continuity and boundedness of  $f_n|_S$  hold uniformly in  $n$  implies the uniformity of the above limit. Moreover, since both limits exist and are uniform, this implies that they are equal to each other, and any limit of the form

$$\lim_{n \rightarrow \infty} b_{nm(n)}$$

for  $m(n) \xrightarrow{n \rightarrow \infty} \infty$  exists and is equal to the limits in Eqs. (4.19) and (4.20).  $\square$

Finally, we obtain the desired convergence theorem. For our applications,  $\eta_{ij} \equiv -2$  suffices. The range of  $\eta_{ij}$ 's for which the lemma is proven is naturally given by the proof technique.

*Proof of Theorem 4.1.* The uniform integrability of  $\mathbf{X}^{(n)} := g(\tilde{\mathbf{A}}^{(n)})$  is the content of Lemma 4.3. Recall that uniform integrability means that

$$\lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \mathbb{E} \left[ |\mathbf{X}^{(n)}| \cdot \mathbb{1}_{\mathcal{E}_K^c}(\mathbf{A}^{(n)}) \right] = 0,$$

where  $\mathcal{E}_K := \{x \in \mathbb{R}^d : \|x\|_\infty \leq K\}$ . Let now  $\varepsilon > 0$  be arbitrary, and  $K < \infty$  be such that the following conditions are true:

$$\sup_{n \in \mathbb{N}} \mathbb{E} \left[ |\mathbf{X}^{(n)}| \cdot \mathbb{1}_{\mathcal{E}_K^c}(\mathbf{A}^{(n)}) \right] \leq \frac{\varepsilon}{3} \quad \mathbb{E} \left[ g(\mathbf{A}) \mathbb{1}_{\mathcal{E}_K^c}(\mathbf{A}) \right] \leq \frac{\varepsilon}{3},$$

where  $\mathbf{A}$  is distributed as the spectrum of a  $\text{GUE}_d^0$  matrix. For the bound on the second expectation value, recall that the density of the eigenvalues  $(\mu_1, \dots, \mu_d)$  of a  $\text{GUE}_d^0$  matrix is proportional to  $\exp(-\sum_{i=1}^d \mu_i^2) \prod_{i < j} (\mu_i - \mu_j)^2$ , and hence decays exponentially in  $\|\mu\|_\infty$ . By Lemma 4.4,  $\lim_{n \rightarrow \infty} \mathbb{E} \left[ \mathbf{X}^{(n)} \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}^{(n)}) \right] = \mathbb{E} \left[ g(\mathbf{A}) \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}) \right]$ . Thus, we can choose  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\left| \mathbb{E} \left[ \mathbf{X}^{(n)} \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}^{(n)}) \right] - \mathbb{E} \left[ g(\mathbf{A}) \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}) \right] \right| \leq \frac{\varepsilon}{3}.$$

Using the above choices, we then have

$$\begin{aligned} \left| \mathbb{E}[\mathbf{X}^{(n)}] - \mathbb{E}[g(\mathbf{A})] \right| &\leq \mathbb{E} \left[ |\mathbf{X}^{(n)}| \cdot \mathbb{1}_{\mathcal{E}_K^c}(\mathbf{A}^{(n)}) \right] + \left| \mathbb{E} \left[ g(\mathbf{A}) \mathbb{1}_{\mathcal{E}_K^c}(\mathbf{A}) \right] \right| \\ &\quad + \left| \mathbb{E} \left[ \mathbf{X}^{(n)} \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}^{(n)}) \right] - \mathbb{E} \left[ g(\mathbf{A}) \mathbb{1}_{\mathcal{E}_K}(\mathbf{A}) \right] \right| \leq \varepsilon \end{aligned}$$

for all  $n \geq n_0$ , proving the desired convergence of the expectation values.  $\square$

From Theorem 4.1 we immediately obtain the following corollary about uniform integrability of the moments of  $\mathbf{A}$ .

**Corollary 4.5.** *Let  $k \in \mathbb{N}$ , let  $j \in \{1, \dots, d\}$ , and, for every  $n$ , let  $\mathbf{A}^{(n)}$  be the random vector defined in (4.3). Then, the sequence of  $k$ th moments  $\{(\mathbf{A}_j^{(n)})^k\}_{n \in \mathbb{N}}$  is uniformly integrable and  $\lim_{n \rightarrow \infty} \mathbb{E}[(\mathbf{A}_j^{(n)})^k] = \mathbb{E}[\mathbf{A}_j^k]$ , where  $\mathbf{A} \sim \text{GUE}_d^0$ .*

## 5. Probabilistic PBT

Our goal in this section is to determine the asymptotics of  $p_d^{\text{EPR}}$  using the formula (3.9) and exploiting our convergence theorem, Theorem 4.1. The main result is the following theorem stated in Sect. 1.2, which we restate here for convenience.

**Theorem 1.3** (Restated). *For probabilistic port-based teleportation in arbitrary but fixed dimension  $d$  with EPR pairs as resource states,*

$$p_d^{\text{EPR}}(N) = 1 - \sqrt{\frac{d}{N-1}} \mathbb{E}[\lambda_{\max}(\mathbf{G})] + o\left(N^{-1/2}\right),$$

where  $\mathbf{G} \sim \text{GUE}_d^0$ .

Previously, such a result was only known for  $d = 2$  following from an exact formula for  $p_2^{\text{EPR}}(N)$  derived in [4]. We show in Lemma C.1 in ‘‘Appendix C’’ that, for  $d = 2$ ,  $\mathbb{E}[\lambda_{\max}(\mathbf{G})] = \frac{2}{\sqrt{\pi}}$ , hence rederiving the asymptotics from [4].

While Theorem 1.3 characterizes the limiting behavior of  $p^{\text{EPR}}$  for large  $N$ , it contains the constant  $\mathbb{E}[\lambda_{\max}(\mathbf{G})]$ , which depends on  $d$ . As  $\mathbb{E}[\mathbf{M}] = 0$  for  $\mathbf{M} \sim \text{GUE}_d$ , it suffices to analyze the expected largest eigenvalue for  $\text{GUE}_d$ . The famous Wigner semicircle law [24] implies immediately that

$$\lim_{d \rightarrow \infty} \frac{\mathbb{E}[\lambda_{\max}(\mathbf{G})]}{\sqrt{d}} = 2,$$

but meanwhile the distribution of the maximal eigenvalue has been characterized in a much more fine-grained manner. In particular, according to [61], there exist constants  $C$  and  $C'$  such that the expectation value of the maximal eigenvalue satisfies the inequalities

$$1 - \frac{1}{C'd^{\frac{2}{3}}} \leq \frac{\mathbb{E}[\lambda_{\max}(\mathbf{G})]}{2\sqrt{d}} \leq 1 - \frac{1}{Cd^{\frac{2}{3}}}.$$

This also manifestly reconciles Theorem 1.3 with the fact that teleportation needs at least  $2 \log d$  bits of classical communication (see Sect. 7), since the amount of classical communication in a port-based teleportation protocol consists of  $\log N$  bits.

*Proof of Theorem 1.3.* We start with Eq. (3.9), which was derived in [21], and which we restate here for convenience:

$$p_d^{\text{EPR}}(N) = \frac{1}{d^N} \sum_{\alpha \vdash N-1} m_{d,\alpha}^2 \frac{d_{\mu^*}}{m_{d,\mu^*}},$$

where  $\mu^*$  is the Young diagram obtained from  $\alpha \vdash N-1$  by adding one box such that  $\gamma_\mu(\alpha) = N \frac{m_{d,\mu} d_\alpha}{m_{d,\alpha} d_\mu}$  is maximal. By Lemma 3.7, we have  $\gamma_\mu(\alpha) = \alpha_i - i + d + 1$  for  $\mu = \alpha + e_i$ . This is maximal if we choose  $i = 1$ , resulting in  $\gamma_{\mu^*}(\alpha) = \alpha_1 + d$ . We therefore obtain:

$$\begin{aligned} p_d^{\text{EPR}}(N) &= \frac{1}{d^N} \sum_{\alpha \vdash N-1} m_{d,\alpha} d_\alpha \frac{m_{d,\alpha} d_{\mu^*}}{m_{d,\mu^*} d_\alpha} \\ &= \frac{1}{d^N} \sum_{\alpha \vdash N-1} m_{d,\alpha} d_\alpha \frac{N}{\gamma_{\mu^*}(\alpha)} \\ &= \frac{1}{d} \mathbb{E} \alpha \left[ \frac{N}{\gamma_{\mu^*}(\alpha)} \right] \\ &= \frac{1}{d} \mathbb{E} \alpha \left[ \frac{N}{\alpha_1^{(N-1)} + d} \right]. \end{aligned}$$

Recall that

$$\alpha^{(N-1)} = (\alpha_1^{(N-1)}, \dots, \alpha_d^{(N-1)}) \sim p_{d,N-1}$$

is a random vector corresponding to Young diagrams with  $N-1$  boxes and at most  $d$  rows, where  $p_{d,N-1}$  is the Schur–Weyl distribution defined in (4.1). We continue by abbreviating  $n = N-1$  and changing to the centered and renormalized random variable  $\mathbf{A}^{(n)}$  from Eq. (4.3). Corollary 4.5 implies that

$$\mathbb{E} \left[ \mathbf{A}_1^{(n)} \right] \xrightarrow{N \rightarrow \infty} \mathbb{E}[\lambda_{\max}(\mathbf{G})] \quad \mathbb{E} \left[ (\mathbf{A}_1^{(n)})^2 \right] \xrightarrow{N \rightarrow \infty} \mathbb{E}[\lambda_{\max}(\mathbf{G})^2]. \quad (5.1)$$

Using the  $\mathbf{A}^{(n)}$  variables from (4.3), linearity of the expectation value and suitable rearranging, one finds that

$$\begin{aligned} \sqrt{N-1} \left( 1 - p_d^{\text{EPR}}(N) \right) &= \mathbb{E} \left[ \sqrt{N-1} - \frac{\sqrt{N-1} N}{\sqrt{d(N-1) \mathbf{A}_1^{(n)} + N-1 + d^2}} \right] \\ &= \mathbb{E} \left[ f_{d,N} \left( \mathbf{A}_1^{(n)} \right) \right], \end{aligned}$$

where we set

$$f_{d,N}(x) := \frac{x\sqrt{d} + \frac{d^2-1}{\sqrt{N-1}}}{1 + \frac{d^2}{N-1} + \frac{x\sqrt{d}}{\sqrt{N-1}}}.$$

Note that, for  $x \geq 0$ ,

$$\begin{aligned} \left| f_{d,N}(x) - x\sqrt{d} \right| &= \left| \frac{\frac{d^2-1}{\sqrt{N-1}} + \frac{xd^{5/2}}{N-1} + \frac{x^2d}{\sqrt{N-1}}}{1 + \frac{d^2}{N-1} + \frac{x\sqrt{d}}{\sqrt{N-1}}} \right| \\ &\leq \frac{d^2-1}{\sqrt{N-1}} + \frac{xd^{5/2}}{N-1} + \frac{x^2d}{\sqrt{N-1}} \\ &\leq \frac{1}{\sqrt{N-1}} \left( K_1 + K_2x + K_3x^2 \right) \end{aligned}$$

for some constants  $K_i$ , where the first inequality follows from the fact that the denominator in the first line is greater than 1 for  $x \geq 0$ . Since both  $\mathbf{A}_1^{(n)} \geq 0$  and  $\lambda_{\max}(\mathbf{G}) \geq 0$ , and using (5.1), it follows that

$$\begin{aligned} &\left| \mathbb{E} \left[ f_{d,N} \left( \mathbf{A}_1^{(n)} \right) \right] - \sqrt{d} \mathbb{E}[\lambda_{\max}(\mathbf{G})] \right| \\ &\leq \left| \mathbb{E} \left[ f_{d,N} \left( \mathbf{A}_1^{(n)} \right) \right] - \sqrt{d} \mathbb{E} \left[ \mathbf{A}_1^{(n)} \right] \right| + \sqrt{d} \left| \mathbb{E} \left[ \mathbf{A}_1^{(n)} \right] - \mathbb{E}[\lambda_{\max}(\mathbf{G})] \right| \\ &\leq \frac{K_1 + K_2 \mathbb{E} \left[ \mathbf{A}_1^{(n)} \right] + K_3 \mathbb{E} \left[ \left( \mathbf{A}_1^{(n)} \right)^2 \right]}{\sqrt{N-1}} + \sqrt{d} \left| \mathbb{E} \left[ \mathbf{A}_1^{(n)} \right] - \mathbb{E}[\lambda_{\max}(\mathbf{G})] \right| \\ &\xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Thus we have shown that, for fixed  $d$  and large  $N$ ,

$$p_d^{\text{EPR}}(N) = 1 - \sqrt{\frac{d}{N-1}} \mathbb{E}[\lambda_{\max}(\mathbf{G})] + o\left(N^{-1/2}\right),$$

which is what we set out to prove.  $\square$

*Remark 5.1.* For the probabilistic protocol with optimized resource state, recall from Eq. (3.8) that

$$p_d^*(N) = 1 - \frac{d^2-1}{d^2-1+N} = 1 - \frac{d^2-1}{N} + o(1/N).$$

For fixed  $d$ , this converges to unity as  $O(1/N)$ , i.e., much faster than the  $O(1/\sqrt{N})$  convergence in the EPR case proved in Theorem 1.3 above.

## 6. Deterministic PBT

The following section is divided into two parts. First, in Sect. 6.1 we derive the leading order of the standard protocol for deterministic port-based teleportation (see Sect. 3, where this terminology is explained). Second, in Sect. 6.2 we derive a lower bound on the leading order of the optimal deterministic protocol. As in the case of probabilistic PBT, the optimal deterministic protocol converges quadratically faster than the standard deterministic protocol, this time displaying an  $N^{-2}$  versus  $N^{-1}$  behavior (as opposed to  $N^{-1}$  versus  $N^{-1/2}$  in the probabilistic case).

6.1. *Asymptotics of the standard protocol.* Our goal in this section is to determine the leading order in the asymptotics of  $F_d^{\text{std}}$ . We do so by deriving an expression for the quantity  $\lim_{N \rightarrow \infty} N(1 - F_d^{\text{std}}(N))$ , that is, we determine the coefficient  $c_1 = c_1(d)$  in the expansion

$$F_d^{\text{std}}(N) = 1 - \frac{c_1}{N} + o(N^{-1}).$$

We need the following lemma that states that we can restrict a sequence of expectation values in the Schur–Weyl distribution to a suitably chosen neighborhood of the expectation value and remove degenerate Young diagrams without changing the limit. Let

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

be the Heaviside step function. Recall the definition of the centered and normalized variables

$$A_i := \frac{\alpha_i - n/d}{\sqrt{n/d}},$$

such that  $\alpha_i = \sqrt{\frac{n}{d}}A_i + \frac{n}{d}$ . In the following it will be advantageous to use both variables, so we use the notation  $A(\alpha)$  and  $\alpha(A)$  to move back and forth between them.

**Lemma 6.1.** *Let  $C > 0$  be a constant and  $0 < \varepsilon < \frac{1}{2}(d-2)^{-1}$  (for  $d=2$ ,  $\varepsilon > 0$  can be chosen arbitrary). Let  $f_N$  be a function on the set of centered and rescaled Young diagrams (see Eq. (4.3)) that grows at most polynomially in  $N$ , and for  $N$  large enough and all arguments  $A$  such that  $\|A\|_1 \leq n^\varepsilon$  fulfills the bound*

$$f_N(A) \leq CN.$$

*Then the limit of its expectation values does not change when removing degenerate and large deviation diagrams,*

$$\lim_{N \rightarrow \infty} \mathbb{E}_\alpha[f_N(\mathbf{A})] = \lim_{N \rightarrow \infty} \mathbb{E}_\alpha[f_N(\mathbf{A})H(n^\varepsilon - \|\mathbf{A}\|_1)\mathbb{1}_{\text{ND}}(\mathbf{A})],$$

where  $\mathbb{1}_{\text{ND}}$  is the indicator function that is 0 if two or more entries of its argument are equal, and 1 else. Moreover we have the stronger statement

$$|\mathbb{E}_\alpha[f_N(\mathbf{A})] - \mathbb{E}_\alpha[f_N(\mathbf{A})H(n^\varepsilon - \|\mathbf{A}\|_1)\mathbb{1}_{\text{ND}}(\mathbf{A})]| = O(N^{-1/2+(d-2)\varepsilon}).$$

*Proof.* The number of all Young diagrams is bounded from above by a polynomial in  $N$ . But  $p_{d,n}(\alpha(A)) = O(\exp(-\gamma\|A\|_1^2))$  for some  $\gamma > 0$  according to Lemma 4.2, which implies that

$$\lim_{N \rightarrow \infty} \mathbb{E}_\alpha[f_N(\mathbf{A})] = \lim_{N \rightarrow \infty} \mathbb{E}_\alpha[f_N(\mathbf{A})H(n^\varepsilon - \|\mathbf{A}\|_1)].$$

Let us now look at the case of degenerate diagrams. Define the set of degenerate diagrams that are also in the support of the above expectation value,

$$\begin{aligned} \Xi &= \left\{ \alpha \vdash_d n : \exists 1 \leq i \leq d-1 \text{ s.t. } \alpha_i = \alpha_{i+1} \wedge \left(\frac{n}{d}\right)^{-1/2} \left\| \alpha - \frac{n}{d} \mathbf{1} \right\|_1 \leq n^\varepsilon \right\} \\ &= \text{ND}(d, n)^c \cap \text{supp}(H(n^\varepsilon - \|A\|_1)). \end{aligned}$$

Here,  $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^d$  is the all-one vector. We write

$$\Xi = \bigcup_{k=1}^{d-1} \Xi_k \tag{6.1}$$

with

$$\Xi_k = \left\{ \alpha \vdash_d n : \alpha_k = \alpha_{k+1} \wedge \left(\frac{n}{d}\right)^{-1/2} \left\| \alpha - \frac{n}{d} \mathbf{1} \right\|_1 \leq n^\varepsilon \right\}.$$

It suffices to show that

$$\lim_{N \rightarrow \infty} \mathbb{E}_\alpha [f(A(\alpha)) H(n^\varepsilon - \|A(\alpha)\|_1) \mathbb{1}_{\Xi_k}(\alpha)] = 0$$

for all  $k = 1, \dots, d-1$ . We can now apply Lemma 4.2 to  $\Xi_k$  and choose the constants  $\gamma_k = 0$  and  $\gamma_i = \frac{1}{2} + \varepsilon$  for  $i \neq k$ . Using (4.11), the 1-norm condition on  $A$  and bounding the exponential function by a constant we therefore get the bound

$$p_{d,n}(\alpha(A)) \leq C_1 n^{-\frac{d+1}{2}}$$

for some constant  $C_1 > 0$ . The cardinality of  $\Xi_k$  is not greater than the number of integer vectors whose entries are between  $n/d - n^{1/2+\varepsilon}$  and  $n/d + n^{1/2+\varepsilon}$  and sum to  $n$ , and for which the  $k$ th and  $(k+1)$ st entries are equal. It therefore holds that

$$|\Xi_k| \leq C_2 n^{(d-2)\left(\frac{1}{2}+\varepsilon\right)}.$$

By assumption,

$$f(A) \leq Cn \quad \text{for all } A \text{ such that } \alpha(A) \in \Xi_k.$$

Finally, we conclude that

$$\begin{aligned} \mathbb{E}_\alpha [f(A(\alpha)) H(n^\varepsilon - \|A(\alpha)\|_1) \mathbb{1}_{\Xi_k}(\alpha)] &\leq C C_1 C_2 n \cdot n^{-\frac{d+1}{2}} n^{(d-2)\left(\frac{1}{2}+\varepsilon\right)} \\ &\leq \tilde{C} n^{(d-2)\varepsilon - \frac{1}{2}}. \end{aligned}$$

This implies that we have indeed that

$$\lim_{N \rightarrow \infty} \mathbb{E}_\alpha [f(A(\alpha)) H(n^\varepsilon - \|A(\alpha)\|_1) \mathbb{1}_{\Xi_k}(\alpha)] = 0.$$

In fact, we obtain the stronger statement

$$\left| \mathbb{E}_\alpha [f(A(\alpha)) H(n^\varepsilon - \|A(\alpha)\|_1) \mathbb{1}_{\Xi_k}(\alpha)] \right| = O(N^{-1/2+(d-2)\varepsilon}).$$

The statement follows now using Eq. (6.1).  $\square$

With Lemma 6.1 in hand, we can now prove the main result of this section, which we stated in Sect. 1.2 and restate here for convenience.



**Theorem 1.2** (Restated). *For arbitrary but fixed  $d$  and any  $\delta > 0$ , the entanglement fidelity of the standard protocol of PBT is given by*

$$F_d^{\text{std}}(N) = 1 - \frac{d^2 - 1}{4N} + O(N^{-\frac{3}{2} + \delta}).$$

*Proof.* We first define  $n = N - 1$  and recall (3.5), which we can rewrite as follows:

$$\begin{aligned} F_d^{\text{std}}(N) &= d^{-N-2} \sum_{\alpha \vdash_{-d} n} \left( \sum_{\mu = \alpha + \square} \sqrt{d_{\mu} m_{d, \mu}} \right)^2 \\ &= d^{-N-2} \sum_{\alpha \vdash_{-d} n} d_{\alpha} m_{d, \alpha} \left( \sum_{\mu = \alpha + \square} \frac{m_{d, \mu}}{m_{d, \alpha}} \sqrt{d_{\mu} m_{d, \alpha}} \right)^2 \\ &= d^{-N-2} \sum_{\alpha \vdash_{-d} n} d_{\alpha} m_{d, \alpha} \left( \sum_{\mu = \alpha + e_i \text{ YD}} \left[ \prod_{j: j \neq i} \frac{\alpha_i - \alpha_j + j - i + 1}{\alpha_i - \alpha_j + j - i} \right] \sqrt{\frac{N}{\mu_i - i + d}} \right)^2 \\ &= \frac{1}{d^3} \mathbb{E}_{\alpha} \left[ \left( \sum_{\mu = \alpha + e_i \text{ YD}} \left[ \prod_{j: j \neq i} \frac{\alpha_i^{(n)} - \alpha_j^{(n)} + j - i + 1}{\alpha_i^{(n)} - \alpha_j^{(n)} + j - i} \right] \sqrt{\frac{N}{\mu_i^{(N)} - i + d}} \right)^2 \right] \\ &= \frac{1}{d^3} \mathbb{E}_{\alpha} \left[ \left( \sum_{\mu = \alpha + e_i \text{ YD}} \left[ \prod_{j: j \neq i} \left( 1 + \frac{1}{\alpha_i^{(n)} - \alpha_j^{(n)} + j - i} \right) \right] \sqrt{\frac{N}{\mu_i^{(N)} - i + d}} \right)^2 \right]. \end{aligned}$$

In the third step, we used Lemma 3.7 for the term  $\frac{d_{\mu} m_{d, \alpha}}{m_{d, \mu} d_{\alpha}}$  and the Weyl dimension formula (2.3) for the term  $\frac{m_{d, \mu}}{m_{d, \alpha}}$ . The expectation value refers to a random choice of  $\alpha \vdash_{-d} n$  according to the Schur–Weyl distribution  $p_{d, n}$ . The sum over  $\mu = \alpha + e_i$  is restricted to only those  $\mu$  that are valid Young diagrams, i.e., where  $\alpha_{i-1} > \alpha_i$ , which we indicate by writing ‘YD’. Hence, we have

$$\begin{aligned} N(1 - F_d^{\text{std}}(N)) &= \frac{N}{d^2} \mathbb{E}_{\alpha} \left[ d^2 - \left( \sum_{\mu = \alpha + e_i \text{ YD}} \left[ \prod_{j \neq i} \left( 1 + \frac{1}{\alpha_i^{(n)} - \alpha_j^{(n)} + j - i} \right) \right] \sqrt{\frac{N/d}{\mu_i^{(N)} - i + d}} \right)^2 \right]. \end{aligned} \quad (6.2)$$

In the following, we suppress the superscript indicating  $n = N - 1$  for the sake of readability. The random variables  $\alpha$ ,  $\mathbf{A}$ , and  $\Gamma_{ij}$ , as well as their particular values  $\alpha$ ,  $A$ , and  $\Gamma_{ij}$ , are all understood to be functions of  $n = N - 1$ .

The function

$$f_N(A) := \frac{N}{d^2} \left( d^2 - \left( \sum_{\mu = \alpha(A) + e_i \text{ YD}} \left[ \prod_{j: j \neq i} \left( 1 + \frac{1}{\alpha_i(A) - \alpha_j(A) + j - i} \right) \right] \sqrt{\frac{N/d}{\mu_i - i + d}} \right)^2 \right)$$

satisfies the requirements of Lemma 6.1. Indeed we have that

$$\frac{1}{\alpha_i(A) - \alpha_j(A) + j - i} \leq 1$$

for all  $i \neq j$ , and clearly

$$\sqrt{\frac{N/d}{\mu_i - i + d}} \leq \sqrt{N}.$$

Therefore we get

$$f_N(A) \leq CN^2$$

for some constant  $C$ . If  $\|A\|_1 \leq n^\varepsilon$ , we have that

$$\sqrt{\frac{N/d}{\mu_i - i + d}} \leq \sqrt{\frac{N/d}{n/d - n^\varepsilon}} \leq \sqrt{N/n} + O\left(n^{-(1-\varepsilon)}\right)$$

and hence

$$f_N(A) \leq CN$$

for  $N$  large enough. We therefore define, using an  $\varepsilon$  in the range given by Lemma 6.1, the modified expectation value

$$\tilde{\mathbb{E}}_\alpha[f(\alpha)] := \mathbb{E}_\alpha[f(\alpha)\mathbb{1}_{\text{ND}(n,d)}(\alpha)H(n^\varepsilon - \|A(\alpha)\|_1)], \quad (6.3)$$

and note that an application of Lemma 6.1 shows that the limit that we are striving to calculate does not change when replacing the expectation value with the above modified expectation value, and the difference between the members of the two sequences is  $O(n^{-\frac{1}{2}+\varepsilon(d-2)})$ .

For a non-degenerate  $\alpha$ , adding a box to any row yields a valid Young diagram  $\mu$ . Hence, the sum  $\sum_{\mu=\alpha+e_i} \text{YD}$  in (6.2) can be replaced by  $\sum_{i=1}^d$ , at the same time replacing  $\mu_i$  with  $\alpha_i + 1$ . The expression in (6.2) therefore simplifies to

$$R_N := \frac{N}{d^2} \tilde{\mathbb{E}}_\alpha \left[ d^2 - \left( \sum_{i=1}^d \left[ \prod_{k:k \neq i} \left( 1 + \frac{1}{\alpha_i - \alpha_k + k - i} \right) \right] \sqrt{\frac{N/d}{\alpha_i + 1 - i + d}} \right)^2 \right].$$

Let us look at the square root term, using the variables  $\mathbf{A}_i$ . For sufficiently large  $n$ , we write

$$\begin{aligned} \sqrt{\frac{N/d}{\alpha_i + 1 - i + d}} &= \sqrt{N/n} \left( 1 + \frac{(1-i+d)d}{n} + \sqrt{\frac{d}{n}} \mathbf{A}_i \right)^{-1/2} \\ &= \sqrt{\frac{N\gamma_{i,d,n}}{n}} \left( 1 + \gamma_{i,d,n} \sqrt{\frac{d}{n}} \mathbf{A}_i \right)^{-1/2} \\ &= \sqrt{\frac{N\gamma_{i,d,n}}{n}} \sum_{r=0}^{\infty} a_r \left( \gamma_{i,d,n} \sqrt{\frac{d}{n}} \mathbf{A}_i \right)^r. \end{aligned}$$

In the second line we have defined

$$\gamma_{i,d,n} = \left( 1 + \frac{(1-i+d)d}{n} \right)^{-1},$$

and in the third line we have written the inverse square root in terms of its power series around 1. This is possible as we have  $\|\mathbf{A}\|_1 \leq n^\varepsilon$  on the domain of  $\tilde{\mathbb{E}}$ , so  $\gamma_{i,d,n}\sqrt{\frac{d}{n}}\mathbf{A}_i = O(n^{-1/2+\varepsilon})$ , i.e., it is in particular in the convergence radius of the power series, which is equal to 1. This implies also that the series converges absolutely in that range. Defining

$$\Gamma_{ik} = -\Gamma_{ki} = \mathbf{A}_i - \mathbf{A}_k + \frac{k-i}{\sqrt{\frac{n}{d}}}$$

as in Sect. 4, we can write

$$\begin{aligned} R_N &= \frac{N}{d^2} \tilde{\mathbb{E}}_\alpha \left[ d^2 - \frac{N}{n} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n}\gamma_{j,d,n}} \left[ \prod_{k:k<i} \left( 1 - \sqrt{\frac{d}{n}} \Gamma_{ki}^{-1} \right) \right] \right. \\ &\quad \left. \left[ \prod_{k:k>i} \left( 1 + \sqrt{\frac{d}{n}} \Gamma_{ik}^{-1} \right) \right] \right. \\ &\quad \left. \left[ \prod_{l:l<j} \left( 1 - \sqrt{\frac{d}{n}} \Gamma_{lj}^{-1} \right) \right] \left[ \prod_{l:l>j} \left( 1 + \sqrt{\frac{d}{n}} \Gamma_{jl}^{-1} \right) \right] \right. \\ &\quad \left. \left( \sum_{r=0}^{\infty} a_r \left( \gamma_{i,d,n} \sqrt{\frac{d}{n}} \mathbf{A}_i \right)^r \right) \left( \sum_{r=0}^{\infty} a_r \left( \gamma_{j,d,n} \sqrt{\frac{d}{n}} \mathbf{A}_j \right)^r \right) \right] \\ &=: \frac{N}{d^2} \tilde{\mathbb{E}}_\alpha \left[ d^2 - \frac{N}{n} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n}\gamma_{j,d,n}} \left( \sum_{s=0}^{2(d-1)} \binom{d}{n}^{\frac{s}{2}} P_{i,j}^{(1,s)}(\Gamma^{-1}) \right) \right. \\ &\quad \left. \left( \sum_{r=0}^{\infty} \binom{d}{n}^{\frac{r}{2}} P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right) \right]. \end{aligned}$$

Here we have defined  $\tilde{\mathbf{A}}$  by  $\tilde{\mathbf{A}}_i = \gamma_{i,d,n}\mathbf{A}_i$  and the polynomials  $P_{i,j}^{(1,s)}$ ,  $P_{i,j}^{(2,r)}$ , for  $s = 0, \dots, 2(d-1)$ ,  $r \in \mathbb{N}$ ,  $i, j = 1, \dots, d$ , which are homogeneous of degree  $r$ , and  $s$ , respectively. In the last equality we have used the absolute convergence of the power series. We have also abbreviated  $\Gamma := (\Gamma_{ij})_{i<j}$ ,  $\Gamma^{-1}$  is to be understood elementwise, and  $P_{i,j}^{(1,s)}$  has the additional property that for all  $k, l \in \{1, \dots, d\}$  it has degree at most 2 in each variable  $\Gamma_{k,l}$ .

By the Fubini-Tonelli Theorem, we can now exchange the infinite sum and the expectation value if the expectation value

$$\tilde{\mathbb{E}}_\alpha \left[ \left( \sum_{s=0}^{2(d-1)} \binom{d}{n}^{\frac{s}{2}} \tilde{P}_{i,j}^{(1,s)}(|\Gamma^{-1}|) \right) \left( \sum_{r=0}^{\infty} \binom{d}{n}^{\frac{r}{2}} \tilde{P}_{i,j}^{(2,r)}(|\tilde{\mathbf{A}}|) \right) \right]$$

exists, where the polynomials  $\tilde{P}_{i,j}^{(1,s)}$  and  $\tilde{P}_{i,j}^{(2,r)}$  are obtained from  $P_{i,j}^{(1,s)}$  and  $P_{i,j}^{(2,r)}$ , respectively, by replacing the coefficients with their absolute value, and the absolute values  $|\Gamma^{-1}|$  and  $|\tilde{\mathbf{A}}|$  are to be understood element-wise. But the power series of the square root we have used converges absolutely on the range of  $\mathbf{A}$  restricted by  $\tilde{\mathbb{E}}$  (see Eq. (6.3)), yielding a continuous function on an appropriately chosen compact interval.

Moreover, if  $A$  is in the range of  $\mathbf{A}$  restricted by  $\tilde{\mathbb{E}}$ , then so is  $|A|$ . The function is therefore bounded, as is  $\tilde{\mathbf{A}}$  for fixed  $N$ , and the expectation value above exists. We therefore get

$$R_N = \frac{N}{d^2} \left[ d^2 - \frac{N}{n} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} \sum_{s=0}^{2(d-1)} \sum_{r=0}^{\infty} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \right].$$

Now note that the expectation values above have the right form to apply Theorem 4.1, so we can start calculating expectation values provided that we can exchange the limit  $N \rightarrow \infty$  with the infinite sum. We can then split up the quantity  $\lim_{N \rightarrow \infty} R_N$  as follows,

$$\begin{aligned} \lim_{N \rightarrow \infty} R_N &= \lim_{N \rightarrow \infty} \frac{N}{d^2} \left[ d^2 - \frac{N}{n} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} \sum_{s=0}^{2(d-1)} \sum_{r=0}^{\infty} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} \right. \\ &\quad \left. \tilde{\mathbb{E}}_{\alpha} \left[ P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \right] \\ &= \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \tilde{\mathbb{E}}_{\alpha} \left[ \frac{d^2 n}{N} - \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,0)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,0)}(\tilde{\mathbf{A}}) \right] \end{aligned} \quad (6.4)$$

$$\begin{aligned} &- \sum_{r,s \in \mathbb{N}: 1 \leq r+s \leq 2} \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{r+s}{2}} \\ &\quad \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \end{aligned} \quad (6.5)$$

$$\begin{aligned} &- \lim_{N \rightarrow \infty} \sum_{\substack{r,s \in \mathbb{N} \\ r+s \geq 3 \\ s \leq 2(d-1)}} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{r+s}{2}} \\ &\quad \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right], \end{aligned} \quad (6.6)$$

provided that all the limits on the right hand side exist. We continue by determining these limits and begin with Eq. (6.6). First observe that, for fixed  $r$  and  $s$  such that  $r+s \geq 3$ ,

$$\lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{r+s}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] = 0. \quad (6.7)$$

This is because the expectation value in Eq. (6.7) converges according to Theorem 4.1 and Lemma 6.1, which in turn implies that the whole expression is  $O(N^{-1/2})$ . In particular, there exists a constant  $K > 0$  such that, for the finitely many values of  $r$  and  $s$  such that  $r \leq r_0 := \lceil (\frac{1}{2} - \varepsilon)^{-1} \rceil$ ,

$$\frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{r+s}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \leq K \quad (\forall N).$$

Now suppose that  $r > r_0$ . On the domain of  $\tilde{\mathbb{E}}$ , we have  $\|\mathbf{A}\|_1 \leq n^\varepsilon$ . Therefore, we can bound

$$\frac{N^2}{nd} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \leq CN^{1+r(\varepsilon-1/2)} \leq C2^{1+r(\varepsilon-1/2)} \quad (\forall N) \quad (6.8)$$

The first step holds because  $\left(\frac{d}{n}\right)^{\frac{s}{2}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1})$  is a polynomial in the variables  $\left(\frac{d}{n}\right)^{\frac{1}{2}} \mathbf{\Gamma}_{ij}^{-1} \leq 1$  with coefficients independent of  $n$ , and in the second step we used that  $1+r(\varepsilon-1/2) < 0$ . We can therefore apply the dominated convergence theorem using the dominating function

$$g(r, s) = \begin{cases} K & r \leq r_0 = \lceil (\frac{1}{2} - \varepsilon)^{-1} \rceil \\ C \cdot 2^{1+r(\varepsilon-1/2)} & \text{else} \end{cases}$$

to exchange the limit and the sum in Eq. (6.6). Thus, Eq. (6.7) implies that Eq. (6.6) is zero.

It remains to compute the limits Eqs. (6.4) and (6.5), i.e., the terms

$$T_{s,r} := \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} \tilde{\mathbb{E}} \alpha \left[ \delta_{s0} \delta_{r0} \frac{d^2 n}{N} - \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right].$$

for  $r+s = 0, 1, 2$ . The first few terms of the power series for the inverse square root are given by

$$(1+x)^{-1/2} = 1 - \frac{x}{2} + \frac{3x^2}{8} + O(x^3).$$

The relevant polynomials to calculate the remaining limits are, using the above and  $-\mathbf{\Gamma}_{ik} = \mathbf{\Gamma}_{ki}$ ,

$$\begin{aligned} P_{i,j}^{(1,0)} &= P_{i,j}^{(1,0)} \equiv 1 \\ P_{i,j}^{(1,1)}(\mathbf{\Gamma}^{-1}) &= - \sum_{k:k<i} \mathbf{\Gamma}_{k:ki}^{-1} + \sum_{k>i} \mathbf{\Gamma}_{ik}^{-1} - \sum_{l:l<j} \mathbf{\Gamma}_{lj}^{-1} + \sum_{l:l>j} \mathbf{\Gamma}_{jl}^{-1} \\ &= \sum_{k:k \neq i} \mathbf{\Gamma}_{ik}^{-1} + \sum_{l:l \neq j} \mathbf{\Gamma}_{jl}^{-1} \\ P_{i,j}^{(2,1)}(\tilde{\mathbf{A}}) &= a_1 a_0 (\tilde{\mathbf{A}}_i + \tilde{\mathbf{A}}_j) = -\frac{1}{2} (\tilde{\mathbf{A}}_i + \tilde{\mathbf{A}}_j) \\ P_{i,j}^{(1,2)}(\mathbf{\Gamma}^{-1}) &= \sum_{k,l:i \neq k \neq l \neq i} \mathbf{\Gamma}_{ik}^{-1} \mathbf{\Gamma}_{il}^{-1} + \sum_{k,l:j \neq k \neq l \neq j} \mathbf{\Gamma}_{ik}^{-1} \mathbf{\Gamma}_{il}^{-1} + \sum_{k,l:i \neq k, l \neq j} \mathbf{\Gamma}_{ik}^{-1} \mathbf{\Gamma}_{il}^{-1} \\ P_{i,j}^{(2,2)}(\tilde{\mathbf{A}}) &= a_2 a_0 (\tilde{\mathbf{A}}_i^2 + \tilde{\mathbf{A}}_j^2) + a_1^2 \tilde{\mathbf{A}}_i \tilde{\mathbf{A}}_j = \frac{3}{8} (\tilde{\mathbf{A}}_i^2 + \tilde{\mathbf{A}}_j^2) + \frac{1}{4} \tilde{\mathbf{A}}_i \tilde{\mathbf{A}}_j. \end{aligned}$$

We now analyze the remaining expectation values using these explicit expressions for the corresponding polynomials.

*Evaluating  $T_{0,0}$ .* Using the power series expansions

$$\sqrt{\gamma_{i,d,n}} = 1 - \frac{d(d+1-i)}{2n} + O(n^{-2}) = 1 - O(n^{-1}),$$

we simplify

$$\begin{aligned} T_{0,0} &= \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \tilde{\mathbb{E}}_{\alpha} \left[ \frac{d^2 n}{N} - \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,0)}(\Gamma^{-1}) P_{i,j}^{(2,0)}(\tilde{\mathbf{A}}) \right] \\ &= \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left( \frac{d^2 n}{N} - \sum_{i,j=1}^d \left( 1 - \frac{d(d+1-i)}{2n} - \frac{d(d+1-j)}{2n} + O(n^{-2}) \right) \right) \\ &= \lim_{N \rightarrow \infty} \left( N - \frac{N^2}{n} + \frac{N^2}{nd^2} \sum_{i,j=1}^d \left( \frac{d(2d+2-i-j)}{2n} + O(n^{-2}) \right) \right) \\ &= \lim_{N \rightarrow \infty} \left( -\frac{n+1}{n} + \frac{N^2}{2n^2} (2d^2 + 2d - d(d+1)) + O(n^{-1}) \right) \\ &= \frac{d(d+1)}{2} - 1. \end{aligned} \tag{6.9}$$

In the second-to-last line we have replaced  $N = n + 1$ .

*Evaluating  $T_{0,1}$  and  $T_{1,0}$ .* We first compute

$$\begin{aligned} &\sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,1)}(\Gamma^{-1}) P_{i,j}^{(2,0)}(\tilde{\mathbf{A}}) \\ &= \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right) \\ &= \sum_{i,j=1}^d (1 + O(n^{-1})) \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right) \\ &= \sum_{i,j=1}^d O(n^{-1}) \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right). \end{aligned}$$

In the last equation we have used that

$$\sum_{i \neq k} \Gamma_{ik}^{-1} = 0,$$

as the summation domain is symmetric in  $i$  and  $k$ , and  $\Gamma_{ik}^{-1} = -\Gamma_{ki}^{-1}$ . But now we can determine the limit,

$$\begin{aligned}
 T_{1,0} &= \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{1}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,1)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,0)}(\tilde{\mathbf{A}}) \right] \\
 &= \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{1}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d O(n^{-1}) \left( \sum_{k:k \neq i} \mathbf{\Gamma}_{ik}^{-1} + \sum_{l:l \neq j} \mathbf{\Gamma}_{jl}^{-1} \right) \right] \\
 &= \sum_{i,j=1}^d \lim_{N \rightarrow \infty} O(n^{-1/2}) \tilde{\mathbb{E}}_{\alpha} \left[ \left( \sum_{k:k \neq i} \mathbf{\Gamma}_{ik}^{-1} + \sum_{l:l \neq j} \mathbf{\Gamma}_{jl}^{-1} \right) \right] = 0. \tag{6.10}
 \end{aligned}$$

Here we have used Theorem 4.1 to see that the sequence of expectation values converges, implying that the expression vanishes due to the  $O(n^{-1/2})$  prefactor.

Similarly, to show that  $T_{0,1}$  vanishes as well, we calculate

$$\begin{aligned}
 &\sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,0)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,1)}(\tilde{\mathbf{A}}) \\
 &= -\frac{1}{2} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} (\tilde{\mathbf{A}}_i + \tilde{\mathbf{A}}_j) \\
 &= -\frac{1}{2} \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} (\gamma_{i,d,n} \mathbf{A}_i + \gamma_{j,d,n} \mathbf{A}_j) \\
 &= -\sum_{i,j=1}^d (1 + O(n^{-1})) (\mathbf{A}_i + \mathbf{A}_j) \\
 &= -\sum_{i,j=1}^d O(n^{-1}) (\mathbf{A}_i + \mathbf{A}_j).
 \end{aligned}$$

Here we have used that  $\sqrt{\gamma_{i,d,n}} = 1 - O(n^{-1}) = \gamma_{i,d,n}$ , and in the last line we used that  $\sum_{i=1}^d \mathbf{A}_i = 0$ . This implies, using the same argument as in Eq. (6.10), that

$$T_{0,1} = \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{1}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,0)}(\mathbf{\Gamma}^{-1}) P_{i,j}^{(2,1)}(\tilde{\mathbf{A}}) \right] = 0. \tag{6.11}$$

Evaluating  $T_{s,r}$  for  $s+r=2$ . For  $s+r=2$ , we first observe that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\Gamma^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \\ &= \lim_{N \rightarrow \infty} \frac{N^2}{n^2 d} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\Gamma^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \\ &= \frac{1}{d} \lim_{N \rightarrow \infty} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\Gamma^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right]. \end{aligned}$$

Therefore we can replace all occurrences of  $\gamma_{i,d,n}$  by 1 using the same argument as in Eq. (6.10). There are three cases to take care of,  $(s,r) = \{(2,0), (1,1), (0,2)\}$ . For  $(s,r) = (2,0)$ , we first look at the term

$$\begin{aligned} \sum_{i,j} \sum_{k,l:i \neq k \neq l \neq i} \Gamma_{ik}^{-1} \Gamma_{il}^{-1} &= d \sum_{i,k,l:i \neq k \neq l \neq i} \Gamma_{ik}^{-1} \Gamma_{il}^{-1} \\ &= d \sum_{i,k,l:i \neq k \neq l \neq i} \frac{1}{(\mathcal{A}_i - \mathcal{A}_k)(\mathcal{A}_i - \mathcal{A}_l)}, \end{aligned}$$

where we have defined  $\mathcal{A}_i = \mathbf{A}_i - i\sqrt{\frac{d}{n}}$ . For fixed  $i_0 \neq k_0 \neq j_0 \neq i_0$ , all permutations of these indices appear in the sum. For these terms with  $i, j, k \in \{i_0, j_0, k_0\}$ ,

$$\begin{aligned} & \sum_{\substack{i,j,k \in \{i_0, j_0, k_0\} \\ i \neq k \neq j \neq i}} \frac{1}{(\mathcal{A}_i - \mathcal{A}_j)(\mathcal{A}_i - \mathcal{A}_k)} \\ &= \frac{2}{(\mathcal{A}_{i_0} - \mathcal{A}_{j_0})(\mathcal{A}_{j_0} - \mathcal{A}_{k_0})(\mathcal{A}_{k_0} - \mathcal{A}_{i_0})} \\ & \quad \left( -(\mathcal{A}_{j_0} - \mathcal{A}_{k_0}) - (\mathcal{A}_{k_0} - \mathcal{A}_{i_0}) - (\mathcal{A}_{i_0} - \mathcal{A}_{j_0}) \right) = 0, \end{aligned}$$

implying

$$\sum_{i,j} \sum_{k,l:i \neq k \neq l \neq i} \Gamma_{ik}^{-1} \Gamma_{il}^{-1} = 0$$

and therefore

$$T_{2,0} = \lim_{N \rightarrow \infty} \frac{N^2}{n^2 d} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,2)}(\Gamma^{-1}) P_{i,j}^{(2,0)}(\tilde{\mathbf{A}}) \right] = 0. \quad (6.12)$$

Moving on to the case  $(s,r) = (0,2)$ , we first note that, as in the previous case  $(s,r) = (2,0)$  and again replacing all occurrences of  $\gamma_{i,d,n}$  by 1, we have



$$\begin{aligned}
 & \lim_{N \rightarrow \infty} \tilde{\mathbb{E}} \left[ \sum_{i,j=1}^d P_{i,j}^{(2,2)}(\tilde{\mathbf{A}}) \right] \\
 &= \sum_{i,j=1}^d \lim_{N \rightarrow \infty} \tilde{\mathbb{E}} [P_{i,j}^{(2,2)}(\mathbf{A})] \\
 &= \sum_{i,j=1}^d \mathbb{E} \left[ \frac{3}{8} (\mathbf{S}_i^2 + \mathbf{S}_j^2) + \frac{1}{4} \mathbf{S}_i \mathbf{S}_j \right] \\
 &= \frac{3d}{4} \mathbb{E} \left[ \sum_{i=1}^d \mathbf{S}_i^2 \right] + \frac{1}{4} \mathbb{E} \left[ \sum_{i,j=1}^d \mathbf{S}_i \mathbf{S}_j \right]. \tag{6.13}
 \end{aligned}$$

Here,  $\mathbf{S} = \text{spec}(\mathbf{G}) \sim \text{GUE}_d^0$  and we have used Lemma 6.1 to switch back to the unrestricted expectation value and Theorem 4.1 in the second equality. First we observe that  $\mathbf{G}$  is traceless, and hence  $\sum_{i=1}^d \mathbf{S}_i = 0$  such that the second term in (6.13) vanishes. For the first term in (6.13), let  $\mathbf{X} \sim \text{GUE}_d$  such that  $\mathbf{G} = \mathbf{X} - \frac{\text{tr}(\mathbf{X})}{d} \mathbf{I} \sim \text{GUE}_d^0$ . We calculate

$$\begin{aligned}
 \sum_i \mathbf{S}_i^2 &= \text{tr}(\mathbf{G}^2) \\
 &= \text{tr}(\mathbf{X}^2) - \frac{1}{d} \text{tr}(\mathbf{X})^2 \\
 &= \sum_{i,j} |\mathbf{X}_{ij}|^2 - \frac{1}{d} \left( \sum_i \mathbf{X}_{ii} \right)^2 \\
 &= \sum_i \mathbf{X}_{ii}^2 + \sum_{i \neq j} |\mathbf{X}_{ij}|^2 - \frac{1}{d} \sum_i \mathbf{X}_{ii}^2 - \frac{2}{d} \sum_{i \neq j} \mathbf{X}_{ii} \mathbf{X}_{jj}. \tag{6.14}
 \end{aligned}$$

We have  $\mathbb{E}[\mathbf{X}_{ii}^2] = \mathbb{V}[\mathbf{X}_{ii}] = 1$ , where  $\mathbb{V}[\cdot]$  denotes the variance of a random variable. Similarly, for  $i \neq j$ ,

$$\mathbb{E}[|\mathbf{X}_{ij}|^2] = \mathbb{V}[\Re(\mathbf{X}_{ij})] + \mathbb{V}[\Im(\mathbf{X}_{ij})] = \frac{1}{2} + \frac{1}{2} = 1,$$

and  $\mathbb{E}[\mathbf{X}_{ii} \mathbf{X}_{jj}] = \mathbb{E}[\mathbf{X}_{ii}] \mathbb{E}[\mathbf{X}_{jj}] = 0$ , since the entries of a  $\text{GUE}_d$ -matrix are independent. Hence, taking expectation values in (6.14) gives

$$\mathbb{E}_\alpha \left[ \sum_i \mathbf{S}_i^2 \right] = d + d(d-1) - 1 = d^2 - 1,$$

and we can calculate

$$T_{0,2} = \lim_{N \rightarrow \infty} \frac{N^2}{n^2 d} \tilde{\mathbb{E}}_\alpha \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,0)}(\Gamma^{-1}) P_{i,j}^{(2,2)}(\tilde{\mathbf{A}}) \right] = \frac{3}{4} (d^2 - 1). \tag{6.15}$$

We finally turn to the only missing case,  $(s, r) = (1, 1)$ . The polynomial  $P_{ij}^{(1,1)}$  is symmetric in  $i$  and  $j$ , therefore we can simplify

$$\begin{aligned}
& \sum_{i,j} P_{ij}^{(1,1)}(\Gamma^{-1}) P_{ij}^{(2,1)}(\Gamma^{-1}) \\
&= -\frac{1}{2} \sum_{i,j} \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right) (\tilde{\mathbf{A}}_i + \tilde{\mathbf{A}}_j) \\
&= -\sum_{i,j} \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right) \tilde{\mathbf{A}}_i \\
&= -\sum_{i,j} \left( \sum_{k:k \neq i} \Gamma_{ik}^{-1} + \sum_{l:l \neq j} \Gamma_{jl}^{-1} \right) \mathbf{A}_i \\
&= -d \sum_{i,k:k \neq i} \Gamma_{ik}^{-1} \mathbf{A}_i, \tag{6.16}
\end{aligned}$$

where we have used in the second-to-last equation that we can replace any occurrence of  $\gamma_{i,d,n}$  by one, and the last equation follows by the same reasoning as used in the case  $(s, r) = (0, 1)$  above. Now observe that for each  $i \neq k$ , both  $\Gamma_{ik}^{-1}$  and  $\Gamma_{ki}^{-1} = -\Gamma_{ik}^{-1}$  occur in the sum. Therefore we can simplify

$$\begin{aligned}
\sum_{i,k:k \neq i} \Gamma_{ik}^{-1} \mathbf{A}_i &= \sum_{i,k:i < k} \Gamma_{ik}^{-1} (\mathbf{A}_i - \mathbf{A}_k) \\
&= \sum_{i,k:i < k} \Gamma_{ik}^{-1} \left( \mathbf{A}_i - \mathbf{A}_k + \frac{k-i}{\sqrt{\frac{n}{d}}} \right) + \sqrt{\frac{d}{n}} \sum_{i,k:i < k} (i-k) \Gamma_{ik}^{-1} \\
&= \sum_{i,k:i < k} \Gamma_{ik}^{-1} \Gamma_{ik} + \sqrt{\frac{d}{n}} \sum_{i,k:i < k} (i-k) \Gamma_{ik}^{-1} \\
&= \frac{d(d-1)}{2} - O(n^{-1/2}) \sum_{i,k:i < k} (i-j) \Gamma_{ik}^{-1}, \tag{6.17}
\end{aligned}$$

where we have used the definition of  $\Gamma_{ij}$  in the last equality. Combining Eqs. (6.16) and (6.17) we arrive at

$$T_{1,1} = \lim_{N \rightarrow \infty} \frac{N^2}{n^2 d} \tilde{\mathbb{E}}_{\alpha} \left[ \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,1)}(\Gamma^{-1}) P_{i,j}^{(2,1)}(\tilde{\mathbf{A}}) \right] = -\frac{d(d-1)}{2}. \tag{6.18}$$

Collecting all the terms  $T_{r,s}$  for  $r+s \leq 2$  that we have calculated in Eqs. (6.9) to (6.12), (6.15) and (6.18), we arrive at

$$\begin{aligned}
 \lim_{N \rightarrow \infty} R_N &= \lim_{N \rightarrow \infty} \sum_{\substack{r,s \in \{0,1,2\} \\ r+s \leq 2}} \frac{N^2}{nd^2} \left(\frac{d}{n}\right)^{\frac{s+r}{2}} \tilde{\mathbb{E}}_{\alpha} \\
 &\quad \left[ \delta_{s0} \delta_{r0} \frac{d^2 n}{N} - \sum_{i,j=1}^d \sqrt{\gamma_{i,d,n} \gamma_{j,d,n}} P_{i,j}^{(1,s)}(\Gamma^{-1}) P_{i,j}^{(2,r)}(\tilde{\mathbf{A}}) \right] \\
 &= T_{0,0} + T_{0,1} + T_{1,0} + T_{0,2} + T_{2,0} + T_{1,1} \\
 &= \frac{d(d+1)}{2} - 1 - \frac{3(d^2-1)}{4} + \frac{d(d-1)}{2} \\
 &= \frac{d^2-1}{4},
 \end{aligned}$$

which implies that

$$\lim_{N \rightarrow \infty} N \left(1 - F_d^{\text{std}}(N)\right) = \frac{d^2-1}{4}.$$

To determine the lower order term, note that in all expressions above we have neglected terms of at most  $O(n^{-1/2+\varepsilon(d-2)})$ . Eq. (6.8) shows that the terms with  $r+s \geq 3$  are  $O(n^{-1/2+3/2\varepsilon})$ , and the difference between  $R_N$  and  $N(1 - F_d^{\text{std}}(N))$  is  $O(n^{-1/2+\varepsilon(d-2)})$  as well. As  $\varepsilon \in (0, (d-2)^{-1})$  was arbitrary we conclude that, for all  $\delta > 0$ ,

$$F_d^{\text{std}}(N) = 1 - \frac{d^2-1}{4N} + O(N^{-3/2+\delta}),$$

which concludes the proof.  $\square$

**6.2. Asymptotics of the optimal protocol.** In this section, our goal is to obtain an asymptotic lower bound on the optimal entanglement fidelity  $F_d^*$  of a deterministic PBT protocol with both the entangled resource state and the POVM optimized. This is achieved by restricting the optimization in Eq. (3.6) to the class of protocol families that use a density  $c_\mu$  such that the probability distribution  $q(\mu) = c_\mu p_{N,d}(\mu)$  converges for  $N \rightarrow \infty$  in a certain sense. We then continue to show that the optimal asymptotic entanglement fidelity within this restricted class is related to the first eigenvalue of the Dirichlet Laplacian on the simplex of ordered probability distributions.

The main result of this section is the following theorem, which we restate from Sect. 1.2 for convenience.

**Theorem 1.4** (Restated). *The optimal fidelity for deterministic port-based teleportation is bounded from below by*

$$F_d^*(N) \geq 1 - \frac{\lambda_1(\text{OS}_{d-1})}{dN^2} - O(N^{-3}),$$

where

$$\text{OS}_{d-1} = \left\{ x \in \mathbb{R}^d \mid \sum_i x_i = 1, x_i \geq x_{i+1}, x_d \geq 0 \right\}$$

is the  $(d-1)$ -dimensional simplex of ordered probability distributions with  $d$  outcomes and  $\lambda_1(\Omega)$  is the first eigenvalue of the Dirichlet Laplacian on a domain  $\Omega$ .

Before commencing the proof of Theorem 1.4, let us build some intuition for the fact that the fidelity formula Eq. (3.6) is related to a Laplacian. Reparametrizing Equation (3.6), we obtain the expression

$$F_d^*(N) = d^{-N-2} \max_p \sum_{\alpha \vdash_d N-1} \left( \sum_{\mu=\alpha+\square} \sqrt{p(\mu)} \right)^2,$$

where the maximization is taken over all probability distributions  $p$  on the set of Young diagrams with  $N$  boxes. Rearranging the sums yields

$$F_d^*(N) = d^{-N-2} \max_p \sum_{\mu, \mu'} \sqrt{p(\mu)} \sqrt{p(\mu')} (1 + (d-1)\delta_{\mu\mu'}) + \{\text{boundary terms}\}, \quad (6.19)$$

where the sum is taken over all pairs  $\mu, \mu' \vdash_d N$  such that  $\mu'$  can be obtained from  $\mu$  by removing one box and adding one box, and the “boundary terms” subsume differences that arise when it is possible to remove a box from  $\mu$  and add one back such that the result is not a Young diagram (this will be made more rigorous below). It is now instructive to equip the set of Young diagrams with a graph structure, where we draw an edge between  $\mu$  and  $\mu' \neq \mu$  precisely whenever the pair  $\mu, \mu'$  is part of the sum in (6.19). This graph is, in fact, the intersection of the root lattice of  $\mathfrak{su}(d)$  with a certain simplex. Observing that  $\sqrt{p}$  is an  $L^2$ -normalized function, we conclude that Eq. (6.19) is equal to the graph Laplacian on the Young diagram lattice we have defined, up to a constant. Normalizing the Young diagrams as done in Sect. 6.1, we see that the graphs for increasing  $N$  are finer and finer discretizations of the simplex of ordered probability distributions. We can thus expect that, when  $p$  is a sufficiently nice function, these graph Laplacians converge to the continuous Laplacian.

For the proof of Theorem 1.4 it will be convenient to switch back and forth between summation over a lattice and integration, which is the content of Lemma 6.2 below. Before stating the lemma, we make a few definitions. For a set  $\Omega$  we define  $d(x, \Omega) := \inf_{y \in \Omega} \|x - y\|_2$ , and for  $\delta \geq 0$  we define

$$\partial_\delta \Omega := \{x \in \Omega : d(x, \partial\Omega) \leq \delta\}.$$

Let  $V_0^{d-1} = \{x \in \mathbb{R}^d \mid \sum_{i=1}^d x_i = 0\}$  and  $\mathbb{Z}_0^d = \mathbb{Z}^d \cap V_0^{d-1}$ . For a vector subspace  $V \subset \mathbb{R}^d$  and lattice  $\Lambda \subset \mathbb{R}^d$ , we denote by  $v + V$  and  $v + \Lambda$  the affine space and affine lattice with the origin shifted to  $v \in \mathbb{R}^d$ , respectively. We denote by  $\{e_i\}_{i=1}^d$  the standard basis in  $\mathbb{R}^d$ . For  $y \in e_1 + \frac{1}{N}\mathbb{Z}_0^d$ , define  $U_N(y) \subset e_1 + V_0^{d-1}$  by the condition

$$x \in U_N(y) \Leftrightarrow \forall y' \in e_1 + \frac{1}{N}\mathbb{Z}_0^d, y' \neq y : \|x - y\|_2 < \|x - y'\|_2.$$

In other words, up to sets of measure zero we have tiled  $e_1 + V_0^{d-1}$  regularly into neighborhoods of lattice points. This also induces a decomposition  $\text{OS}_{d-1} \subset e_1 + V_0^{d-1}$  via intersection,  $U_N^{\text{OS}}(y) = U_N(y) \cap \text{OS}_{d-1}$ . We define the function  $g_N : e_1 + V_0^{d-1} \rightarrow e_1 + \frac{1}{N}\mathbb{Z}_0^d$  via  $g_N(x) = y$  where  $y$  is the unique lattice point such that  $x \in U_N(y)$ , if such a point exists. On the measure-zero set  $\left(\bigcup_{y \in e_1 + \frac{1}{N}\mathbb{Z}_0^d} U_N(y)\right)^c$ , the function  $g_N$  can be set to an arbitrary value.

**Lemma 6.2.** *Let  $f \in C^1(\text{OS}_{d-1}) \cap C(\mathbb{R}^d)$  be such that  $f(x) = O(d(x, \partial\text{OS}_{d-1})^p)$  for some  $p \geq 1$ , and  $f \equiv 0$  on  $\mathbb{R}^d \setminus \text{OS}_{d-1}$ . Then,*

$$(i) \quad \left| \frac{1}{N^{d-1}} \sum_{y \in \text{OS}_{d-1} \cap \frac{1}{N}\mathbb{Z}^d} f(y) - \int_{\text{OS}_{d-1}} f(g_N(x)) dx \right| \leq O(N^{-p-2});$$

$$(ii) \quad \left| \int_{\text{OS}_{d-1}} f(g_N(x)) dx - \int_{\text{OS}_{d-1}} f(x) dx \right| \leq O(N^{-1}).$$

If furthermore  $f \in C^2(\text{OS}_{d-1})$ , then

$$(iii) \quad \int_{\text{OS}_{d-1}} f(g_N(x))(-\Delta f)(g_N(x)) dx = \int_{\text{OS}_{d-1}} f(x)(-\Delta f)(x) dx + O(N^{-1}).$$

*Proof.* Throughout the proof we set  $\Lambda := \text{OS}_{d-1} \cap \frac{1}{N}\mathbb{Z}^d$ . Observe first that the largest radius of the cell  $U_N(y)$  around  $y \in \frac{1}{N}\mathbb{Z}^d$  is equal to half the length  $\frac{\sqrt{d}}{N}$  of a main diagonal in a  $d$ -dimensional hypercube of length  $\frac{1}{N}$ . Setting  $c := \frac{\sqrt{d}}{2}$ , it follows that  $g_N^{-1}(y) \subseteq \text{OS}_{d-1}$  for all  $y \in \Lambda$  with

$$d(y, \partial\text{OS}_{d-1}) > \frac{c}{N}. \quad (6.20)$$

Hence, we can write

$$\int_{\text{OS}_{d-1}} f(g_N(x)) dx = \sum_{y \in \Lambda} \omega(y) f(y),$$

where  $\omega(y)$  assigns the weight  $N^{-d+1}$  to all  $y \in \Lambda$  satisfying (6.20), and  $0 \leq \omega(y) \leq N^{-d+1}$  for all  $y \in \partial_{c/N}\text{OS}_{d-1}$  to compensate for i) the fact that in this region  $g_N$  maps some  $x \in \text{OS}_{d-1}$  to a lattice point outside of  $\text{OS}_{d-1}$ , and ii) the fact that for some lattice points in  $y \in \text{OS}_{d-1}$ , not all of the neighborhood of  $y$  is contained in  $\text{OS}_{d-1}$ , i.e.,  $U_N(y) \setminus \text{OS}_{d-1} \neq \emptyset$ .

We bound

$$\begin{aligned} & \left| \sum_{y \in \Lambda} N^{-d+1} f(y) - \int_{\text{OS}_{d-1}} f(g_N(x)) dx \right| \\ &= \left| \sum_{y \in \Lambda} (N^{-d+1} - \omega(y)) f(y) \right| \\ &\leq \sum_{y \in \partial_{c/N}\text{OS}_{d-1}} N^{-d+1} |f(y)| \\ &\leq \sum_{y \in \partial_{c/N}\text{OS}_{d-1}} N^{-d+1} \left( \frac{c}{N} \right)^p \\ &\leq \frac{c}{N} C_d N^{d-2} N^{-d+1} \left( \frac{c}{N} \right)^p \\ &= O(N^{-p-2}), \end{aligned}$$

where in the second inequality we used the assumption  $f \in O(d(x, \partial\text{OS}_{d-1})^p)$ , and in the third inequality we used that there are at most  $\frac{c}{N} C_d N^{d-2}$  lattice points in  $\partial_{c/N}\text{OS}_{d-1}$  for some constant  $C_d$  that only depends on  $d$ . This proves (i).

In order to prove (ii), we first develop  $f(g_N(x))$  into a Taylor series around a point  $x$ :

$$f(g_N(x)) = f(x) + (g_N(x) - x)^T \nabla f(x) + O(N^{-1})$$

where we used the bound  $\|g_N(x) - x\|_2 \leq \frac{c}{N}$  for some constant  $c$  for the remainder term in the Taylor series. Hence, we have

$$\begin{aligned} \left| \int_{\text{OS}_{d-1}} f(g_N(x)) dx - \int_{\text{OS}_{d-1}} f(x) dx \right| &\leq \int_{\text{OS}_{d-1}} \left| (g_N(x) - x)^T \nabla f(x) \right| dx + O(N^{-1}) \\ &\leq \int_{\text{OS}_{d-1}} \|g_N(x) - x\|_2 \|\nabla f(x)\|_2 dx + O(N^{-1}) \\ &\leq \frac{c}{N} K \text{vol}(\text{OS}_{d-1}) + O(N^{-1}) \\ &= O(N^{-1}), \end{aligned}$$

where the second inequality follows from the Cauchy-Schwarz inequality, and in the third inequality we used the fact that by assumption  $\|\nabla f(x)\|_2$  is a continuous function on the compact domain  $\text{OS}_{d-1}$  and therefore bounded by a constant  $K$ , proving (ii).

Finally, we prove assertion (iii). We denote by  $\partial_{ij} f := (e_i - e_j)^T \nabla f$  the partial derivative of  $f$  in the direction  $e_{ij} := e_i - e_j$ . We approximate  $\partial_{ij} f(x)$  using a central difference  $D_{ij}[f(x)] := f(x + \frac{h}{2}e_{ij}) - f(x - \frac{h}{2}e_{ij})$ , where  $h > 0$  is to be chosen later. To this end, consider the Taylor expansions

$$\begin{aligned} f(x + \frac{h}{2}e_{ij}) &= f(x) + \frac{h}{2}e_{ij}^T \nabla f(x) + O(h^2) \\ f(x - \frac{h}{2}e_{ij}) &= f(x) - \frac{h}{2}e_{ij}^T \nabla f(x) + O(h^2). \end{aligned}$$

Subtracting the second expansion from the first and rearranging gives

$$\partial_{ij} f(x) = \frac{1}{h} D_{ij}[f(x)] + O(h). \quad (6.21)$$

It is easy to see that

$$\sum_{i,j=1}^d e_{ij} e_{ij}^T = 2d \mathbf{1}_{V_0^{d-1}},$$

and hence, for the Laplacian  $\Delta = \text{tr}(H(\cdot))$  on  $V_0^{d-1}$  with  $H(\cdot)$  the Hessian matrix, we have

$$\begin{aligned} \Delta f(x) &= \text{tr}(H(f)(x)) \\ &= \text{tr}\left(\mathbf{1}_{V_0^{d-1}} H(f)(x)\right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2d} \sum_{i,j=1}^d e_{ij}^T H(f)(x) e_{ij} \\
&= \frac{1}{2d} \sum_{i,j=1}^d \partial_{ij}^2 f(x).
\end{aligned} \tag{6.22}$$

Similarly, denoting by  $\langle \cdot, \cdot \rangle_{V_0^{d-1}}$  the inner product on  $V_0^{d-1}$ , we have

$$\begin{aligned}
\langle \nabla f(x), \nabla f(x) \rangle_{V_0^{d-1}} &= \frac{1}{2d} \sum_{i,j=1}^d \langle \nabla f(x), e_{ij} e_{ij}^T \nabla f(x) \rangle_{V_0^{d-1}} \\
&= \frac{1}{2d} \sum_{i,j=1}^d \left( e_{ij}^T \nabla f(x) \right)^2 \\
&= \frac{1}{2d} \sum_{i,j=1}^d (\partial_{ij} f(x))^2.
\end{aligned} \tag{6.23}$$

We now calculate, abbreviating  $\sum'_{y \in \Lambda} = \sum_{y \in \Lambda} \omega(y)$ :

$$\begin{aligned}
&\int_{\text{OS}_{d-1}} f(g_N(x)) (-\Delta f)(g_N(x)) dx \\
&= \sum'_{y \in \Lambda} f(y) (-\Delta f)(y) \\
&= -\frac{1}{2d} \sum'_{y \in \Lambda} \sum_{i,j=1}^d f(y) \partial_{ij}^2 f(y) \\
&= \frac{1}{2d} \sum'_{y \in \Lambda} \sum_{i,j=1}^d (\partial_{ij} f(y))^2 - \frac{1}{2d} \sum'_{y \in \Lambda} \sum_{i,j=1}^d \partial_{ij} [f(y) \partial_{ij} f(y)] \\
&= \sum'_{y \in \Lambda} \langle \nabla f(y), \nabla f(y) \rangle_{V_0^{d-1}} - \frac{1}{2dh} \sum_{i,j=1}^d \sum'_{y \in \Lambda} D_{ij} [f(y) \partial_{ij} f(y)] + O(h),
\end{aligned} \tag{6.24}$$

where we used (6.22) in the second equality, and (6.23) and (6.21) in the last equality.

For the first term in (6.24), we have

$$\begin{aligned}
\sum'_{y \in \Lambda} \langle \nabla f(y), \nabla f(y) \rangle_{V_0^{d-1}} &= \int_{\text{OS}_{d-1}} \langle \nabla f(g_N(x)), \nabla f(g_N(x)) \rangle_{V_0^{d-1}} dx \\
&= \int_{\text{OS}_{d-1}} \langle \nabla f(x), \nabla f(x) \rangle_{V_0^{d-1}} dx + O(N^{-1}) \\
&= \int_{\text{OS}_{d-1}} f(x) (-\Delta f(x)) dx + O(N^{-1}),
\end{aligned}$$

where the second equality follows from (ii), and the third equality is ordinary integration by parts. For the second term in (6.24), we use the definition of  $D_{ij}$  to obtain

$$\begin{aligned} & \frac{1}{2dh} \sum_{i,j=1}^d \sum'_{y \in \Lambda} D_{ij} [f(y) \partial_{ij} f(y)] + O(h) \\ &= \frac{1}{2dh} \sum_{i,j=1}^d \sum'_{y \in \Lambda} \left( f(y + \frac{h}{2} e_{ij}) \partial_{ij} f(y + \frac{h}{2} e_{ij}) - f(y - \frac{h}{2} e_{ij}) \partial_{ij} f(y - \frac{h}{2} e_{ij}) \right) + O(h). \end{aligned} \tag{6.25}$$

We choose  $h = O(N^{-1})$  such that  $y \pm \frac{h}{2} e_{ij} \in \Lambda$  for all  $y \in \Lambda$  sufficiently far away from the boundary of  $\Lambda$ . Then all terms in (6.25) cancel except for those terms involving evaluations of  $f$  on  $\partial_h \text{OS}_{d-1}$  or outside  $\text{OS}_{d-1}$ . But these terms in turn are  $O(h) = O(N^{-1})$ , which can be seen using the same arguments as those in the proof of (ii). It follows that, with the above choice of  $h = O(N^{-1})$ ,

$$\frac{1}{2dh} \sum_{i,j=1}^d \sum'_{y \in \Lambda} D_{ij} [f(y) \partial_{ij} f(y)] + O(h) = O(N^{-1}).$$

In summary, we have shown that

$$\int_{\text{OS}_{d-1}} f(g_N(x)) (-\Delta f)(g_N(x)) dx = \int_{\text{OS}_{d-1}} f(x) (-\Delta f)(x) dx + O(N^{-1}),$$

which is what we set out to prove.

We are now ready to prove Theorem 1.4:

*Proof of Thm. 1.4.* Fix a dimension  $d$ , and let  $a \in C^2(\text{OS}_{d-1})$  be twice continuously differentiable<sup>5</sup> such that  $a|_{\partial \text{OS}_{d-1}} \equiv 0$ ,  $a(x) \geq 0$  for all  $x \in \text{OS}_{d-1}$ , and  $\|a\|_2 = 1$ , where  $\|\cdot\|_2$  is the  $L_2$ -norm on  $\text{OS}_{d-1}$ . As  $d$  is fixed throughout the proof, we omit indicating any dependence on  $d$  except when we would like to emphasize the dimension of an object. Note that clearly  $a \in L_2(\text{OS}_{d-1})$  as  $a$  is continuous and  $\text{OS}_{d-1}$  is compact.

We use the square of a scaled version of  $a$  as a candidate probability distribution  $q$  on Young diagrams  $\mu$  with  $N$  boxes and at most  $d$  rows,

$$q(\mu) = \frac{\eta_N}{N^{d-1}} a^2 \left( \frac{\mu}{N} \right). \tag{6.26}$$

Here  $\eta_N$  is a normalization constant which is close to one. Roughly speaking, this is due to the fact that the normalization condition for  $q(\mu)$  is essentially proportional to a Riemann sum for the integral that calculates the  $L_2$ -norm of  $a$ , which is equal to unity by assumption. Indeed, since  $a^2$  satisfies the assumptions of Lemma 6.2 with  $p = 1$ , we have

$$\begin{aligned} 1 &= \sum_{\mu \vdash_d N} q(\mu) \\ &= \frac{\eta_N}{N^{d-1}} \sum_{\mu \in \mathbb{Z}^d \cap N \text{OS}_{d-1}} a^2 \left( \frac{\mu}{N} \right) \end{aligned}$$

<sup>5</sup> The second derivative is continuous and its limit for the argument approaching the boundary exists.



$$\begin{aligned}
 &= \frac{\eta_N}{N^{d-1}} \sum_{y \in \left(\frac{1}{N}\mathbb{Z}^d\right) \cap \text{OS}_{d-1}} a^2(y) \\
 &= \eta_N \left( \int_{\text{OS}_{d-1}} a^2(g_N(y)) dy + O(N^{-3}) \right) \\
 &= \eta_N \left( \int_{\text{OS}_{d-1}} a^2(y) dy + O(N^{-1}) \right) \\
 &= \eta_N \left( 1 + O(N^{-1}) \right),
 \end{aligned}$$

where the fourth and fifth equality follow from Lemma 6.2(i) and (ii), respectively, and the last equality follows from  $\|a\|_2 = 1$ . Hence,  $\eta_N = 1 + O(N^{-1})$ .

Before we proceed, we restate the fidelity formula in (3.6) for the optimal deterministic protocol for the reader's convenience:

$$F_d^*(N) = d^{-N-2} \max_{c_\mu} \sum_{\alpha \vdash_d N-1} \left( \sum_{\mu=\alpha+\square} \sqrt{c_\mu d_\mu m_{d,\mu}} \right)^2. \quad (6.27)$$

We bound this expression from below by choosing  $c_\mu = q(\mu)/p(\mu)$ , where  $q(\mu)$  is defined as in (6.26) and  $p(\mu) = \frac{d_\mu m_{d,\mu}}{d^N}$  is the Schur–Weyl distribution. The choice of  $c_\mu$  in (6.27) corresponds to a particular PBT protocol whose entanglement fidelity we denote by  $F_a$  in the following. It will be convenient to rewrite the sums over Young diagrams  $\alpha \vdash_d N-1$  and  $\mu = \alpha + \square$  in (6.27) as a sum over Young diagrams  $\mu \vdash_d N$  and  $i, j = 1, \dots, d$ , requiring that both  $\mu + e_i - e_j$  and  $\mu - e_j$  be Young diagrams themselves. Using this trick, the quantity  $\frac{d^2}{\eta_N} F_a$  can be expressed as

$$\begin{aligned}
 \frac{d^2}{\eta_N} F_a &= N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d \mathbb{1}_{\text{YD}}(\mu + e_i - e_j) \mathbb{1}_{\text{YD}}(\mu - e_j) a\left(\frac{\mu + e_i - e_j}{N}\right) \\
 &= N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d a\left(\frac{\mu + e_i - e_j}{N}\right) \\
 &\quad + N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d \mathbb{1}_{\text{YD}}(\mu + e_i - e_j) \mathbb{1}_{\text{YD}}(\mu - e_j) a\left(\frac{\mu + e_i - e_j}{N}\right) \\
 &\quad - N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d a\left(\frac{\mu + e_i - e_j}{N}\right). \quad (6.28)
 \end{aligned}$$

We first argue that up to order  $N^{-2}$  we only need to consider the first term in the above expression. To this end, we rewrite the sum in the second term as an integral,

$$\begin{aligned}
 &N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d f_{i,j}(\mu) a\left(\frac{\mu + e_i - e_j}{N}\right) \\
 &= \int_{\text{OS}_{d-1}} h_N(x) a(g_N(x)) \sum_{i,j=1}^d f_{i,j}(x) a\left(g_N(x) + \frac{e_i - e_j}{N}\right) dx,
 \end{aligned}$$

where  $f_{i,j}(x) := \mathbb{1}_{\text{YD}}(Ng_N(x) + e_i - e_j) \mathbb{1}_{\text{YD}}(Ng_N(x) - e_j)$ . The function  $h_N(x) \in [0, 1]$  takes care of normalization around the boundaries of  $\text{OS}_{d-1}$ , that is,  $h_N(x) = 1$  except in a region  $\partial_{c_1/N} \text{OS}_{d-1}$  for some constant  $c_1$  that only depends on  $d$ . Note that the same statement is true for the function  $f_{i,j}(x)$ , and therefore, this also holds for the product  $h_N(x) f_{i,j}(x)$ . Using Lemma 6.2(i) for the third term in (6.28) gives

$$\begin{aligned} & N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d a\left(\frac{\mu + e_i - e_j}{N}\right) \\ &= \int_{\text{OS}_{d-1}} a(g_N(x)) \sum_{i,j=1}^d a\left(g_N(x) + \frac{e_i - e_j}{N}\right) dx + O(N^{-3}). \end{aligned}$$

Hence, for the difference of the second and third term in (6.28), we obtain

$$\begin{aligned} & N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d [\mathbb{1}_{\text{YD}}(\mu + e_i - e_j) \mathbb{1}_{\text{YD}}(\mu - e_j) - 1] a\left(\frac{\mu + e_i - e_j}{N}\right) \\ &= \int_{\text{OS}_{d-1}} a(g_N(x)) \sum_{i,j=1}^d [h_N(x) f_{i,j}(x) - 1] a\left(g_N(x) + \frac{e_i - e_j}{N}\right) dx + O(N^{-3}) \\ &\leq \frac{c_2}{N^2} \int_{\partial_{c_1/N} \text{OS}_{d-1}} [h_N(x) \mathbb{1}_{\text{YD}}(Ng_N(x) + e_i - e_j) \mathbb{1}_{\text{YD}}(Ng_N(x) - e_j) - 1] dx + O(N^{-3}) \\ &\leq \frac{c_3}{N^2} \text{vol}(\partial_{c_1/N} \text{OS}_{d-1}) + O(N^{-3}) \\ &= O(N^{-3}) \end{aligned}$$

for some constants  $c_2$  and  $c_3$ . Here, the first inequality is obtained by a Taylor expansion of the different occurrences of  $a$  around the respective closest boundary point and using the fact that  $a$  vanishes on the boundary by assumption. The second inequality follows since  $h_N$  is bounded uniformly in  $N$ .<sup>6</sup>

We now turn to the first term in (6.28), applying Lemma 6.2(i) once more to obtain

$$\begin{aligned} & N^{-d+1} \sum_{\mu \vdash_d N} a\left(\frac{\mu}{N}\right) \sum_{i,j=1}^d a\left(\frac{\mu + e_i - e_j}{N}\right) \\ &= \int_{\text{OS}_{d-1}} a(g_N(x)) \sum_{i,j=1}^d a\left(g_N(x) + \frac{e_i - e_j}{N}\right) dx + O(N^{-3}). \end{aligned}$$

Expanding  $a\left(g_N(x) + \frac{e_i - e_j}{N}\right)$  into a Taylor series gives

$$\begin{aligned} a\left(g_N(x) + \frac{e_i - e_j}{N}\right) &= a(g_N(x)) + \frac{1}{N} \langle e_i - e_j, (\nabla a)(g_N(x)) \rangle_{V_0^{d-1}} \\ &\quad + \frac{1}{2N^2} \text{tr} \left[ (e_i - e_j)(e_i - e_j)^T (H(a))(g_N(x)) \right] + O(N^{-3}), \end{aligned}$$

<sup>6</sup> Observe that a constant fraction of  $U_N(y)$  of each lattice point  $y \in \text{OS}_{d-1} \cap \frac{1}{N} \mathbb{Z}^d$  lies inside  $\text{OS}_{d-1}$ . This fraction is not uniformly bounded in  $d$ , as the solid angle of the vertices of  $\text{OS}_{d-1}$  decreases with  $d$ . However, this does not concern us, since we are only interested in the limit  $N \rightarrow \infty$  for fixed  $d$ .

where  $\langle \cdot, \cdot \rangle_{V_0^{d-1}}$  is the standard inner product on  $V_0^{d-1}$  and  $H(a)$  denotes the Hessian of  $a$  on  $V_0^{d-1}$ . Summing over  $i$  and  $j$  yields

$$\sum_{i,j=1}^d e_i - e_j = 0 \quad \sum_{i,j=1}^d (e_i - e_j)(e_i - e_j)^T = 2d1_{V_0^{d-1}}. \quad (6.29)$$

It follows that

$$\begin{aligned} & \int_{\text{OS}_{d-1}} a(g_N(x)) \sum_{i,j=1}^d a\left(g_N(x) + \frac{e_i - e_j}{N}\right) dx + O(N^{-3}) \\ &= \int_{\text{OS}_{d-1}} a(g_N(x)) \left( d^2 a(g_N(x)) + \frac{d}{N^2} (\Delta a)(g_N(x)) \right) dx + O(N^{-3}) \\ &= \frac{d^2}{N^{d-1}} \sum_{y \in \text{OS}_{d-1} \cap \frac{1}{N} \mathbb{Z}^d} a^2(y) - \frac{d}{N^2} \int_{\text{OS}_{d-1}} a(g_N(x)) (-\Delta a)(g_N(x)) dx + O(N^{-3}) \\ &= \frac{d^2}{\eta_N} - \frac{d}{N^2} \int_{\text{OS}_{d-1}} a(g_N(x)) (-\Delta a)(g_N(x)) dx + O(N^{-3}) \\ &= \frac{d^2}{\eta_N} - \frac{d}{N^2} \int_{\text{OS}_{d-1}} a(x) (-\Delta a)(x) dx + O(N^{-3}), \end{aligned}$$

where in the first equality the  $N-1$  term vanishes due to (6.29), and we defined the Laplace operator  $\Delta(a) = \text{tr } H(a)$  on  $V_0^{d-1}$ . In the second equality we used Lemma 6.2(i) to switch back to discrete summation, in the third equality we used the normalization of  $a$ , and in the fourth equality we used Lemma 6.2(iii).

Putting together everything we have derived so far, we obtain

$$F_a = 1 - \frac{1}{dN^2} \int_{\text{OS}_{d-1}} a(x) (-\Delta a)(x) dx + O(N^{-3}).$$

In equation Eq. (3.6), the fidelity is maximized over all densities  $c_\mu$ . The above expression shows, that restricting to the set of densities  $c_\mu$  that stem from a function  $a$  on  $\text{OS}_{d-1}$  makes the problem equivalent to minimizing the expression

$$\int_{\text{OS}_{d-1}} a(x) (-\Delta a)(x) dx.$$

When taking the infimum over  $a \in H^2(\text{OS}_{d-1})$ , where  $H^2(\text{OS}_{d-1})$  is the Sobolev space of twice weakly differentiable functions, instead of  $a \in C^2(\text{OS}_{d-1})$ , this is exactly one of the variational characterizations of the first Dirichlet eigenvalue of the Laplace operator on  $\text{OS}_{d-1}$ . This is because the eigenfunction corresponding to the first eigenvalue of the Dirichlet Laplacian can be chosen positive (see, e.g., [62]). But  $C^2(\text{OS}_{d-1})$  is dense in  $H^2(\text{OS}_{d-1})$ , which implies that

$$\sup_a F_a = 1 - \frac{\lambda_1(\text{OS}_{d-1})}{dN^2} + O(N^{-3}),$$

where the supremum is taken over all non-negative functions  $a \in C^2(\text{OS}_{d-1})$ .  $\square$

Upper and lower bounds for the first Dirichlet eigenvalue of the Laplacian on a sufficiently well-behaved domain readily exist.

**Theorem 6.3.** [37,63] *For the first Dirichlet eigenvalue  $\lambda_1(\Omega)$  on a bounded convex domain  $\Omega \subset \mathbb{R}^d$ , the following inequalities hold,*

$$\lambda_1(\Omega) \geq \lambda_1(B_1) \left( \frac{\text{vol}(B_1)}{\text{vol}(\Omega)} \right)^{\frac{2}{d}}, \text{ and}$$

$$\lambda_1(\Omega) \leq \lambda_1(B_1) \frac{\text{vol}(\partial\Omega)}{dr_\Omega \text{vol}(\Omega)},$$

where  $B_1 \subset \mathbb{R}^d$  is the unit ball and  $r_\Omega$  is the inradius of  $\Omega$ .

The inradius of  $\text{OS}_{d-1}$  is equal to  $1/d^2$ . This can be seen by guessing the center of the inball  $\hat{x} = ((2d-1)/d^2, (2d-3)/d^2, \dots, 1/d^2)$  and checking that the distance to each facet is  $1/d^2$ . Therefore we get the following lower bound on the optimal PBT fidelity. This theorem is stated in Sect. 1.2, and restated here for convenience.

**Theorem 1.5** (Restated). *For the optimal fidelity of port-based teleportation with arbitrary but fixed input dimension  $d$  and  $N$  ports, the following bound holds,*

$$F_d^*(N) \geq 1 - \frac{d^5 + O(d^{9/2})}{4\sqrt{2}N^2} + O(N^{-3}).$$

*Proof.* Theorem 1.4 gives us the bound

$$F_d^*(N) \geq 1 - \frac{\lambda_1(\text{OS}_{d-1})}{dN^2} + O(N^{-3}).$$

Using Theorem 6.3 and Lemma D.5 we bound

$$\begin{aligned} \lambda_1(\text{OS}_{d-1}) &\leq \lambda_1(B_1^{d-1}) \frac{\text{vol}(\partial\Omega)}{dr_\Omega \text{vol}(\Omega)} \\ &\leq \lambda_1(B_1^{d-1}) d^2 \left( \frac{d(d-1)}{\sqrt{2}} + \sqrt{d(d-1)} + \sqrt{2} \right). \end{aligned}$$

The first eigenvalue of the Dirichlet Laplacian on the  $(d-1)$ -dimensional Ball is given by

$$\lambda_1(B_1^{d-1}) = j_{\frac{d-3}{2},1}^2,$$

where  $j_{\nu,l}$  is the  $l$ th root of the Bessel function of the first kind with parameter  $\nu$ . This is, in turn, bounded as [64]

$$j_{\nu,1} \leq \sqrt{\nu+1}(\sqrt{\nu+2}+1).$$

Putting the inequalities together we arrive at

$$\lambda_1(B_1^{d-1}) \leq \frac{d-1}{2} \left( \sqrt{\frac{d+1}{2}} + 1 \right)^2,$$

$$\begin{aligned}\lambda_1(\text{OS}_{d-1}) &\leq \frac{d-1}{2} \left( \sqrt{\frac{d+1}{2}} + 1 \right)^2 d^2 \left( \frac{d(d-1)}{\sqrt{2}} + \sqrt{d(d-1)} + \sqrt{2} \right), \text{ and hence} \\ F_d^*(N) &\geq 1 - \frac{\frac{d-1}{2} \left( \sqrt{\frac{d+1}{2}} + 1 \right)^2 d \left( \frac{d(d-1)}{\sqrt{2}} + \sqrt{d(d-1)} + \sqrt{2} \right)}{N^2} + O(N^{-3}) \\ &= 1 - \frac{d^5 + O(d^{9/2})}{4\sqrt{2}N^2} + O(N^{-3}).\end{aligned}$$

In the appendix, we provide a concrete protocol in Theorem B.1 that achieves the same asymptotic dependence on  $N$  and  $d$ , with a slightly worse constant.

Intuitively it seems unlikely that a “wrinkly” distribution, i.e. a distribution that does not converge against an  $L_1$  density on OS, is the optimizer in Eq. (3.6). Supposing that the optimizer comes from a function  $a$  as described above, we can also derive a converse bound for the asymptotics of the entanglement fidelity  $F_d^*(N)$  using Theorem 6.3.

*Remark 6.4.* Let  $P_a^N$  be the PBT protocol with  $c_\mu = N^{d-1}a^2(\mu/N)/P(\mu)$  for some function  $a \in L_2(\text{OS}_{d-1})$ . For the asymptotic fidelity of such protocols for large  $N$  the following converse bound holds,

$$F_a(N) \leq 1 - \frac{\pi d^4 + O(d^3)}{8e^3 N^2} + O(N^{-3}).$$

This can be seen as follows. From Theorem 1.4 we have that

$$F_a \leq 1 - \frac{\lambda_1(\text{OS}_{d-1})}{dN^2} + O(N^{-3}).$$

Theorem 6.3 together with Lemma D.5 yields

$$\begin{aligned}\lambda_1(\text{OS}_{d-1}) &\geq \lambda_1(B_1) \left( \frac{\text{vol}(B_1^{d-1})}{\text{vol}(\text{OS}_{d-1})} \right)^{\frac{2}{d}} \\ &= \lambda_1(B_1) \left( \frac{\pi^{\frac{d-1}{2}} \sqrt{d} ((d-1)!)^2}{\Gamma(\frac{d-1}{2} + 1)} \right)^{\frac{2}{d}} \\ &\geq \pi^{1-1/d} \lambda_1(B_1) \left( \frac{((d-1)!)^2}{\Gamma(\frac{d-1}{2} + 1)} \right)^{\frac{2}{d}}\end{aligned}$$

where in the second line we have used the volume of the  $(d-1)$ -dimensional Ball,

$$\text{vol}(B_1^{d-1}) = \frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2} + 1)},$$

and  $\Gamma(x)$  is the gamma function. Using bound versions of Stirling’s approximation we obtain

$$\lambda_1(\text{OS}_{d-1}) \geq O(1) \lambda_1(B_1) \left( \frac{d-1}{e} \right)^{3(1-1/d)}.$$

Using a lower bound for the first zero of the Bessel function of the first kind [65] we bound

$$\lambda_1(B_1^{d-1}) \geq \left(\frac{d}{2} + c\right)^2$$

for some constant  $c$ , so we finally arrive at

$$F_a = 1 - \frac{\pi d^4 + O(d^3)}{8e^3 N^2} + O(N^{-3}).$$

This bound has the nice property that  $N \propto d^2$  if the error of the PBT protocol is fixed, which is what we expect from information theoretic insights (see Sect. 7).

## 7. Converse Bound

We begin by deriving a lower bound on the communication requirements for approximate quantum teleportation of any kind, i.e., not only for PBT. Such a result could be called folklore, but has, to the best of our knowledge, not appeared elsewhere.<sup>7</sup>

For the proof we need the converse bound for one-shot quantum state splitting that was given in [67] in terms of the smooth max-mutual information  $I_{\max}^\varepsilon(E : A)_\rho$ . To define this quantity, let  $D_{\max}(\rho \|\sigma) = \min\{\lambda \in \mathbb{R} \mid 2^\lambda \sigma \geq \rho\}$  be the max-relative entropy [68], and let  $P(\rho, \sigma) := \sqrt{1 - F(\rho, \sigma)}$  be the purified distance. Furthermore, let  $B_\varepsilon(\rho) := \{\bar{\rho} : \bar{\rho} \geq 0, \text{tr } \bar{\rho} \leq 1, P(\rho, \bar{\rho}) \leq \varepsilon\}$  be the  $\varepsilon$ -ball of subnormalized states around  $\rho$  with respect to the purified distance. The smooth max-mutual information is defined as

$$I_{\max}^\varepsilon(E : A)_\rho := \min_{\bar{\rho} \in B_\varepsilon(\rho)} I_{\max}(E : A)_{\bar{\rho}},$$

where  $I_{\max}(E : A)_{\bar{\rho}} := \min_{\sigma_A} D_{\max}(\bar{\rho}_{AE} \|\sigma_A \otimes \bar{\rho}_E)$  with the minimization over normalized quantum states  $\sigma_A$ .

**Lemma 7.1.** *Let*

$$|\phi^+\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$$

*be the  $d \times d$ -dimensional maximally entangled state. Then*

$$2 \log \left[ d(1 - \varepsilon^2) \right] \geq I_{\max}^\varepsilon(A : B)_{\phi^+} \geq 2 \log \left( d(1 - \varepsilon^2) \right).$$

*Proof.* Let  $\rho \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$  be a quantum state such that  $I_{\max}^\varepsilon(A : B)_{\phi^+} = I_{\max}(A : B)_\rho$ , and let  $|\gamma\rangle_{ABE}$  be a purification of  $\rho$ . Uhlmann's Theorem ensures that there exists a pure quantum state  $|\alpha\rangle_E$  such that

$$\sqrt{1 - \varepsilon^2} \leq \sqrt{F(\phi^+, \rho)} = \langle \phi^+ |_{AB} \langle \alpha |_E |\gamma\rangle_{ABE}. \quad (7.1)$$

<sup>7</sup> Except in the PhD thesis of one of the authors [66].

This holds without taking the absolute value because any phase can be included in  $|\alpha\rangle$ . Let

$$|\gamma\rangle_{ABE} = \sum_{i=0}^{d-1} \sqrt{p_i} |\phi_i\rangle_A \otimes |\psi_i\rangle_{BE} \quad (7.2)$$

be the Schmidt decomposition of  $|\gamma\rangle$  with respect to the bipartition  $A : BE$ . Let further  $U_A$  be the unitary matrix such that  $U_A|i\rangle_A = |\phi_i\rangle_A$ . Using the Mirror Lemma D.1 we get

$$\begin{aligned} |\phi^+\rangle_{AB} &= U_A U_A^\dagger |\phi^+\rangle_{AB} \\ &= U_A \bar{U}_B |\phi^+\rangle_{AB} \\ &= \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |\phi_i\rangle_A |\xi_i\rangle_B, \end{aligned}$$

where  $\bar{U}$  is the complex conjugate in the computational basis and  $|\xi_i\rangle_B = \bar{U}_B|i\rangle_B$ . With this we obtain from (7.1) that

$$\begin{aligned} 1 - \varepsilon^2 &\leq (\langle \phi^+ |_{AB} \langle \alpha |_E |\gamma\rangle_{ABE})^2 \\ &= (\Re \langle \phi^+ |_{AB} \langle \alpha |_E |\gamma\rangle_{ABE})^2 \\ &= \left( \sum_{i=0}^{d-1} \sqrt{\frac{p_i}{d}} \Re \langle \xi_i |_B \langle \alpha |_E |\psi_i\rangle_{BE} \right)^2 \\ &\leq \frac{1}{d} \sum_{i=0}^{d-1} (\Re \langle \xi_i |_B \langle \alpha |_E |\psi_i\rangle_{BE})^2 \\ &\leq \frac{1}{d} \sum_{i=0}^{d-1} \Re \langle \xi_i |_B \langle \alpha |_E |\psi_i\rangle_{BE}. \end{aligned}$$

The second inequality is the Cauchy-Schwarz inequality and the third inequality follows from  $\Re \langle \xi_i |_B \langle \alpha |_E |\psi_i\rangle_{BE} \leq 1$ .

The next step is to bound the max-mutual information of  $\rho$ . Let

$$\lambda = I_{\max}(A : B)_\rho = I_{\max}^\varepsilon(A : B)_{\phi^+}.$$

By the definition of  $I_{\max}$  there exists a quantum state  $\sigma_B$  such that

$$2^\lambda = \left\| \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} \rho_{AB} \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} \right\|_\infty.$$

Here,  $X^{-1}$  denotes the pseudo-inverse of a matrix  $X$ , i.e.,  $X^{-1}X = XX^{-1}$  is equal to the projector onto the support of  $X$ . Let  $|\phi_\sigma\rangle = \sqrt{d}\sigma_B^{1/2}|\phi^+\rangle$  be the standard purification of  $\sigma$ . We bound

$$\begin{aligned} 2^\lambda &= \left\| \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} \rho_{AB} \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} \right\|_\infty \\ &\geq \langle \phi_\sigma | \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} \rho_{AB} \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} | \phi_\sigma \rangle \end{aligned}$$

$$\begin{aligned}
&= \text{tr} \langle \phi_\sigma | \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} | \gamma \rangle \langle \gamma |_{ABE} \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} | \phi_\sigma \rangle \\
&\geq \langle \phi_\sigma |_{AB} \langle \alpha |_E \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} | \gamma \rangle \langle \gamma |_{ABE} \rho_A^{-\frac{1}{2}} \otimes \sigma_B^{-\frac{1}{2}} | \phi_\sigma \rangle_{AB} | \alpha \rangle_E \\
&= d \left| \langle \phi^+ |_{AB} \rho_A^{-\frac{1}{2}} \langle \alpha |_E | \gamma \rangle_{ABE} \right|^2 \\
&= \left| \sum_i \langle \xi_i |_B \langle \alpha |_E | \psi_i \rangle_{BE} \right|^2 \\
&\geq \left( \sum_i \Re \langle \xi_i |_B \langle \alpha |_E | \psi_i \rangle_{BE} \right)^2 \\
&\geq d^2 (1 - \varepsilon^2)^2,
\end{aligned}$$

where we used the particular form of  $|\phi_\sigma\rangle$  in the third equality, and (7.2) in the fourth equality, together with the fact that  $\{p_i\}_i$  are the eigenvalues of  $\rho_A$ . This proves the claimed upper bound on  $I_{\max}^\varepsilon(A : B)_{\phi^+}$ .

In order to prove the lower bound, let  $r = \lceil d(1 - \varepsilon^2) \rceil$  and

$$|\phi_r^+\rangle = \frac{1}{\sqrt{r}} \sum_{i=0}^{r-1} |ii\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

Then we have

$$\begin{aligned}
I_{\max}(A : B)_{\phi_r^+} &= 2 \log r = 2 \log \lceil d(1 - \varepsilon^2) \rceil \\
|\langle \phi^+ | \phi_r^+ \rangle|^2 &= r/d \geq 1 - \varepsilon^2.
\end{aligned}$$

The observation that  $|\phi_r^+\rangle \langle \phi_r^+|$  is a point in the minimization over  $\sigma$  finishes the proof.  $\square$

Using the special case of state merging/splitting with trivial side information and the converse bound from [67], we can bound the necessary quantum communication for simulating the identity channel with a given entanglement fidelity.

**Corollary 7.2.** *Let  $\mathcal{E}_{AA' \rightarrow B}$ ,  $\mathcal{D}_{BB' \rightarrow A}$  be quantum (encoding and decoding) channels with  $\dim \mathcal{H}_A = d$  and  $\dim \mathcal{H}_B = d'$  such that there exists a resource state  $\rho_{A'B'}$  achieving*

$$F(\mathcal{D} \circ \mathcal{E}((\cdot) \otimes \rho_{A'B'})) = 1 - \varepsilon^2.$$

Then the following inequality holds:

$$d' \geq d (1 - \varepsilon^2).$$

*Proof.* Using Lemma 7.1, this follows from applying the lower bound on the communication cost of one-shot state splitting from [67] to the special case where Alice and the reference system share a maximally entangled state.  $\square$

Together with superdense coding this implies a lower bound on approximate teleportation.

**Corollary 7.3.** *If in the above corollary  $\mathcal{E}$  is a qc-channel, then*

$$d' \geq d^2 (1 - \varepsilon^2)^2.$$



*Proof.* This follows as any protocol with a lower classical communication in conjunction with superdense coding would violate Corollary 7.2.  $\square$

For the special case of port-based teleportation, this implies a lower bound on the number of ports.

**Corollary 7.4.** *Any port-based teleportation protocol with input dimension  $d$  and  $N$  ports has entanglement fidelity at most*

$$F_d^*(N) \leq \frac{\sqrt{N}}{d}.$$

*Proof.* In port-based teleportation, the only information that is useful to the receiver is which port to select. More precisely, given a protocol  $P$  for PBT in which Alice sends a message that is not a port number, we can construct a modified protocol  $P$  where Alice applies the procedure that Bob uses in  $P$  to deduce the port to select and then sends the port number instead. For a given entanglement fidelity  $F$ , having fewer than  $(dF)^2$  ports would therefore violate the bound from Corollary 7.3.  $\square$

The converse bound on the amount of quantum communication in Corollary 7.2 holds for arbitrary protocols implementing a simulation of the identity channels, and Corollary 7.3 puts a lower bound on the classical communication of any (approximate) teleportation scheme. We continue to derive a converse bound specifically for port-based teleportation that is nontrivial for all combinations of  $d$  and  $N$ . Let us consider a general port-based teleportation scheme, given by POVMs  $\{E_{A^N}^{(i)}\}$  and a resource state  $\rho_{A^N B^N}$ , where  $A_0 \cong \mathbb{C}^d$  and  $B_1, \dots, B_N \cong \mathbb{C}^d$ . We would like to upper-bound the entanglement fidelity

$$F_d^*(N) = F \left( \sum_{i=1}^N (I_{B_0} \otimes I_{B_i \rightarrow B_1}) \text{tr}_{(B_0 B_i)^c} [(E_A^{(i)} \otimes I_B) (\rho_{A^N B^N} \otimes \phi_{A_0 B_0}^+)] \right. \\ \left. (I_{B_0} \otimes I_{B_i \rightarrow B_1}^\dagger), \phi_{B_0 B_1}^+ \right), \quad (7.3)$$

where  $B_0 \cong \mathbb{C}^d$  and  $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$  is the fidelity. This fidelity corresponds to the special case of Alice using an arbitrary PBT protocol to teleport half of a maximally-entangled state to Bob, who already possesses the other half. An upper bound for this fidelity then directly implies an upper bound for the entanglement fidelity of the PBT protocol. We prove the following

**Theorem 7.5.** *For any port-based teleportation scheme, the entanglement fidelity (7.3) can be bounded from above as*

$$F_d^*(N) \leq 1 - \frac{d^2 - 1}{8N^2} \frac{1}{1 + \frac{d^2 - 2}{2N}}. \quad (7.4)$$

*Asymptotically, this bound becomes*

$$F_d^*(N) \leq 1 - \frac{d^2 - 1}{8} \frac{1}{N^2} + O(N^{-3}). \quad (7.5)$$

*Proof.* Note first that for a pure state  $|\psi\rangle$  we have  $F(\psi, \tau) = \langle \psi | \tau | \psi \rangle$  for any mixed state  $\tau$ , and hence  $\tau \mapsto F(\psi, \tau)$  is linear for any  $\tau$ . Since  $\phi_{B_0 B_1}^+$  is pure, the entanglement fidelity (7.3) can hence be rewritten as

$$\begin{aligned} F_d^*(N) &= \sum_{i=1}^N p(i) F \left( \frac{1}{p(i)} \operatorname{tr}_{(B_0 B_i)^c} [((E^{(i)})_A^{1/2} \otimes I_B)(\rho_{A^N B^N} \otimes \phi_{A_0 B_0}^+) \right. \\ &\quad \left. ((E^{(i)})_A^{1/2} \otimes I_B), \phi_{B_0 B_i}^+ \right) \\ &= \sum_{i=1}^N p(i) F \left( \frac{1}{p(i)} ((E^{(i)})_A^{1/2} \otimes I_B)(\rho_{A^N B^N} \otimes \phi_{A_0 B_0}^+) \right. \\ &\quad \left. ((E^{(i)})_A^{1/2} \otimes I_B), \phi_{B_0 B_i}^+ \otimes \sigma_{(B_0 B_i)^c}^{(i)} \right) \end{aligned}$$

for suitable  $\sigma_{(B_0 B_i)^c}^{(i)}$  whose existence is guaranteed by Uhlmann's Theorem. Here we have introduced  $p(i) = \operatorname{tr}[(E^{(i)})_A^{1/2}(\rho_{A^N B^N} \otimes \tau_{A_0})(E^{(i)})_A^{1/2}]$ . Abbreviating  $\sqrt{F}(\cdot, \cdot) \equiv \sqrt{F(\cdot, \cdot)}$ , we now have for any  $j \in \{1, \dots, N\}$  that

$$\begin{aligned} F_d^*(N) &\leq \sum_{i=1}^N p(i) \sqrt{F} \left( \frac{1}{p(i)} ((E^{(i)})_A^{1/2} \otimes I_B)(\rho_{A^N B^N} \otimes \phi_{A_0 B_0}^+) \right. \\ &\quad \left. ((E^{(i)})_A^{1/2} \otimes I_B), \phi_{B_0 B_i}^+ \otimes \sigma_{(B_0 B_i)^c}^{(i)} \right) \\ &\leq \sqrt{F} \left( \rho_{B_j} \otimes \tau_{B_0}, p(j) \phi_{B_0 B_j}^+ + (1 - p(j)) \tau_{B_0} \otimes \sigma_{B_j} \right) \end{aligned}$$

where the second step uses joint concavity of the root fidelity, and we trace out all systems but  $B_0 B_j$ , with  $\sigma_{B_j}$  being some appropriate state. Now, the fact that  $\langle \phi_{AB}^+ | X_A \otimes \tau_B | \phi_{AB}^+ \rangle = \frac{1}{d^2} \operatorname{tr}(X_A)$  for any operator  $X_A$  and data processing inequality with respect to the binary measurement  $\{\phi_{B_0 B_j}^+, I - \phi_{B_0 B_j}^+\}$  gives

$$F_d^*(N) \leq \sqrt{f} \left( \frac{1}{d^2}, p(j) + (1 - p(j)) \frac{1}{d^2} \right),$$

where  $\sqrt{f}(x, y) = \sqrt{xy} + \sqrt{(1-x)(1-y)}$  is the binary root fidelity. Note that  $f(q, p + (1-p)q)$  is monotonically increasing as  $p$  decreases from 1 to 0. Now, one of the  $N$  probabilities  $p(j)$  has to be  $\geq 1/N$ . Thus,

$$F_d^*(N) \leq \sqrt{f} \left( \frac{1}{d^2}, \frac{1}{N} + \left(1 - \frac{1}{N}\right) \frac{1}{d^2} \right). \quad (7.6)$$

To derive the non-asymptotic bound (7.4), Equation (7.6) can be rearranged as

$$\begin{aligned} F_d^*(N) &\leq \frac{1}{d^2} \left[ (d^2 - 1) \left(1 - \frac{1}{2N}\right) \sqrt{1 - \frac{1}{(1 - 2N)^2}} + \left(\frac{d^2 - 1}{2N} + 1\right) \right. \\ &\quad \left. \sqrt{1 - \frac{(d^2 - 1)^2}{(d^2 + 2N - 1)^2}} \right]. \end{aligned}$$

We bound the square roots using  $\sqrt{1+a} \leq 1 + a/2$  for any  $a \geq -1$  to obtain

$$\begin{aligned} F_d^*(N) &\leq \frac{1}{d^2} \left[ (d^2 - 1) \left(1 - \frac{1}{2N}\right) \left(1 - \frac{1}{2(1 - 2N)^2}\right) + \left(\frac{d^2 - 1}{2N} + 1\right) \right. \\ &\quad \left. \left(1 - \frac{(d^2 - 1)^2}{2(d^2 + 2N - 1)^2}\right) \right] \\ &= 1 - \frac{d^2 - 1}{8N^2} \frac{1}{\left(1 - \frac{1}{2N}\right) \left(1 + \frac{d^2 - 1}{2N}\right)} \\ &\leq 1 - \frac{d^2 - 1}{8N^2} \frac{1}{1 + \frac{d^2 - 2}{2N}}, \end{aligned}$$

which is (7.4). For  $N \rightarrow \infty$  this implies

$$F_d^*(N) \leq 1 - \frac{d^2 - 1}{8} \frac{1}{N^2} + O(N^{-3}),$$

which is (7.5) and concludes the proof.  $\square$

Combining Theorem 7.5 with Corollary 7.4 above yields a simplified bound as a corollary, that we stated as Corollary 1.6 in Sect. 1.2 as one of our main results. We restate it below for convenience, and in Fig. 4 we compare the quality of this bound for  $N > d^2/2$  with the converse bound (3.4) derived in [19].

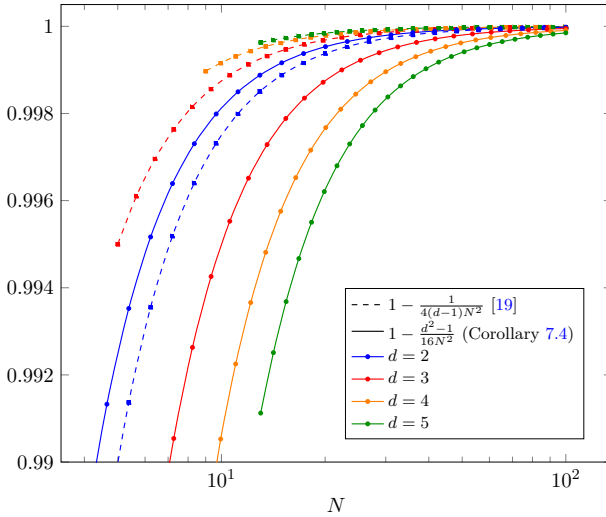
**Corollary 1.6** (Restated). *For a general port-based teleportation scheme with input dimension  $d$  and  $N$  ports, the entanglement fidelity  $F_d^*$  and the diamond norm error  $\varepsilon_d^*$  can be bounded as*

$$F_d^*(N) \leq \begin{cases} \frac{\sqrt{N}}{d} & \text{if } N \leq \frac{d^2}{2} \\ 1 - \frac{d^2 - 1}{16N^2} & \text{otherwise} \end{cases} \quad \varepsilon_d^*(N) \geq \begin{cases} 2\left(1 - \frac{\sqrt{N}}{d}\right) & \text{if } N \leq \frac{d^2}{2} \\ 2\frac{d^2 - 1}{16N^2} & \text{otherwise.} \end{cases}$$

## 8. Conclusion

In this paper, we completed the picture of the asymptotic performance of port-based teleportation (PBT) in the important regime when the input dimension is fixed while the number of ports tends to infinity. In particular, we determined the asymptotic performance of deterministic PBT in the fully optimized setting, showing that the optimal infidelity decays as  $\Theta(1/N^2)$  with the number of ports  $N$ . We also determined the precise asymptotics of the standard protocol for deterministic PBT (which uses EPR pairs and the ‘pretty good’ measurement) as well as probabilistic PBT using EPR pairs. The asymptotics for probabilistic PBT in the fully optimized setting had been determined previously in [22].

While our work closes a chapter in the study of PBT, it opens several interesting avenues for further investigation, both in the finite and in the asymptotic regime. Note that the limit  $d \rightarrow \infty$  for fixed  $N$  is not very interesting, as the error tends to one in this regime. However, it would be natural to consider limits where both  $N$  and  $d$  tend to infinity. In particular, the fidelity  $F_d^*(N)$  plausibly has a nontrivial limit when the



**Fig. 4.** Comparison of the converse bound  $F_d^*(N) \leq 1 - \frac{1}{4(d-1)N^2}$  derived in [19] and the converse bound  $F_d^*(N) \leq 1 - \frac{d^2-1}{16N^2}$  derived in Corollary 7.4, valid for  $N > d^2/2$

ratio  $N/d^2$  remains fixed. Given the import of PBT to, e.g., instantaneous non-local quantum computation, it would be desirable to determine the limiting value. Finally, we also mention the problem of determining the exact functional dependence on  $d$  of the leading order coefficient  $\lim_{N \rightarrow \infty} N^2(1 - F_d^*(N))$  in fully optimized deterministic PBT. Furthermore, we hope that our mathematical tools will be helpful for determining the asymptotics of other quantum information-theoretic tasks that can also be characterized in terms of representation-theoretic data, such as quantum state purification.

*Acknowledgements.* We acknowledge interesting discussions with Charles Bordenave, Benoît Collins, Marek Mozrzyms, Māris Ozols, Jan Philip Solovej, Sergii Strelchuk, and Michał Studziński. MC and FS acknowledge financial support from the European Research Council (ERC Grant Agreement No. 337603) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059). MC further acknowledges the Quant-ERA project Quantalgo and the hospitality of the Center for Theoretical Physics at MIT, where part of this work was done. FL and GS are supported by National Science Foundation (NSF) Grant No. PHY 1734006. FL appreciates the hospitality of QuSoft, CWI, and the University of Amsterdam, where part of this work was done. CM was supported by a Netherlands Organisation for Scientific Research (NWO) VIDI Grant (639.022.519). GS is supported by the NSF Grant No. CCF 1652560. MW thanks JILA for hospitality, where this work was partly initiated. MW acknowledges financial support by the NWO through Veni Grant No. 680-47-459.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### A. Proof of Lemma 3.6

The following lemma was first derived in [21]. In this section we give an alternative proof. Our proof is elementary and only uses the Schur–Weyl duality and the Pieri rule.

**Lemma 3.6.** (Restated) *The eigenvalues of the operator*

$$T(N)_{AB^N} = \frac{1}{N} \left( \phi_{AB_1}^+ \otimes I_{B_1^c} + \dots + \phi_{AB_N}^+ \otimes I_{B_N^c} \right)$$

on  $(\mathbb{C}^d)^{\otimes(1+N)}$  are given by the numbers

$$\frac{1}{dN} \gamma_\mu(\alpha) = \frac{1}{d} \frac{d_\alpha m_{d,\mu}}{d_{\mu} m_{d,\alpha}},$$

where  $\alpha \vdash_d N - 1$ , the Young diagram  $\mu \vdash_d N$  is obtained from  $\alpha$  by adding a single box, and  $\gamma_\mu(\alpha)$  is defined in Eq. (3.10).

*Proof.* We note that the operator  $T(N)$  commutes with the action of  $U(d)$  by  $\bar{U} \otimes U^{\otimes N}$  as well as with the action of  $S_N$  that permutes the systems  $B_1, \dots, B_N$ . Let us work out the corresponding decomposition of  $(\mathbb{C}^d)^{1+N}$ : We first consider the action of  $U(d) \times U(d)$  by  $\bar{U} \otimes V^{\otimes N}$  together with the  $S_N$ . By Schur–Weyl duality,

$$(\mathbb{C}^d)^{1+N} \cong (\mathbb{C}^d)^* \otimes \bigoplus_{\mu \vdash_d N} V_\mu^d \otimes W_\mu.$$

The notation means that  $\mu$  runs over all Young diagrams with  $N$  boxes and no more than  $d$  rows (i.e.,  $\mu_1 \geq \dots \geq \mu_d \geq 0$  and  $\sum_j \mu_j = N$ ). We write  $V_\mu^d$  for the irreducible  $U(d)$ -representation with highest weight  $\mu$ , and  $W_\mu$  for the irreducible  $S_N$ -representation corresponding to the partition  $\mu$ .

The dual representation  $(\mathbb{C}^d)^*$  is not polynomial; its highest weight is  $(-1, 0, \dots, 0)$ . However,  $(\mathbb{C}^d)^* \cong V_{(1, \dots, 1, 0)}^d \otimes \det^{-1}$ . The (dual) Pieri rule tells us that  $V_{(1, \dots, 1, 0)}^d \otimes V_\mu^d$  contains all irreducible representations whose highest weight can be obtained by adding 1's to all but one of the rows (with multiplicity one). Tensoring with the determinant amounts to subtracting  $(-1, \dots, -1)$ , so the result of tensoring with  $(\mathbb{C}^d)^*$  amounts to subtracting 1 from one of the rows:

$$(\mathbb{C}^d)^* \otimes V_\mu^d = \bigoplus_{i: \mu_i > \mu_{i+1}} V_{\mu - \epsilon_i}^d,$$

where we write  $\epsilon_i$  for the  $i$ th standard basis vector. (Note that we set  $\mu_{d+1} = -\infty$ , so that  $i = d$  is always a valid choice in the direct sum above; hence,  $\mu - \epsilon_i$  is always a highest weight, but does not need to be a Young diagram.) Thus, we obtain the following multiplicity-free decomposition into  $U(d) \times S_N$ -representations:

$$(\mathbb{C}^d)^{1+N} \cong \bigoplus_{\mu \vdash_d N} \bigoplus_{i: \mu_i > \mu_{i+1}} V_{\mu - \epsilon_i}^d \otimes W_\mu.$$

The operator  $T(N)$  can be decomposed accordingly:

$$T(N) = \bigoplus_{\mu, i} t_{\mu, i} \cdot I_{V_{\mu - \epsilon_i}^d} \otimes I_{W_\mu}$$

for some  $t_{\mu,i} \geq 0$ . To determine the  $t_{\mu,i}$ , let us denote by  $P_\mu$  the isotypical projectors for the  $S_N$ -action on  $(\mathbb{C}^d)^{\otimes N}$  and by  $Q_\alpha$  the isotypical projectors for the  $U(d)$  action by  $\bar{U} \otimes U^{\otimes N}$  (they commute). Then:

$$\mathrm{tr} T(N)(I_A \otimes P_\mu) Q_{\mu-\epsilon_i} = t_{\mu,i} \dim(V_{\mu-\epsilon_i}^d) \dim(W_\mu). \quad (\text{A.1})$$

On the other hand:

$$\mathrm{tr} T(N)(I_A \otimes P_\mu) Q_{\mu-\epsilon_i} = \mathrm{tr} \phi_{AB_1}^+(I_A \otimes P_\mu) Q_{\mu-\epsilon_i}.$$

The maximally entangled state  $\phi_{AB_1}^+$  is invariant under  $\bar{U} \otimes U$ . This means that on the range of the projector  $\phi_{AB_1}^+$ , the actions of  $\bar{U} \otimes U^{\otimes N}$  and  $I_{AB_1} \otimes U^{\otimes(N-1)}$  agree! Explicitly:

$$\mathbb{C}|\phi^+\rangle_{AB_1} \otimes (\mathbb{C}^d)^{\otimes(N-1)} \cong \mathbb{C}|\phi^+\rangle_{AB_1} \otimes \bigoplus_{\alpha \vdash_d N-1} V_\alpha^d \otimes [\alpha] \cong \bigoplus_{\alpha \vdash_d N-1} V_\alpha^d \otimes [\alpha].$$

It follows that

$$\phi_{AB_1}^+ Q_\alpha = \phi_{AB_1}^+(I_{AB_1} \otimes Q'_\alpha)$$

where  $Q'_\alpha$  refers to the action of  $U(d)$  by  $U^{\otimes(N-1)}$  on  $B_2 \dots B_n$ , and so

$$\mathrm{tr} \phi_{AB_1}^+(I_A \otimes P_\mu) Q_{\mu-\epsilon_i} = \begin{cases} \mathrm{tr} \phi_{AB_1}^+(I_A \otimes P_\mu)(I_{AB_1} \otimes Q'_\alpha) & \text{if } \alpha := \mu - \epsilon_i \text{ is a partition,} \\ 0 & \text{otherwise.} \end{cases}$$

We can now trace over the A-system:

$$\mathrm{tr} \phi_{AB_1}^+(I_A \otimes P_\mu)(I_{AB_1} \otimes Q'_\alpha) = \frac{1}{d} \mathrm{tr} P_\mu(I_{B_1} \otimes Q'_\alpha).$$

The remaining trace is on  $(\mathbb{C}^d)^{\otimes N}$ . The operator  $P_\mu$  refers to the  $S_N$ -action, while  $Q'_\alpha$  refers to the  $U(d)$ -action by  $U^{\otimes(N-1)}$  on  $B_2 \dots B_n$ . Equivalently, we can define  $Q'_\alpha$  with respect to the  $S_{N-1}$  action by permuting the last  $N-1$  tensor factors. Using Schur–Weyl duality and the branching rule for restricting  $S_N$  to  $S_1 \times S_{N-1}$ :

$$(\mathbb{C}^d)^{\otimes N} = \bigoplus_{\mu} V_\mu^d \otimes W_\mu = \bigoplus_{\mu} V_\mu^d \otimes \bigoplus_{i: \alpha = \mu - \epsilon_i \text{ partition}} W_\alpha$$

And hence

$$\frac{1}{d} \mathrm{tr} P_\mu(I_{B_1} \otimes Q'_\alpha) = \frac{1}{d} \dim(V_\mu^d) \dim(W_\alpha)$$

in the case of interest. Comparing this with Eq. (A.1), we obtain the following result:

$$t_{\mu,i} = \frac{1}{d} \frac{\dim(V_\mu^d) \dim(W_\alpha)}{\dim(V_\alpha^d) \dim(W_\mu)}$$

if  $\alpha = \mu - \epsilon_i$  is a partition, and otherwise zero. These are the desired eigenvalues of  $T(N)$ .  $\square$

## B. A Family of Explicit Protocols for Deterministic PBT

Guessing a good candidate density  $c_\mu$  with a simple functional form for the optimization in Eq. (3.6) yields a protocol with performance close to the achievability bound Theorem 1.5.

**Theorem B.1.** *For fixed but arbitrary dimension  $d$ , there exists a concrete protocol for deterministic PBT with entanglement fidelity*

$$F \geq 1 - \frac{d^4(d+3)}{2N^2} - O(N^{-3})$$

*Proof.* Assume that  $N/d^2$  is an integer (otherwise use only the first  $d^2 \lfloor \frac{N}{d^2} \rfloor$  ports). Let  $c_\mu$  be defined such that

$$q(\mu) = c_\mu p(\mu) = \begin{cases} \eta_N (R^2 - r(\mu)^2)^2 & r \leq R \\ 0 & \text{else,} \end{cases}$$

with

$$\begin{aligned} r(\mu) &= \|\mu - \hat{\mu}\|_2, \\ \hat{\mu} &= \left( (2d-1)\frac{N}{d^2}, (2d-3)\frac{N}{d^2}, \dots, \frac{N}{d^2} \right), \\ R &= \sqrt{2}\frac{N}{d^2}, \end{aligned}$$

and

$$\eta_N = \left( \sum_{\substack{\mu \in \hat{\mu} + \Lambda_d \\ r(\mu) \leq R}} (R^2 - r(\mu)^2)^2 \right)^{-1}$$

is a normalization factor that ensures that  $q$  is a probability distribution.  $\hat{\mu}$  has Euclidean distance  $R$  from the boundary of the set of Young diagrams, i.e. all vectors  $\mu \in \hat{\mu} + \Lambda_d$  such that  $\|\mu - \hat{\mu}\|_2 \leq R$  are Young diagrams. We extend the probability distribution  $q$  to be defined on all  $v \in \hat{\mu} + \Lambda_d$  for convenience. Let  $B_L^{\Lambda_d}(v_0) = \{v \in v_0 + \Lambda_d \mid \|v - v_0\|_2 \leq L\}$ . We now look at the PBT-fidelity for the protocol using the density  $c_\mu$ . First note that the formula Eq. (3.6) can be rearranged in the following way,

$$\begin{aligned} d^2 F &= \sum_{\alpha \vdash_{dN-1}} \left( \sum_{\mu=\alpha+\square} \sqrt{q(\mu)} \right)^2 \\ &= \sum_{\alpha \vdash_{dN-1}} \sum_{\mu, \mu'=\alpha+\square} \sqrt{q(\mu)q(\mu')} \\ &= \sum_{\mu \vdash_{dN}} \sum_{\mu'=\mu+\square-\square} \sqrt{q(\mu)q(\mu')}. \end{aligned}$$

In the last line, the notation  $\mu' = \mu + \square - \square$  means summing over all possibilities to remove a square from  $\mu$  and adding one, including removing and adding the same square. Noting that all vectors in  $B_R^{\Lambda_d}(\hat{\mu})$  are Young diagrams, we can write

$$\begin{aligned}
d^2 F &= \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \sqrt{q(\mu)q(\mu + e_i - e_j)} \\
&= \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \sqrt{\frac{q(\mu + e_i - e_j)}{q(\mu)}} \\
&= \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \left(1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}}\right) \\
&= \sum_{\mu \in B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \left(1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}}\right) \\
&\quad + \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \left(1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}}\right).
\end{aligned} \tag{B.1}$$

Here we have defined the functions

$$g_{ij}(\mu) = \mu_j - \hat{\mu}_j - \mu_i + \hat{\mu}_i$$

and

$$f(\mu) = \left(R^2 - r(\mu)^2\right)^2.$$

The last equation holds because  $\|e_i - e_j\|_2 = (1 - \delta_{ij})\sqrt{2}$ , i.e. for all  $\mu \in B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})$  and all  $1 \leq i, j \leq d$  we have  $\mu + e_i - e_j \in B_R^{\Lambda_d}(\hat{\mu})$ .

We can bound the normalization constant as follows. Denote by  $\mathcal{P}(\Lambda_d)$  the unit cell of  $\Lambda_d$  with smallest diameter,  $\ell$ . The volume of the unit cell is  $\sqrt{d}$ , which can be seen as follows.<sup>8</sup> A basis for the lattice  $\Lambda_d = \{v \in \mathbb{Z}^d : \sum_{i=1}^d v_i = 0\}$  is given by  $B = \{b_i\}_{i=1}^{d-1}$ , where  $b_i = e_1 - e_{i+1}$ . It follows that  $M = B^T B$  is a  $(d-1) \times (d-1)$ -matrix with all diagonal elements equal to 2 and all off-diagonal elements equal to 1. The matrix  $M$  has one eigenvalue  $d$  corresponding to the eigenvector  $\sum_{i=1}^{d-1} e_i$ , and  $d-2$  eigenvalues 1 corresponding to the eigenvectors  $e_1 - e_{i+1}$ , respectively. Hence,  $\det(\Lambda_d) = \sqrt{d}$ .

Let further  $g : \mathbb{R}\Lambda_d \rightarrow \Lambda_d$  be the function such that for all  $x \in \mathbb{R}\Lambda^d$  there exist  $\gamma_i \in (-1/2, 1/2]$ ,  $i = 1, \dots, d-1$  such that

$$x = g_N(x) + \sum_{i=1}^{d-1} \gamma_i a_i.$$

<sup>8</sup> For a general lattice  $\mathcal{L}_d \subset \mathbb{R}^d$  with basis  $B = \{b_1, \dots, b_m\}$  (where  $m \leq d$ ), the volume of the unit cell of  $\mathcal{L}^d$  is equal to  $\det(\mathcal{L}_d) = \sqrt{\det(B^T B)}$ .



Heuristically,  $g$  is the function that maps every point in the  $(d - 1)$ -dimensional subspace  $\Lambda^d$  lives in to the lattice point  $v$  in whose surrounding unit cell it lies, where the surrounding unit cell is here the set  $\{v + \sum_{i=1}^{d-1} \gamma_i a_i | \gamma_i \in (-1/2, 1/2]\}$ , i.e. the point lies in the center of the cell. As  $f$  is nonnegative, we have with  $l$  as defined above that

$$\begin{aligned}
 \eta_N^{-1} &= \sum_{\mu \in B_R^{\Lambda^d}(\hat{\mu})} f(\mu) \\
 &\leq \frac{1}{\sqrt{d}} \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} f(g_N(x)) dx \\
 &\leq \frac{1}{\sqrt{d}} \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} f(x) dx \\
 &\quad + \frac{1}{\sqrt{d}} \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} \frac{\ell}{2} \max_{x': \|x-x'\|_2 \leq \ell/2} \|(\nabla f)(x')\|_2 dx
 \end{aligned} \tag{B.2}$$

The gradient of  $f$  is given by

$$(\nabla f)(x) = -4(R^2 - \|x\|_2^2)x.$$

We can bound

$$|4(R^2 - (r \pm l)^2)(r \pm l)| \leq 4(R^2 - (r - l)^2)(r + l),$$

so

$$\begin{aligned}
 &\frac{1}{\sqrt{d}} \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} \frac{\ell}{2} \max_{x': \|x-x'\|_2 \leq \ell/2} \|(\nabla f)(x')\|_2 dx \\
 &\leq \frac{1}{\sqrt{d}} \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} \frac{\ell}{2} 4(R^2 - (r(x) - \ell/2)^2)(r(x) + \ell/2) dx \\
 &= \frac{2\ell \text{vol}(\mathbb{S}_{d-2})}{\sqrt{d}} \int_0^{R+\ell/2} r^{d-2} (R^2 - (r - \ell/2)^2)(r + \ell/2) dr \\
 &= \frac{2\ell \text{vol}(\mathbb{S}_{d-2})}{\sqrt{d}} \left( \frac{1}{d} - \frac{1}{d+2} \right) (R + \ell/2)^{d+2} + O(R^{d+1}).
 \end{aligned}$$

Here, we changed into spherical coordinates with origin in  $\hat{\mu}$  in the third line, and  $\text{vol}(\mathbb{S}_{d-2})$  is the volume of the  $(d - 2)$ -dimensional sphere. Turning to the first term in Eq. (B.2), we calculate

$$\begin{aligned}
 \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} f(x) dx &= \int_{B_{R+\ell/2}^{\Lambda^d}(\hat{\mu})} (R^2 - r(x)^2)^2 dx \\
 &= \text{vol}(\mathbb{S}_{d-2}) \int_0^{R+\ell/2} r^{d-2} (R^2 - r^2)^2 dr \\
 &= \text{vol}(\mathbb{S}_{d-2}) (R + \ell/2)^{d+3} \left( \frac{1}{d-1} - \frac{2}{d+1} + \frac{1}{d+3} \right) \\
 &= \frac{8 \text{vol}(\mathbb{S}_{d-2}) (R + \ell/2)^{d+3}}{d^3 + 3d^2 - d - 3}.
 \end{aligned}$$

Combining the last two equations, expanding the polynomials of the form  $(R + \ell/2)^k$  and using the power series expansion of  $1/(1+x)$  we finally arrive at

$$\eta_N = \frac{\sqrt{d}(d^3 + 3d^2 - d - 3)}{8 \operatorname{vol}(\mathbb{S}_{d-2})} R^{-(d+3)} + O(R^{-(d+4)}).$$

Returning to equation Eq. (B.1), let us first bound the magnitude of the last term. To this end, observe that for  $r(\mu) \geq R - \sqrt{2}$ , we have

$$\begin{aligned} \sqrt{f(\mu)} &= (R^2 - r(\mu)^2) \\ &\leq 2\sqrt{2}R. \end{aligned}$$

Furthermore we have that

$$\mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \left( 1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}} \right) \leq 1 + 2 \frac{2R}{\sqrt{f(\mu)}},$$

and hence

$$\begin{aligned} &\sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \left( 1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}} \right) \\ &\leq d^2 \eta_N \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} (f(\mu) + 2\sqrt{f(\mu)}R) \\ &\leq d^2 \eta_N \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} \left( (2\sqrt{2}R)^2 + 4\sqrt{2}R^2 \right) \\ &\leq 4(2 + \sqrt{2})d^2 R^2 \eta_N \left| B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu}) \right|. \end{aligned}$$

To bound the number of lattice points in the spherical shell  $B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})$ , note that i) each lattice point is surrounded by its own unit cell, and ii) these cells have diameter  $\ell$ . Therefore all these unit cells are disjoint subsets of a shell of width  $\sqrt{2} + \ell$ , and hence we have the bound

$$\left| B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu}) \right| \leq \operatorname{vol} \mathbb{S}_{d-2} (R + \ell)^{d-2} (\ell + \sqrt{2}).$$

Combining the bounds we arrive at

$$\sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \mathbb{1}_{B_R(\hat{\mu})}(\mu + e_i - e_j) \left( 1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}} \right) = O(R^{-3})$$

Turning to the first expression on the right hand side of Eq. (B.1), we observe that both the set  $B_{R-\sqrt{2}}^{\Lambda_d}$  and the distribution  $q$  are invariant under the map  $\mu \mapsto 2\mu - \hat{\mu}$ , i.e.

central reflection about  $\hat{\mu}$ . Therefore the sum over  $g_{ij}(\mu)$ , which is linear in  $\mu - \hat{\mu}$ , vanishes, i.e.

$$\begin{aligned}
 & \sum_{\mu \in B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \left( 1 + 2 \frac{g_{ij}(\mu) - 1}{\sqrt{f(\mu)}} \right) \\
 &= \sum_{\mu \in B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \left( 1 - 2 \frac{1}{\sqrt{f(\mu)}} \right) \\
 &= \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \left( 1 - 2 \frac{1}{\sqrt{f(\mu)}} \right) - \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} q(\mu) \sum_{i,j=1}^d \left( 1 - 2 \frac{1}{\sqrt{f(\mu)}} \right) \\
 &\geq d^2 - \eta_N \left( 2 \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} \sqrt{f(\mu)} + \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} \sqrt{f(\mu)} (\sqrt{f(\mu)} - 2) \right)
 \end{aligned}$$

Using the same argument as for bounding  $\eta_N$ , we find

$$\begin{aligned}
 \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu})} \sqrt{f(\mu)} &\leq \frac{2\ell \text{vol}(\mathbb{S}_{d-2})}{\sqrt{d}} \int_0^{R+\ell/2} r^{d-2} (R^2 - (r - \ell/2)^2) dr \\
 &= \frac{\text{vol}(\mathbb{S}_{d-2}) R^{d+1}}{(d^2 - 1)\sqrt{d}} + O(R^d).
 \end{aligned}$$

The second term is bounded in the same way as the spherical shell sum above, yielding

$$\eta_N d^2 \sum_{\mu \in B_R^{\Lambda_d}(\hat{\mu}) \setminus B_{R-\sqrt{2}}^{\Lambda_d}(\hat{\mu})} \sqrt{f(\mu)} (\sqrt{f(\mu)} - 2) = O(R^{-3}).$$

Combining all bounds, we arrive at

$$F \geq 1 - \frac{d^3 + 3d^2 - d - 3}{d^2 - 1} R^{-2} + O(R^{-3}) = (d + 3)R^{-2} + O(R^{-3}).$$

Using  $R = \frac{N}{d^2}$  we obtain the final bound

$$F \geq 1 - \frac{d^4(d + 3)}{2N^2} + O(R^{-3}).$$

□

### C. The Maximal Eigenvalue of a $2 \times 2$ GUE<sub>0</sub> Matrix

The maximal eigenvalue  $\lambda_{\max}(\mathbf{G})$  of a  $2 \times 2$  GUE<sub>0</sub> matrix  $\mathbf{G}$  can be easily analyzed, as  $\lambda_{\max}(\mathbf{G}) = \sqrt{\frac{1}{2} \text{tr} \mathbf{G}^2}$ .

**Lemma C.1.** *For  $\mathbf{X} \sim \text{GUE}_0(2)$ ,  $\sqrt{2}\lambda_{\max}(\mathbf{G}) \sim \chi_3$ , where  $\chi_3$  is the chi-distribution with three degrees of freedom.<sup>9</sup> Consequently,  $\mathbb{E}[\lambda_{\max}(\mathbf{G})] = \frac{2}{\sqrt{\pi}}$ .*

<sup>9</sup> This distribution is also known as the Maxwell–Boltzmann distribution.

*Proof.* By definition, the probability density of  $\text{GUE}(d)$  is

$$p_{\text{GUE}}(M) = (2\pi)^{-\frac{d^2}{2}} \exp\left(-\frac{\text{tr } M^2}{2}\right),$$

and therefore we get

$$p_{\text{GUE}_0}(G) = (2\pi)^{-\frac{d^2-1}{2}} \exp\left(-\frac{\text{tr } G^2}{2}\right),$$

for the density of  $\text{GUE}_0$ . Writing  $\mathbf{G} = \sum_{i=1}^3 \mathbf{x}_i \sigma_i$  with the Pauli matrices  $\sigma_i$ ,  $i = 1, 2, 3$ , we see that the  $\mathbf{x}_i$  are independent normal random variables with variance  $1/2$ , and

$$\lambda_{\max}(\mathbf{G}) = \sqrt{\frac{\text{tr } \mathbf{G}^2}{2}} = \sqrt{\sum_{i=1}^3 \mathbf{x}_i^2},$$

proving the claim.  $\square$

## D. Technical Lemmas

The following ‘‘mirror lemma’’, also called ‘‘transpose trick’’, is well known in the literature, and can be proven in a straightforward way:

**Lemma D.1** (Mirror lemma, transpose trick). *Let  $\{|i\rangle\}_{i=1}^d$  be a basis and  $|\gamma\rangle = \sum_{i=1}^d |i\rangle|i\rangle$  be the unnormalized maximally entangled state. For any operator  $X$ ,*

$$I \otimes X |\gamma\rangle = X^T \otimes I |\gamma\rangle,$$

where  $X^T$  denotes transposition of  $X$  with respect to the basis  $\{|i\rangle\}_{i=1}^d$ .

The maximization in the definition of the diamond norm can be carried out explicitly for the distance of two unitarily covariant channels. This is the statement of the following lemma, which is a special case of a more general result about generalized divergences proven in [69].

**Lemma D.2** [69]. *Let  $\Lambda_{A \rightarrow A}^{(i)}$  for  $i = 1, 2$  be unitarily covariant maps. Then the maximally entangled state  $|\phi^+\rangle_{AA'}$  is a maximizer for their diamond norm distance, i.e.,*

$$\left\| \Lambda_{A \rightarrow A}^{(1)} - \Lambda_{A \rightarrow A}^{(2)} \right\|_{\diamond} = \left\| \left( \Lambda_{A \rightarrow A}^{(1)} - \Lambda_{A \rightarrow A}^{(2)} \right) (\phi_{AA'}^+) \right\|_1.$$

The following Lemma from Ref. [5] shows that the entanglement fidelity and the diamond norm distance to the identity channel are even in a 1-1 relation for unitarily covariant channels.

**Lemma D.3** [5]. *For a unitarily covariant channel  $\Lambda: A \rightarrow A$ ,*

$$\| \text{id}_A - \Lambda \|_{\diamond} = 2 \left( 1 - \sqrt{F(\Lambda)} \right).$$

We need an explicit limit of certain Riemann sums. The proof of the following can, e.g., be found in [70].

**Lemma D.4.** Let  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be nonincreasing such that the (proper or improper) Riemann integral

$$\int_a^b f(x)dx$$

exists for all  $a, b \in [0, \infty]$  with  $a < b$ . Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{gn} f\left(\frac{c+i}{n}\right) = \int_0^g f(x)dx$$

for all  $c \geq 0$  and  $g \in [0, \infty]$ .

The following lemma provides the volume of the simplex of ordered probability distributions as well as the volume of its boundary.

**Lemma D.5.** Let

$$\text{OS}_{d-1} = \left\{ x \in \mathbb{R}^d \mid \sum_i x_i = 0, x_i \geq x_{i+1}, x_d \geq 0 \right\}$$

be the simplex of ordered probability distributions. The volume of this simplex, and the volume of its boundary, are given by

$$\begin{aligned} \text{vol}(\text{OS}_{d-1}) &= \frac{1}{\sqrt{d}((d-1)!)^2}, \text{ and} \\ \text{vol}(\partial\text{OS}_{d-1}) &= \text{vol}(\text{OS}_{d-1}) \left( \frac{d(d-1)^2}{\sqrt{2}} + \sqrt{d}(d-1)^{3/2} + \sqrt{2}(d-1) \right), \end{aligned} \quad (\text{D.1})$$

respectively.

*Proof.*  $\text{OS}_{d-1}$  is given in its dual description above, let us therefore begin by finding its extremal points. These are clearly given by

$$v_i = \left( \frac{1}{i}, \dots, \frac{1}{i}, 0, \dots, 0 \right),$$

i.e. the  $i$ th extremal point has  $i$  entries  $\frac{1}{i}$  and  $d-i$  entries 0. The supporting (affine) hyperplanes  $H_i$  of the facets  $F_i, i = 1, \dots, d$  of  $\text{OS}_{d-1}$  in  $V_0^{(d-1)} = \{x \in \mathbb{R}^d \mid \sum_i x_i = 0\}$  are given by the normalized normal vectors

$$\begin{aligned} n_i &= \frac{e_i - e_{i+1}}{\sqrt{2}}, \quad i = 1, \dots, d-1, \text{ and} \\ n_d &= \frac{1}{\sqrt{d(d-1)}}(1, \dots, 1, -d+1). \end{aligned}$$

Now note that the facet  $F_d = \{x \in \text{OS}_{d-1} \mid x_d = 0\}$  is equal to  $\text{OS}_{d-2}$ , and the volume of a  $(d-1)$ -dimensional pyramid is given by the product of the volume of its base and its height, divided by  $d-1$ . Therefore we get the recursive formula

$$\text{vol}(\text{OS}_{d-1}) = \frac{1}{d-1} \text{vol}(\text{OS}_{d-2})h_d,$$

where we have defied the distance  $h_i$  between  $v_i$  and  $H_i$ . Let us calculate  $h_d$ . This can be done by taking the difference of  $v_d$  and any point in  $H_i$  and calculating the absolute value of its inner product with  $n_d$ . We thus get

$$\begin{aligned} h_d &= |\langle n_d, v_d - v_1 \rangle| \\ &= \frac{1}{\sqrt{d(d-1)}} \left| -\frac{d-1}{d} + (d-2)\frac{1}{d} - \frac{d-1}{d} \right| \\ &= \frac{1}{\sqrt{d(d-1)}}. \end{aligned}$$

The recursion therefore becomes

$$\text{vol}(\text{OS}_{d-1}) = \sqrt{\frac{d-1}{d}} \frac{1}{(d-1)^2} \text{vol}(\text{OS}_{d-2}).$$

The claimed formula for the volume is now proven by induction.  $\text{OS}_2$  is just the line from  $(1, 0)$  to  $(1/2, 1/2)$ , so its volume is clearly

$$\text{vol}(\text{OS}_2) = \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}(1!)^2},$$

proving Eq. (D.1) for  $d = 2$ . For the induction step, assume that the formula Eq. (D.1) holds for  $d = k - 1$ . Then we have

$$\begin{aligned} \text{vol}(\text{OS}_{k-1}) &= \sqrt{\frac{k-1}{k}} \frac{1}{(k-1)^2} \text{vol}(\text{OS}_{k-2}) \\ &= \sqrt{\frac{k-1}{k}} \frac{1}{(k-1)^2} \frac{1}{\sqrt{k-1}((k-2)!)^2} \\ &= \frac{1}{\sqrt{k}((k-1)!)^2}. \end{aligned}$$

For the boundary volume, we can use the pyramid volume formula again to obtain

$$\text{vol}(\text{OS}_{d-1}) = \frac{1}{d-1} \text{vol}(F_i) h_i,$$

i.e. we obtain the formula

$$\begin{aligned} \text{vol}(\partial\text{OS}_{d-1}) &= \sum_{i=1}^d \text{vol}(F_i) \\ &= (d-1) \text{vol}(\text{OS}_{d-1}) \sum_{i=1}^d \frac{1}{h_i}. \end{aligned}$$

We calculate the heights  $h_i$  for  $i \neq d$ . For  $1 < i < d$  we get in the same way as above for  $i = d$ ,

$$h_i = |\langle n_i, v_i - v_1 \rangle|$$

$$= \frac{1}{i\sqrt{2}}.$$

for  $i = 1$  we calculate

$$\begin{aligned} h_1 &= |\langle n_1, v_1 - v_2 \rangle| \\ &= \frac{1}{2\sqrt{2}}. \end{aligned}$$

Therefore we get the boundary volume

$$\begin{aligned} \text{vol}(\partial\text{OS}_{d-1}) &= (d-1) \text{vol}(\text{OS}_{d-1}) \left( 2\sqrt{2} + \sqrt{d(d-1)} + \sqrt{2} \sum_{i=2}^{d-1} i \right) \\ &= (d-1) \text{vol}(\text{OS}_{d-1}) \left( \sqrt{2} + \sqrt{d(d-1)} + \frac{d(d-1)}{\sqrt{2}} \right) \\ &= \text{vol}(\text{OS}_{d-1}) \left( \frac{d(d-1)^2}{\sqrt{2}} + \sqrt{d}(d-1)^{3/2} + \sqrt{2}(d-1) \right). \end{aligned}$$

## References

1. Bennett, C.H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A., Wootters, W.K.: Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **70**(13), 1895 (1993)
2. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935)
3. Ishizaka, S., Hiroshima, T.: Asymptotic teleportation scheme as a universal programmable quantum processor. *Phys. Rev. Lett.* **101**(24), 240501 (2008). [arXiv:0807.4568](#) [quant-ph]
4. Ishizaka, S., Hiroshima, T.: Quantum teleportation scheme by selecting one of multiple output ports. *Phys. Rev. A* **79**(4), 042306 (2009). [arXiv:0901.2975](#) [quant-ph]
5. Pirandola, S., Laurenza, R., Lupo, C., Pereira, J.L.: Fundamental limits to quantum channel discrimination. *npj Quant. Inf.* **5**(1), 50 (2019). [arXiv:1803.02834](#) [quant-ph]
6. Beigi, S., König, R.: Simplified instantaneous non-local quantum computation with applications to position-based cryptography. *New J. Phys.* **13**(9), 093036 (2011). [arXiv:1101.1065](#) [quant-ph]
7. Buhrman, H., Chandran, N., Fehr, S., Gelles, R., Goyal, V., Ostrovsky, R., Schaffner, C.: Position-based quantum cryptography: impossibility and constructions. *SIAM J. Comput.* **43**(1), 150–178 (2014). [arXiv:1009.2490](#) [quant-ph]
8. Chandran, N., Goyal, V., Moriarty, R., Ostrovsky, R.: Cryptography, position based. In: Halevi, S. (ed.) *Advances in Cryptology—CRYPTO*, vol. 2009, pp. 391–407. Springer, Berlin (2009)
9. Malaney, R.: The quantum car. *IEEE Wirel. Commun. Lett.* **5**(6), 624–627 (2016). [arXiv:1512.03521](#) [quant-ph]
10. Unruh, D.: Verification, quantum position, in the random oracle model. In: Garay, J.A., Gennaro, R. (eds.) *Advances in Cryptology—CRYPTO*, vol. 2014, pp. 1–18. Springer, Berlin (2014)
11. Yu, L.: Fast controlled unitary protocols using group or quasigroup structures. *arXiv preprint* (2011). [arXiv:1112.0307](#) [quant-ph]
12. Li, Y., Griffiths, R.B., Cohen, S.M.: Fast protocols for local implementation of bipartite nonlocal unitaries. *Phys. Rev. A* **85**(1), 012304 (2012). [arXiv:1109.5013](#) [quant-ph]
13. Broadbent, A.: Popescu–Rohrlich correlations imply efficient instantaneous nonlocal quantum computation. *Phys. Rev. A* **94**(2), 022318 (2016). [arXiv:1512.04930](#) [quant-ph]
14. Speelman, F.: Instantaneous non-local computation of Low T-depth quantum circuits. In: *11th Conference on the Theory of Quantum Computation, Communication and Cryptography—TQC* (2016). [arXiv:1511.02839](#) [quant-ph]
15. Tomamichel, M., Fehr, S., Kaniewski, J., Wehner, S.: A monogamy-of-entanglement game with applications to device-independent quantum cryptography. *New J. Phys.* **15**(10), 103002 (2013). [arXiv:1210.4359](#) [quant-ph]

16. Ribeiro, J., Grosshans, F.: A tight lower bound for the BB84-states quantumposition- verification protocol. arXiv preprint (2015). [arXiv:1504.07171](https://arxiv.org/abs/1504.07171) [quant-ph]
17. Strelchuk, S., Horodecki, M., Oppenheim, J.: Generalized teleportation and entanglement recycling. Phys. Rev. Lett. **110**(1), 010505 (2013). [arXiv:1209.2683](https://arxiv.org/abs/1209.2683) [quant-ph]
18. Pitalúa-García, D.: Deduction of an upper bound on the success probability of port-based teleportation from the no-cloning theorem and the no-signaling principle. Phys. Rev. A **87**(4), 040303 (2013). [arXiv:1206.4836](https://arxiv.org/abs/1206.4836) [quant-ph]
19. Ishizaka, S.: Some remarks on port-based teleportation. arXiv preprint (2015). [arXiv:1506.01555](https://arxiv.org/abs/1506.01555) [quant-ph]
20. Wang, Z.-W., Braunstein, S.L.: Higher-dimensional performance of port-based teleportation. Sci. Rep. **6** (2016)
21. Studziński, M., Strelchuk, S., Mozrzykas, M., Horodecki, M.: Portbased teleportation in arbitrary dimension. Sci. Rep. **7**, 10871 (2017). [arXiv:1612.09260](https://arxiv.org/abs/1612.09260) [quant-ph]
22. Mozrzykas, M., Studziński, M., Strelchuk, S., Horodecki, M.: Optimal port-based teleportation. New J. Phys. **20**(5), 053006 (2018). [arXiv:1707.08456](https://arxiv.org/abs/1707.08456) [quant-ph]
23. Supplementary Python code. <https://github.com/amsqi/port-based>
24. Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions I. In: Mehra, J. (ed.) The Collected Works of Eugene Paul Wigner, pp. 524–540. Springer, New York (1993)
25. Keyl, M., Werner, R.F.: Estimating the spectrum of a density operator. Phys. Rev. A **64**(5), 052311 (2001). [arXiv:quant-ph/0102027](https://arxiv.org/abs/quant-ph/0102027)
26. Hayashi, M., Matsumoto, K.: Quantum universal variable-length source coding. Phys. Rev. A **66**(2), 022311 (2002). [arXiv:quant-ph/0202001](https://arxiv.org/abs/quant-ph/0202001)
27. Christandl, M., Mitchison, G.: The spectra of quantum states and the Kronecker coefficients of the symmetric group. Commun. Math. Phys. **261**(3), 789–797 (2006)
28. O’Donnell, R., Wright, J.: Quantum spectrum testing. In: Proceedings of the 47th Annual ACM Symposium on Theory of Computing. STOC 2015. ACM, pp. 529–538 (2015). [arXiv:1501.05028](https://arxiv.org/abs/1501.05028) [quant-ph]
29. Haah, J., Harrow, A.W., Ji, Z., Xiaodi, W., Nengkun, Yu.: Sampleoptimal tomography of quantum states. IEEE Trans. Inf. Theory **63**(9), 5628–5641 (2017). [arXiv:1508.01797](https://arxiv.org/abs/1508.01797) [quant-ph]
30. O’Donnell, R., Wright, J.: Efficient quantum tomography. In: Proceedings of the 48th Annual ACM Symposium on Theory of Computing. STOC 2016. ACM, pp. 899–912 (2016). [arXiv:1508.01907](https://arxiv.org/abs/1508.01907) [quant-ph]
31. O’Donnell, R., Wright, J.: Efficient quantum tomography II. In: Proceedings of the 49th Annual ACM Symposium on Theory of Computing. STOC 2017. ACM, pp. 962–974 (2017). [arXiv:1612.00034](https://arxiv.org/abs/1612.00034) [quant-ph]
32. Christandl, M., Doran, B., Walter, M.: Computing multiplicities of Lie group representations. In: 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science. FOCS 2012. IEEE, pp. 639–648 (2012). [arXiv:1204.4379](https://arxiv.org/abs/1204.4379) [cs.CC]
33. Christandl, M., Doran, B., Kousidis, S., Walter, M.: Eigenvalue distributions of reduced density matrices. Commun. Math. Phys. **332**(1), 1–52 (2014). [arXiv:1204.0741](https://arxiv.org/abs/1204.0741) [quant-ph]
34. Christandl, M., Şahinoğlu, M.B., Walter, M.: Recoupling coefficients and quantum entropies. Ann. Henri Poincaré **19**(2), 385–410 (2018). [arXiv:1210.0463](https://arxiv.org/abs/1210.0463) [quant-ph]
35. Christandl, M., Harrow, A.W., Mitchison, G.: Nonzero Kronecker coefficients and what they tell us about spectra. Commun. Math. Phys. **270**(3), 575–585 (2007). [arXiv:quant-ph/0511029](https://arxiv.org/abs/quant-ph/0511029)
36. Johansson, K.: Discrete orthogonal polynomial ensembles and the Plancherel measure. Ann. Math. **153**(1), 259–296 (2001). [arXiv:math/9906120](https://arxiv.org/abs/math/9906120) [math.CO]
37. Freitas, P., Krejčířik, D.: A sharp upper bound for the first Dirichlet eigenvalue and the growth of the isoperimetric constant of convex domains. Proc. Am. Math. Soc. **136**(8), 2997–3006 (2008). [arXiv:0710.5475](https://arxiv.org/abs/0710.5475) [math.SP]
38. Kubicki, A.M., Palazuelos, C., Pérez-García, D.: Resource quantification for the no-programming theorem. Phys. Rev. Lett. **122**(8), 080505 (2019). [arXiv:1805.00756](https://arxiv.org/abs/1805.00756) [quant-ph]
39. Fulton, W.: Young Tableaux: With Applications to Representation Theory and Geometry, vol. 35. Cambridge University Press, Cambridge (1997)
40. Simon, B.: Representations of Finite and Compact Groups. Graduate Studies in Mathematics, vol. 10. American Mathematical Society, Providence (1996)
41. Nielsen, M.A., Chuang, I.L.: Programmable quantum gate arrays. Phys. Rev. Lett. **79**(2), 321 (1997). [arXiv:quant-ph/9703032](https://arxiv.org/abs/quant-ph/9703032)
42. Nielsen, B.M., Neergaard-Nielsen, J.S., Polzik, E.S.: Time gating of heralded single photons for atomic memories. Opt. Lett. **34**(24), 3872–3874 (2009). [arXiv:0909.0646](https://arxiv.org/abs/0909.0646) [quant-ph]
43. Belavkin, V.P.: Optimal multiple quantum statistical hypothesis testing. Stoch. Int. J. Probab. Stoch. Process. **1**(1–4), 315–345 (1975)
44. Hausladen, P., Wootters, W.K.: A ‘pretty good’ measurement for distinguishing quantum states. J. Mod. Opt. **41**(12), 2385–2390 (1994)



45. Pérez-García, D.: Optimality of programmable quantum measurements. *Phys. Rev. A* **73**(5), 052315 (2006). [arXiv:quant-ph/0602084](https://arxiv.org/abs/quant-ph/0602084)
46. Hillery, M., Ziman, M., Bužek, V.: Approximate programmable quantum processors. *Phys. Rev. A* **73**(2), 022345 (2006). [arXiv:quant-ph/0510161](https://arxiv.org/abs/quant-ph/0510161)
47. Renner, R.: Security of quantum key distribution. Ph.D. thesis. ETH Zürich (2005). [arXiv:quant-ph/0512258](https://arxiv.org/abs/quant-ph/0512258)
48. Christandl, M., König, R., Renner, R.: Postselection technique for quantum channels with applications to quantum cryptography. *Phys. Rev. Lett.* **102**, 020504 (2009). [arXiv:0809.3019](https://arxiv.org/abs/0809.3019) [quant-ph]
49. Gross, D., Nezami, S., Walter, M.: Schur–Weyl Duality for the Clifford Group with Applications: Property Testing, a Robust Hudson Theorem, and de Finetti Representations. arXiv preprint (2017). [arXiv:1712.08628](https://arxiv.org/abs/1712.08628) [quant-ph]
50. Seymour, P.D., Zaslavsky, T.: Averaging sets: a generalization of mean values and spherical designs. *Adv. Math.* **52**(3), 213–240 (1984)
51. Kane, D.: Small designs for path-connected spaces and path-connected homogeneous spaces. *Trans. Am. Math. Soc.* **367**(9), 6387–6414 (2015). [arXiv:1112.4900](https://arxiv.org/abs/1112.4900) [math.CO]
52. Mozrzykas, M., Studziński, M., Horodecki, M.: A simplified formalism of the algebra of partially transposed permutation operators with applications. *J. Phys. A Math. Theor.* **51**(12), 125202 (2018). [arXiv:1708.02434](https://arxiv.org/abs/1708.02434) [quant-ph]
53. Johnson, P.D., Viola, L.: Compatible quantum correlations: extension problems for werner and isotropic states. *Phys. Rev. A* **88**(3), 032323 (2013). [arXiv:1305.1342](https://arxiv.org/abs/1305.1342) [quant-ph]
54. Alicki, R., Rudnicki, S., Sadowski, S.: Symmetry properties of product states for the system of  $N$   $n$ -level atoms. *J. Math. Phys.* **29**(5), 1158–1162 (1988)
55. Duffield, N.G.: A large deviation principle for the reduction of product representations. *Proc. Am. Math. Soc.* **109**(2), 503–515 (1990)
56. Tracy, C.A., Widom, H.: On the distributions of the lengths of the longest monotone subsequences in random words. *Probab. Theory Relat. Fields* **119**(3), 350–380 (2001). [arXiv:math/9904042](https://arxiv.org/abs/math/9904042) [math.CO]
57. Kuperberg, G.: Random words, quantum statistics, central limits, random matrices. arXiv preprint (1999). [arXiv: math/9909104](https://arxiv.org/abs/math/9909104) [math.PR]
58. Pinsker, M.S.: Information and Information Stability of Random Variables and Processes. Holden-Day, San Francisco (1964)
59. Durrett, R.: Probability: Theory and Examples, 4th edn. Cambridge University Press, Cambridge (2010)
60. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. 2. Wiley, New York (2008)
61. Ledoux, M.: Deviation Inequalities on Largest Eigenvalues. Geometric Aspects of Functional Analysis, pp. 167–219. Springer, New York (2007)
62. Grebenkov, D.S., Nguyen, B.-T.: Geometrical structure of Laplacian eigenfunctions. *SIAM Rev.* **55**(4), 601–667 (2013). [arXiv:1206.1278](https://arxiv.org/abs/1206.1278) [math.AP]
63. Krahn, E.: Über Minimaleigenschaften der Kugel in drei und mehr Dimensionen. *Mattiesen* (1926)
64. Chambers, L.G.: An upper bound for the first zero of Bessel functions. *Math. Comput.* **38**(158), 589–591 (1982)
65. Breen, S.: Uniform upper and lower bounds on the zeros of Bessel functions of the first kind. *J. Math. Anal. Appl.* **196**(1), 1–17 (1995)
66. Majenz, C.: Entropy in Quantum Information Theory—Communication and Cryptography. Ph.D. thesis. Department of Mathematical Sciences, Faculty of Science, University of Copenhagen (2017)
67. Berta, M., Christandl, M., Renner, R.: The quantum reverse Shannon theorem based on one-shot information theory. *Commun. Math. Phys.* **306**(3), 579 (2011). [arXiv:0912.3805](https://arxiv.org/abs/0912.3805) [quant-ph]
68. Datta, N.: Min-and max-relative entropies and a new entanglement monotone. *IEEE Trans. Inf. Theory* **55**(6), 2816–2826 (2009). [arXiv:0803.2770](https://arxiv.org/abs/0803.2770) [quant-ph]
69. Leditzky, F., Kaur, E., Datta, N., Wilde, M.M.: Approaches for approximate additivity of the Holevo information of quantum channels. *Phys. Rev. A* **97**(1), 012332 (2018). [arXiv:1709.01111](https://arxiv.org/abs/1709.01111) [quant-ph]
70. Convergence of Riemann sums for improper integrals. <https://math.stackexchange.com/questions/1744250/convergence-of-riemann-sums-for-improper-integrals>. Accessed 30 Apr 2018