

Distributing Multipartite Entanglement over Noisy Quantum Networks

Luís Bugalho,^{1,2} Bruno C. Coutinho,² Francisco A. Monteiro,^{2,3} and Yasser Omar^{1,2}

¹*Instituto Superior Técnico, Universidade de Lisboa, Portugal*

²*Instituto de Telecomunicações, Physics of Information and Quantum Technologies Group, Portugal*

³*ISCTE - Instituto Universitário de Lisboa, Portugal*

(Dated: November 2, 2021)

A quantum internet aims at harnessing networked quantum technologies, namely by distributing bipartite entanglement between distant nodes. However, multipartite entanglement between the nodes may empower the quantum internet for additional or better applications for communications, sensing, and computation. In this work, we present an algorithm for generating multipartite entanglement between different nodes of a quantum network with noisy quantum repeaters and imperfect quantum memories, where the links are entangled pairs. Our algorithm is optimal for GHZ states with 3 qubits, maximising simultaneously the final state fidelity and the rate of entanglement distribution. Furthermore, we determine the conditions yielding this simultaneous optimality for GHZ states with a higher number of qubits, and for other types of multipartite entanglement. Our algorithm is general also in the sense that it can optimise simultaneously arbitrary parameters. This work opens the way to optimally generate multipartite quantum correlations over noisy quantum networks, an important resource for distributed quantum technologies.

Introduction.— Quantum technologies hold the promise of faster computing, securer private communications [1–3], and more precise sensing and metrology [4, 5]. Quantum networks open the possibility to explore these applications in distributed scenarios, allowing for increased performance and/or tasks involving multiple parties. In fact, quantum networks where the links correspond to quantum entanglement between the nodes, can be thought over long distances, *e.g.* inter-city, such as in the case of a quantum Internet [6, 7], or locally, *e.g.* inside a laboratory or a quantum local-area network (QLAN) [8]. The immediate question that arises on a quantum network is how to optimally generate bipartite entanglement between two user-nodes, as a function of the relevant metrics (fidelity, rate,...). This encompasses selecting the appropriate protocols for entanglement generation and entanglement swapping [9], and algorithms that find the optimal way to execute them over a quantum network [10–15].

However, a number of applications go beyond the two-party paradigm and require multipartite entanglement, a framework with no classical analogue. Important examples of these applications are quantum sensor networks [16–19], multi-party quantum communication [20–22] and distributed quantum computation [23, 24]. For the distribution of multipartite entanglement, theoretical upper bounds derived from the communication capacities have been studied [15, 25–29] and several distribution schemes have already been developed [30–32].

In this Letter we aim at finding the optimal way to distribute multipartite entanglement in noisy quantum networks, under a given distribution scheme. This has particular relevance for applications where noise and the distribution of the state [33] impacts the application itself. To that end, we introduce a new methodology that allows to maximise two different objectives – the rate of distribution and the fidelity of the distributed state – even though our approach is easily generalizable to in-

clude more. We develop an algorithm with tools from classical routing theory [34, 35] that finds the optimal way of distributing a 3-qubit GHZ state, providing that the metrics that describe its distribution follow a set of properties, which we determine. We also find the conditions that yield optimality of our algorithm when considering a higher number of qubits, under one of the possible schemes. Moreover, our methodology is adaptable to different underlying physical implementations of the constituents of a quantum internet and its different stages of development [7].

Quantum Network Description.— Let us first describe our model of a quantum network. Structurally, it is characterised by a graph $G(V, E)$ where each node is denoted by a letter $j \in V$ and the link connecting nodes i and j is denoted by $i : j \in E$. Each path is a sequence of links and we identify them by their initial and final node $m : n$. The set of nodes $\mathcal{T} \subset V$ we want to distribute the state to is called the terminal set, where $|\mathcal{T}| = T$ is the number of terminal nodes, or equivalently, the number of qubits of the distributed state. Each link is a noisy quantum channel, accompanied by a classical channel for signalling success and sending corrections, connecting neighbouring nodes which individually hold qubits in imperfect quantum memories. Each quantum channel is capable of establishing entanglement between its two nodes and each quantum node must be able to realize one and two-qubits unitary gates, as well as performing measurements on its qubits. The protocols for entanglement generation and entanglement swapping are considered to be probabilistic, following a geometric distribution in the number of trials before the first success [36]. Moreover, each entangled pair of the network is modelled by a Werner state [37] $\rho_W = \gamma |\phi^+\rangle \langle \phi^+| + (1 - \gamma)\mathbb{1}$, with $\gamma = (4F - 1)/3$, and F being the fidelity of the state. In the bipartite case, there is an equivalence between a Werner state and a representation with a depolarising channel having an equal amount of bit-flip, phase-flip and phase-bit-flip er-

rors: $\rho_W = \mathcal{D}_1^F(|\phi^+\rangle\langle\phi^+|) = \mathcal{D}_2^F(|\phi^+\rangle\langle\phi^+|)$. This quantum channel is described in two alternative forms, one of them using the partial transposition operator Λ_i on qubit i :

$$\begin{aligned} \mathcal{D}_i^p(\rho) &= p\rho + \frac{1-p}{3}(\hat{X}_i\rho\hat{X}_i^\dagger + \hat{Y}_i\rho\hat{Y}_i^\dagger + \hat{Z}_i\rho\hat{Z}_i^\dagger) \\ &= \frac{1+2p}{3}\rho + \frac{2(1-p)}{3}\Lambda_i(\hat{Y}_i\rho\hat{Y}_i^\dagger). \end{aligned} \quad (1)$$

where the parameter p controls the amount of error in the state and $\hat{X}_i, \hat{Y}_i, \hat{Z}_i$ are the usual Pauli gates acting on qubit i . This equivalence is significant to find the final form of the distributed state. In addition, this trace-preserving map has important properties, among them linearity on the main argument ρ , commutativity for every index i , and invariance under unitary operations applied to single qubits [38].

Distribution of 3-Qubit GHZ states.— Similarly to what has been done in [30], to distribute GHZ states one would have to first find the Steiner tree connecting the set of terminal nodes and then perform some measurements over the Steiner nodes. We start with the case of 3 qubits, and later generalise for an higher number of qubits. The reason is that the Steiner tree connecting 3 terminal nodes is always a star-graph making the

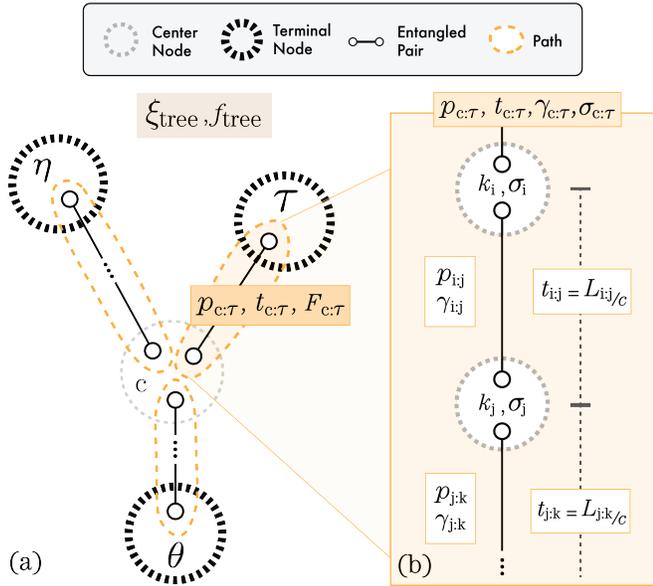


FIG. 1. (a) Star composed by 3 different paths, connecting terminal nodes τ , η and θ to the center node c . Notice that the rate ξ_{tree} depends on each branch probability of success $p_{c:\tau}$ and communications time $t_{c:\tau}$ (Eq. 2). Moreover, the fidelity of the final state f_{tree} depends on each branch fidelity $F_{c:\tau}$ (Eq. 3). (b) Branch composed by an arbitrary number of links. Each branch final fidelity $F_{c:\tau}$ depends on the path values $\gamma_{i:j}$, communications time $t_{i:j}$, and memory decoherence times $\sigma_{i:j}$. The probability of success $p_{c:\tau}$ depends on each link probability $p_{i:j}$ and the entanglement swapping probabilities k_j . The communications time $t_{c:\tau}$ depends on the lengths and the velocity of light $L_{i:j}/c$.

scheme equivalent to first distributing bipartite entanglement between every terminal and a center node and then performing a set of operations on the center node. Thus, the central idea is finding the center node for which this results in the optimal solution. To define optimality, we must concern the metrics for the distribution of this entangled state, namely the rate of distribution and the final state fidelity.

Consider that the center node, c , is capable of coordinating every process. There is an initial phase to signal the beginning of bipartite entanglement distribution, followed by the phase of creating bipartite entanglement in every branch connecting to the center node and finally performing the set of operations that result in the desired state distribution. Following [10], we can provide a measure for the average rate of distribution, ξ_{tree} , considering the different steps in this scheme. This average rate is given by the inverse of the average time, t_{tree} , it takes to complete the distribution. Considering a scheme of attempts until the first success, if any branch fails to create bipartite entanglement, every branch starts over, the final distribution of the probability of the first success will follow a geometric distribution. In such case, the average time it takes to completely distribute the multipartite entangled state is the communications time multiplied by the average number of attempts before the first success:

$$t_{tree} = \frac{2 \cdot \max_{\tau \in \mathcal{T}} \{t_{c:\tau}\}}{\prod_{\tau \in \mathcal{T}} p_{c:\tau}}, \quad \xi_{tree} = \frac{1}{t_{tree}} \quad (2)$$

where $p_{c:\tau}$ and $t_{c:\tau}$ correspond, respectively, to the probability of success and communications time of the path $c : \tau$. If the operations performed at each Steiner node are not deterministic, this time will increase by multiplying the expected value of the stochastic process that characterizes the global set of operations.

The fidelity can be calculated by applying a depolarising channel to each terminal node substituting the p value in Eq. 1 by the fidelity $F_{c:\tau} \equiv F_{\tau}$, of each branch connecting the center node c to the terminal τ . Using the second description of the depolarising channel, the result is:

$$f_{tree} = \frac{1}{2} \left[\prod_{\tau \in \mathcal{T}} \frac{1+2F_{\tau}}{3} + \prod_{\tau \in \mathcal{T}} \frac{2(1-F_{\tau})}{3} + \prod_{\tau \in \mathcal{T}} \frac{4F_{\tau}-1}{3} \right] \quad (3)$$

Under Eqs. 2-3, we can map the $|\mathcal{T}|$ paths' parameters to the tree's rate of distribution, ξ_{tree} , and fidelity of the final state, f_{tree} :

$$\left\{ [p_{c:\tau}, t_{c:\tau}, F_{c:\tau}] \right\}_{\tau \in \mathcal{T}} \mapsto [\xi_{tree}, f_{tree}] \quad (4)$$

In each branch connecting nodes m to n , we must consider a distribution of bipartite entanglement depending on four parameters: *i*) the probability of success of distributing end-to-end entanglement $p_{m:n}$, *ii*) the classical communications time $t_{m:n}$, *iii*) the fidelity expressed in γ

values $\gamma_{m:n}$, and *iv*) the memory decoherence time $\sigma_{m:n}$ (detailed description is provided in Supplemental Material). These four parameters are then contracted to three:

$$[p_{m:n}, t_{m:n}, \gamma_{m:n}, \sigma_{m:n}] \mapsto [p_{m:n}, t_{m:n}, F_{m:n}], \quad (5)$$

subject to $(4F_{m:n} - 1)/3 = \gamma_{m:n}e^{-t_{m:n}/\sigma_{m:n}}$. Eq. 5 is then the input to Eq. 4, under a proper choice of nodes m and n (see Fig. 1). Nonetheless, the four parameters that are the input of Eq. 5 must be routed separately for the optimal paths (and consequently the optimal star) to be found, under appropriate algorithms.

Algorithms for Optimal Multipartite Entanglement Distribution.— Now that we can characterize both the rate of distribution and fidelity of the final state, we need to find the optimal star. Here, the definition of optimal is crucial. In the case we only deal with one parameter, *e.g.* the fidelity, there only exists one optimal solution, unless several solutions have the same fidelity. However, if we want to consider more parameters to optimize, a multi-objective approach is necessary. In [39], multi-objective shortest-paths (MOSP) algorithms are introduced by defining the dominance relation and using the Pareto optimality definition. This results in a set of optimal solutions that is usually larger than one – the Pareto front. The same relation is also valid and central for finding the optimal way to distribute a 3-qubit GHZ state. In [34, 40], a more fundamental definition for these metrics for routing problems is introduced using algebras for routing, which we use inadvertently to refer to metrics as well. They contain a structured algebraic definition of the metric using sets to describe the space of the parameters, binary operations to define how paths are extended and total-orders to define the ordering of the parameters, *i.e.*, which one is the better:

Definition 1. Algebra for Routing is an ordered septet $(W, \preceq, L, \Sigma, \phi, \oplus, f)$ comprised as follows: W a set of weights, \preceq a total order, L a set of labels, Σ a set of signatures, ϕ a special signature, \oplus a binary operation that maps pairs of labels and signatures into a signature, and a function f that maps signatures into weights.

Using this definition, some properties can be defined [34], guaranteeing optimality when finding the shortest-path with an appropriate algorithm:

Definition 2. (Monotonicity) an algebra for routing is called monotone if: $\forall l \in L; \alpha \in \Sigma : f(\alpha) \preceq f(\alpha \oplus l)$.

Definition 3. (Isotonicity) an algebra for routing is called isotone if: $\forall l \in L; \alpha, \beta \in \Sigma : f(\alpha) \preceq f(\beta) \Rightarrow f(\alpha \oplus l) \preceq f(\beta \oplus l)$.

In the same manner, an algebra for trees can also be created. In this case, the labels define paths, rather than of edges ($L \rightarrow \Sigma$), and signatures define trees, rather than of paths ($\Sigma \rightarrow \Xi$). Using these algebras for trees, another property arises, enabling that the shortest-tree necessarily contains the possible shortest-paths if the schemes for creating paths and trees are different. This property is what we defined as label-isotonicity:

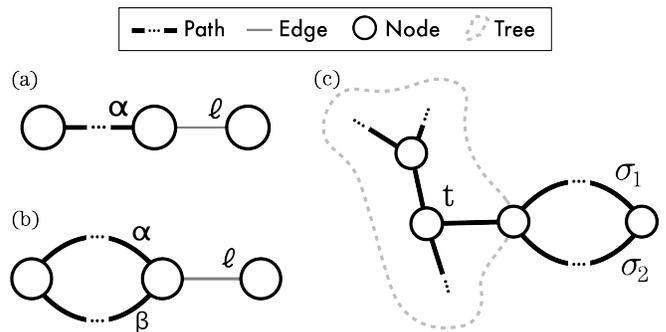


FIG. 2. Illustration of the context in the algebras' properties as found on the definitions of: (a) **Monotonicity**: extending a path α with a link l , results in a worse path; (b) **Isotonicity**: if path α is better than path β , then any extension with any link l , maintains the order; (c) **Label-isotonicity**: if path σ_1 is better than path σ_2 , then any tree t extended by path σ_1 is better than extending with path σ_2 .

Definition 4. (Label-isotonicity) An algebra for trees $(W, \preceq, \Sigma, \Xi, \phi, \oplus, f)$ is said to be label-isotone if [41]: $\forall \sigma_1, \sigma_2 \in \Sigma; t \in \Xi : \sigma_1 \preceq \sigma_2 \Rightarrow f(t \oplus \sigma_1) \preceq f(t \oplus \sigma_2)$

From the scheme for creating a 3-qubit GHZ state, the main problem is finding the correct center node, such that the corresponding star is optimal:

Algorithm 1 Algorithm for T-Star

```

1: procedure T-STAR EXACT(terminal)
2:    $A :=$  Set of solutions initialized empty;
3:   for  $node \in terminal$  do
4:     procedure SHORTEST-PATH(node);
5:    $Nodes_{reach} \leftarrow$  Nodes reachable from every terminal;
6:   for  $node \in Nodes_{reach}$  do
7:     for Tree  $T = \cup_{i,j} Path^j(node, terminal_i)$  do
8:       if  $T$  is non-dominated by every tree in  $A$  then
9:         Add  $T$  to  $A$ ;

```

Proposition 1. For the shortest-star with 3 terminals, the paths connecting the center node to the terminals must be the shortest-paths, if the underlying algebras for trees are label-isotone.

Proposition 1 provides that Algorithm 1 converges to the set of optimal solutions for the 3-qubit GHZ state problem. Considering that the solutions for the MOSP algorithms are optimal, it can also be verified that the rate of distribution and fidelity of the final state are both monotone and label-isotone for any number of qubits in the star scheme. This ensures the optimality of Algorithm 1. The necessary proofs are detailed in Supplemental Material.

Distribution of N-Qubit GHZ states.— When considering a higher number of qubits, there are two possible schemes to distribute this state. The first is the *star scheme*, which is an extension of the one introduced resembling a star-graph, hence its name. Identically, it consists of generating bipartite entanglement between every

terminal node and a center node, which can ultimately be any of the terminal nodes, and then projecting every qubit of the center node in the desired state. This scheme is capable of distributing any multipartite entangled state, besides GHZ states. Using *Algorithm 1* also leads to optimality under the same constraints on the algebras, if the same link can be used more than once. If the latter is untrue and this condition is only imposed *a posteriori*, *i.e.* from the set of solutions from the algorithm, taking only the ones where each link is used only once, then we can only extract a part of the set of optimal solutions and flag if there might exist other optimal solutions or not. Note that for three qubits, this condition did not matter since, if there was an overlap of links used, then there is always another better star, with the center node located elsewhere.

When we do not restrict ourselves to finding the best star, but instead try to find the best tree, then we should use the *tree scheme* [30], detailed in section D of Supplemental Material, alongside with a multi-objective Steiner tree algorithm. Although this is always either equal or better at distributing GHZ states, the algorithm is much more complex. It is more efficient when the number of nodes grows larger, since it almost surely reduces the number of entangled pairs consumed. However, it cannot be used to distribute some states, such as a W state.

Simulations.— In *Figure 3* we present the scaling of the complexity of this algorithm in two different types of networks, namely Erdős-Rényi networks [42, 43], which capture some properties of a quantum internet, namely the small-world property, and random geometric networks [46], which convey the limitation in the length of each individual link in a quantum network [47]. Every link has a value of fidelity uniformly distributed in $[f_{\min}, 1)$, a probability of successful entanglement generation and probability of successful entanglement swapping distributed uniformly in $[p_{\min}, 1)$. The communications time and memory coherence time are distributed uniformly in $[t_{\min}, t_{\max}]$ and $[\sigma_{\min}, \sigma_{\max}]$, respectively. The values of f_{\min} scale approximately through a power law $f_{\min} \sim [f_{\text{trunc}}]^{a/d_{\max}}$ [45] that guarantee each path contains entanglement and the final distributed state fidelity $f_{\text{trunc}}^{\text{GHZ}} > 1/2$. Notice that f_{trunc} is usually $1/2$, the minimum value for which entanglement is known to be present. However, in our case, to guarantee that most of the simulations would result in finding optimal stars with $f_{\text{trunc}}^{\text{GHZ}} > 1/2$, we set the value for each path $f_{\text{trunc}}^{\text{path}} = 0.9$.

Conclusions.— With this work, we provided a methodology and an algorithm capable of optimising multipartite entanglement distribution over noisy quantum networks. To achieve this, we calculated the necessary metrics: (i) the fidelity of the final state, depending on the noise and decoherence of the quantum memories, and (ii) the rate of distribution, which depends on the communication time required and the stochastic behavior of the processes involved. Our methodology can be used to optimize additional parameters, as long as the metrics verify the properties of monotonicity, isotonicity and

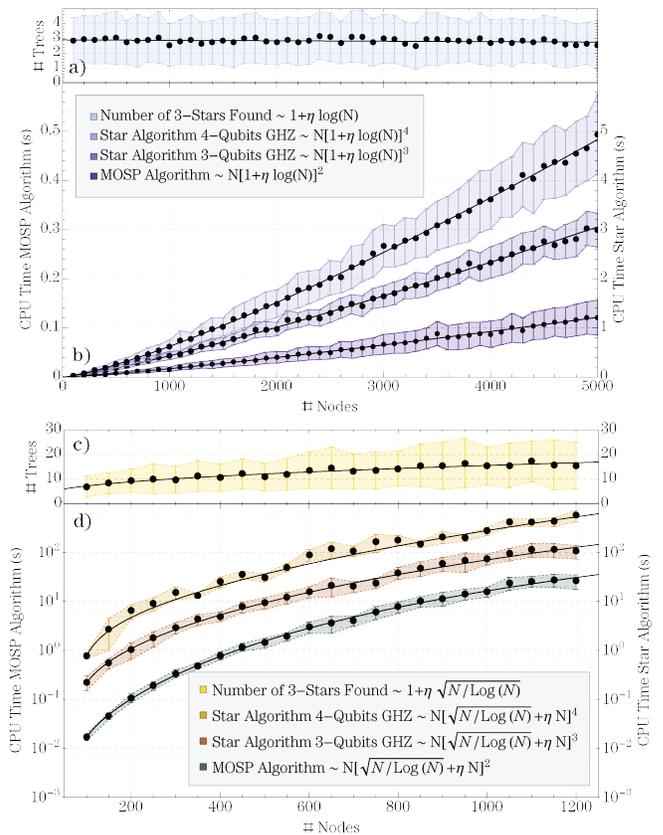


FIG. 3. Simulations for MOSP algorithm and *Algorithm 1* for 3 and 4 qubits GHZ state distribution in (b) Erdős-Rényi networks with average degree $\langle \lambda \rangle = 3$ and (d) random geometric networks with average degree $\langle \lambda \rangle = 8$, displaying as well the average number of solutions found for *Algorithm 1* applied to the 3-qubits problem in (a) Erdős-Rényi networks and (c) random geometric networks. The parameters utilized are: $p_{\min} = 0.5$, $t_{\min} = 1$, $t_{\max} = 100$, $\sigma_{\min} = 10^4$, $\sigma_{\max} = 10^5$, $f_{\text{trunc}}^{\text{GHZ}} = 0.5$, $f_{\text{trunc}}^{\text{path}} = 0.9$. For a Erdős-Rényi network $a = 2$ and $d_{\max} = \log N / \log \langle \lambda \rangle$ [42, 43]. For random geometric networks $a = 2$ and $d_{\max} = \sqrt{N / \log N}$ [44].

label-isotonicity. When extending to higher numbers of qubits, under a star-scheme, our algorithm is still optimal if we allow the links to be used more than once. Furthermore, by separating our approach into two equivalent layers, described by distinct algebras for routing and for trees, we allow bipartite entanglement distribution to be different from multipartite schemes. This opens the possibility for different protocols or algorithms within each layer and a condition that guarantees they fit together: label-isotonicity.

An additional important detail, which is exemplified by the chosen metrics for bipartite entanglement distribution, is that, while a set of parameters might not be isotonic, a decomposition into another set of isotonic parameters might still be possible. If this decomposition results in a larger number of parameters, as it did in our case, by increasing the number of objectives of opti-

mization, the algorithms runtime also increases. This is partly derived from the fact that the set of optimal solutions grows. Nonetheless, doing so is compatible with our methodology, still providing a guarantee of optimality.

This same methodology could then be applied to different distribution schemes [48], introducing entanglement purification rounds in the bipartite distribution scheme [32], or even different quantum repeater protocols [49], opening the way to optimally generate multipartite entanglement over noisy quantum networks.

Given the development of networked quantum technologies, both over a future quantum internet, and over QLANs, optimising the distribution of multipartite entanglement will naturally impact the deployment of the communications, computation and sensing applications that use this important resource [50, 51]. Future directions encompass using this algorithm with other metrics adapted to their physical realization of quantum networks or using this description of the metrics to find and verify the optimality of new algorithms for the multi-objective Steiner tree problem. While optimality might not always be the practical goal, these tools can be used

to find a measure of closeness to optimality and possibly new routing protocols that are close to optimal. The broadness of this methodology should make for a robust tool to deal with routing on quantum networks.

ACKNOWLEDGMENTS

Acknowledgments.— The authors thank D. Markham and E. Z. Cruzeiro for useful comments on the work. Furthermore, the authors thank the support from Fundação para a Ciência e a Tecnologia (Portugal), namely through project UIDB/50008/2020 and QuantSat-PT, as well as from projects TheBlinQC and QuantHEP supported by the EU H2020 QuantERA ERA-NET Co-fund in Quantum Technologies and by FCT (QuantERA/0001/2017 and QuantERA/0001/2019, respectively), and from the EU H2020 Quantum Flagship projects QIA (820445) and QMiCS (820505). B. Coutinho also thanks the support from FCT through project CEECINST/00117/2018/CP1495

-
- [1] C. H. Bennett and G. Brassard, *Theoretical Computer Science* **560**, 7 (2014).
 - [2] A. I. Nurhadi and N. R. Syambas, *Proceeding of 2018 4th International Conference on Wireless and Telematics, ICWT 2018*, 18 (2018).
 - [3] A. Broadbent, J. Fitzsimons, and E. Kashefi, *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, 517 (2009), arXiv:0807.4154.
 - [4] I. L. Chuang, *Phys. Rev. Lett.* **85**, 2006 (2000).
 - [5] D. Gottesman, T. Jennewein, and S. Croke, *Phys. Rev. Lett.* **109**, 070503 (2012).
 - [6] M. Pompili, S. L. N. Hermans, S. Baier, H. K. C. Beukers, P. C. Humphreys, R. N. Schouten, R. F. L. Vermeulen, P. J. Tiggeleman, L. dos Santos Martins, B. Dirkse, S. Wehner, and R. Hanson, *Science* **372**, 259 (2021).
 - [7] S. Wehner, D. Elkouss, and R. Hanson, *Science* **362**, 6412 (2018).
 - [8] M. Alshoukan, B. P. Williams, P. G. Evans, N. S. V. Rao, E. M. Simmerman, H.-h. Lu, N. B. Lingaraju, A. M. Weiner, C. E. Marvinney, Y.-y. Pai, B. J. Lawrie, N. A. Peters, and J. M. Lukens, (2021), arXiv:2102.13596.
 - [9] W. Munro, K. Azuma, K. Tamaki, and K. Nemoto, *IEEE J. Sel. Top. Quantum Electron.* **21**, no. 3, p. 78-90 (2015).
 - [10] M. Caleffi, *IEEE Access* **5**, 22299 (2017).
 - [11] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner, (2019), arXiv:1907.11630.
 - [12] S. Shi and C. Qian, (2019), arXiv:1909.09329.
 - [13] C. Li, T. Li, Y.-X. Liu, and P. Cappellaro, *npj Quantum Inf.* **7**, 10 (2021).
 - [14] W. Dai, T. Peng, and M. Z. Win, *IEEE Journal on Selected Areas in Communications* **38**, 540 (2020).
 - [15] S. Bäuml, K. Azuma, G. Kato, and D. Elkouss, *Commun. Phys.* **3**, 55 (2020).
 - [16] C. Ren and H. F. Hofmann, *Phys. Rev. A* **86**, 014301 (2012).
 - [17] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, *Phys. Rev. A* **100**, 022316 (2019).
 - [18] Z. Eldredge, M. Foss-Feig, J. A. Gross, S. L. Rolston, and A. V. Gorshkov, *Phys. Rev. A* **97**, 042337 (2018).
 - [19] T. Qian, J. Bringewatt, I. Boettcher, P. Bienias, and A. V. Gorshkov, (2020), arXiv:2011.01259.
 - [20] M. Hillery, V. Bužek, and A. Berthiaume, *Phys. Rev. A* **59**, 1829 (1999).
 - [21] C. Zhu, F. Xu, and C. Pei, *Sci. Rep.* **5**, 17449 (2015).
 - [22] G. Murta, F. Grasselli, H. Kampermann, and D. Bruf, *Adv. Quantum Technol.* **3**, no. 11, p. 2000025 (2020).
 - [23] E. D’Hondt and P. Panangaden, *Quantum Inf. Comput.* **6**, 173 (2004).
 - [24] R. Raussendorf and H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188 (2001).
 - [25] R. Laurenza and S. Pirandola, *Phys. Rev. A* **96**, 032318 (2017).
 - [26] S. Pirandola, *Commun. Phys.* **2**, 51 (2019).
 - [27] S. Pirandola, *Quantum Sci. Technol.* **4**, 045006 (2019).
 - [28] S. Pirandola, *IET Quantum Commun.* **1**, 22 (2020).
 - [29] S. Das, S. Bäuml, M. Winczewski, and K. Horodecki, arXiv, (2019), arXiv:1912.03646.
 - [30] C. Meignant, D. Markham, and F. Grosshans, *Phys. Rev. A* **100**, 052333 (2019).
 - [31] J. Wallnöfer, A. Pirker, M. Zwerger, and W. Dür, *Sci. Rep.* **9**, 314 (2019).
 - [32] K. Goodenough, D. Elkouss, and S. Wehner, *Phys. Rev. A* **103**, 032610 (2021).
 - [33] S. N. Filippov, A. A. Melnikov, and M. Ziman, *Phys. Rev. A* **88**, 062328 (2013).
 - [34] J. L. Sobrinho, *IEEE/ACM Trans. Netw.* **13**, 1160 (2005).
 - [35] S. Demeyer, J. Goedgebeur, P. Audenaert, M. Pickavet, and P. Demeester, *4OR - Q. J. Oper.* **11**, 323 (2013).
 - [36] S. Brand, T. Coopmans, and D. Elkouss, *IEEE J. Sel.*

- Areas Commun. **38**, 619 (2020).
- [37] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
- [38] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van Den Nest, and H. J. Briegel, *Proceedings of the International School of Physics "Enrico Fermi"*, **162**, 115 (2006).
- [39] E. Q. V. Martins, Eur. J. Oper. Res. **16**, 236 (1984).
- [40] J. L. Sobrinho, in *Comput. Commun. Rev.* **33**, 49–60 (2003).
- [41] By $\sigma_1 \preceq \sigma_2$ we mean $f(\sigma_1 \oplus e_\Sigma) \preceq f(\sigma_2 \oplus e_\Sigma)$ where e_Σ is the neutral element of Σ w.r.t \oplus .
- [42] A.-L. Barabási and M. Pósfai, *Network science* (Cambridge University Press, Cambridge, 2016).
- [43] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
- [44] R. B. Ellis, J. L. Martin, and C. Yan, *Algorithmica* **47**, 421 (2007).
- [45] B. C. Coutinho, W. J. Munro, K. Nemoto, and Y. Omar, (2021) arXiv:2103.03266.
- [46] J. Dall and M. Christensen, Phys. Rev. E **66**, 016121 (2002).
- [47] T. Inagaki, N. Matsuda, O. Tadanaga, M. Asobe, and H. Takesue, *Optics Express* **21**, 23241 (2013).
- [48] J. Wallnöfer, M. Zwerger, C. Muschik, N. Sangouard, and W. Dür, Phys. Rev. A **94**, 052307 (2016).
- [49] T. Satoh, K. Ishizaki, S. Nagayama, and R. Van Meter, Phys. Rev. A **93**, 032302 (2016).
- [50] P. Sekatski, S. Wölk, and W. Dür, Phys. Rev. Research **2**, 023052 (2020).
- [51] N. Shettell, W. J. Munro, D. Markham, and K. Nemoto, (2021), arXiv:2101.02823.
- [52] Y. neng Guo, Q. long Tian, K. Zeng, and Z. da Li, *Quantum Inf. Process.* **16**, 12 (2017).
- [53] X. Wang, *Exact algorithms for the Steiner Tree Problem*, (2008).
- [54] G. Robins and A. Zelikovsky, *SIAM J. Discrete Math.* **19**, 1 (2005).

Appendix A: Bipartite Entanglement Distribution

In this section, we detail the used metrics for bipartite entanglement distribution across a chain of quantum links, or path. We do this in a way that guarantees each metric is both monotonic and isotonic, so to ensure that a general MOSP algorithm converges to the optimal solution.

Starting with the probability of success, denote by $p_{i:j}$ the probability of successful entanglement generation between neighbouring nodes i and j at the first attempt and by k_j the probability of successful entanglement swapping on node j also at the first attempt. Since every stochastic process is independent, the probability of success of generating end-to-end entanglement across a chain of quantum nodes will also be geometrically distributed with a probability of success at the first attempt given by:

$$p_{m:n} = \prod_{i:j \in m:n} p_{i:j} \prod_{j \in m:n \setminus m,n} k_j. \quad (\text{A1})$$

The classical communications time will also play an important role in the rate and each branch fidelity. We

denote by $t_{i:j} = L_{i:j}/c$ the time it takes to send a classical message between neighbouring nodes i and j , distanced by $L_{i:j}$. In our simplified model, there are two rounds of classical communication across the chain: *i*) one for communicating successful entanglement generation across every link of the chain, and *ii*) another to communicate successful entanglement swapping in every node of the chain, with correspondent corrections. This takes two times the sum of each individual time across the chain to complete each successful round:

$$t_{m:n} = 2 \sum_{i:j \in m:n} t_{i:j}. \quad (\text{A2})$$

As for the fidelity across a chain after the entanglement swapping was completed, taking advantage of the γ change of variables, for Werner states, described in the main text $\gamma = (4F - 1)/3$, results in a simple multiplication of each γ value along the chain. When using γ we have a threshold value at $1/3$ (correspondent to $F = 1/2$), meaning that below this threshold, entanglement does no longer exist and the path cannot be used:

$$\gamma_{m:n} = \begin{cases} \prod_{i:j \in m:n} \gamma_{i:j}, & \gamma_{m:n} \geq 1/3 \\ 0, & \gamma_{m:n} < 1/3. \end{cases} \quad (\text{A3})$$

In order to include the quantum memory coherence time, similarly to what has been done in [36, 52], we adapted the contributions to our simpler scheme. Consider that, for Bell states, the effect of memory decoherence in the fidelity of the state verifies $\gamma \mapsto \gamma \cdot e^{-t_{\text{wait}}/\sigma}$, where σ is the quantum memory coherence time. In the first round, if the generation is successful, entanglement is stored between every neighbouring node of the chain. Therefore, the memory decoherence factor of each node, apart from the first and the last, contributes twice to the final state distribution. On the second round, if the swapping is successful, the entanglement is held only within the first and the last node. This results in every node contributing the same to the decoherence factor, given that the communications time, t_{wait} , is identical:

$$\frac{1}{\sigma_{m:n}} = \sum_{i \in m:n} \frac{2}{\sigma_i}. \quad (\text{A4})$$

By routing these four parameters [$p_{m:n}, t_{m:n}, \gamma_{m:n}, \sigma_{m:n}$] independently, we can guarantee that the contraction from these four parameters into only three parameters detailed in the main text provides the optimal solutions. This is only true since every of these metrics are both monotone and isotone, which can be verified in the following set of inequalities for the:

1. Probability of success metric

$$\begin{aligned} p_{m:n} &\geq p_{m:n} \oplus (p_{n:o}, k_n) = p_{m:n} \cdot p_{n:o} \cdot k_n \\ p_{m:n}^{(1)} &\geq p_{m:n}^{(2)} \implies p_{m:n}^{(1)} \oplus (p_{n:o}, k_n) \geq p_{m:n}^{(2)} \oplus (p_{n:o}, k_n) \\ &\forall p_{m:n}, p_{n:o}, k_n \in [0, 1]. \end{aligned} \quad (\text{A5})$$

2. Communications time metric

$$\begin{aligned} t_{m:n} &\leq t_{m:n} \oplus t_{n:o} = t_{m:n} + 2t_{n:o} \\ t_{m:n}^{(1)} &\leq t_{m:n}^{(2)} \implies t_{m:n}^{(1)} \oplus t_{n:o} \leq t_{m:n}^{(2)} \oplus t_{n:o} \end{aligned} \quad (\text{A6})$$

$$\forall t_{m:n}, t_{n:o} \in \mathbb{R}_0^+.$$

3. Link fidelity metric

$$\begin{aligned} \gamma_{m:n} &\geq \gamma_{m:n} \oplus \gamma_{n:o} = \gamma_{m:n} \cdot \gamma_{n:o} \\ \gamma_{m:n}^{(1)} &\geq \gamma_{m:n}^{(2)} \implies \gamma_{m:n}^{(1)} \oplus \gamma_{n:o} \geq \gamma_{m:n}^{(2)} \oplus \gamma_{n:o} \end{aligned} \quad (\text{A7})$$

$$\forall \gamma_{m:n}, \gamma_{n:o} \in [0, 1].$$

4. and Decoherence time metric

$$\begin{aligned} 1/\sigma_{m:n} &\leq 1/(\sigma_{m:n} \oplus \sigma_{n:o}) = 1/\sigma_{m:n} + 1/\sigma_{n:o} \\ 1/\sigma_{m:n}^{(1)} &\leq 1/\sigma_{m:n}^{(2)} \implies 1/(\sigma_{m:n}^{(1)} \oplus \sigma_{n:o}) \leq 1/(\sigma_{m:n}^{(2)} \oplus \sigma_{n:o}) \end{aligned} \quad (\text{A8})$$

$$\forall \sigma_{m:n}, \sigma_{n:o} \in \mathbb{R}^+.$$

In all the previous inequalities, the first corresponds to monotonicity and the second to isotonicity. When the metric is separable, as in all these cases, both properties are usually verified.

Appendix B: Dominance Relation and Proposition 1 Proof

The dominance relation is essential when dealing with multi-objective routing. This comes from the fact that, while one path might be better for some parameters, another path might be better for other parameters. So, before stepping onto the proof for *Proposition 1*, let us first dive into the dominance relation, so to generalise this result for any set of objectives. Given a set of algebras for routing $\{(W^i, \preceq^i, L^i, \Sigma^i, \phi^i, \oplus^i, f^i)\}$, as introduced in [39], the dominance relation is given by the following definition:

Definition 5. (*Dominance*) let ω and ν be two different signatures in $\{\Sigma^i\}$. We call the relation D dominance, and we say ω dominates $\nu \equiv \omega D \nu$, if $f^j(\omega^j) \preceq^j f^j(\nu^j) \forall j \in \{1, \dots, k\}$ and the strict order holds at least once.

We will first present the proof for the case of one objective of routing, or only one algebra, using only its total order and then generalize for the case of an arbitrary number of objectives using the dominance relation. Moreover, we will do this for a star with an arbitrary number of terminals.

Proposition 1. *For the shortest-star with 3 terminals, the paths connecting the center node and the terminals must be the shortest-paths, if the underlying algebras for trees are label-isotone.*

Proof. Consider the more general case in which the center node is connected by n paths (this does not happen if the center node is one of the terminals, but for that case consider the star composed of $n - 1$ paths), indexed by a number between 1 and n : $path_1, path_2, \dots, path_n \in \Sigma$. Each path is connected to one of n terminals. Fix all paths but $path_1$. Let $t \in \Xi$ be correspondent to the tree formed by $path_2 \cup path_3 \cup \dots \cup path_n$. Now consider there $\exists \overline{path_1} : \overline{path_1} \preceq path_1$, due to label-isotonicity of the algebra for trees, then if $\overline{path_1} \preceq path_1 \implies f(t \oplus \overline{path_1}) \preceq f(t \oplus path_1)$ and the shortest tree would be $path_1 \cup path_2 \cup path_3 \cup \dots \cup path_n$. Doing this for every other path, we get that the shortest-star is the one with every branch being the shortest-path between the center node and the terminals. \square

Remark. *The same applies for the multi-objective shortest-star problem, with the paths being the set of Pareto-optimal paths and requiring that every algebra for trees is label-isotone.*

Proof. Consider that $path_1 \notin X_1$ where X_1 is the set of Pareto-optimal paths between the node 1 and the center node. Then, $\exists \overline{path_1} \in X_1$ such that $\overline{path_1} D path_1$ and from here the star containing $\overline{path_1}$ is better than the one containing $path_1$. This comes from the way the dominance relation is defined. The rest of the proof is identical to the previous one. \square

Moreover, further refinements in the search process can be obtained, taking advantage of properties of the algebras. Namely, if the algebra for trees is monotone, then a necessary, but not sufficient, condition for existence of a shortest-tree connecting the set of terminal can be obtained: if there is no possible path connecting some pair of terminals, then there is no solution of the problem. From the structure of the algorithm, this condition can be implemented while finding the shortest-paths from each terminal node.

One important consideration is that the algebras for trees and the algebras for routing (parameters for trees and parameters for paths) do not need to be equal in number nor in form. As long as each algebra for routing is individually monotone and isotone, then the set of paths are optimal under an appropriate algorithm, which we will exemplify in *Appendix C*. And if each algebra for trees is label-isotone with respect to each individual algebra for routing it depends on, then this optimality is guaranteed.

Notice that in this proof we do not concern with different paths having intersections, as for the case with only three terminals, $T = 3$, this is never the case. For $T > 3$ our approach is only optimal if we allow intersections, which is equivalent to letting the same link be used more than once. If that is not the case, we can only extract at least part of the solutions and flag if there might exist more or not. If after running the algorithm and finding all possible solutions, with none of them having intersections, then all the optimal solutions were found.

However, if some have intersections, then we can discard them (this is what we mean when we say *a posteriori*) but we cannot guarantee all optimal solutions were found, since we ignored some solutions with non-optimal paths that did not have intersections. This can be translated in the following way: consider there exists a star $s = s_1 \oplus p_1$ such that $s_1 \cap p_1 \neq \emptyset$, this is, they share a common link, and that $s \in SA \equiv$ algorithm solutions. Now, consider there $\exists \tilde{p}_1 : p_1 \supset \tilde{p}_1$ and $\tilde{s} = s_1 \oplus \tilde{p}_1$ and $s_1 \cap \tilde{p}_1 = \emptyset$, there is a chance of \tilde{s} being optimal since it has no intersections, but never ends up in the algorithm solutions because it contains a dominated path \tilde{p}_1 . If however, every solution in SA has no intersections and the algorithm is optimal, it would mean that any star with a non-optimal path, intersecting or not, would be dominated by at least one of the solutions and therefore would not be optimal.

Appendix C: MOSP Algorithm

The multi-objective shortest-path (MOSP) algorithm [39] is a powerful tool to find the optimal path in the Pareto sense, which is then the input for our algorithm of distributing multipartite entanglement. We implemented it in the following way, finding every optimal path from the source to every other node of the network:

Algorithm 2 Multi-Objective Shortest-Path (MOSP)

- 1: **procedure** SHORTEST-PATH(*source*) ▷ Finds the shortest path to every node from the source
 - 2: *Nodes* := Set of nodes of the network, each with underlying list of paths $Paths_u$ initialized as empty;
 - 3: *A* := Set of visited nodes of the network initialized as empty;
 - 4: *B* := Set of nodes to visit ordered as a priority queue data structure, with priority defined by the dominance relation;
 - 5: Initialize *source* $\leftarrow \{e_{\Sigma_i}\}$;
 - 6: Add *source* to *B*;
 - 7: **while** *B* \neq empty **do**
 - 8: *node* \leftarrow Top(*B*)
 - 9: Remove *node* from *B* and add to *A*;
 - 10: **for** $v \in$ neighbours(*node*) **do**
 - 11: $Paths^{add} \leftarrow$ possible paths from $\{Paths_{node}^{(i)} \oplus Edge(node, v)\}$;
 - 12: **if** $Paths_v =$ empty **then**
 - 13: $Paths_v \leftarrow Paths^{add}$;
 - 14: Add *v* to *B*;
 - 15: **if** $Paths_v \neq$ empty **then**
 - 16: $Paths^p \leftarrow$ non-dominated paths of $Paths^{add} \cup Paths_v$;
 - 17: **if** $Paths^p \neq Paths_v$ **then**
 - 18: $Paths_v \leftarrow Paths^p$
 - 19: **if** $v \in A$ **then**
 - 20: Add *v* to *B* and remove from *A*;
 - 21: **if** $v \notin A$ **then**
 - 22: Update *v* in *B*;
-

where $\{e_{\Sigma_i}\}$ are the neutral elements of Σ_i w.r.t \oplus_i . The

main reason to separate the fidelity metric

$$\frac{4F_{m:n} - 1}{3} = \gamma_{m:n} e^{-t_{m:n}/\sigma_{m:n}} \quad (C1)$$

into three different metrics is that, while this parameters are individually monotonic and isotonic, together they are not. Then using this type of algorithm would fail at converging to the optimal solution that maximizes $F_{m:n}$. However, by separating them and treating them independently using the dominance relation, one can clearly see that, if one path dominates the other, then:

$$\left\{ \begin{array}{l} \gamma_{m:n}^{(1)} \geq \gamma_{m:n}^{(2)} \\ t_{m:n}^{(1)} \leq t_{m:n}^{(2)} \\ \sigma_{m:n}^{(1)} \geq \sigma_{m:n}^{(2)} \end{array} \right\} \implies F_{m:n}^{(1)} \geq F_{m:n}^{(2)}$$

If not all partial relations on the left hand side are true, then the paths do not dominate each other and we cannot guarantee that, at every extension, the relation of the final fidelities would maintain. This requires keeping track of all the non-dominated paths – which constitute the set of Pareto optimal paths.

Appendix D: Tree-Scheme

The tree scheme [30] consists of a generalisation of the star scheme, when we allow the topology of distribution to be any tree, not just a star-graph. Starting from a tree-graph connecting all terminal nodes (see *Figure 4b*), choosing a leaf node and iteratively applying the star expansion protocol [30] for every node will result in GHZ state distributed across the terminal nodes. Under this distribution scheme, finding the optimal way to distribute entanglement across the network is equivalent to finding the shortest-tree connecting the terminal nodes, under a given set of parameters. This is translatable to solving the multi-objective Steiner tree algorithm with the corresponding metrics. The regular Steiner tree problem is known to be *NP*-Hard [53, 54] and adding the multi-objective setup quickly escalates the problem complexity. On the other hand, finding the shortest-star connecting a set of terminal nodes is a more tractable problem, and a reason why for 3-qubits, this problem is feasible.

To find the metrics expression for the n -qubits GHZ state distributed in a tree-scheme, we considered one center of coordination in order to minimize the communications time. This center of coordination, in the case of the star scheme was trivially the center node. Now, since the coordinating node is not necessarily one branch away from every terminal node, we have to perform a minimisation over the possible locations for this node:

$$t_{tree} = 2 \cdot \frac{\min_{s \in \mathcal{S}} \left\{ \max_{\tau \in \mathcal{T}} \{t_{s:\tau}\} \right\}}{\prod_{m:n \in \text{branches}} p_{m:n}}, \quad \xi_{tree} = 1/t_{tree} \quad (D1)$$

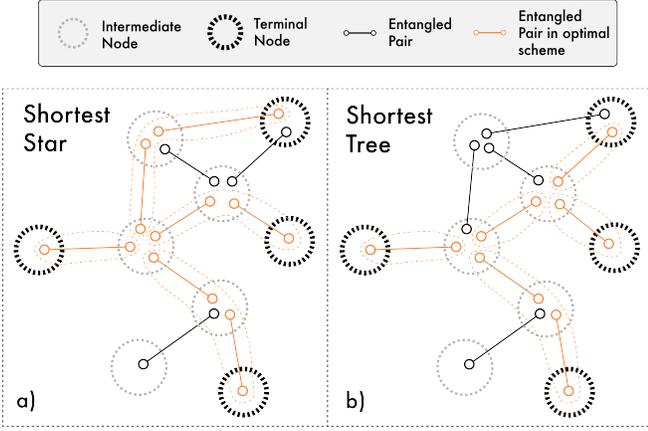


FIG. 4. A example of what a shortest tree and a shortest star are (highlighted in orange). Note that groups, or paths, surrounded by a traced orange line are equivalent to an entangled pair after bipartite entanglement distribution.

where \mathcal{S} is the set of the Steiner nodes, *i.e.*, the set of nodes that belong to the tree, but are not terminal nodes. In our case, they are always intersections of branches. Moreover, notice that the numerator of t_{tree} is the function that, given a tree, outputs the radius of the tree, under the communications time metric.

Moving to the fidelity and density matrix of the distributed state: we must start from one terminal node, τ_0 , making the result dependent from which terminal we apply the procedure. This does not invalidate the scheme described when calculating the rate, since this only affects which operations are made in each node. Starting from an initial node, after performing the star-expansion protocol across all of the intermediary nodes of the tree, the final form of the state is equivalent, up to qubit swap operations which leave the GHZ part invariant and the fidelity unaffected, to applying a depolarising channel to each terminal node except the initial:

$$\mathcal{D}_{\tau_0}(\cdot) = \bigcirc_{\tau \in \mathcal{T} \setminus \tau_0} \mathcal{D}_{\tau}^{F_{\tau}}(\cdot) = \mathcal{D}_{\tau_1}^{F_{\tau_1}} \circ \mathcal{D}_{\tau_2}^{F_{\tau_2}} \circ \dots \quad (\text{D2})$$

and applying, for each Steiner node, a depolarising channel to the initial node:

$$\mathcal{D}_{s_0}(\cdot) = \bigcirc_{s \in \mathcal{S}} \mathcal{D}_{\tau_0}^{F_s}(\cdot) = \mathcal{D}_{\tau_0}^{F_{s_0}} \circ \mathcal{D}_{\tau_0}^{F_{s_1}} \circ \dots \quad (\text{D3})$$

Making the final result

$$\mathcal{D}_{\tau_0} \circ \mathcal{D}_{s_0}(|GHZ\rangle\langle GHZ|). \quad (\text{D4})$$

In all the above, \circ stands for composition and $\bigcirc_{a \in \mathcal{A}}$ represents the enumerated composition for every a in the set \mathcal{A} . The fidelities of each depolarising channel correspond to the fidelities of each branch connecting either the terminals or Steiner nodes, respectively. The final fidelity of this state can be calculated, with the second description of the depolarising channel, given in the main text,

revealing again to be particularly useful. Separating the calculations by the different entries of the density matrix of an GHZ state, let us start on the diagonal terms, dismissing the non-GHZ entries:

$$\begin{aligned} \mathcal{D}_{\tau_0}^{F_{s_0}} \circ \mathcal{D}_{\tau_0}(|0\rangle\langle 0|) &= \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{1 + 2F_{\tau}}{3} |0\rangle\langle 0| + \dots \\ &\dots + \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{2(1 - F_{\tau})}{3} |1\rangle\langle 1| \end{aligned} \quad (\text{D5})$$

$$\begin{aligned} \mathcal{D}_{\tau_0}^{F_{s_0}} \circ \mathcal{D}_{\tau_0}(|1\rangle\langle 1|) &= \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{1 + 2F_{\tau}}{3} |1\rangle\langle 1| + \dots \\ &\dots + \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{2(1 - F_{\tau})}{3} |0\rangle\langle 0| \end{aligned} \quad (\text{D6})$$

with $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ denoting the density matrix entries $|00\dots 0\rangle\langle 00\dots 0|$ and $|11\dots 1\rangle\langle 11\dots 1|$, respectively. The off-diagonal terms are actually eigenvectors, with eigenvalue $(4F_i - 1)/3$, of the depolarising channel, making the calculations very straightforward:

$$\mathcal{D}_{\tau_0}^{F_{s_0}} \circ \mathcal{D}_{\tau_0}(|0\rangle\langle 1|) = \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{4F_{\tau} - 1}{3} |0\rangle\langle 1|, \quad (\text{D7})$$

$$\mathcal{D}_{\tau_0}^{F_{s_0}} \circ \mathcal{D}_{\tau_0}(|1\rangle\langle 0|) = \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup s_0} \frac{4F_{\tau} - 1}{3} |1\rangle\langle 0|. \quad (\text{D8})$$

The next step in the calculation is adding the Steiner nodes' depolarising channels. For this, let us first introduce two complementary functions, $E(\mathcal{S})$ and $O(\mathcal{S})$, that translate even and odd applications of $2(1 - F_i)/3 \cdot \Lambda_{\tau_0}(\hat{Y}_{\tau_0} \rho \hat{Y}_{\tau_0}^{\dagger})$. Throughout the rest of this section, let \mathcal{S} stand for the vector of fidelities of branches connecting to the Steiner nodes $\{F_{s_1}, F_{s_2}, \dots\}$ and adding branches is equivalent to $\{F_{s_1}, F_{s_2}, \dots, F_{s_n}\} \oplus \{F_s\} = \{F_{s_1}, F_{s_2}, \dots, F_{s_n}, F_s\}$. Also, for simplicity define, $\mathcal{F}_i := (1 + 2F_i)/3$ and $\bar{\mathcal{F}}_i := 1 - \mathcal{F}_i = 2(1 - F_i)/3$.

A simple example of how to calculate this can be made, using Eq. D3 with only two Steiner nodes:

$$\begin{aligned} \mathcal{D}_{s_0}(\rho) &= \mathcal{D}_{\tau_0}^{F_{s_1}} \circ \mathcal{D}_{\tau_0}^{F_{s_2}}(\rho) \\ &= \mathcal{D}_{\tau_0}^{F_{s_1}} \left(\mathcal{F}_{s_2} \rho + \bar{\mathcal{F}}_{s_2} \Lambda_0(\hat{Y}_0 \rho \hat{Y}_0^{\dagger}) \right) \\ &= (\mathcal{F}_{s_1} \cdot \mathcal{F}_{s_2} + \bar{\mathcal{F}}_{s_1} \cdot \bar{\mathcal{F}}_{s_2}) \rho + \\ &\quad + (\mathcal{F}_{s_1} \cdot \bar{\mathcal{F}}_{s_2} + \bar{\mathcal{F}}_{s_1} \cdot \mathcal{F}_{s_2}) \Lambda_0(\hat{Y}_0 \rho \hat{Y}_0^{\dagger}) \\ &\triangleq E(\{F_{s_1}, F_{s_2}\}) \rho + O(\{F_{s_1}, F_{s_2}\}) \Lambda_0(\hat{Y}_0 \rho \hat{Y}_0^{\dagger}) \end{aligned} \quad (\text{D9})$$

We can completely define the functions recursively by:

$$\begin{cases} E(\{F_{s_0}\}) = \mathcal{F}_{s_0} \\ O(\mathcal{S}) = 1 - E(\mathcal{S}) \\ E(\mathcal{S} \oplus s) = E(\mathcal{S}) \cdot \mathcal{F}_s + O(\mathcal{S}) \cdot \bar{\mathcal{F}}_s \\ O(\mathcal{S} \oplus s) = O(\mathcal{S}) \cdot \mathcal{F}_s + E(\mathcal{S}) \cdot \bar{\mathcal{F}}_s \end{cases}, \quad (\text{D10})$$

and find some of its properties, namely:

$$\begin{cases} E(\mathcal{S}) \geq O(\mathcal{S}) & , \text{if } \forall i : F_{s_i} > 1/2 \\ E(\mathcal{S}) - E(\mathcal{S} \oplus s) = [E(\mathcal{S}) - O(\mathcal{S})] \cdot \bar{\mathcal{F}}_s & \\ O(\mathcal{S} \oplus s) - O(\mathcal{S}) = [E(\mathcal{S}) - O(\mathcal{S})] \cdot \bar{\mathcal{F}}_s & \end{cases}, \quad (\text{D11})$$

where $s \equiv \{F_s\}$. Adding to *Eqs. D5, D6, D7, D8*, the missing depolarising channels of *Eq. D4*, they become:

$$\begin{aligned} \mathfrak{D}_{\mathcal{T}_0} \circ \mathfrak{D}_{\mathcal{S}_0}(|\mathbf{0}\rangle \langle \mathbf{0}|) &= E(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{1+2F_\tau}{3} |\mathbf{0}\rangle \langle \mathbf{0}| + \dots \\ &\dots + O(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{2(1-F_\tau)}{3} |\mathbf{1}\rangle \langle \mathbf{1}| \end{aligned} \quad (\text{D12})$$

$$\begin{aligned} \mathfrak{D}_{\mathcal{T}_0} \circ \mathfrak{D}_{\mathcal{S}_0}(|\mathbf{1}\rangle \langle \mathbf{1}|) &= E(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{1+2F_\tau}{3} |\mathbf{1}\rangle \langle \mathbf{1}| + \dots \\ &\dots + O(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{2(1-F_\tau)}{3} |\mathbf{0}\rangle \langle \mathbf{0}| \end{aligned} \quad (\text{D13})$$

$$\mathfrak{D}_{\mathcal{T}_0} \circ \mathfrak{D}_{\mathcal{S}_0}(|\mathbf{0}\rangle \langle \mathbf{1}|) = \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup \mathcal{S}} \frac{4F_\tau - 1}{3} |\mathbf{0}\rangle \langle \mathbf{1}| \quad (\text{D14})$$

$$\mathfrak{D}_{\mathcal{T}_0} \circ \mathfrak{D}_{\mathcal{S}_0}(|\mathbf{1}\rangle \langle \mathbf{0}|) = \prod_{\tau \in \mathcal{T} \setminus \tau_0 \cup \mathcal{S}} \frac{4F_\tau - 1}{3} |\mathbf{1}\rangle \langle \mathbf{0}| \quad (\text{D15})$$

Hence, projecting the final state over a GHZ state, the final result for the fidelity becomes:

$$(\{E(\mathcal{S}) \cdot a, O(\mathcal{S}) \cdot b, c\}, F_i) \mapsto \begin{cases} \{E(\mathcal{S}) \cdot a \cdot \mathcal{F}_i, O(\mathcal{S}) \cdot b \cdot \bar{\mathcal{F}}_i, c \cdot (\mathcal{F}_i - \bar{\mathcal{F}}_i)\}, & \text{branch connects to a terminal} \\ \{E(\mathcal{S} \oplus \{F_i\}) \cdot a, O(\mathcal{S} \oplus \{F_i\}) \cdot b, c \cdot (\mathcal{F}_i - \bar{\mathcal{F}}_i)\}, & \text{branch connects two Steiner nodes} \end{cases} \quad (\text{E1})$$

where $\{E(\mathcal{S}) \cdot a, O(\mathcal{S}) \cdot b, c\} \in (0; 1)^3$ is a general signature of a tree. Looking at *Eq. D16* and considering that fidelities below 1/2 can be discarded, $h(\cdot)$ is given by:

$$h(\{a, b, c\}) = \begin{cases} \frac{a+b+c}{2} & , \text{if } \frac{a+b+c}{2} \geq 1/2 \\ 0 & , \text{if } \frac{a+b+c}{2} < 1/2 \end{cases} \quad (\text{E2})$$

Now that we can fully characterize the algebra (or metric), let us prove the monotonicity property of this algebra. Considering one general tree, given by the following signature:

$$\{E(\mathcal{S}) \cdot a, O(\mathcal{S}) \cdot b, c\}.$$

After performing an extension with another path, as stated in *Eq. E1*, there are two possible options. Either

$$\begin{aligned} f &= \frac{1}{2} \left[E(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{1+2F_\tau}{3} + O(\mathcal{S}) \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{2(1-F_\tau)}{3} + \right. \\ &\quad \left. + \prod_{s \in \mathcal{S}} \frac{4F_s - 1}{3} \cdot \prod_{\tau \in \mathcal{T} \setminus \tau_0} \frac{4F_\tau - 1}{3} \right], \end{aligned} \quad (\text{D16})$$

Moreover, since every star is a tree, these metrics also apply for the *star-scheme* under the correct assumptions. It is easy to verify that in a star scheme, there is only one Steiner node (the center node) which translates in the functions $E(\mathcal{S})$ and $O(\mathcal{S})$ taking the values $(1 + 2F_{\tau_0})/3$ and $2(1 - F_{\tau_0})/3$ respectively.

Appendix E: Monotonicity

In this section we will prove that every algebra for trees used is monotone, for both the *star-scheme* and the *tree-scheme* detailed previously. Starting with the fidelity for the more general GHZ state, under any of the possible schemes which are detailed in the main text, we have to prove that the addition of any path to the tree results always in a worse fidelity.

By looking at the three different products in *Eq. D16*, and separating them into a vector of length three, corresponding to the signature, one can then describe the correspondent algebra for trees:

$$\begin{aligned} f_{\text{GHZ}} &: \left([1/2; 1] \cup \{0\}, \geq, (1/2; 1), (0; 1)^3, 0, \oplus_{\text{GHZ}}, h \right) \\ \text{with } \oplus_{\text{GHZ}} &: (0; 1)^3 \times (1/2; 1) \longrightarrow (0; 1)^3 \text{ given by:} \end{aligned}$$

the path connects to a terminal node τ or to a Steiner node s . In the case it connects to a terminal node, the metric is trivially monotonic since every element in the signature would be multiplied by a value smaller than one:

$$\left\{ \begin{array}{c} E(\mathcal{S}) \cdot a \\ O(\mathcal{S}) \cdot b \\ c \end{array} \right\} \mapsto \left\{ \begin{array}{c} E(\mathcal{S}) \cdot a \cdot \mathcal{F}_\tau \\ O(\mathcal{S}) \cdot b \cdot \bar{\mathcal{F}}_\tau \\ c \cdot (\mathcal{F}_\tau - \bar{\mathcal{F}}_\tau) \end{array} \right\}$$

In the case the path connects to a Steiner node then we need to use the know properties for $E(\cdot)$ and $O(\cdot)$ from *Eq. D11* and prove that:

$$\begin{cases} E(\mathcal{S} \oplus s) \cdot a + O(\mathcal{S} \oplus s) \cdot b \leq E(\mathcal{S}) \cdot a + O(\mathcal{S}) \cdot b \\ c \cdot (\mathcal{F}_s - \bar{\mathcal{F}}_s) \leq c \end{cases} \quad (\text{E3})$$

The first expression is always true, considering the properties of $E(\cdot)$ and $O(\cdot)$ and that for any tree made up of n possible paths (paths with a fidelity $F_i > 1/2$), $0 \leq b \leq (1/3)^n < (2/3)^n \leq a \leq 1$ and $F_s - \overline{F}_s \leq 1$. The second expression is also always true as $\mathcal{F}_s - \overline{\mathcal{F}}_s \leq 1$. This proves the monotonicity of this metric.

Moving to the rate metric, consider again the more general case for the GHZ state under any of the possible schemes. The first thing to notice is that while performing a minimization over all nodes for placing the coordination center, one can never actually reduce the maximum of the communications time while adding another path. This comes from the fact that in a tree, if after adding another path, this coordination center position changes, it must change to a place that still has a larger communications time, or else the previous position was not the one for the minimum to be attainable. The corresponding algebra for trees then becomes:

$\xi_{\text{GHZ}} : (\mathbb{R}_0^+, \geq, (0, 1) \times \mathbb{R}^+, (0, 1) \times \mathbb{R}^+, 0, \oplus_\xi, g)$ with $\oplus_\xi : ((0, 1) \times \mathbb{R}^+) \times ((0, 1) \times \mathbb{R}^+) \rightarrow (0, 1) \times \mathbb{R}^+$ given by

$$\begin{aligned} (\{p_{tree}, t_{tree}\}, \{p_{m:n}, t_{m:n}\}) \mapsto \\ \{p_{tree} \cdot p_{m:n}, \text{Radius}(t_{tree} \oplus t_{m:n})\}, \end{aligned} \quad (\text{E4})$$

where $\{p_{tree}, t_{tree}\} \in (0, 1) \times \mathbb{R}^+$ is a general signature and $\text{Radius}(t_{tree})$ is a function that retrieves the radius of a tree under the metric t , which in this case is the communications time metric. Then, using *Eq. D1*, the algebra weight function, $g(\cdot)$, is given by

$$g(\{p_{tree}, t_{tree}\}) = \frac{p_{tree}}{2t_{tree}}. \quad (\text{E5})$$

To prove the monotonicity, let us start from a general signature for the rate metric $\{p_{tree}, t_{tree}\} \in (0, 1) \times \mathbb{R}^+$, adding any path with a probability $0 \leq p_{m:n} \leq 1$ and a waiting time $t_{m:n} > 0$ would result in:

$$\begin{aligned} \xi = \frac{p_{tree}}{2t_{tree}} \mapsto \xi' = \frac{p_{tree} \cdot p_{m:n}}{2 \cdot \text{Radius}(t_{tree} \oplus t_{m:n})} \\ \leq \frac{p_{tree}}{2 \cdot \text{Radius}(t_{tree} \oplus t_{m:n})} \\ \leq \frac{p_{tree}}{2 \cdot t_{tree}}. \end{aligned} \quad (\text{E6})$$

In either possible case the rate always decreases. This proves that both the fidelity metric and the rate metric for a GHZ state, under any of the possible schemes, is monotonic.

Appendix F: Label-Isotonicity

In this section we will focus on the properties of the algebras for trees in a GHZ state distribution under the *star-scheme*, namely in determining whether they are label-isotone with respect to every single algebra for routing they depend on. Starting with the fidelity metric, the

correspondent algebra is identical to the one described in the previous section, but in a *star-scheme*, since there is only one Steiner node, every branch connects to a terminal. From the fidelity metric on the main text, by separating the three different identical parts in a vector, we can arrive at the algebra:

$f_{\text{GHZ}} : ([1/2; 1] \cup \{0\}, \geq, (1/2; 1), (0; 1)^3, 0, \oplus_{\text{GHZ}}, h)$ with $\oplus_{\text{GHZ}} : (0; 1)^3 \times (1/2; 1) \rightarrow (0; 1)^3$ given by:

$$\begin{aligned} (\{a, b, c\}, F_{m:n}) \mapsto \\ \left\{ a \cdot \frac{1 + 2F_{m:n}}{3}, b \cdot \frac{2(1 - F_{m:n})}{3}, c \cdot \frac{4F_{m:n} - 1}{3} \right\}, \end{aligned} \quad (\text{F1})$$

and, using the same arguments as in *Eq. E2*, $h(\cdot)$ is given by

$$h(\{a, b, c\}) = \begin{cases} \frac{a+b+c}{2} & , \text{ if } \frac{a+b+c}{2} \geq 1/2 \\ 0 & , \text{ if } \frac{a+b+c}{2} < 1/2 \end{cases}. \quad (\text{F2})$$

Now, to prove the label-isotonicity itself, consider a tree t with corresponding fidelity signature $\{a, b, c\} \in (0, 1)^3$. Moreover, consider two different paths with fidelity signatures correspondent to σ_1, σ_2 such that $\sigma_1 \preceq \sigma_2$, i.e., σ_1 has a fidelity given by $F_{m:n}^{(1)}$ and σ_2 has a fidelity given by $F_{m:n}^{(2)} \leq F_{m:n}^{(1)}$. Then:

$$\begin{aligned} F_{m:n}^{(1)} \geq F_{m:n}^{(2)} \Rightarrow \\ \Rightarrow a \cdot \frac{1 + 2F_{m:n}^{(1)}}{3} + b \cdot \frac{2(1 - F_{m:n}^{(1)})}{3} + c \cdot \frac{4F_{m:n}^{(1)} - 1}{3} \\ = \frac{a + 2b - c}{3} + \frac{2a - 2b + 4c}{3} \cdot F_{m:n}^{(1)} \\ \geq \frac{a + 2b - c}{3} + \frac{2a - 2b + 4c}{3} \cdot F_{m:n}^{(2)}, \end{aligned} \quad (\text{F3})$$

$\forall a, b, c \in (0, 1), a > b; F_{m:n}^{(1)}, F_{m:n}^{(2)} \in (1/2, 1)$, which is always the case. This proves the label-isotonicity of the fidelity of a GHZ state.

As for the rate, the corresponding algebra for trees, under a star scheme, is given by:

$\xi_{\text{GHZ}} : (\mathbb{R}_0^+, \geq, (0, 1) \times \mathbb{R}^+, (0, 1) \times \mathbb{R}^+, 0, \oplus_\xi, g)$ with $\oplus_\xi : ((0, 1) \times \mathbb{R}^+) \times ((0, 1) \times \mathbb{R}^+) \rightarrow (0, 1) \times \mathbb{R}^+$ given by

$$\begin{aligned} (\{p_{star}, t_{star}\}, \{p_{m:n}, t_{m:n}\}) \mapsto \\ \{p_{star} \cdot p_{m:n}, \max(t_{star}, t_{m:n})\}, \end{aligned} \quad (\text{F4})$$

where $\{p_{star}, t_{star}\} \in (0, 1) \times \mathbb{R}^+$ are again a general signature. Then, using similar arguments as in *Eq. E5*, g in this case is given by

$$g(\{p_{star}, t_{star}\}) = \frac{p_{star}}{2t_{star}}. \quad (\text{F5})$$

As one can see, this algebra actually depends on two parameters from each path: the probability of success

and the communications time. For this reason, we have to guarantee that the algebra is label-isotone with respect to each parameter in order for *Proposition 1* to work. Starting with the probability of success, consider two paths with corresponding probabilities $p_{m:n}^{(1)}$ and $p_{m:n}^{(2)}$ such that:

$$\begin{aligned} p_{m:n}^{(1)} \geq p_{m:n}^{(2)} &\Rightarrow \\ \Rightarrow \frac{p_{star} \cdot p_{m:n}^{(1)}}{2t_{star}} &\geq \frac{p_{star} \cdot p_{m:n}^{(2)}}{2t_{star}}, \end{aligned} \quad (\text{F6})$$

$$\forall p_{star} \in (0, 1), \forall t_{star} \in \mathbb{R}^+.$$

Moving to the communications time, consider two paths with corresponding communication times $t_{m:n}^{(1)}$ and $t_{m:n}^{(2)}$ such that:

$$\begin{aligned} t_{m:n}^{(1)} \leq t_{m:n}^{(2)} &\Rightarrow \\ \Rightarrow \frac{p_{star}}{2 \max(t_{star}, t_{m:n}^{(1)})} &\geq \frac{p_{star}}{2 \max(t_{star}, t_{m:n}^{(2)})}, \end{aligned} \quad (\text{F7})$$

$$\forall p_{star} \in (0, 1), \forall t_{star} \in \mathbb{R}^+.$$

This guarantees that both metrics are label-isotonic with respect to the corresponding algebras for routing, allowing the dominance relation used for paths to be cohesive with the dominance relation used for trees and ensuring the optimality of our algorithm.