Metrology-assisted entanglement distribution in noisy quantum networks

Simon Morelli[1,2], David Sauerwein[3], Michalis Skotiniotis[4], and Nicolai Friis[1,2]

[1]Institute for Quantum Optics and Quantum Information — IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

[2]Atominstitut, Technische Universität Wien, 1020 Vienna, Austria

[3]Amazon Web Services Europe, Zürich, Switzerland

[4]Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

arXiv:2110.15627v2 [quant-ph] 23 May 2022

We consider the distribution of high-dimensional entangled states to multiple parties via noisy channels and the subsequent probabilistic conversion of these states to desired target states using stochastic local operations and classical communication. We show that such state-conversion protocols can be enhanced by embedded channel-estimation routines at no additional cost in terms of the number of copies of the distributed states. The defining characteristic of our strategy is the use of those copies for which the conversion was unsuccessful for the estimation of the noise, thus allowing one to counteract its detrimental effect on the successfully converted copies. Although this idea generalizes to various more complex situations, we focus on the realistic scenario, where only finitely many copies are distributed and where the parties are not required to process multiple copies simultaneously. In particular, we investigate the performance of so-called one-successful-branch protocols, applied sequentially to single copies and an adaptive Bayesian estimation strategy. Finally, we compare our strategy to more general but less easily practically implementable strategies involving distillation and the use of quantum memories to process multiple copies simultaneously.

1 Introduction

Entanglement between two or more parties is an important ingredient for the development of many quantum technologies such as, e.g., fault-tolerant quantum computation [1–4], quantum simulation [5–7], and, in particular, communication via quantum networks [8–12]. A landmark goal of quantum technologies is the establishment of a 'quantum internet' [13–15]—a highly interconnected quantum network able to distribute and manipulate entangled quantum states via fibres and optical links. Recent efforts [16] go in the direction of understanding the resources, challenges and opportunities that come along with such an endeavour. In this context, entanglement is a crucial resource that allows spatially separated parties to overcome the restriction of local operations and classical communication (LOCC) thereby implementing classically impossible tasks, such as quantum teleportation [17–19]. However, the technological requirements for quantum processors and networks on a large scale are immense and will have to involve significant improvements in the efficient manipulation and control of quantum systems. A deeper comprehension of the possibilities and limitations for multipartite entanglement distribution and its manipulation via LOCC is hence necessary. The present work aims to contribute to this endeavour by exploring the synergies between stochastic LOCC (SLOCC) protocols—that aim to convert multiparty quantum states—and quantum parameter estimation protocols whose goal is to improve entanglement distribution by precisely identifying the noise suffered by such states when distributed through quantum networks.

On the theoretical side, a lot of progress has been made with regards to the conversion of multipartite entanglement via LOCC (cf. [20–26]), despite the fact that LOCC is notoriously difficult to handle mathematically [27], while practical developments towards efficient entanglement certification techniques (see, e.g., [28–33]) and real-world quantum networks are rapidly progressing (cf. [34]). Yet, basic theoretical questions about multipartite entanglement structures arising in networks have only begun to be explored [35–39]. In particular, apart from recent examples [40, 41], mostly transformations of pure states in the single-copy regime or the asymptotic limit of infinitely many copies have been considered.

Here, we consider a scenario that more closely resembles situations expected to be encountered in future real-world quantum networks: the distribution of multiple but finitely many copies of multipartite quantum states via imperfect or varying channels available only for a restricted period of time. There, one central node, the vendor, is assumed to be capable of implementing the entangling operations necessary to create a highly entangled quantum state that is subsequently distributed to spatially separated parties via imperfect but local quantum channels.

Simon Morelli: simon.morelli@oeaw.ac.at
David Sauerwein: sauerwein.david@gmail.com, This work was done prior to joining AWS.
Michalis Skotiniotis: michail.skoteiniotis@uab.cat
Nicolai Friis: nicolai.friis@univie.ac.at

In general, the states provided by such a vendor will not match the target state that the parties might wish to employ in their communication or computation protocol of choice. For multiple parties, there is no unique resource state that can be deterministically converted to any other state via LOCC; in fact, the so-called maximally entangled set (MES) of states from which any state can be reached in this way has infinitely many elements [22]. In addition, the vendor might only be able to create a limited variety of such states and may therefore be unable to offer the desired state. Indeed, this might even be in the interest of the users, as they may not want to divulge information on the specific form (and hence the intended use) of their desired target state. Since deterministic LOCC transformations between multi-party pure states (with fixed local dimensions) almost never exist [26, 42], the distributed state will have to be converted to the target state *probabilistically*, i.e., via SLOCC. Two states that can be converted into each other via SLOCC are said to belong to the same *entanglement class*.

Optimal SLOCC conversion protocols will generally require acting on many copies simultaneously. Beside the fact that little is known about the structure of such protocols yet, the implementation of complex multi-copy operations also represents a substantial technological challenge. Therefore we focus on sequential strategies that process a single copy at a time. Compared to general protocols the difficulty to implement a sequential strategy does not increase with the number of copies. To employ a finite number, k_s, of copies for multipartite quantum communication or computation, users must purchase, on average, a larger number $k = k_\mathrm{s}/p_\mathrm{s}$ of copies from the vendor to compensate for the nonzero failure probability $p_\mathrm{f} = 1 - p_\mathrm{s} > 0$ of the conversion protocol.

Here, we demonstrate that this apparent disadvantage has important redeeming qualities in the presence of noise. We show that, for certain 'benign' types of noise, those copies for which the state conversion was unsuccessful can nevertheless be useful for obtaining information about the noise and hence partially compensate its effects on the successfully converted copies. Our results thus establish that multipartite entanglement distribution protocols can be assisted by quantum metrology at no additional resource costs in terms of additionally distributed states, and thus feature built-in noise robustness 'for free'.

To reach this conclusion, we first consider a particular situation as described above, i.e., a family of specific noisy multipartite state conversion protocols and show how the use of parameter estimation tools leads to an improvement in the quality of the resulting final states. Based on this exemplary situation, we then discuss the applicability of this approach to more general entanglement distribution protocols and contrast it with alternative strategies based on distillation. Finally, we discuss the merits of metrology-assisted entanglement conversion protocols and argue for their default integration in future quantum networks.

2 Noisy entanglement distribution

We consider a situation where N parties require several copies of a highly entangled N-qudit state ρ_target for a specific task. They purchase a number of copies from a vendor, where the purchased state ρ_vendor differs from the target state ρ_target, but belongs to the same entanglement class and can thus be converted to the target state via SLOCC. The vendor distributes the state ρ_vendor to the N parties via local quantum channels represented by completely-positive, trace-preserving (CPTP) maps, $\{\Lambda_i\}_{i=1}^N$, so that the state shared by the N parties is given by

$$\rho_\mathrm{received} = \Lambda(\rho_\mathrm{vendor}) = \bigotimes_{i=1}^{N} \Lambda_i(\rho_\mathrm{vendor}). \tag{1}$$

The parties now wish to transform the received state ρ_received into the target state ρ_target within the same entanglement class. This can be done either deterministically or probabilistically. We focus on transformations of fully entangled (i.e., with full local rank) pure states via a so-called one-successful-branch protocol (OSBP) [43, 44]. The OSBP is in fact the optimal SLOCC protocol for almost all N-qudit pure-state transformations [44]. However, the idea of using unsuccessful branches of SLOCC protocols for parameter estimation presented in this paper applies in equal measure to any other probabilistic transformations of pure and mixed states of arbitrary dimensions.

In an OSBP that transforms $|\Psi\rangle$ into $|\Phi\rangle$, each party k performs a single two-outcome measurement $\{M_\mathrm{s}^k, M_\mathrm{f}^k\}$ with $\sum_{i=\mathrm{s,f}} (M_i^k)^\dagger M_i^k = \mathbb{1}$. The successful outcome s leads to a state that can still be transformed to the final state, while the failure outcome f leads to a state that is no longer fully entangled and can therefore no longer be transformed into $|\Phi\rangle$. Hence, the final state is only reached if the successful measurement outcome is realised for all parties. That is, only one branch of the OSBP is successful, as suggested by its name, and realized with some probability p_s.

The transformation performed by the N parties to convert the received state into the target state depends on the channel used to distribute the state. For the general case when the channel is not sufficiently known, the task for the parties becomes more challenging. They have to estimate the channel, i.e., identify the channel Λ from a discrete or continuous family of possible candidates, before transforming the received state ρ_received. The goal of the N parties is now to obtain the largest number of copies ρ_corr that are ϵ-close in trace distance to the desired target state ρ_target, i.e.,

$$D(\rho_\mathrm{corr}, \rho_\mathrm{target}) = \frac{1}{2}\mathrm{Tr}\left[\sqrt{(\rho_\mathrm{corr} - \rho_\mathrm{target})^2}\right] \leq \epsilon. \tag{2}$$

If an asymptotically large number of copies of the state is available, the parties can always sacrifice a finite number of such states to estimate the channel to any desired precision before commencing with the state transformation. Consequently, the incurred cost for estimation has no

bearing on the final yield of the target state. On the other hand, if only a finite number of copies is available, then the above strategy can be quite wasteful, and protocols that estimate the channel whilst implementing the OSBP at the same time may provide better yields. This is due to the observation that the states resulting from the failure branches of the OSBP potentially carry information about the noisy channel and can thus be used to estimate the latter at no extra cost.

A particular class of channels for which simultaneous channel estimation and state conversion may be beneficial is the family of channels that admit a decomposition into Kraus operators that commute with the POVM corresponding to the OSBP. Such channels include dephasing and phase drift [45] noise which are typical types of noise in quantum communication using optical fibres [46]. Alternatively, such channels also describe a situation often encountered in satellite communication where the sender (vendor) and receivers (N parties) do not share a common phase reference during the transmission stage [47]. In this latter case the receivers know that the state they receive is local-unitarily (LU) equivalent to the original state of the vendor and thus is in the same entanglement class. However, the lack of knowledge of the relative phase between the sender's and receivers' phase reference makes channel estimation indispensable, especially if the task they have in mind requires a particular state ρ_{target}, and not any other local-unitarily equivalent state.

In what follows, we shall consider the above class of channels as a special case of more general parameter estimation problems and employ a Bayesian estimation strategy to estimate the parameter (or set of parameters) θ that characterizes the channel, illustrated in Fig. 1. Such a strategy involves one or more rounds of subsequent measurements on different copies of the received states. The knowledge of the parameter, θ, is updated in every round depending on the measurement outcome, m, according to Bayes' law, i.e.,

$$p(\theta|m) = \frac{p(m|\theta)\, p(\theta)}{p(m)}. \tag{3}$$

Based on the posterior distribution of θ one can then assign an estimator $\hat{\theta}(m)$ after each round and adapt the next measurement to this estimate. This makes Bayesian parameter estimation more flexible and better suited for a limited number of probes compared to frequentist estimation. For an increasing number of copies, also a frequentist strategy can be employed to estimate the channel from the failure branches of the OSBP. Here, however, we will focus on scenarios with a very restricted number of copies. In particular, we consider the exemplary situation of local dephasing noise, to provide quantitative support for our strategy of metrology-assisted entanglement distribution.

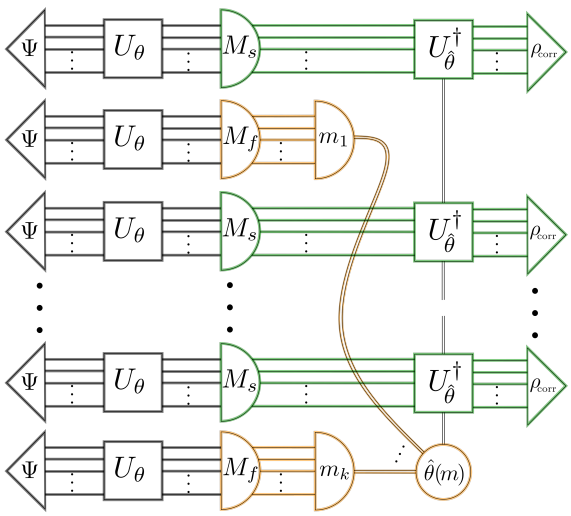

Figure 1: Schematic of the metrology-assisted entanglement distribution protocol. Each copy of the distributed multi-party state $|\Psi\rangle$ is subject to a (local) noise channel (here represented by an unknown unitary U_θ) parameterized by θ, which is subsequently measured by an OSBP with measurement operators M_{s} and M_{f}, corresponding to successful (green) and failed (orange) state conversion, respectively. The copies in the failure branch can be used to perform further measurements with outcomes m_i to estimate the noise channel, and apply corresponding corrections (represented here by $U_{\hat{\theta}}^\dagger$) to obtain partially corrected states ρ_{corr}.

3 Entanglement distribution with dephasing noise

Assume the state distributed by the vendor is a pure N-qudit state from the GHZ class of the form

$$|\Psi\rangle = \sum_{l=0}^{d-1} \sqrt{\lambda_l}\, |l\rangle^{\otimes N}, \tag{4}$$

where we assume without loss of generality that $\lambda_m \geq \lambda_{m+1}$. Let us further assume that the vendor distributes these states via N local dephasing channels, i.e., each qudit undergoes a unitary transformation parameterized by an unknown angle θ,

$$U(\theta) = \sum_{l=0}^{d-1} e^{\mathrm{i}l\theta}\, |l\rangle\langle l|. \tag{5}$$

The parties wish to obtain the maximally-entangled qudit state

$$|\mathrm{GHZ}_d\rangle = \sqrt{\tfrac{1}{d}} \sum_{l=0}^{d-1} |l\rangle^{\otimes N}. \tag{6}$$

To this end, they estimate the unknown phase θ of the dephasing noise acting on the received states and recover the target state via SLOCC-transformations from the received state.

More specifically, let us assume the vendor produces the

pure two-parameter qutrit state

$$|\Psi(\alpha,\beta)\rangle = \sin\alpha\cos\beta\,|0\rangle^{\otimes N} + \sin\alpha\sin\beta\,|1\rangle^{\otimes N} + \cos\alpha\,|2\rangle^{\otimes N}, \tag{7}$$

where $\tan^{-1}(\csc\beta) \leq \alpha \leq \pi/2$ and $0 \leq \beta \leq \pi/4$. This choice of angles ensures that $\sin\alpha\cos\beta \geq \sin\alpha\sin\beta \geq \cos\alpha \geq 0$. The parties wish to obtain the maximally entangled state of Eq. (6), for which the optimal OSBP comprises measurement operators

$$M_{\mathrm{s}} = \mathrm{diag}(\cot\alpha\sec\beta, \cot\alpha\csc\beta, 1), \tag{8}$$
$$M_{\mathrm{f}} = \mathrm{diag}(\sqrt{1-\cot^2\alpha\sec^2\beta}, \sqrt{1-\cot^2\alpha\csc^2\beta}, 0)$$

for one party and the identity for the remaining parties. The conversion is successful with probability $p_{\mathrm{s}} = 3\cos^2\alpha$, which is already the optimal conversion rate for any protocol acting on a single copy of the state. This follows from the fact that the optimal conversion rate from a bipartite pure state $|\Psi\rangle$ to another bipartite pure state $|\tilde{\Psi}\rangle$ can be calculated [48] from the respective Schmidt coefficients λ_i and $\tilde{\lambda}_j$ (real, non-negative,and in decreasing order) as $P(\Psi \to \tilde{\Psi}) = \min_j \sum_{l=j} \lambda_l^2 / \sum_{l'=j} \tilde{\lambda}_{l'}^2$, and from the observation that the optimal conversion rate for any splitting of the N parties into two groups gives an upper bound on the conversion rate. By splitting the state in Eq. (7) into one party versus the remaining $N-1$ parties, one immediately sees that the conversion rate for a single state is bounded by $P(\Psi(\alpha,\beta) \to \mathrm{GHZ}_3) = 3\cos^2\alpha$ for the parameter range we have chosen for α and β.

With probability $p_{\mathrm{f}} = 1 - 3\cos^2\alpha$ the conversion protocol fails and results in the state

$$|\Phi_{\mathrm{f}}\rangle = a\,|0\rangle^{\otimes N} + \sqrt{1-a^2}\,|1\rangle^{\otimes N}, \tag{9}$$

where

$$a = \frac{\sqrt{\sin^2\alpha\cos^2\beta - \cos^2\alpha}}{\sqrt{1-3\cos^2\alpha}}. \tag{10}$$

Observing that $[M_x, U(\theta)] = 0,\ \forall\, x \in \{\mathrm{s},\mathrm{f}\}$, it follows that the resulting states for the successful and failure branches of the OSBP are given by

$$|\mathrm{GHZ}_3(\theta)\rangle = \sqrt{\frac{1}{3}}\sum_{l=0}^{2} e^{\mathrm{i}lN\theta}\,|l\rangle^{\otimes N}, \tag{11a}$$
$$|\Phi_{\mathrm{f}}(\theta)\rangle = a\,|0\rangle^{\otimes N} + \sqrt{1-a^2}e^{\mathrm{i}N\theta}\,|1\rangle^{\otimes N}, \tag{11b}$$

respectively, with the corresponding probabilities invariant under the presence of dephasing noise.

In the case that the OSBP is successful, the parties obtain the state

$$\Lambda[\rho_{\mathrm{GHZ}_3}] = \int_0^{2\pi} p(\theta)U(\theta)^{\otimes N}\rho_{\mathrm{GHZ}_3}U(\theta)^{\dagger\otimes N}\mathrm{d}\theta,$$

where $\rho_{\mathrm{GHZ}_3} = |\mathrm{GHZ}_3\rangle\langle\mathrm{GHZ}_3|$ and $p(\theta)$ is a probability distribution over $\theta \in [0, 2\pi)$ describing the belief/knowledge the parties have of[1] θ. Loosely speaking, the more peaked the distribution becomes, the closer the state gets to the target state. The goal of the N-parties now is to apply correction operations to the successfully converted copies to obtain the highest yield of N-partite highly entangled qutrit states ρ_{corr}, that are ϵ-close in trace distance to the ideal state $|\mathrm{GHZ}_3\rangle$.

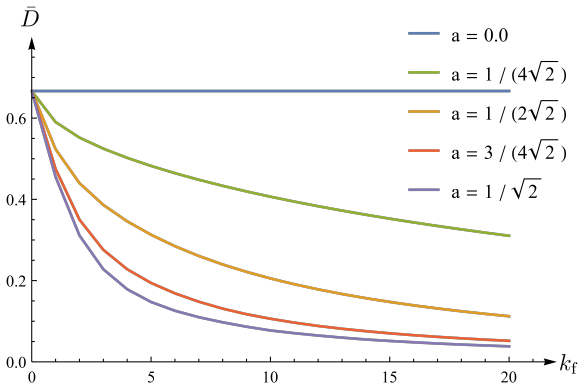

Figure 2: Average distance to the target state. The curves show the average trace distance (vertical axis) between the target state and the successfully converted copies for the exemplary metrology-assisted entanglement distribution protocol discussed in Sec. 3 for different values of a as a function of the number k_{f} of copies (horizontal axis) in the failure branch.

In the case of a flat prior, $p(\theta) = (2\pi)^{-1}$, we can analytically calculate (see Appendix A for details) the average trace distance to the target state after the first measurement to be

$$\bar{D}(\rho_{\mathrm{corr}}, \rho_{\mathrm{target}}) = \tfrac{1}{6}\left(1 + \sqrt{9 + 8a^2(1-a^2) - 16a\sqrt{1-a^2}}\right). \tag{12}$$

The average distance after one measurement hence only depends on the parameter a. Since $a \in [0,1]$, the maximal value $\frac{2}{3}$ is obtained for $a = 0$ or $a = 1$. This coincides with the distance based on the prior, i.e., ignoring measurement data. The minimal value of $\bar{D} = \frac{1}{6}(1+\sqrt{3})$ is obtained for $a = 1/\sqrt{2}$.

Given a particular measurement outcome the posterior distribution is updated using Bayes' law, [Eq. (3)]. In general the parties can optimize their measurement directions at every step but this amount of generality quickly renders the problem analytically intractable. We thus resort to numerical analysis henceforth, and we focus on a strategy where the measurement direction on the Bloch sphere of the two-dimensional subspace relevant for the copies in the failure branch (see Appendix A) is chosen on the equator but perpendicular to the estimator of θ of the previous round. This measurement is particularly sensitive in the direction of $\hat{\theta}$, since both outcomes are equally likely in

[1] Actually the estimation only gives us information about $N\theta$, which allows us to draw conclusions about θ only up to a shift by $2\pi/N$. But this is not a problem, since the final state also only depends on $N\theta$.

this case. The simplest strategy of fixed local measurements is treated in Appendix B, where also strategies that change the measurement direction according to some pre-established rule are considered.

In Fig. 2 we see how the trace distance $\bar{D}$ of the corrected state resulting from the successful branch to the target state reduces with the number of failure branches used for the estimation. After 20 measurement rounds the average trace distance drops to $\bar{D} \approx 0.038$ for $a = 1/\sqrt{2}$. This is a remarkable improvement considering that the failure states would be discarded in the naive protocol that we will discuss in Sec. 4.

One could also think of more sophisticated strategies, where a swap operation is applied to the state by the vendor before it is distributed, changing the basis vectors $|1\rangle \leftrightarrow |2\rangle$. Since the state $|2\rangle$ reacts more sensitive to the transformation, i.e., picks up a phase of $e^{2\mathrm{i}N\theta}$, this can be advantageous for estimation. Note, however, that in this case one can draw conclusions about θ only up to a π/N phase shift, whereas the successful branch still depends on $N\theta$. Thus, this strategy can only be employed once the interval of θ is sufficiently narrowed down.

Although we focus on the distribution of qutrit GHZ states under local dephasing noise in our example, we want to emphasize that this situation can be straightforwardly generalized to the distribution of any d-dimensional state in the GHZ entanglement class of the form of Eq. (4) with the same noise model. Here an OSBP converting the received state to the GHZ state in Eq. (6) leaves the failure-branch in a subspace of dimension $2 \leq d_\mathrm{f} < d$ and one can use estimation strategies similar to that used in Appendix C. Additionally, the GHZ class is broad in the sense that, according to Ref. [20], almost all pure multipartite entangled three-qubit states belong to this entanglement class, whereas states from the W class form a subset of measure zero. While this may not necessarily hold for higher dimensions or more parties, our example remains relevant as states in the GHZ class are often used (see, e.g., [49–52]) and have desirable properties such as their symmetry with respect to the exchange of parties.

4 Comparison with 'naive' protocols

In this section we compare the performance of our strategy to what we call 'naive' strategies. These are strategies that first sacrifice a number of copies of the received state for the estimation of the LU-channel and subsequently perform the transformation to the target state for all following copies. For this comparison we are interested in determining the number k of copies of received states required to obtain a certain number k_s of successfully converted states whose average trace distance from the desired target state is below a given value ϵ. Moreover, to make the comparison fair, first note that our previous metrology-assisted entanglement conversion protocol processes the received multipartite systems *sequentially*, i.e., one copy of each multipartite state at a time, and only uses local measurements and operations both for the estimation and the state transformation. Although this drastically simplifies potential practical implementations, it is known that processing multiple copies simultaneously — via the use of quantum memories — can lead to a higher success rate for entanglement distillation [53, 54]. For a fair comparison we therefore impose the same restrictions on the 'naive' strategy, i.e. we assume that also any 'naive' protocol is restricted to process each copy of the multipartite state sequentially via LOCC.

Within this framework for the 'naive' approach we consider a Bayesian strategy that estimates the phase parameter based on measurements performed on a certain number k_e of individual received states, while the $k - k_\mathrm{e}$ subsequent copies are individually transformed via the previously considered SLOCC protocol. We calculate the average trace distance of the successfully converted copies from the target state as a function of the initially sacrificed k_e copies used for estimation. Since the chosen Bayesian strategy might not be optimal in general, we supplement this analysis with an estimate of the lower bound on the achievable trace distance based on the Bayesian quantum Cramér-Rao bound. Comparing these results to the performance of our previously presented metrology-assisted protocol, we conclude that there exist combinations of the parameters k and ϵ for which our protocol outperforms any 'naive' protocol since the latter always sacrifices some of the copies that can be successfully converted within an approach that solely uses the failure-branch states for estimation.

After these 'fair' comparisons, we also relax the condition of sequential processing and consider protocols based on distillation where each party can use quantum memories and is able to process multiple copies at the same time. For this scenario, we derive upper bounds on the obtainable number of successfully converted copies valid for any distillation strategy and conclude that such strategies potentially outperform our sequential strategy. As these bounds are only asymptotically tight, however, it remains unclear how well the optimal strategy actually performs on a finite number of copies.

Let us now consider the task of the previous example, i.e., k copies of the state in Eq. (7) are distributed via a local dephasing channel, described in Eq. (5), and the goal of the parties is to obtain the highest number of states (at least) ϵ-close to $|\mathrm{GHZ}_3\rangle$. They therefore use k_e copies of $|\Psi(\alpha,\beta)\rangle$ to estimate the phase θ and assign an estimator $\hat{\theta}(\mathbf{m})$ based on the measurement outcome $\mathbf{m} = (m_1, \ldots, m_{k_\mathrm{e}})$. Using a Bayesian estimation strategy as before, we have to choose an appropriate measurement. As our probe is a three-level system now, we perform a three-outcome projective measurement. Since we assume not to have any knowledge of θ initially, we consider a flat prior, for which a reasonable choice of measurement basis presents itself in the form of any basis that is mutually unbiased with respect to the eigenbasis of the generating Hamiltonian of the local dephasing noise $U(\theta)$. Since no further restrictions seem to apply to the measurement basis we therefore use the discrete Fourier transform

of the eigenbasis of the generator [55]. However, note that for the measurement on the second and subsequent copies, we rotate the measurement basis by $U(\hat{\theta}+\pi/3)$ to take into account our estimate $\hat{\theta}$. The obtained basis still is mutually unbiased with respect to the eigenbasis of the encoding Hamiltonian. A more detailed calculation of this estimation can be found in Appendix C.

Let us compare this strategy to the numerical results of the previously employed estimation of the failure branch for the exemplary parameters $k = 100$, $\alpha = \arccos(2/\sqrt{15})$ and $\beta = \pi/4$. This leads to $p_\mathrm{s} = 4/5$, and so to $k_\mathrm{s} = 80$ and $k_\mathrm{f} = 20$ successful and failure branches, respectively, on average. Twenty rounds of estimation with the failure branch lead to an average trace distance of $\bar{D} \approx 0.038$, see Fig. 2. To achieve a comparable distance with the received state requires $k_\mathrm{e} = 9$ copies on average, see Fig. 4 in Appendix C. The observation that fewer copies of the received state are needed to estimate the phase to the same precision as compared to the failure-branch state comes as no surprise, since any processing of the state after the transmission can only reduce (on average) the amount of information about θ encoded in the system. This means that the received states will generally be more advantageous for the estimation than states in the failure branch. While this is true, it is essentially the same strategy that was also used for the failure branch in the previous protocol adapted to the different probe state. Hence we argue that this is indeed a fair comparison, as we do not know the optimal strategy for finitely many copies.

This notwithstanding, we can calculate a lower bound on the number of copies needed for any estimation strategy to reach a given trace distance based on the Bayesian quantum Cramér-Rao bound, see Appendix D for a detailed calculation. Comparing this to the numerical results of the previously employed strategy for the same exemplary parameters $k = 100$, $\alpha = \arccos(2/\sqrt{15})$ and $\beta = \pi/4$, one finds that achieving a comparable trace distance, any estimation needs on average at least $k_\mathrm{e} \geq 7.5$ copies of the received state, see Fig. 4 in Appendix C. However, note that $k_\mathrm{e} \geq 7.5$ is only a lower bound, and it is not clear if a strategy exists that saturates this bound for finitely many copies.

Once the phase θ has been estimated sufficiently well, the states resulting from the successful branches can be transformed via SLOCC. We have already argued in the previous section that any sequential strategy transforming individual copies of the state in Eq. (7) to the state $|\,\mathrm{GHZ}_3\,\rangle$ is bounded by $p_\mathrm{s} \leq 3\cos^2\alpha$. For the same exemplary parameters considered before, this leads to a success rate of $p_\mathrm{s} = 4/5$. For the 'naive' Bayesian strategy we hence obtain $(k-k_\mathrm{e})p_\mathrm{s} = 91\,\frac{4}{5} = 72.8$ successfully converted copies on average from the initial 100 copies of the distributed state. Additionally, the lower bound based on the Bayesian quantum Cramér-Rao bound informs us that any 'naive' sequential strategy obtains at most $(k-k_\mathrm{e})p_\mathrm{s} = 92.5\,\frac{4}{5} = 74$ copies of the target state on average, proving our failure-branch estimation considerably superior with 80 successfully converted copies on average in this example.

Indeed, for any noise particular SLOCC conversion protocol with success probability p_s (in which a channel is estimated whose Kraus operators commute with the SLOCC protocol) one may find pairs of k and ϵ for which our strategy outperforms naive protocols since the latter always sacrifice some of the copies that can otherwise be successfully converted. If the average trace distance $\bar{D}$ after k copies is equal to or smaller than the desired distance ϵ, then our protocol cannot be beaten by a naive strategy for this combination of parameters since any naive strategy would have to sacrifice some number $k_\mathrm{e} > 0$ of the received states, thus ultimately being left with only $p_\mathrm{s}(k-k_\mathrm{e}) < k_\mathrm{s} = p_\mathrm{s}k$ copies within the desired distance ϵ. However, note that this does not imply that our strategy outperforms naive protocols for all values of k and ϵ. There also exist pairs of k and ϵ where our failure branch estimation is not sufficiently precise and good copies need to be sacrificed for estimation to reach the desired precision. In such a case the performance of such a 'hybrid' strategy compared to a 'naive' strategy depends on the specific parameters of the considered example.

In addition, we note that naive protocols based on distillation, i.e., which act jointly on all $k-k_\mathrm{e}$ states available after the initial estimation could in principle perform better than sequential protocols based on failure-branch estimation only, even in cases where the latter outperform sequential naive protocols. For pure bipartite states the asymptotic distillation rate from $|\,\psi\,\rangle$ to $|\,\phi\,\rangle$ is bounded by the entropy ratio of the reduced states for any party i, i.e.,

$$\mathcal{D}(|\,\psi\,\rangle \to |\,\phi\,\rangle) = \frac{S(|\psi\rangle\langle\psi|_i)}{S(|\phi\rangle\langle\phi|_i)}, \tag{13}$$

and this bound is tight [56]. For multipartite states this immediately gives the, generally not tight, asymptotic bound

$$\mathcal{D}(|\,\psi\,\rangle \to |\,\phi\,\rangle) = \min_i\Big(\frac{S(|\psi\rangle\langle\psi|_i)}{S(|\phi\rangle\langle\phi|_i)}\Big). \tag{14}$$

In our case, however, the latter bound is tight asymptotically [57, 58] and we can calculate

$$\begin{aligned}\mathcal{D}(|\,\Psi(\alpha,\beta)\,\rangle \to |\,\mathrm{GHZ}\,\rangle) = &-\cos^2\alpha\log_3(\cos^2\alpha) \\ &-\sin^2\alpha\cos^2\beta\log_3(\sin^2\alpha\cos^2\beta) \\ &-\sin^2\alpha\sin^2\beta\log_3(\sin^2\alpha\sin^2\beta).\end{aligned} \tag{15}$$

For our choice of parameters this means there might exist a distillation protocol acting on all available copies at once able to successfully distill (on average) 91.6 copies of the desired GHZ-state out of the $k-k_\mathrm{e} = 92.5$ remaining copies after an optimal estimation. This clearly exceeds the 80 successful states obtained by our adaptive strategy. Here we again want to emphasise that these 91.6 copies are only an upper bound. Both an optimal estimation and distillation strategy are only known to exist asymptotically and it is not clear that any strategy can attain this bound for finitely many copies. Indeed, even if such a strategy exists, it would generally require acting on all remaining

copies simultaneously. Even if the presented metrology-assisted protocol does not saturate these 'optimal' bounds, it nonetheless shows the potential of practically easily implementable adaptive strategies acting on single copies. In particular, the task of designing and implementing optimal strategies for estimation and especially distillation becomes extremely challenging for larger numbers of copies and may be out of reach for experimental realizations in the medium term. Therefore it is essential to find feasible strategies that perform well and whose implementation remains tractable with growing number of copies.

5 Metrology-assisted protocols beyond dephasing noise

A key assumption in our previous example was that the parties know that the measurement operators of the OSBP commute with the Kraus operators of the noisy channel. If this does not hold, as is the case for depolarizing, channels with high loss, or for dephasing channels with unknown rotation axis, the order in which the POVM measurements of the OSBP and those for the estimation protocol are carried out matter. If the channel is not known to commute with the POVM operators of the OSBP, this impacts the performance of the presented strategy. Nonetheless one could design strategies relying on the failure branch for the channel estimation. We now present a protocol that integrates the estimation of general unknown unitary transformations $\mathrm{SU}(d)$ into an OSBP. Possible strategies to estimate a general unitary transformation are given by [59, 60]. We focus on how an estimation strategy can incorporate the failure branch of an OSBP for a potentially more efficient estimation, without restricting ourselves to a specific strategy.

The naive way to perform this task would again be to first use a number of copies to estimate the unknown unitary sufficiently well, such that the received state can be corrected up to within trace distance $\bar{D} \leq \epsilon$ of the target. After that, some distillation procedure is applied to recover the desired state.

In our approach we combine both tasks into one. The idea is to construct the POVM operators for the OSBP based on the knowledge of the unknown unitary at that specific stage of the protocol. We start with the POVM $\{\mathbb{1}, 0\}$, that will always result in the success branch. Then we check if the state is sufficiently close to the target state, based on our knowledge of the unitary. This will not be the case if the distribution $p(U)$ is sufficiently broad. If the state is not close enough to the target, we make a measurement and update our believe and the estimator. With increasing knowledge of the unitary transformation, also the POVM of the OSBP will change and converge towards the optimal POVM for the received state. During this process the copies in the successful branch close to the target (according to the knowledge at the time) are stored, while all the others are used for estimation. With an increasingly peaked distribution $p(U)$ also the ratio of stored copies to

Algorithm 1: Metrology-integrated OSBP

input : k copies of the state $|\psi\rangle$
output: $p_{(k)}(U)$, $\hat{U}_{(k)}$ and all saved copies $|\tilde{\psi}_s^{(i)}\rangle$
begin
  start with uniform distribution $p_{(0)}(U)$ over $SU(d)$
  set estimator $\hat{U}_{(0)} = \mathbb{1}$
  $i = 1$
  while $i \leq k$, *i.e., there are copies left* **do**
    send copy number i of the state $|\psi\rangle$
    obtain $U^{\otimes N}|\psi\rangle$
    correct to $|\tilde{\psi}^{(i)}\rangle = (\hat{U}_{(i-1)}^{\dagger} U)^{\otimes N}|\psi\rangle$
    construct the POVM for the OSBP:
      - define M_{s}^{V} as the operator s.t. $M_{\mathrm{s}}^{V} V^{\otimes N} |\psi\rangle = |\mathrm{GHZ}_d\rangle$
      - $M_{\mathrm{s}}^{(i)} = \int_{SU(d)} p_{(i-1)}(V) M_{\mathrm{s}}^{\hat{U}_{(i-1)}^{\dagger} V} \mathrm{d}V$
      - $M_{\mathrm{f}}^{(i)}$ s.t. $(M_{\mathrm{s}}^{(i)})^{\dagger} M_{\mathrm{s}}^{(i)} + (M_{\mathrm{f}}^{(i)})^{\dagger} M_{\mathrm{f}}^{(i)} = \mathbb{1}$
    perform OSBP with $\{M_{\mathrm{s}}^{(i)}, M_{\mathrm{f}}^{(i)}\}$
    if *success* **then**
      get $|\tilde{\psi}_s^{(i)}\rangle = M_{\mathrm{s}}^{(i)}|\tilde{\psi}^{(i)}\rangle$
      if $\int_{SU(d)} p_{(i-1)}(U) |\langle \mathrm{GHZ}_d | M_{\mathrm{s}}^{(i)} (\hat{U}_{(i-1)}^{\dagger} U)^{\otimes N} |\psi\rangle|^2 \mathrm{d}U \geq 1 - \epsilon$ **then**
        keep $|\tilde{\psi}_s^{(i)}\rangle$ and save corresponding $M_{\mathrm{s}}^{(i)}$ and $\hat{U}_{(i-1)}$
        leave probability distribution and estimator of the unitary unchanged
        $p_{(i)}(U) = p_{(i-1)}(U)$
        $\hat{U}_{(i)} = \hat{U}_{(i-1)}$
      else
        perform a measurement on $|\tilde{\psi}_s^{(i)}\rangle$ to estimate U
        update distribution $p_{(i)}(U)$ and estimator $\hat{U}_{(i)}$
      end
    else
      perform a measurement on $|\tilde{\psi}_f^{(i)}\rangle = M_{\mathrm{f}}^{(i)}|\tilde{\psi}^{(i)}\rangle$ to estimate U
      update distribution $p_{(i)}(U)$ and estimator $\hat{U}_{(i)}$
    end
    $i++$
  end
  return $p_{(k)}(U)$, $\hat{U}_{(k)}$ and all saved copies $|\tilde{\psi}_s^{(i)}\rangle$
end

discarded copies (used for estimation) will increase. Given that enough copies were purchased, we can assume that the unitary transformation is known sufficiently well at the end of this protocol. For all stored copies it is checked if they are indeed close enough to the target according to the updated knowledge of U. If not, they are transformed into the target state via SLOCC. The exact estimation and updating procedure are left unspecified here on purpose because these will strongly depend on the exact state and noise model. In Algorithm 1 we consider the exemplary situation where the desired state is an N-partite GHZ-state of local dimension d and the distributed state is of the form of Eq. (4) with corresponding dimension and party number. But the idea works for general scenarios and Algorithm 1 can be adapted for broader classes of vendor and target states.

Once the first round is completed, all states that were saved, i.e., the states with a successful OSBP that were sufficiently close to the target state (with the information at the time), are checked to be sufficiently close with the final distribution. If not, one may try to transform them into the target via unitary transformations and an OSBP. The eventual failure branch copies can be used for estimation.

For the first rounds this protocol will do almost the same as the naive protocol. Since the distribution $p(U)$ is broad, the POVM will be close to $\{\mathbb{1}, 0\}$, resulting most likely in the success branch (not altering the state). But as the state will be far from the target, it will be used for estimation. After many rounds the distribution $p(U)$ will be very peaked, all copies in the successful branch are now close to the target state and only the failure branches are used for estimation, resulting in a protocol similar to the one presented above.

6 Conclusion

Potential future implementations of large-scale quantum networks, such as the envisioned 'quantum internet' [13–15], can be expected to merge idealized protocols for quantum communication and information processing with real-world conditions arising from noise and technological constraints which will necessitate the development of pragmatic and flexible solutions. Here, we have considered a quantum communication protocol that will be relevant in this context: the distribution and conversion of high-dimensional entangled states among multiple parties using noisy channels. Procedures for quantum state-conversion employed under these circumstances will involve probabilistic transformations via SLOCC. As we have shown, such scenarios are naturally amenable to enhancements via embedded channel estimation routines.

While the ideas presented here apply quite generally, we have illustrated the improvements one may expect in such metrology-assisted entanglement distribution by focusing on a particular example: probabilistic conversion of finitely many copies of N-qutrit states distributed via a local dephasing channels to GHZ target states. For this example, we have shown that already practically easily implementable Bayesian estimation strategies based on local measurements of individual copies can lead to excellent performance comparable to optimal distillation procedures. The latter, however, might involve significantly more complex and technologically challenging joint operations on multiple copies.

We believe that metrology-assisted entanglement distribution has the potential to become a relevant factor in future quantum networks. We thus expect the study and optimization of such protocols for different target states and noisy models, as well as their comparison to relevant alternative protocols (e.g., based on distillation) to open up a broad range of interesting questions beyond this initial pilot study.

Acknowledgments

We thank Cornelia Spee for feedback on the manuscript. S.M. acknowledges support from the Austrian Science Fund (FWF) through projects Y879-N27 (START) and P 31339-N27 (Stand-Alone). N.F. acknowledges support from the Austrian Science Fund (FWF) through the projects Y879-N27 (START) and P 31339-N27 (Stand-Alone), and the joint Czech-Austrian project MultiQUEST (I 3053-N27 and GF17-33780L). M.S. ackowledges support from Spanish MINECO reference FIS2016-80681-P (with the support of AEI/FEDER,EU), the Catalan Government for the project QuantumCAT 001-P-001644 (RIS3CAT comunitats) co-financed by the European Regional Development Fund (FEDER), and the Generalitat de Catalunya, project CIRIT 2017-SGR-1127.

References

[1] Andrew J. Scott, *Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions,* Phys. Rev. A **69**, 052330 (2004), arXiv:quant-ph/0310137.

[2] Tohya Hiroshima and Masahito Hayashi, *Entanglement and Quantum Error Correction,* in *Quantum Computation and Information*, edited by Hiroshi Imai and Masahito Hayashi (Springer, Berlin, Heidelberg, 2006) Chap. 5, pp. 111–132.

[3] Wolfgang Dür and Hans J. Briegel, *Entanglement purification and quantum error correction,* Rep. Prog. Phys. **70**, 1381 (2007), arXiv:0705.4165.

[4] Todd A. Brun and Min-Hsiu Hsieh, *Entanglement-assisted quantum error-correcting codes,* in *Quantum Error Correction*, edited by Daniel A. Lidar and Todd A. Brun (Cambridge University Press, Cambridge, U.K., 2013) Chap. 7, pp. 181–200, arXiv:1610.04013.

[5] Oliver Marty, Marcus Cramer, and Martin B. Plenio, *Practical Entanglement Estimation for Spin-System Quantum Simulators,* Phys. Rev. Lett. **116**, 105301 (2016), arXiv:1504.03572.

[6] Nicolai Friis, Oliver Marty, Christine Maier, Cornelius Hempel, Milan Holzäpfel, Petar Jurcevic, Martin B. Plenio, Marcus Huber, Christian Roos, Rainer Blatt, and Ben Lanyon, *Observation of Entangled States of a Fully Controlled 20-Qubit System,* Phys. Rev. X **8**, 021012 (2018), arXiv:1711.11092.

[7] Marcello Dalmonte, Benoît Vermersch, and Peter Zoller, *Quantum simulation and spectroscopy of entanglement Hamiltonians,* Nat. Phys. **14**, 827 (2018), arXiv:1707.04455.

[8] Michael Epping, Hermann Kampermann, Chiara Macchiavello, and Dagmar Bruß, *Multi-partite entanglement can speed up quantum key distribution in networks,* New J. Phys. **19**, 093012 (2017), arXiv:1612.05585.

[9] Stefan Bäuml and Koji Azuma, *Fundamental limitation on quantum broadcast networks,* Quantum Sci. Technol. **2**, 024004 (2017), arXiv:1609.03994.

[10] Matej Pivoluska, Marcus Huber, and Mehul Malik, *Layered quantum key distribution,* Phys. Rev. A **97**, 032312 (2018), arXiv:1709.00377.

[11] Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner, *Fully device-independent conference key agreement,* Phys. Rev. A **97**, 022307 (2018), arXiv:1708.00798.

[12] Matteo Pompili, Sophie L. N. Hermans, Simon Baier, Hans K. C. Beukers, Peter C. Humphreys, Raymond N. Schouten, Raymond F. L. Vermeulen, Marijn J. Tiggelman, Laura dos Santos Martins, Bas Dirkse, Stephanie Wehner, and Ronald Hanson, *Realization of a multinode quantum network of remote solid-state qubits,* Science **372**, 259 (2021), arXiv:2102.04471.

[13] H. Jeff Kimble, *The quantum internet,* Nature **453**, 1023 (2008), arXiv:0806.4195.

[14] Stephanie Wehner, David Elkouss, and Ronald Hanson, *Quantum internet: A vision for the road ahead,* Science **362**, eaam9288 (2018).

[15] Angela Sara Cacciapuoti, Marcello Caleffi, Francesco Tafuri, Francesco Saverio Cataliotti, Stefano Gherardini, and Giuseppe Bianchi, *Quantum Internet: Networking Challenges in Distributed Quantum Computing,* IEEE Network **34**, 137 (2020), arXiv:1810.08421.

[16] Wolfgang Dür, Raphael Lamprecht, and Stefan Heusler, *Towards a quantum internet,* Eur. J. Phys. **38**, 043001 (2017).

[17] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, *Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels,* Phys. Rev. Lett. **70**, 1895–1899 (1993).

[18] Dik Bouwmeester, Jian-Wei Pan, Klaus Mattle, Manfred Eibl, Harald Weinfurter, and Anton Zeilinger, *Experimental Quantum Teleportation,* Nature **390**, 575–579 (1997), arXiv:1901.11004.

[19] Daniel Llewellyn, Yunhong Ding, Imad I. Faruque, Stefano Paesani, Davide Bacco, Raffaele Santagati, Yan-Jun Qian, Yan Li, Yun-Feng Xiao, Marcus Huber, Mehul Malik, Gary F. Sinclair, Xiaoqi Zhou, Karsten Rottwitt, Jeremy L. O'Brien, John G. Rarity, Qihuang Gong, Leif K. Oxenlowe, Jianwei Wang, and Mark G. Thompson, *Chip-to-chip quantum teleportation and multi-photon entanglement in silicon,* Nat. Phys. **16**, 148 (2020), arXiv:1911.07839.

[20] Wolfgang Dür, Guifré Vidal, and Juan Ignacio Cirac, *Three qubits can be entangled in two inequivalent ways,* Phys. Rev. A **62**, 062314 (2000), arXiv:quant-ph/0005115.

[21] Antonio Acín, Dagmar Bruß, Maciej Lewenstein, and Anna Sanpera, *Classification of Mixed Three-Qubit States,* Phys. Rev. Lett. **87**, 040401 (2001), arXiv:quant-ph/0103025.

[22] Julio I. de Vicente, Cornelia Spee, and Barbara Kraus, *Maximally Entangled Set of Multipartite Quantum States,* Phys. Rev. Lett. **111**, 110502 (2013), arXiv:1305.7398.

[23] Katharina Schwaiger, David Sauerwein, Martí Cuquet, Julio I. de Vicente, and Barbara Kraus, *Operational Multipartite Entanglement Measures,* Phys. Rev. Lett. **115**, 150502 (2015), arXiv:1503.00615.

[24] Julio I. de Vicente, Cornelia Spee, David Sauerwein, and Barbara Kraus, *Entanglement manipulation of multipartite pure states with finite rounds of classical communication,* Phys. Rev. A **95**, 012323 (2017), arXiv:1607.05145.

[25] Cornelia Spee, Julio I. de Vicente, David Sauerwein, and Barbara Kraus, *Entangled Pure State Transformations via Local Operations Assisted by Finitely Many Rounds of Classical Communication,* Phys. Rev. Lett. **118**, 040503 (2017), arXiv:1606.04418.

[26] David Sauerwein, Nolan R. Wallach, Gilad Gour, and Barbara Kraus, *Transformations among Pure Multipartite Entangled States via Local Operations are Almost Never Possible,* Phys. Rev. X **8**, 031020 (2018), arXiv:1711.11056.

[27] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter, *Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask),* Commun. Math. Phys. **328**, 303 (2014), arXiv:1210.4583.

[28] Jessica Bavaresco, Natalia Herrera Valencia, Claude Klöckl, Matej Pivoluska, Paul Erker, Nicolai Friis, Mehul Malik, and Marcus Huber, *Measurements in two bases are sufficient for certifying high-dimensional entanglement,* Nat. Phys. **14**, 1032 (2018), arXiv:1709.07344.

[29] He Lu, Qi Zhao, Zheng-Da Li, Xu-Fei Yin, Xiao Yuan, Jui-Chen Hung, Luo-Kan Chen, Li Li, Nai-Le Liu, Cheng-Zhi Peng, Yeong-Cherng Liang, Xiongfeng Ma, Yu-Ao Chen, and Jian-Wei Pan, *Entanglement Structure: Entanglement Partitioning in Multipartite Systems and Its Experimental Detec-*

tion Using Optimizable Witnesses, Phys. Rev. X **8**, 021072 (2018), arXiv:1711.01784.

[30] Nicolai Friis, Giuseppe Vitagliano, Mehul Malik, and Marcus Huber, *Entanglement Certification From Theory to Experiment,* Nat. Rev. Phys. **1**, 72 (2019), arXiv:1906.10929.

[31] You Zhou, Qi Zhao, Xiao Yuan, and Xiongfeng Ma, *Detecting multipartite entanglement structure with minimal resources,* npj Quantum Inf. **5**, 83 (2019), arXiv:1904.05001.

[32] Natalia Herrera Valencia, Vatshal Srivastav, Matej Pivoluska, Marcus Huber, Nicolai Friis, Will McCutcheon, and Mehul Malik, *High-Dimensional Pixel Entanglement: Efficient Generation and Certification,* Quantum **4**, 376 (2020), arXiv:2004.04994.

[33] Gary J. Mooney, Gregory A. L. White, Charles D. Hill, and Lloyd C. L. Hollenberg, *Whole-Device Entanglement in a 65-Qubit Superconducting Quantum Computer,* Adv. Quantum Technol. **4**, 2100061 (2021), arXiv:2102.11521.

[34] Raju Valivarthi, Samantha Davis, Cristian Pena, Si Xie, Nikolai Lauk, Lautaro Narvaez, Jason P. Allmaras, Andrew D. Beyer, Yewon Gim, Meraj Hussein, George Iskander, Hyunseong Linus Kim, Boris Korzh, Andrew Mueller, Mandy Rominsky, Matthew Shaw, Dawn Tang, Emma E. Wollman, Christoph Simon, Panagiotis Spentzouris, Neil Sinclair, Daniel Oblak, and Maria Spiropulu, *Teleportation Systems Towards a Quantum Internet,* PRX Quantum **1**, 020317 (2020), arXiv:2007.11157.

[35] Miguel Navascues, Elie Wolfe, Denis Rosset, and Alejandro Pozas-Kerstjens, *Genuine Network Multipartite Entanglement,* Phys. Rev. Lett. **125**, 240505 (2020), arXiv:2002.02773.

[36] Tristan Kraft, Sébastien Designolle, Christina Ritz, Nicolas Brunner, Otfried Gühne, and Marcus Huber, *Quantum entanglement in the triangle network,* Phys. Rev. A **103**, L060401 (2021), arXiv:2002.03970.

[37] Tristan Kraft, Cornelia Spee, Xiao-Dong Yu, and Otfried Gühne, *Characterizing quantum networks: Insights from coherence theory,* Phys. Rev. A **103**, 052405 (2021), arXiv:2006.06693.

[38] Johan Åberg, Ranieri Nery, Cristhiano Duarte, and Rafael Chaves, *Semidefinite tests for quantum network topologies,* Phys. Rev. Lett. **125**, 110505 (2020), arXiv:2002.05801.

[39] Cornelia Spee and Tristan Kraft, *Transformations in quantum networks via local operations assisted by finitely many rounds of classical communication,* arXiv.2105.01090 [quant-ph] (2021).

[40] Hayata Yamasaki, Simon Morelli, Markus Miethlinger, Jessica Bavaresco, Nicolai Friis, and Marcus Huber, *Activation of genuine multipartite entanglement: beyond the single-copy paradigm of entanglement characterisation,* Quantum **6**, 695 (2022), arXiv:2106.01372.

[41] Antoine Neven, David Gunn, Martin Hebenstreit, and Barbara Kraus, *Local Transformations of Multiple Multipartite States,* SciPost Phys. **11**, 42 (2021), arXiv:2007.06256.

[42] Gilad Gour, Barbara Kraus, and Nolan R. Wallach, *Almost all multipartite qubit quantum states have trivial stabilizer,* J. Math. Phys. **58**, 092204 (2017), arXiv:1609.01327.

[43] Antonio Acín, Enric Jané, Wolfgang Dür, and Guifre Vidal, *Optimal Distillation of a Greenberger-Horne-Zeilinger State,* Phys. Rev. Lett. **85**, 4811 (2000), arXiv:quant-ph/0007042.

[44] David Sauerwein, Katharina Schwaiger, and Barbara Kraus, *Discrete and differentiable entanglement transformations,* arXiv:1808.02819 [quant-ph] (2018).

[45] Marcin Jarzyna, Konrad Banaszek, and Rafał Demkowicz-Dobrzański, *Dephasing in coherent communication with weak signal states,* J. Phys. A: Math. Theor. **47**, 275302 (2014), arXiv:1307.6871.

[46] Keith H. Wanser, *Fundamental phase noise limit in optical fibres due to temperature fluctuations,* Electron. Lett. **28**, 53 (1992).

[47] Marco Fanizza, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Vittorio Giovannetti, *Squeezing-enhanced communication without a phase reference,* Quantum **5**, 608 (2021), arXiv:2006.06522.

[48] Guifré Vidal, *Entanglement of Pure States for a Single Copy,* Phys. Rev. Lett. **83**, 1046 (1999), arXiv:quant-ph/9902033.

[49] Xi-Lin Wang, Yi-Han Luo, He-Liang Huang, Ming-Cheng Chen, Zu-En Su, Chang Liu, Chao Chen, Wei Li, Yu-Qiang Fang, Xiao Jiang, Jun Zhang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan, *18-Qubit Entanglement with Six Photons' Three Degrees of Freedom,* Phys. Rev. Lett. **120**, 260502 (2018), arXiv:1801.04043.

[50] Ming Gong, Ming-Cheng Chen, Yarui Zheng, Shiyu Wang, Chen Zha, Hui Deng, Zhiguang Yan, Hao Rong, Yulin Wu, Shaowei Li, Fusheng Chen, Youwei Zhao, Futian Liang, Jin Lin, Yu Xu, Cheng Guo, Lihua Sun, Anthony D. Castellano, Haohua Wang, Chengzhi Peng, Chao-Yang Lu, Xiaobo Zhu, and Jian-Wei Pan, *Genuine 12-Qubit Entanglement on a Superconducting Quantum Processor,* Phys. Rev. Lett. **122**, 110501 (2019), arXiv:1811.02292.

[51] Gary J. Mooney, Gregory A. L. White, Charles D. Hill, and Lloyd C. L. Hollenberg, *Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer,* J. Phys. Commun. **5**, 095004 (2021), arXiv:2101.08946.

[52] Ivan Pogorelov, Thomas Feldker, Christian D. Marciniak, Georg Jacob, Verena Podlesnic, Michael Meth, Vlad Negnevitsky, Martin Stadler, Kirill Lakhmanskiy, Rainer Blatt, Philipp Schindler, and Thomas Monz, *Compact Ion-Trap Quantum Com-*

puting Demonstrator, PRX Quantum **2**, 020343 (2021), arXiv:2101.11390.

[53] Nilanjana Datta and Felix Leditzky, *Second-Order Asymptotics for Source Coding, Dense Coding, and Pure-State Entanglement Conversions,* IEEE Trans. Inf. Theory **61**, 582 (2015), arXiv:1403.2543.

[54] Kun Fang, Xin Wang, Marco Tomamichel, and Runyao Duan, *Non-Asymptotic Entanglement Distillation,* IEEE Trans. Inf. Theory **65**, 6454 (2019), arXiv:1706.06221.

[55] Stephen Brierley, Stefan Weigert, and Ingemar Bengtsson, *All Mutually Unbiased Bases in Dimensions Two to Five,* Quantum Inf. Comput. **10**, 0803 (2009), arXiv:0907.4097.

[56] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher, *Concentrating partial entanglement by local operations,* Phys. Rev. A **53**, 2046 (1996), arXiv:quant-ph/9511030.

[57] John A. Smolin, Frank Verstraete, and Andreas Winter, *Entanglement of assistance and multipartite state distillation,* Phys. Rev. A **72**, 052317 (2005), arXiv:quant-ph/0505038.

[58] Alexander Streltsov, Clément Meignant, and Jens Eisert, *Rates of Multipartite Entanglement Transformations,* Phys. Rev. Lett. **125**, 080502 (2020), arXiv:1709.09693.

[59] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Massimiliano Federico Sacchi, *Efficient Use of Quantum Resources for the Transmission of a Reference Frame,* Phys. Rev. Lett. **93**, 180503 (2004), arXiv:quant-ph/0405095.

[60] Jonas Kahn, *Fast rate estimation of a unitary operation in SU(d),* Phys. Rev. A **75**, 022326 (2007), arXiv:quant-ph/0603115.

[61] Richard D. Gill and Boris Y. Levit, *Applications of the van Trees Inequality: A Bayesian Cramér-Rao Bound,* Bernoulli **1**, 59 (1995), https://projecteuclid.org/euclid.bj/1186078362.

[62] Samuel L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states,* Phys. Rev. Lett. **72**, 3439 (1994).

[63] Rafał Demkowicz-Dobrzański, Marcin Jarzyna, and Janek Kołodyński, *Quantum limits in optical interferometry,* Prog. Optics **60**, 345 (2015), arXiv:1405.7703.

[64] Nicolai Friis, Davide Orsucci, Michalis Skotiniotis, Pavel Sekatski, Vedran Dunjko, Hans J. Briegel, and Wolfgang Dür, *Flexible resources for quantum metrology,* New J. Phys. **19**, 063044 (2017), arXiv:1610.09999.

[65] Simon Morelli, Ayaka Usui, Eliza Agudelo, and Nicolai Friis, *Bayesian parameter estimation using Gaussian states and measurements,* Quantum Sci. Technol. **6**, 025018 (2021), arXiv:2009.03709.

A Details on single measurement

We consider an initial measurement where each of the N parties measures in the basis $\{|\pm\rangle_i = \frac{1}{\sqrt{2}}(|0\rangle_i \pm |1\rangle_i)\}$ for $i = 1, 2, \ldots, N$ of the two-dimensional subspace relevant for the copies in the failure branch. Since the setup is symmetric with respect to the exchange of any of the N parties, only the parity of the overall measurement matters, i.e., if one obtains an even or odd number of "$-$" outcomes. To see this note that the probability to get an outcome "n" corresponding to a possible operator E_n is

$$\begin{aligned}\mathrm{Tr}(|\Phi_\mathrm{f}(\theta)\rangle\langle\Phi_\mathrm{f}(\theta)|\,E_n) &= \langle\Phi_\mathrm{f}(\theta)|\,E_n\,|\Phi_\mathrm{f}(\theta)\rangle \\ &= a^2\,\langle 0|^{\otimes N} E_n\,|0\rangle^{\otimes N} + (1-a^2)\,\langle 1|^{\otimes N} E_n\,|1\rangle^{\otimes N} \\ &\quad + a\sqrt{1-a^2}\,\mathrm{e}^{-i\theta N}\,\langle 1|^{\otimes N} E_n\,|0\rangle^{\otimes N} \\ &\quad + a\sqrt{1-a^2}\,\mathrm{e}^{i\theta N}\,\langle 0|^{\otimes N} E_n\,|1\rangle^{\otimes N}\,,\end{aligned} \tag{16}$$

while E_n is one of 2^N products of projectors $|\pm\rangle\langle\pm|$, for instance $E_{+-+--} = |+\rangle\langle+| \otimes |-\rangle\langle-| \otimes |+\rangle\langle+| \otimes |-\rangle\langle-| \otimes |-\rangle\langle-|$. But only the parity of the number of "$-$" projector matters. Out of all 2^N combinations of $|+\rangle\langle+|$ and $|-\rangle\langle-|$ projectors for N qubits, half have an even number of "$-$" projectors, so we have

$$p(\text{even/odd}|\theta) = \tfrac{1}{2} \pm a\sqrt{1-a^2}\cos(N\theta) \tag{17}$$

Given a prior $p(\theta)$, we then assign a probability

$$\begin{aligned}p(\text{even/odd}) &= \int_0^{2\pi} p(\theta)\,p(\text{even/odd}|\theta)d\theta \\ &= \tfrac{1}{2} \pm a\sqrt{1-a^2}\int_0^{2\pi} p(\theta)\cos(N\theta)d\theta\end{aligned} \tag{18}$$

and for a flat prior $p(\theta) = \frac{1}{2\pi}$ this becomes $p(\text{even/odd}) = \frac{1}{2}$. We now use Bayes' law to obtain a posterior distribution. For a flat prior this becomes

$$p(\theta|\text{even/odd}) = \frac{1}{2\pi}[1 \pm 2a\sqrt{1-a^2}\cos(N\theta)]. \tag{19}$$

Next, we nominate a suitable estimator for $N\theta$. As an estimator we use

$$\hat{\theta}_N^{(m)} = \arg(\int_0^{2\pi} p(\theta|m)\,\mathrm{e}^{iN\theta}\,d\theta), \tag{20}$$

and for a flat prior we have

$$\hat{\theta}_N^{(\text{even/odd})} = \arg(\mathrm{e}^{\frac{1\mp1}{2}i\pi}), \tag{21}$$

and so $\hat{\theta}_N^{(\text{even})} = 0$ and $\hat{\theta}_N^{(\text{odd})} = \pi$. We are then interested in evaluating the trace distance between the updated and corrected success-branch state

$$\rho_\text{corr}^{(m)} = \int_0^{2\pi} p(\theta|m)\,|\psi_\mathrm{s}(\theta - \tfrac{\hat{\theta}_N^{(m)}}{N})\rangle\langle\psi_\mathrm{s}(\theta - \tfrac{\hat{\theta}_N^{(m)}}{N})|\,, \tag{22}$$

with

$$|\,\psi_{\rm s}(\theta)\,\rangle = \frac{1}{\sqrt{3}}\sum_{l=0}^{2} {\rm e}^{ilN\theta}\,|\,l\,\rangle^{\otimes N} \tag{23}$$

and the desired target state $|\,\psi_{\rm s}(0)\,\rangle$. The trace distance is

$$D(\rho_{\rm corr}^{(m)},\rho_{\rm target}) = \frac{1}{2}{\rm Tr}\left(\sqrt{(\rho_{\rm corr}^{(m)}-\rho_{\rm target})^2}\right), \tag{24}$$

and we can calculate

$$\rho_{\rm corr}^{(m)} - \rho_{\rm target} = \frac{1}{3}\sum_{j\neq l} c_{jl}^{(m)}(|j\rangle\langle l|)^{\otimes N}, \tag{25}$$

with

$$c_{jl}^{(m)} = \int_0^{2\pi} p(\theta|m)\,{\rm e}^{i(N\theta-\hat{\theta}_N^{(m)})(j-l)}d\theta - 1. \tag{26}$$

For $\hat{\theta}_N^{(m)} = 0$ and $\hat{\theta}_N^{(m)} = \pi$ the coefficients $c_{jl}^{(m)}$ are real and therefore

$$D(\rho_{\rm corr}^{(m)},\rho_{\rm target}) = \frac{1}{6}(|c_{02}| + \sqrt{8c_{01}^2 + c_{02}^2}). \tag{27}$$

For the flat prior we find

$$c_{0,1}^{(\rm even/odd)} = c_{1,2}^{(\rm even/odd)} = a\sqrt{1-a^2} - 1, \tag{28a}$$

$$c_{0,2}^{(\rm even/odd)} = -1, \tag{28b}$$

and with this the trace distance becomes

$$D(\rho_{\rm corr}^{(m)},\rho_{\rm target}) = \tfrac{1}{6}\left(1+\sqrt{9+8a^2(1-a^2)-16a\sqrt{1-a^2}}\right), \tag{29}$$

and since this is independent of the outcome it already matches the average trace distance $\bar{D}(\rho_{\rm corr},\rho_{\rm target}) = \sum_m p(m)D(\rho_{\rm corr}^{(m)},\rho_{\rm target})$.

B Multiple copies without updating

Now let us see how measurements on more than one copy, in particular on k copies, can improve this value by considering a simple scenario where the measurement directions are not updated. This means, each copy is locally measured in the basis $\{|\pm\rangle_i\}$ by every party, resulting in k individual even/odd measurement outcomes. Since there is no updating in-between, the particular order of these outcomes is irrelevant and only the overall number of even/odd outcomes matter. We can thus view the entire procedure as one $(k+1)$-outcome measurement with outcomes $m = 0, 1, 2, \ldots, k$, i.e., the number of "even" outcomes. We hence assign the probabilities

$$p(m|\theta) = p({\rm even}|\theta)^m p({\rm odd}|\theta)^{k-m}\binom{k}{m}. \tag{30}$$

The corresponding unconditional probabilities are

$$p(m) = \frac{1}{2^k}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\int_0^{2\pi} p(\theta)\,\cos^{i+j}(N\theta)d\theta. \tag{31}$$

For a flat prior we calculate

$$p(m) = \frac{1}{2^k}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\,\frac{(-1)^i+(-1)^j}{2^{i+j+1}}\binom{i+j}{\frac{i+j}{2}}. \tag{32}$$

The posterior given outcome m is then

$$p(\theta|m) = \frac{1}{2^k p(m)}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j} \times (-1)^j\,(2a\sqrt{1-a^2})^{i+j}\,p(\theta)\,\cos^{i+j}(N\theta). \tag{33}$$

For the estimator we have to calculate

$$\bar{z}_{N,k}^{(m)} = \int_0^{2\pi} p(\theta|m){\rm e}^{iN\theta}d\theta. \tag{34}$$

For the flat prior the estimator becomes the argument of

$$\bar{z}_{N,k}^{(m)} = \frac{1}{2^k p(m)}\binom{k}{m}\sum_{i=0}^{m}\sum_{j=0}^{k-m}\binom{m}{i}\binom{k-m}{j}(-1)^j \times (2a\sqrt{1-a^2})^{i+j}\,\frac{(-1)^i+(-1)^j}{2^{i+j+1}}\binom{i+j}{\frac{i+j}{2}}. \tag{35}$$

We can distinguish two cases, as long as $m < k/2$ we have $\bar{z}_{N,k}^{(m)} < 0$ and therefore $\arg(\bar{z}_{N,k}^{(m)}) = \pi = \hat{\theta}_{N,m}^{(m)}$. For $m \geq k/2$ we have $\bar{z}_{N,k}^{(m)} \geq 0$ and therefore $\arg(\bar{z}_{N,k}^{(m)}) = 0 = \hat{\theta}_{N,m}^{(m)}$. What remains is calculating the distance to our target state, and the result is illustrated in Fig. 3.

In the same figure we numerically investigate the performance of different measurement-updating strategies. We notice that, perhaps unsurprisingly, the strategy with constant measurement direction performs worst of all considered strategies, whereas strategies with changing measurement directions perform considerably better. More specifically, we compare an alternating strategy that switches between two mutually unbiased measurements, and a rotating strategy, where the measurement direction is rotated by a fixed angle, to the adaptive strategy. In the latter strategy, which is also considered in the main text, the measurement direction is updated based on the outcome of the previous measurement rounds and chosen perpendicular

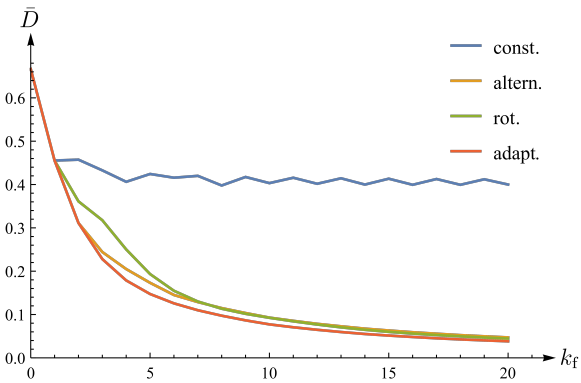

Figure 3: Average distance to target state for different ways of adapting the measurement direction. The curves show the average trace distance between the target state and the successfully converted copies for different measurement-updating strategies for the exemplary metrology-assisted entanglement distribution protocol discussed in Sec. 3 for fixed value of $a = \frac{1}{\sqrt{2}}$ as a function of the number k_f of copies in the failure branch. The blue curve represents the strategy with constant measurement directions discussed in Appendix B, for the yellow curve measurements are alternated between σ_x and σ_y, for green the measurement directions are rotated by a constant angle $\pi/8$ after each measurement, and the red curve corresponds to an adaptive strategy updating the measurement direction using the maximum likelihood estimator from Eq. (21) of the previous round.

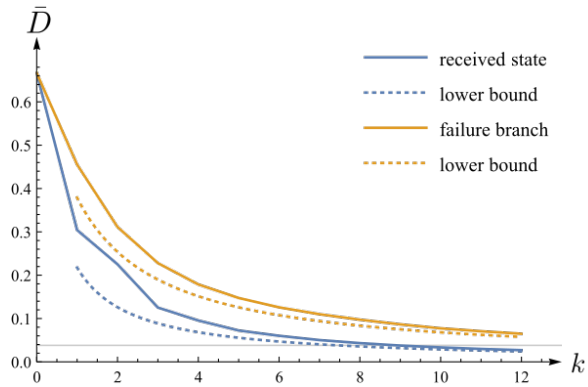

Figure 4: Average distance to target state for estimation with received states. The curve shows the average trace distance between the target state and the successfully converted copies as a function of the number k of copies used. The blue line shows the 'naive' protocol discussed in Sec. 4, where the received states $|\Psi(\alpha,\beta)\rangle$ from Eq. (7) are used directly for the estimation. The yellow curve shows the estimation performed on copies of the failure branch described by Eq. (9). The gray line shows the trace distance of $\epsilon \approx 0.038$ after measuring 20 copies of the failure branch and we see that 9 copies of the received state are necessary to achieve the same distance. The dotted line represents the lower bound calculated in Appendix D based on the Bayesian Cramér-Rao bound for the respective probe state. The parameters used are $\alpha = \arccos(2/\sqrt{15})$ and $\beta = \pi/4$.

(in Bloch representation) to the estimator for the angle θ from the previous round. While the adaptive strategy performs best, as expected, also the strategies that change the measurement direction according to some pre-established rule perform comparably well. This gives viable alternatives that are possibly easier to implement, as they do not depend on previous outcomes.

C Estimation with received copies

In this appendix we briefly explain how to estimate the phase θ using the state $|\Psi(\alpha,\beta)\rangle$ given in Eq. (7) along with a Bayesian estimation strategy. The strategy is essentially the same as that employed before, adapted to the qutrit probe states. First, we choose an appropriate three-outcome projective measurement in a basis that is mutually unbiased with respect to the eigenbasis of the generating Hamiltonian of the transformation $U(\theta)$ from Eq. (5), e.g. the discrete Fourier transform of the standard basis

$$|\xi_0\rangle = \tfrac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \tag{36a}$$

$$|\xi_1\rangle = \tfrac{1}{\sqrt{3}}(|0\rangle + \omega|1\rangle + \omega^2|2\rangle), \tag{36b}$$

$$|\xi_2\rangle = \tfrac{1}{\sqrt{3}}(|0\rangle + \omega^2|1\rangle + \omega|2\rangle), \tag{36c}$$

where $\omega = \mathrm{e}^{2\pi i/3}$. Each party performs a measurement in this basis. Let E_n denote the 3^N possible products of N projectors $|\xi_j\rangle\langle\xi_j|$. We then calculate

$$\begin{aligned}\mathrm{Tr}(|\Psi(\alpha,\beta)\rangle\langle\Psi(\alpha,\beta)|\,E_n) &= \langle\Psi(\alpha,\beta)|E_n|\Psi(\alpha,\beta)\rangle \\ &= \tfrac{1}{3^N}\big[1 + 2\sin^2\alpha\cos\beta\sin\beta\cos(\tfrac{2\pi}{3}[n_1+2n_2] - \theta N) \\ &\quad + 2\sin\alpha\cos\alpha\cos\beta\cos(\tfrac{2\pi}{3}[2n_1+n_2] - 2\theta N) \\ &\quad + 2\sin\alpha\sin\beta\cos\alpha\cos(\tfrac{2\pi}{3}[n_1-n_2] - \theta N)\big],\end{aligned} \tag{37}$$

where n_j denotes the number of projectors $|\xi_j\rangle\langle\xi_j|$ in E_n. The probabilities for the different outcomes of the measurement hence only depend on the respective number of projectors constituting the measurement, which also follows from the symmetry with respect to the exchange of the N parties. Moreover, the probabilities of the outcomes only depend on $n_1 + 2n_2$, $2n_1 + n_2$ and $n_1 - n_2$ modulo 3. We therefore only need to distinguish between three possible measurement outcomes $(0,0,0)$, $(1,2,1)$, and $(2,1,2)$, which we can rewrite as $m^l = (l, 2l, l)$. It is clear that there are $3^{(N-1)}$ sequences resulting in a given outcome, and hence it follows that

$$\begin{aligned}p(m^l|\theta) = \tfrac{1}{3}\big[&1 + 2\sin^2\alpha\cos\beta\sin\beta\cos(\tfrac{2\pi}{3}l - \theta N) \\ &+ 2\sin\alpha\cos\alpha\cos\beta\cos(\tfrac{4\pi}{3}l - 2\theta N) \\ &+ 2\sin\alpha\sin\beta\cos\alpha\cos(\tfrac{2\pi}{3}l - \theta N)\big].\end{aligned} \tag{38}$$

Here, we again want to mention that we actually estimate θN and that the estimation precision is therefore independent of the number of parties. By Bayes' law we can now calculate $p(\theta|m)$ and use the estimator from Eq. (20). In

each subsequent measurement round we adapt the measurement direction based on the outcome of the previous round, that is we rotate our measurement basis by $U(\hat{\theta}+\pi/3)$. In Fig. 4 we compare the resulting average trace distance for this strategy to that obtained from an estimation strategy based only on the failure-branch states.

D A lower bound for the number of copies for any estimation protocol

In this appendix we derive a lower bound for the number of copies needed for any estimation to achieve a given trace distance. Since we assume that no prior information about the phase is available initially, we rely on Bayesian estimation with a flat prior distribution. After k measurement rounds the posterior distribution of the parameter θ is $p(\theta|m_1,\ldots,m_k)=p(\theta|\mathbf{m})$ and the corrected state becomes

$$\rho_{\text{corr}}^{(\mathbf{m})}=\int_0^{2\pi} p(\theta|\mathbf{m})\,|\,\psi_{\text{s}}(\theta-\tfrac{\hat{\theta}_N^{(\mathbf{m})}}{N})\rangle\langle\psi_{\text{s}}(\theta-\tfrac{\hat{\theta}_N^{(\mathbf{m})}}{N})\,|\,d\theta, \tag{39}$$

for some estimator $\hat{\theta}_N^{(\mathbf{m})}$. Now we can bound the average trace distance from below using its convexity, i.e., using the notation $\rho_{\text{target}}=|\,\psi_{\text{s}}(0)\rangle\langle\psi_{\text{s}}(0)\,|$ and $\rho_{\text{s}}(\theta)=|\,\psi_{\text{s}}(\theta)\rangle\langle\psi_{\text{s}}(\theta)\,|$ we have

$$\begin{aligned}\bar{D}(\rho_{\text{corr}}^{(\mathbf{m})},\rho_{\text{target}})&=\sum_{\mathbf{m}}p(\mathbf{m})D(\rho_{\text{corr}}^{(\mathbf{m})},\rho_{\text{target}}) \qquad (40)\\ &\geq D(\sum_{\mathbf{m}}p(\mathbf{m})\rho_{\text{corr}}^{(\mathbf{m})},\rho_{\text{target}})\\ &=D\Big(\int_0^{2\pi}\sum_{\mathbf{m}}p(\mathbf{m})p(\theta|\mathbf{m})\rho_{\text{s}}(\theta-\tfrac{\hat{\theta}_N^{(\mathbf{m})}}{N})d\theta,\ \rho_{\text{target}}\Big).\end{aligned}$$

To continue the calculation and to actually compute a useful lower bound we need to make an approximation. We replace the average posterior distribution by a Gaussian distribution with the same mean μ and variance σ^2. Since the trace distance $D(|\,\psi_{\text{s}}(\theta)\rangle,|\,\psi_{\text{s}}(0)\rangle$ grows sublinearly with θ, we expect that the difference in the higher-order moments of the distributions will not influence the result significantly. Additionally, after some number of measurements the estimation is sufficiently precise such that the numerical support of the distribution is within the interval $[0,2\pi)$ and we can therefore integrate over the real line. With this we finally conclude

$$\begin{aligned}&\bar{D}(\rho_{\text{corr}}^{(\mathbf{m})},\rho_{\text{target}})\geq D(\sum_{\mathbf{m}}p(\mathbf{m})\rho_{\text{corr}}^{(\mathbf{m})},\rho_{\text{target}})\\ &\approx D\Big(\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\mu)^2}{2\sigma^2}}\rho_{\text{s}}(\theta-\tfrac{\hat{\theta}_N^{(\mathbf{m})}}{N})d\theta,\ \rho_{\text{target}}\Big)\\ &\geq D\Big(\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(\theta-\hat{\theta}_N^{(\mathbf{m})}/N)^2}{2\sigma^2}}\rho_{\text{s}}(\theta-\tfrac{\hat{\theta}_N^{(\mathbf{m})}}{N})d\theta,\ \rho_{\text{target}}\Big)\\ &=\frac{\mathrm{e}^{-2N^2\sigma^2}}{6}\big(\mathrm{e}^{\frac{N^2\sigma^2}{2}}-1\big)\Big[1+\mathrm{e}^{\frac{N^2\sigma^2}{2}}+\mathrm{e}^{N^2\sigma^2}+\mathrm{e}^{\frac{3N^2\sigma^2}{2}}\\ &\quad+\sqrt{8\mathrm{e}^{3N^2\sigma^2}+(1+\mathrm{e}^{N^2\sigma^2/2})^2(1+\mathrm{e}^{N^2\sigma^2})^2}\Big], \qquad (41)\end{aligned}$$

where we have used the fact that the trace distance is minimal if the mean of the Gaussian distribution equals the value of the estimator, i.e., using the mean of the posterior as estimator. The average variance of any estimator is bounded from below by the van Trees inequality $V[\hat{\theta}(\mathbf{m})]\geq 1/(I[p(\theta)]+\bar{I}[p(\mathbf{m}|\theta)])$, where $I[p(\theta)]$ is the Fisher information (FI) of the prior, and $\bar{I}[p(\mathbf{m}|\theta)])$ is the average FI of the likelihood over all values of θ [61]. Now we bound the FI from above by the quantum FI (QFI), $p(\mathbf{m}|\theta)]\geq\mathcal{I}[|\,\psi(\theta)\rangle]$, and notice that the QFI is independent of θ for a unitary evolution and hence equals its average [62, 63]. One can thus get a Bayesian analogue of the quantum Cramér-Rao bound for the mean of the posterior, bounding the average posterior variance by $\bar{V}[p(\theta|\mathbf{m})]\geq 1/(I[p(\theta)]+\mathcal{I}[|\,\psi(\theta)\rangle])$ (see, e.g., [64, 65]). For the state in Eq. (7) the FI is

$$\begin{aligned}\mathcal{I}\big[|\,\psi(\theta)\rangle\big]&=4N^2[\sin^2\alpha\sin^2\beta+4\cos^2\alpha\\ &\quad-(\sin^2\alpha\sin^2\beta+2\cos^2\alpha)^2] \qquad (42)\end{aligned}$$

and for a flat prior the FI vanishes. Since the expression in the last line of (41) increases monotonically in σ^2 we can further bound the average trace distance from below by inserting $\sigma^2=1/(k\,\mathcal{I}\big[|\,\psi(\theta)\rangle\big])$. This leaves us with an expression that is decreasing monotonically with the number k of copies used for estimation but which is independent of the number of parties N. With this we can determine a lower bound on the number of copies needed to reach a given trace distance.