Spatial search by continuous-time quantum walks on renormalized Internet networks

Malmi, Joonas

2022-12-13

Spatial search by continuous-time quantum walks on renormalized Internet networks

Joonas Malmi [1,2,3,*] Matteo A. C. Rossi [3,4,5] Guillermo García-Pérez [1,2,3,6] and Sabrina Maniscalco[1,2,3,4,5]

[1]*QTF Centre of Excellence, Department of Physics, Faculty of Science, University of Helsinki, FI-00014 Helsinki, Finland*
[2]*InstituteQ - the Finnish Quantum Institute, University of Helsinki, FI-00014 Helsinki, Finland*
[3]*Algorithmiq Ltd, Kanavakatu 3 C, FI-00160 Helsinki, Finland*
[4]*QTF Centre of Excellence, Department of Applied Physics, Aalto University, FI-00076 Aalto, Finland*
[5]*InstituteQ - the Finnish Quantum Institute, Aalto University, FI-00076 Aalto, Finland*
[6]*Complex Systems Research Group, Department of Mathematics and Statistics, University of Turku, FI-20014 Turun Yliopisto, Finland*

We study spatial search with continuous-time quantum walks on real-world complex networks. We use smaller replicas of the Internet network obtained with a recent geometric renormalization method introduced by García-Pérez *et al.*, Nat. Phys. **14**, 583 (2018). This allows us to infer for the first time the behavior of a quantum spatial search algorithm on a real-world complex network. By simulating numerically the dynamics and optimizing the coupling parameter, we study the optimality of the algorithm and its scaling with the size of the network, showing that on average it is considerably better than the classical scaling $\mathcal{O}(N)$, but it does not reach the ideal quadratic speedup $\mathcal{O}(\sqrt{N})$ that can be achieved, e.g. in complete graphs. However, the performance of the search algorithm strongly depends on the degree of the nodes and, in fact, the scaling is found to be very close to optimal when we consider the nodes below the 99th percentile ordered according to the degree.

I. INTRODUCTION

Continuous-time quantum walks (CTQWs), initially proposed in Ref. [1], are the quantum analogs of continuous-time classical random walks, which describe the propagation of a particle over a discrete set of positions. Together with their discrete-time counterpart [2], they have received a lot of attention for their applications in quantum information processing [3,4], quantum computation [5], and quantum transport [6]. In recent years, great progress has been made in the experimental implementation of CTQWs on different topologies, such as 2D lattices [7,8] and fractal graphs [9], and in their simulation on quantum devices [10,11].

One of the most interesting algorithmic applications of CTQWs is spatial search [12], which is the problem of finding a marked element in a structured database. Essentially, this is the generalization, or the "analog analogue" [13], of Grover's algorithm, where the spatial structure of a connected database has to be taken into account, and is usually encoded in a set of nodes and links between them, i.e., a graph.

Childs and Goldstone showed that a spatial search algorithm based on CTQWs [12] can solve the spatial search problem in certain regular graphs with a quadratic speedup $[T = O(\sqrt{N})]$ when compared to the best possible classical

algorithms $[T = O(N)]$ (with T being the total time and N being the number of nodes in the graph).

Later research has shown that spatial search by CTQW is optimal for other graph topologies such as the star graph [14,15], graphs with broken links [14], fractal graphs [16], Erdős-Rényi graphs [17], and 1D graphs with long-range interactions [18]. However, general conditions for a graph to be efficiently searchable are not known. It was shown that global symmetry and high connectivity are not necessary conditions for optimality of spatial search [19,20]. Recently, an approximate prediction of the search time and success rate has been derived based on spectral properties of the graph [21] (although as we will show later, the approximation may not hold for all graph topologies).

The networks found in real systems (e.g., biological systems, communications systems, social networks, etc.) are characterized by nontrivial topological features that depart from completely regular as well as purely random graphs and they are called complex networks. They have recently received large attention from the quantum physics community [22], and, in particular, they are studied in the context of CTQWs for coherent transport [6], for example, in biological systems [23,24], community detection [25], and link prediction [26,27].

While spatial search has been extensively studied for regular and random graphs, quite little work has been done on studying its efficiency on complex networks, with very recent results obtained for synthetic networks, such as Kronecker graphs [28] and Bollobás networks [29], that reproduce some characteristics of complex networks, such as being small world and scale-free. While synthetic models allow for fine control of the parameters of the network and to derive

---

analytical results, they often fail at capturing the properties of real-world networks.

A crucial aspect when studying spatial search is the scaling of the average search time with the size of the network. While synthetic models allow for the generation of networks with an arbitrary number of nodes, with real networks one is limited to the observed ones. To obtain replicas of the network with different sizes, in this paper we employ a recently proposed technique based on the geometric renormalization group [30,31]. The method takes the original complex network and uses geometric scaling to renormalize it while preserving the overall structure of the network. A real-space renormalization group [32], originally used to study phase transitions in statistical mechanics, has been used to research complexity bounds of a spatial search algorithm in fractal networks, albeit with coin-based discrete-time quantum walks [33]. However, this is different from the geometric renormalization group used in this paper, which is targeted toward complex networks.

In this paper, we study the effectiveness of spatial search by CTQWs on real networks by considering renormalized replicas of manageable size, namely, we consider the Internet network at the level of autonomous systems [34,35], and compare the dynamics on its renormalized replicas with the one on synthetic scale-free graphs (using the Barabási-Albert model [36]) and random graphs (using the Erdös-Rényi model [37]) of similar sizes. Note that, while here we focus on the Internet as a paradigmatic example, the approach is general, as demonstrated in Ref. [30]. Remarkably, it has been shown that the same replicas can be used to study the inverse process, namely, scaling up of real networks, through a process called geometric branching growth [31]. Specifically, in our case, the optimality and scaling analysis on replicas allow us to infer the CTQW behavior on the original Internet network and predict its properties as the Internet grows.

Summarizing, the geometric renormalization of complex networks, used here in the quantum context, has two advantages. On the one hand, it solves the issue of having just one instance of the network, as opposed to the case of synthetic networks. On the other hand, it overcomes the problem of the network size, which would be generally too large to be realistically studied. In this sense, our results pave the way for the investigation of scaling and optimality of CTQW, and its algorithmic and transport applications, to any real network.

In the following, we show that, for the Internet network, the efficacy of the spatial search depends on the degree of the target node, with large-degree nodes being characterized by the lowest success probability. Also, the optimal value of the coupling constant, leading to the best scaling of the spatial search, depends on the degree of the target node, and specifically it is inversely proportional to it. When averaged over all target nodes, quantum spatial search on the Internet exhibits a better scaling than the classical $O(N)$, but does not reach the optimal quadratic speedup $O(\sqrt{N})$. However, when averaged over 99% of the smallest-degree nodes, the scaling becomes very close to optimal.

This paper is structured as follows: In Sec. II, we review the relevant models used in this paper, namely, the spatial search by CTQW and the geometric renormalization technique, as well as the numerical methods used in the simulations. In

Sec. III, we present our results on the behavior of the spatial search algorithm on different layers of the renormalized Internet network and the corresponding mimic graphs. Section IV concludes the paper with a discussion of the results and an outlook on future research.

II. THE MODEL

In this section, we discuss the spatial search by CTQW and the geometric renormalization technique for complex networks. Numerical methods used in the calculations of the results are also considered.

A. Spatial search by CTQW

Given a network described by an undirected graph $G = (V, E)$ with $|V| = N$ nodes, a set E of edges and no self-loops, the adjacency matrix A is defined as

$$A_{jk} = \begin{cases} 1, & (j, k) \in E \\ 0, & (j, k) \notin E, \end{cases}$$

where (j, k) describes an edge between nodes j and k. We define the *Laplacian* matrix of the graph as $L = D - A$, where D is the diagonal matrix $D = \mathrm{diag}(k_1, ..., k_N)$, k_i being the degree of node i, $k_i = \sum_j A_{ij}$.

The spatial search algorithm is aimed at finding a given node w of the network by considering a CTQW on the N-dimensional Hilbert space spanned by the basis states $\{|1\rangle, ..., |N\rangle\}$. The walker is prepared in some initial state $|\psi(0)\rangle$ and then letting it evolve according to the Hamiltonian

$$H = \gamma L + H_w, \tag{1}$$

where $H_w = -|w\rangle\langle w|$ is the *oracle Hamiltonian*, which causes the amplitude to accumulate to the target site $|w\rangle$, and γ is a suitable coupling constant that is independently chosen. The value of γ regulates the rate at which the probability amplitude flows between nodes per unit time. We have chosen to neglect a constant in the oracle Hamiltonian with units of inverse time, which also makes γ a dimensionless parameter.

The state of the walker at time t is the solution of the Schrödinger equation:

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle.$$

In spatial search, the initial state of the walker is usually set to be an equal superposition between all the nodes, i.e.,

$$|\psi(0)\rangle = |s\rangle = \frac{1}{\sqrt{N}} \sum_j |j\rangle. \tag{2}$$

The state of the walker at time t is then

$$|\psi(t)\rangle = e^{-iHt}|s\rangle. \tag{3}$$

The spatial search is successful if, when measuring the position of the quantum walker at a certain time t, we find the walker in the target state $|w\rangle$ with high probability, the probability being $p_w(t) = |\langle w|\psi(t)\rangle|^2$. One thus needs to optimize this probability with respect to both t and γ so the success probability becomes as close to one as quickly as possible. These optimized values of t and γ, namely, T and
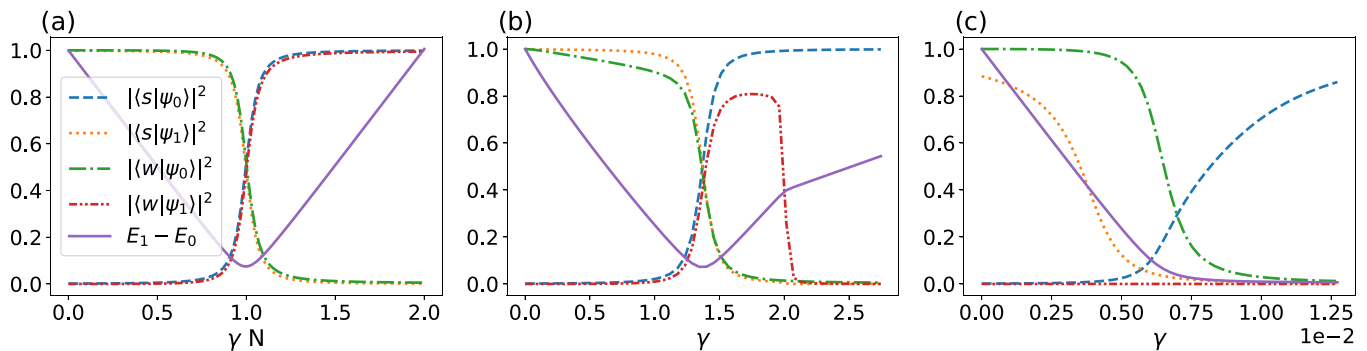
FIG. 1. The overlaps of the ground and first excited states ($|\psi_0\rangle$ and $|\psi_1\rangle$) of the Hamiltonian with the initial and target states ($|s\rangle$ and $|w\rangle$, respectively), as well as the differences between the ground (E_0) and excited state energies (E_1) in different scenarios. Plot (a) shows the case for the complete graph, where a clear point for the largest overlap can be seen for $\gamma = 1/N$. Plot (b) corresponds to a node of degree 1 and plot (c) to the highest-degree node in the smallest replica of the Internet network considered in this paper

γ_{opt}, respectively, may, in general, depend on properties of the chosen target state and of the network.

We write the success probability as

$$p_{\mathrm{succ}} = \max_{t} p_w(t), \qquad (4)$$

and indicate with T the smallest value of t that optimizes the success probability. We say that the search algorithm is optimal if $p_{\mathrm{succ}} = O(1)$ in time $T = O(\sqrt{N})$, where N is the number of nodes in the graph. When the success probability is not close to one, multiple repetitions of the algorithm may be necessary to have a high confidence of finding the target node. One should thus look at the search time, defined as [15]

$$T_w = \frac{T}{p_{\mathrm{succ}}}. \qquad (5)$$

Following Ref. [12], we recall that, for any graph, $|s\rangle$ is the ground state of the Laplacian and, therefore, for $\gamma \to \infty$ the ground state $|\Psi_0\rangle$ of the Hamiltonian in Eq. (1) is close to $|s\rangle$. On the other hand, for $\gamma \to 0$, $|\Psi_0\rangle$ is close to the target state $|w\rangle$. Moreover, since the overlap between the initial and target state is inversely proportional to $\sqrt{N}$, degenerate perturbation theory shows that the first excited state $|\Psi_1\rangle$ will be close to $|s\rangle$ for $\gamma \to 0$ in the large N limit. For many graphs, there exists some intermediate range of values of γ for which the ground state changes from the target to initial state, and at the same time the states $|\Psi_0\rangle$ and $|\Psi_1\rangle$ have substantial overlap with both $|w\rangle$ and $|s\rangle$. Under these conditions, the Hamiltonian drives transitions between these two states, and thus rotates the initial state $|s\rangle$ to a state with substantial overlap with the target state in a time inversely proportional to the energy gap $E_1 - E_0$, where E_0 is the ground-state energy and E_1 is the first excited-state energy. As we can see from Fig. 1(a), for the complete graph, this is indeed what happens. In this case, there exists a value of γ for which the energy gap approximately closes and almost perfect oscillations between the initial and target states occur. In this case, the spatial search is optimal, showing a quadratic speedup [12].

Real networks lack the symmetry of complete graphs and have very different topologies, so we expect dependence on the target node. To begin grasping the main properties of spatial search on real networks, we show in Figs. 1(b) and 1(c) the overlaps of the lowest two eigenstates of the search

Hamiltonian with both the target and the initial state, as well as the energy gap, for a renormalized replica of the Internet and two exemplary nodes, namely, a node with degree one and the node with the largest degree. The behavior of the plotted quantities for the node of degree, one retains some features seen in the complete graph, while the case of the highest degree node shows clear differences. Specifically, it is clear that, for increasingly large values of γ, the first excited eigenstate does not coincide with the target state any longer, so while the gap closes, the dynamics are not anymore confined to the two-dimensional subspace generated by $|s\rangle$ and $|w\rangle$ and we expect the probability of transition to the target state to be significantly reduced.

In this paper, we will find the optimal values γ_{opt} and T numerically. We will use the approximate values derived in Ref. [21] as one set of the initial conditions for the optimization, as well as a benchmark for our results. Following the derivation of Ref. [21] with our choice of Hamiltonian given in Eq. (1), we derive the initial approximate optimal values for γ and time. We can recast the search Hamiltonian as $H_{\mathrm{search}} = \gamma_{\mathrm{norm}} H_{\mathrm{norm}} - H_w$, where $H_{\mathrm{norm}} = I - L/\lambda_L$ is a normalized Hamiltonian. Here λ_L is the largest eigenvalue of L. The eigenvalues of H_{norm} now lie in the interval $[0,1]$ as $0 = \lambda_1 \leqslant \lambda_2 \leqslant \dots \leqslant \lambda_N = 1$ with the corresponding eigenstates of λ_N and λ_{N-1} being the ground and first excited states of L, respectively. The spectral gap is given by $\Delta = 1 - \lambda_{N-1}$. The eigenvalues λ_i and corresponding eigenstates $|v_i\rangle$ of H_{norm}, i.e., $H_{\mathrm{norm}}|v_i\rangle = \lambda_i |v_i\rangle$, can be used to express first the target node in the basis of the eigenstates as

$$|w\rangle = \sum_{i=1}^{N} a_i |v_i\rangle, \qquad (6)$$

and then to define the parameters:

$$S_k = \sum_{i=1}^{N-1} \frac{|a_i|^2}{(1 - \lambda_i)^k}. \qquad (7)$$

The approximate optimal solutions for γ and T derived for generic networks in Ref. [21] are given by

$$\gamma \approx S_1/\lambda_L, \qquad T = O\left(\frac{1}{\sqrt{\epsilon}} \frac{\sqrt{S_2}}{S_1}\right), \qquad (8)$$

where $\epsilon = 1/N$. The authors also impose a *spectral condition* for the regime of validity in which the approximations are accurate,

$$\sqrt{\epsilon} \leqslant c \min \left\{ \frac{S_1 S_2}{S_3}, \Delta \sqrt{S_2} \right\}, \tag{9}$$

where c is a sufficiently small positive constant.

B. Geometric renormalization of complex networks

In this subsection, we give a short overview of the state-of-the-art and key concepts in the geometric renormalization of complex networks.

Despite their inherent complexity, real-world networks belonging to different domains, from the Internet to metabolic networks, exhibit certain common characteristic traits, such as degree heterogeneity (that is, most nodes have a low degree while a few nodes are connected to a macroscopic fraction of the system), high levels of clustering (high density of connected triples of nodes), and the small-world property (pairs of nodes lie a few hops away from each other), among others.

Naturally, since the early days of network theory, one of the main goals has been to explain these properties across domains. Yet, the origin of clustering remained elusive until the proposal of so-called geometric models [38]. In this theoretical framework, clustering is considered to be a reflection of similarity between nodes, that is, nodes of relatively low *popularity* (low-degree nodes) can nevertheless be connected to one another with high probability if they share some other common traits.

For example, in the Internet, similarity space is found to be congruent with geographic distance [39]. In the paradigmatic network geometric models, such as $\mathbb{S}^1$ [38] and the $\mathbb{H}^2$ [40], similarity space is mathematically represented in terms of an underlying metric space (often one-dimensional for simplicity), and the connection probability between nodes depends on the distance between them in the metric space. Interestingly, these models enable the so-called embedding of the network: one can find the coordinates of the nodes of a real network in the abstract similarity space by means of a likelihood maximization approach [39,41]. The resulting network embeddings not only reveal the underlying similarity between nodes, but they can even be used to navigate the network efficiently [30,39].

Network geometric models excel at generating networks that resemble real-world networks with remarkable accuracy, including the prediction of highly nontrivial topological properties of real systems (see Ref. [42] for a recent review). One such result is the prediction that real-world graphs must be self-similar with respect to a coarse-graining transformation in similarity space, which was confirmed in real systems [30].

This discovery was further exploited to develop a technique to produce *smaller-scale replicas* of real networks, which have very similar topologies to the original networks despite their reduced size. In Ref. [30], it was also found that dynamical processes running on the replicas exhibit the same behavior as in the original network, which enables us not only to study dynamics on smaller substrates (with reduced computational costs) but also to analyze the dependence of dynamical processes on the system size even for real-world systems, for

which one typically only has access to a single, fixed-size instance [31].

While the details of the renormalization can be found in Ref. [30], we give here an intuitive explanation of how it works. Once the network has been embedded in a geometric space, we define groups of r nodes with close coordinates in the metric space, and we then replace them with *supernodes*. Each supernode is then placed within the angular region defined by the corresponding block so the order of nodes is preserved. A link between two supernodes is added if there are one or more links between the underlying groups of nodes. This operation can be iterated to obtain many renormalized network layers l, starting from the original network (which we identify as layer $l = 0$). Each layer l is then r^l times smaller than the original one. Here we set $r = 2$, that is, we group nodes in pairs. This transformation produces smaller networks with similar topological properties except for the average degree. In fact, the renormalization flow of the average degree signals the small-worldness of the network under consideration. To obtain smaller-scale replicas on which dynamical processes resemble those running on the original network, links must therefore be removed to match the original average degree. This is done stochastically according to the underlying $\mathbb{S}^1$ model in a so-called *pruning* procedure that preserves the topological features of the graph [30].

In this paper, we exploit this technique to analyze the efficiency of spatial search by CTQW on the Internet autonomous system network [34,35] as a function of the system size, hence addressing the question of whether spatial search is optimal for real-world communication network topologies.

C. Numerical methods

A geometric embedding algorithm for networks is used to obtain coordinates in a hidden metric space for the network nodes, and to then renormalize the network according to the method described in Sec. II B. The algorithm is available in the open-source package Mercator [41], which allows us to embed the complex networks in an $\mathbb{S}^1/\mathbb{H}^2$ space. Mercator uses a combination of Laplacian eigenmaps and maximum likelihood estimation methods to find the coordinates and other model parameters of every node in the underlying metric space for which the congruency between the $\mathbb{S}^1$ model and the observed real-world complex network is maximized. The $\mathbb{S}^1$ model can be mapped to the $\mathbb{H}^2$ model, which lives in a disk in hyperbolic space. In Fig. 2 , we use the $\mathbb{H}^2$ representation to depict the networks.

The optimal values of the success probability, Eq. (4), in terms of search time t and the parameter γ are obtained via numerical maximization, starting from two possible initial guesses: one is the approximate solution of Eq. (8) while the other one is $t = 0$ and $\gamma = 1/k$, where k is the degree of the target node. A standard implementation of the gradient-free Nelder-Mead algorithm is used in the optimization. These numerical optimizations are performed for each node as the target node separately in the case of each graph. For each node, the initial guess that gives the optimal result is chosen.

In principle, the calculation of the state of the walker, Eq. (3), requires diagonalizing the Hamiltonian, which is a costly (or even unfeasible) task for the largest networks, even

FIG. 2. Visualizations of the Internet network replicas, layers $l = 1$, $l = 3$, and $l = 5$ from left to right, in the hyperbolic embedding $\mathbb{H}^2$. In this embedding, the radial coordinate of a node is related to the logarithm of its degree, which is also represented by the size of the node. The color indicates the optimal success probability of the spatial search for a given target node. We notice that nodes with larger degrees (hubs) tend to have a lower success probability than small-degree nodes (see also Fig. 3).

when sparse diagonalization algorithms are used. There are, however, algorithms that allow us to efficiently evaluate the action of the exponential of a sparse matrix on a vector. One of such algorithms is Expokit, introduced in Ref. [43]. The algorithm makes use of Krylov subspace projection methods to iteratively evaluate the action of a large sparse matrix exponential on a vector, without having to calculate the exponential directly.

III. RESULTS

In this paper, we consider different renormalized replicas of a snapshot of the Internet network at the level of autonomous systems taken in 2009 [30,34,35]. The full network contains 23 748 nodes and 58 414 edges. We show results that have been calculated for five different pruned layers of the Internet, and the number of nodes and edges can be seen in Table I. Generally, the pruned networks are used to assess dynamics for different system sizes, while the renormalized ones tell us about the structure of the network at different scales (see Ref. [30] for details). Since we are interested in CTQW, we focus on the pruned replicas. Note that only the giant component is considered if the network has some disconnected components. In particular, the pruned Internet networks have from two to five disconnected components with two to three nodes each, which are left out.

For clarity, in the rest of the paper we will use the term *renormalized* in place of *pruned*, since the process of renor-

malization is the exact same in both cases, with only the stochastic pruning of the edges as an addition to the latter.

Erdös-Renyi (ER) [37] and Barabási-Albert (BA) [36] graphs are also analyzed to compare to the different renormalized replicas of the Internet. The synthetic graphs are generated in such a way that they have approximately the same number of nodes and edges of the Internet replicas [44]. Figure 2 shows three layers of renormalization of the Internet network in the $\mathbb{H}^2$ hyperbolic space, after pruning. The effect of the renormalization is easily visualized: the number of nodes and links is reduced, but the structure of the network is preserved. In the figure, the optimal probability of success for a given target node is indicated by the color, while the size indicates the degree; the larger the degree, the bigger the node. We can easily see that p_{succ} is higher for small-degree nodes and much lower for large-degree nodes.

The leftmost plot in Fig. 3 shows the optimal values γ_{opt} of the search parameter for different layers of the renormalized Internet networks as a function of the degrees, in logarithmic scale. The values of γ_{opt} follow the $1/k$ relation quite closely, with more variance in the low-degree nodes of larger networks. The middle panel of Fig. 3 shows the optimal success probability versus the degree of all nodes for the various layers of the Internet network. The inset highlights the success probabilities in the small-degree region. This panel shows that for a small-degree node p_{succ} tends to be larger on average, and for large-degree nodes it tends to be smaller. This can be intuitively understood by considering that if the target state is a small-degree node, then in the Hamiltonian $H = \gamma L - |w\rangle\langle w|$ the oracle causes a more significant change to the corresponding diagonal element (since we consider the dynamics in the node basis) than if the target state was a large-degree node. Correspondingly, the largest-degree nodes, or more accurately their states, correspond almost exactly to eigenstates of the Laplacian.

The fact that spatial search algorithm is more optimal for the peripheral small-degree nodes than for the hubs is in contrast to the case of the star graph: there, the central node is the easiest to find for the algorithm [15]. However, the case of the star graph is anomalous in that it is highly symmetric, with one central node, and with all other nodes being connected only to that node. In this case with the spatial search, regardless

TABLE I. The number of nodes (N) and edges (m) in different layers of the Internet network replicas. Only the giant component is considered for each layer.

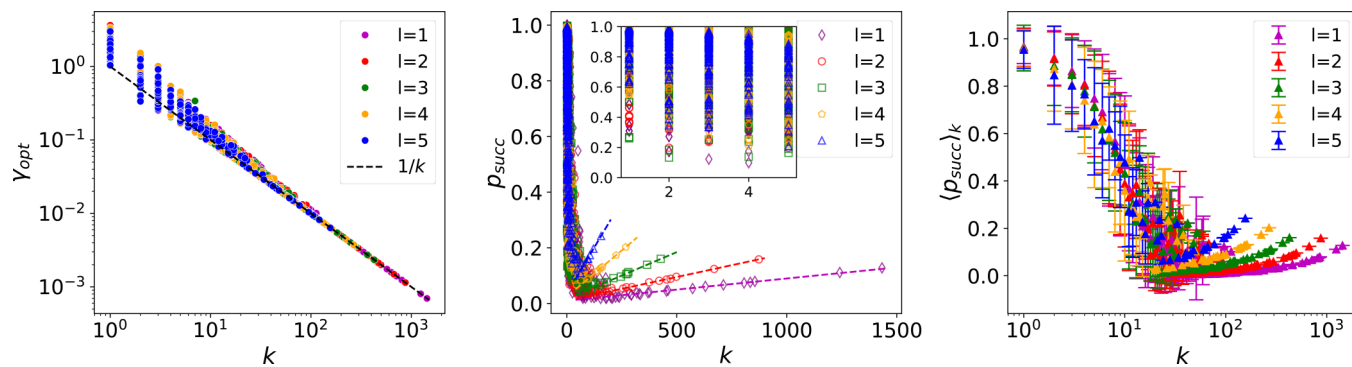| Layer	N	m
$l = 1$	11 626	28 880
$l = 2$	5729	14 660
$l = 3$	2846	7439
$l = 4$	1409	3679
$l = 5$	697	1817

FIG. 3. The leftmost plot shows the optimal values of γ_{opt} of the Internet replicas with respect to the degree k in log-log scale. The dashed black line shows the function $1/k$, which is closely followed by γ_{opt}, especially for large-degree nodes, while there are some fluctuations for small-degree nodes. The middle plot shows optimal success probabilities p_{succ} of the different layers of the renormalized Internet networks with respect to the degree k, as well as linear functions fitted to the success probabilities of the larger-degree nodes The constants of the linear function are given in Table II. The inset shows a close-up of small-degree nodes. The rightmost plot shows the averages of the optimal success probabilities over degree classes $\langle p_{\text{succ}} \rangle_k$, as well as their standard deviations.

of whether the target is the central or an outer node, the eigenspectrum of the search Hamiltonian is $N-2$ degenerate, which greatly affects the dynamics of the walk in comparison to complex networks. Furthermore, in the computational limit $N \gg 1$ there is no significant difference for the optimality of the search in the two different types of targets in the star graph [15].

For the nodes with the largest degrees, there seems to be a linear dependence between p_{succ} and the degree. To assess this dependence quantitatively, we fitted a linear function $p_{\text{succ}}(k) = ak + b$ to the data points of the largest degree nodes for each layer. The coefficients are reported in Table II. We can see that the slope nearly doubles between consecutive layers. Figure 3 shows that for nodes with intermediate connectivity, there is no clear dependence between degree k and success probability p_{succ}. We investigated the possible dependence of p_{succ} on other quantities related to the local topology of nodes, such as the weighted degree, eigenvector centrality, and the local clustering coefficient, but a clear connection was not found in any of these cases.

The rightmost plot in Fig. 3 shows the averages of optimal success probabilities with the same degree and their standard deviations in logarithmic scale. We also checked that the scaling of the optimal time T is $O(\sqrt{N})$ for all nodes with p_{succ} of the order $O(1)$ (results are not shown here). Later in the paper we show that the scaling of the optimal time

TABLE II. The constants for the fitted linear functions of the optimal success probabilities of the Internet network replicas given with respect to the degree as seen in the middle plot of Fig. 3.

$p_{\text{succ}} = ak + b$	Renormalized	
	a	b
$l = 1$	0.8×10^{-4}	0.010
$l = 2$	1.6×10^{-4}	0.014
$l = 3$	2.8×10^{-4}	0.018
$l = 4$	6.2×10^{-4}	0.025
$l = 5$	10.0×10^{-4}	0.062

when taking the average is also $O(\sqrt{N})$. This highlights the fact that for small-degree nodes the search is always optimal, and only suboptimal for mid- to large-degree nodes, which make up only a small fraction of all the nodes in the network. Let us now look at the optimal search times and how they depend on the characteristics of the target node. Figure 4 shows the optimal search times for all nodes of different layers of renormalization of the Internet, given as a function of the degrees in logarithmic scale. As one would expect, in general, the smaller the size of the replica, i.e., the larger the renormalization layer index l, the shorter the optimal search time. For the largest-degree nodes, the search time T_w and the degree k are correlated, as it was for the success probabilities depicted in Fig. 3. However, the scaling is now given by $T_w \sim k^{-1/2}$. One would then expect, for consistency, the optimal time T of
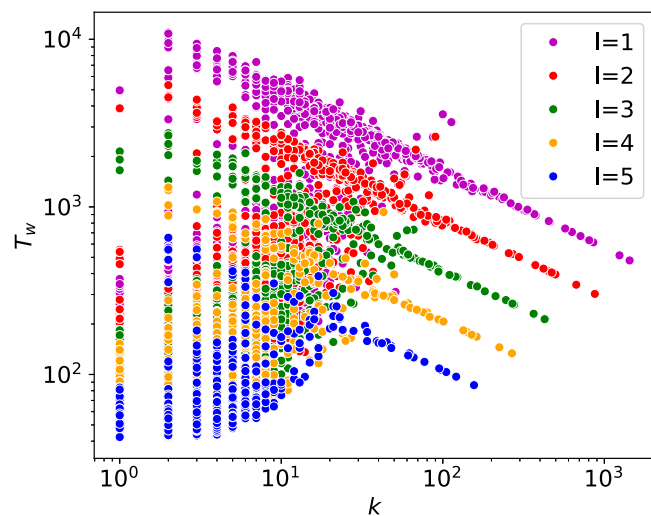

FIG. 4. Log-log plot of the optimal search time T_w versus the degree k for the nodes of the five renormalized layers of the Internet network. As one can expect, the optimal search time increases with the size of the network. For the largest-degree nodes, the search time T_w and the degree k are correlated, with $T_w \sim k^{-1/2}$.
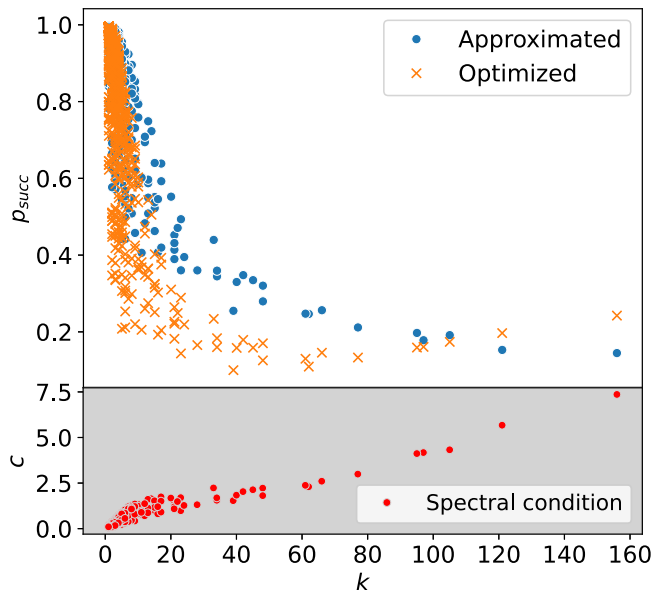
FIG. 5. The calculated approximated and optimized values of optimal success probabilities for the $l = 5$ network, as well as the constant c with respect to the degree, which indicates a decrease in the accuracy of the prediction.

the largest-degree nodes to scale roughly as $T \sim k^{1/2}$, which is indeed the case (not shown).

We now assess to which extent the approximated results of Eq. (8) capture the main features of the quantum spatial search on a real network like the Internet, or more specifically its smaller-scale replicas. Figure 5 compares the approximated values of p_{succ} obtained with Eq. (8) to the numerical results for layer $l = 5$ of the renormalized Internet network. From the plot it is evident that the approximation is quite accurate for small-degree nodes, but deviates from the numerical optimum for large-degree nodes. This can be explained by calculating the value of the constant c from Eq. (9), which is shown in the lower part of the plot; we notice that c is almost linearly dependent on k. As a consequence, the accuracy of the approximation decreases as k increases. Indeed, in the upper plot we do see that the approximate and optimized values overlap well only for small-degree nodes. Although the two values are very similar in the region of nodes with degrees roughly 100,

the slopes are different, so the approximate solutions do not predict the overall behavior for large-degree nodes very well.

As noted in Sec. I, when assessing the performance of spatial search, one is interested in the dependence of the search time and success probability as functions of the size of the graph. In the ideal case, the search time will grow as $O(\sqrt{N})$. Since in nonregular graphs the search time and probability depend on the individual target node, we look at these quantities averaged over all the nodes of the network. The results are shown in Fig. 6 for the various layers of the Internet replicas, as well as for ER and BA graphs of similar sizes.

We see that the average success probability $\overline{p}_{\text{succ}}$ for the renormalized Internet graphs increases with increasing N, which indicates that when the average degree is held constant while scaling down the network, the optimal success probability of the walker gets worse, on average. In the case of the ER and BA graphs, on the other hand, $\overline{p}_{\text{succ}}$ remains constant as a function of N. Hence, the average success probability of spatial search on a real network scales better with the system size than on equivalent synthetic graphs.

The average optimal time $\overline{T}$ for all graphs follows the optimal $O(\sqrt{N})$. The averages of the search times $\overline{T}_w$, defined in Eq. (5), increases more rapidly with increasing N for the Internet networks. The synthetic graphs mimicking the network replicas, on the other hand, show almost optimal scaling, being very close to $O(\sqrt{N})$. To substantiate these interpretations of the plots, we fitted a logarithmic function to the data points, which we discuss in detail in the last two paragraphs of this section. When this figure of merit is taken into account, the performance of spatial search on a real-world complex network is better than classical, but not optimal. Note that the average times $\overline{T}$ have a similar scaling for the Internet and mimic replicas while the success probability is on average higher for the renormalized Internet. Nonetheless, $\overline{T}_w$ for the Internet replicas is always larger than for the mimic ER and BA graphs. Recalling Eq. (5), this suggests a nontrivial correlation between success probability and optimal search time for the nodes in the network. Indeed, we find an anti-correlation between these two quantities that increases with the system size. Under these circumstances, it is reasonable to expect $\overline{T}_w$ to become increasingly larger than $\overline{T}/\overline{p}_{\text{succ}}$ as N increases.
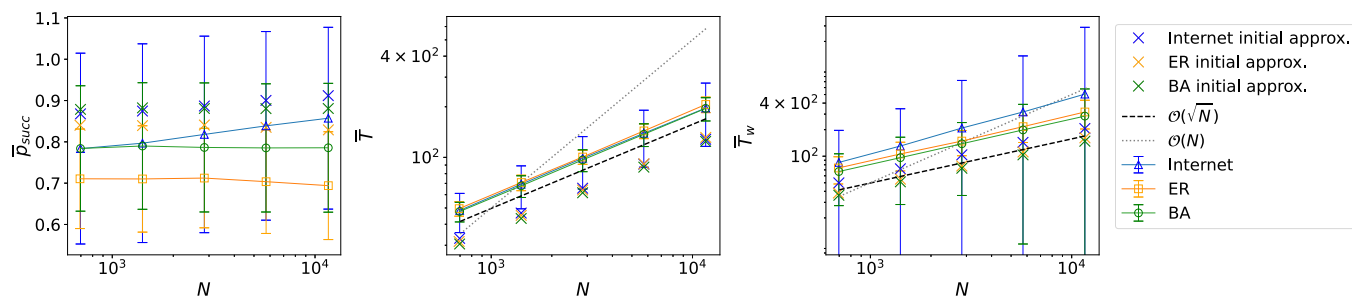

FIG. 6. The averages and standard deviations of optimal success probabilities, times and search times of the renormalized Internet networks, as well as those of the ER and BA graphs, as functions of the number of nodes N. The optimal quantum scaling $O(\sqrt{N})$ is highlighted by a black dashed line, and the classical scaling $O(N)$ is highlighted by a gray dotted line. The lines show the same quantities calculated using the approximations in Eq. (8).

TABLE III. The constants for the fitted logarithmic function ($\ln \overline{T}_w = c + x \ln N$) on the average search times of the Internet network replicas as well as the ER and BA mimic graphs as seen on the rightmost plot of Fig. 6. For the exponential constant x, the standard error is also provided

$\ln \overline{T}_w = c$	All nodes		99% of nodes in order of ascending degree	
$+x \ln N$	c	x	c	x
Internet	0.290	0.633 ± 0.006	1.115	0.511 ± 0.015
ER	0.914	0.516 ± 0.007	0.915	0.516 ± 0.006
BA	0.838	0.514 ± 0.001	0.864	0.504 ± 0.009

The averages of the approximated values in Fig. 6 show that for all graphs the approximations overestimate the efficiency of the spatial search algorithm, since the approximated $\overline{p}_{\text{succ}}$ is always larger than the actual one, and the approximated $\overline{T}$ and $\overline{T}_w$ are always lower. The difference between the averages of approximated values and the actual values remains the same with changing N, for all plotted quantities, with the exceptions of $\overline{T}_w$ and $\overline{p}_{\text{succ}}$ of the Internet replicas, for the first of which the averages of approximated values scale much better than the corresponding average calculated values, and for the latter where for smaller N the difference is larger than for larger N. The better scaling for approximated $\overline{T}_w$ is caused by the approximate values overestimating the optimal success probabilities for the mid-degree nodes, which also have the largest search times.

The standard deviations in Fig. 6 also show that the optimal times are quite uniform, the optimal success probabilities have more variance but the standard deviations stay similar between layers, and the standard deviations in the search times are much greater due to the standard deviations in both $\overline{p}_{\text{succ}}$ and $\overline{T}$. We can also see that the variance in Internet replicas is larger in all cases than the variance in ER and BA mimic graphs. For $\overline{p}_{\text{succ}}$, the changes in averages between layers are also always within the standard deviations, so not much can be interpreted about the scaling of $\overline{p}_{\text{succ}}$ from this.

To determine the scaling of $\overline{T}_w$, we fitted a logarithmic function of the form $c + x \ln N$ to the logarithm of the data points of the optimized results. The values for the constants can be seen in Table III. The obtained exponent is $x \approx 0.63$ for the Internet replicas, so the scaling is indeed better than classical, but not optimal. For the ER and BA graphs, the scaling exponents are approximately 0.52 and 0.51, respectively, so the spatial search is already almost optimal for those graph types.

We recall that the approximated quantities well describe the case in which the target state is a node with a small degree, lying at the periphery of the hyperbolic graphs of Fig. 2. One may then argue that, if one is interested in searching such nodes—or, equivalently, in efficient transport by CTQW to such nodes—one observes a quadratic speedup, with optimal scaling $O(\sqrt{N})$. Indeed, if we only take into consideration the first 99% of the nodes in each network arranged in the order of ascending degree (essentially only discarding the hubs) and fit the logarithmic function again to the average optimized

search times, then the scaling for the Internet replicas becomes approximately $N^{0.51}$, a noticeable improvement and very close to optimal (indeed, within one standard deviation). For the ER and BA graphs, the scaling remains practically the same as in the previous case.

IV. CONCLUSIONS

In this paper, we studied the efficacy of spatial search by CTQW on a real-world complex network, namely, the Internet. Given the large size of the graph, its nontrivial structure, and spectral properties, exact analytical solutions are not possible and numerical simulation is computationally intensive for full-sized graphs. We thus applied in the quantum information field a recently introduced geometric renormalization technique for complex networks to bring the graph to a tractable size while also preserving its structural properties. This also allowed us to study the scaling of spatial search with the size of the network.

We found numerically the success probability, optimal coupling constant, and optimal search time for each node of various renormalized versions of the Internet, possessing between 700 and 12 000 nodes, approximately. We showed that the optimal value of the coupling constant is inversely proportional to the degree of the target node. The success probability is, in general, larger and close to one for small-degree target nodes; it decreases abruptly for nodes of intermediate degree and grows linearly with the degree for large-degree nodes. We also found that recent approximate results [21] on the success probability and search time do not hold well for the Internet and its renormalized replicas, in particular, for large-degree nodes and, in general, tend to overestimate the performance of the algorithm.

Finally, we found that the scaling of the average optimal time with the number of nodes in the network is quite close to the ideal $O(\sqrt{N})$, showing an advantage over classical search algorithms. However, if one does take into account that the average success probability is significantly smaller than one, and thus factors in the time required to repeat the search for a sufficient number of times, the scaling reduces to approximately $O(N^{0.63})$. In general, this is due to the fact that the performance of the spatial search depends sensibly on the choice of the target node, and some of the nodes, such as the hubs in the Internet graph, show very low success probability. Indeed, the scaling gets much closer to the optimal one if we disregard 1% of the nodes with the largest degrees.

In this paper, we have considered the Childs and Goldstone algorithm for spatial search, where the state preparation does not assume knowledge of the topology of the graph. Other refined algorithms, such as the one recently introduced in Ref. [45], could give better performance, although they generally require preparing the walker in an optimized state, which usually requires prior knowledge of the topology.

Complex network science combined with quantum physics is a very new field with much left to discover. In particular, in further research we could look at the quantum speedup of the search time with respect to the hitting time of a classical random walk, rather than assuming a uniform search time

$O(N)$, since different nodes in these complex networks can have very different hitting times. In this paper, we opted to follow the approach that is customary in the literature for checking the quantum speedup of the Childs and Goldstone quantum spatial search algorithm [12].

Another interesting aspect to the research is the performance of the algorithm with multiple marked nodes [46] on real-world network. Since no analytical bounds are currently known for the Childs and Goldstone algorithm when there are multiple marked nodes, this could offer valuable insight into the behavior of the spatial search algorithm.

Finally, a crucial topic to address in the future is the performance of spatial search by CTQW on complex networks in the presence of noise [15]. While, in general, detrimental for quantum information applications, noise has been shown to enhance transport, for example, in complex biological networks [24,47].

The above aspects are particularly relevant in the context of quantum biology, and specifically interesting for the lively debate on whether quantum effects, described in terms of transport via CTQW and spatial search, may have a role in biological systems, including living organisms [48]. To properly address this question, one needs to analyze the dynamics on real biological networks. The tools and methods developed in this paper are ideally suited to such an endeavor.

In conclusion, we have shown how to meaningfully describe quantum spatial search, and more, in general, assess the performance of CTQW, on real-world complex networks. In this sense, our results bridge a gap between complex network theory and quantum science, and are therefore expected to lead to pathways in quantum biology and communication, as well as quantum simulation and algorithms.

[1] E. Farhi and S. Gutmann, Quantum computation and decision trees, Phys. Rev. A **58**, 915 (1998).

[2] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, Phys. Rev. A **48**, 1687 (1993).

[3] J. Kempe, Quantum random walks: An introductory overview, Contemp. Phys. **44**, 307 (2003).

[4] S. E. Venegas-Andraca, Quantum walks: A comprehensive review, Quant. Info. Proc. **11**, 1015 (2012).

[5] A. M. Childs, Universal Computation by Quantum Walk, Phys. Rev. Lett. **102**, 180501 (2009).

[6] O. Mülken and A. Blumen, Continuous-time quantum walks: Models for coherent transport on complex networks, Phys. Rep. **502**, 37 (2011).

[7] K. Poulios, R. Keil, D. Fry, J. D. A. Meinecke, J. C. F. Matthews, A. Politi, M. Lobino, M. Gräfe, M. Heinrich, S. Nolte, A. Szameit, and J. L. O'Brien, Quantum Walks of Correlated Photon Pairs in Two-Dimensional Waveguide Arrays, Phys. Rev. Lett. **112**, 143604 (2014).

[8] H. Tang, X.-F. Lin, Z. Feng, J.-Y. Chen, J. Gao, K. Sun, C.-Y. Wang, P.-C. Lai, X.-Y. Xu, Y. Wang, L.-F. Qiao, A.-L. Yang, and X.-M. Jin, Experimental two-dimensional quantum walk on a photonic chip, Sci. Adv. **4**, eaat3174 (2018).

[9] X.-Y. Xu, X.-W. Wang, D.-Y. Chen, C. M. Smith, and X.-M. Jin, Quantum transport in fractal networks, Nat. Photonics **15**, 703 (2021).

[10] X. Qiang, Y. Wang, S. Xue, R. Ge, L. Chen, Y. Liu, A. Huang, X. Fu, P. Xu, T. Yi, F. Xu, M. Deng, J. B. Wang, J. D. A. Meinecke, J. C. F. Matthews, X. Cai, X. Yang, and J. Wu, Implementing graph-theoretic quantum algorithms on a silicon photonic quantum walk processor, Sci. Adv. **7**, eabb8375 (2021).

[11] D. Qu, S. Marsh, K. Wang, L. Xiao, J. Wang, and P. Xue, Deterministic Search on Star Graphs via Quantum Walks, Phys. Rev. Lett. **128**, 050501 (2022).

[12] A. M. Childs and J. Goldstone, Spatial search by quantum walk, Phys. Rev. A **70**, 022314 (2004).

[13] E. Farhi and S. Gutmann, Analog analogue of a digital quantum computation, Phys. Rev. A **57**, 2403 (1998).

[14] L. Novo, S. Chakraborty, M. Mohseni, H. Neven, and Y. Omar, Systematic dimensionality reduction for quantum walks: Optimal spatial search and transport on non-regular graphs, Sci. Rep. **5**, 13304 (2015).

[15] M. Cattaneo, M. A. C. Rossi, M. G. A. Paris, and S. Maniscalco, Quantum spatial search on graphs subject to dynamical noise, Phys. Rev. A **98**, 052347 (2018).

[16] E. Agliari, A. Blumen, and O. Mülken, Quantum-walk approach to searching on fractal structures, Phys. Rev. A **82**, 012305 (2010).

[17] S. Chakraborty, L. Novo, A. Ambainis, and Y. Omar, Spatial Search by Quantum Walk is Optimal for Almost All Graphs, Phys. Rev. Lett. **116**, 100501 (2016).

[18] D. Lewis, A. Benhemou, N. Feinstein, L. Banchi, and S. Bose, Optimal Quantum Spatial Search with One-Dimensional Long-Range Interactions, Phys. Rev. Lett. **126**, 240502 (2021).

[19] J. Janmark, D. A. Meyer, and T. G. Wong, Global Symmetry is Unnecessary for Fast Quantum Search, Phys. Rev. Lett. **112**, 210502 (2014).

[20] D. A. Meyer and T. G. Wong, Connectivity is a Poor Indicator of Fast Quantum Search, Phys. Rev. Lett. **114**, 110503 (2015).

[21] S. Chakraborty, L. Novo, and J. Roland, Optimality of spatial search via continuous-time quantum walks, Phys. Rev. A **102**, 032214 (2020).

[22] J. Biamonte, M. Faccin, and M. De Domenico, Complex networks from classical to quantum, Commun. Phys. **2**, 53 (2019).

[23] C. Benedetti, M. A. C. Rossi, and M. G. A. Paris, Continuous-time quantum walks on dynamical percolation graphs, Europhys. Lett. **124**, 60001 (2019).

[24] D. A. Chisholm, G. García-Pérez, M. A. C. Rossi, G. M. Palma, and S. Maniscalco, Stochastic collision model approach to transport phenomena in quantum networks, New J. Phys. **23**, 033031 (2021).

[25] M. Faccin, P. Migdał, T. H. Johnson, V. Bergholm, and J. D. Biamonte, Community Detection in Quantum Complex Networks, Phys. Rev. X **4**, 041012 (2014).

[26] J. P. Moutinho, A. Melo, B. Coutinho, I. A. Kovács, and Y. Omar, Quantum link prediction in complex networks, arXiv:2112.04768.

[27] M. Goldsmith, G. García-Pérez, J. Malmi, M. A. C. Rossi, H. Saarinen, and S. Maniscalco, Link prediction with continuous-time classical and quantum walks, arXiv:2208.11030.

[28] T. G. Wong, K. Wünscher, J. Lockhart, and S. Severini, Quantum walk search on Kronecker graphs, Phys. Rev. A **98**, 012338 (2018).

[29] T. Osada, B. Coutinho, Y. Omar, K. Sanaka, W. J. Munro, and K. Nemoto, Continuous-time quantum-walk spatial search on the Bollobás scale-free network, Phys. Rev. A **101**, 022310 (2020).

[30] G. García-Pérez, M. Á. Serrano, and M. Boguñá, Multiscale unfolding of real networks by geometric renormalization, Nat. Phys. **14**, 583 (2018).

[31] M. Zheng, G. García-Pérez, M. Boguñá, and M. Serrano, Scaling up real networks by geometric branching growth, Proc. Natl. Acad. Sci. USA **118**, e2018994118 (2021).

[32] R. K. Pathria, *Statistical Mechanics*, 2nd ed. (Elsevier Butterworth Heinemann, Oxford, UK, 1996).

[33] S. Boettcher, S. Li, T. D. Fernandes, and R. Portugal, Complexity bounds on quantum search algorithms in finite-dimensional networks, Phys. Rev. A **98**, 012320 (2018).

[34] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, CATCH, http://www.caida.org/projects/ark/ (2009).

[35] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, Internet mapping: From art to science, in *2009 Cybersecurity Applications Technology Conference for Homeland Security* (NW Washington, DC, USA, 2009), pp. 205–211.

[36] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47 (2002).

[37] P. Erdös and A. Rényi, On Random Graphs I, Publ. Math. Debrecen **6**, 290 (1959).

[38] M. A. Serrano, D. Krioukov, and M. Boguñá, Self-Similarity of Complex Networks and Hidden Metric Spaces, Phys. Rev. Lett. **100**, 078701 (2008).

[39] M. Boguñá, F. Papadopoulos, and D. Krioukov, Sustaining the internet with hyperbolic mapping, Nat. Commun. **1**, 62 (2010).

[40] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, Hyperbolic geometry of complex networks, Phys. Rev. E **82**, 036106 (2010).

[41] G. García-Pérez, A. Allard, M. Á. Serrano, and M. Boguñá, Mercator: Uncovering faithful hyperbolic embeddings of complex networks, New J. Phys. **21**, 123033 (2019).

[42] M. Boguñá, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. Á. Serrano, Network geometry, Nat. Rev. Phys. **3**, 114 (2021).

[43] R. Sidje, Expokit: A software package for computing matrix exponentials, ACM Trans. Math. Softw. **24**, 130 (1998).

[44] In particular, for the BA graphs, the number of nodes is exactly the same, while the number of edges is slightly different, since the generation model only allows for multiples of the number of nodes. The number of edges was then decided by looking at the average degree $\langle k \rangle$ of the Internet network and the rounding to the closest integer to get the multiple of the number of nodes. The number 3 was chosen for all BA graphs since the average degree of the renormalized versions is around 2.5 for all of them. The ER graphs were generated simply by generating a random graph with the same amount of nodes and edges as the corresponding Internet network. When generating ER graphs similar to the Internet replicas, the low average degree gives rise to many disconnected components. We chose to generate graphs with a higher initial number of nodes to offset the removal of the smaller disconnected components, so in these cases the generated graphs are not fully random.

[45] S. Apers, S. Chakraborty, L. Novo, and J. Roland, Quadratic Speedup for Spatial Search by Continuous-Time Quantum Walk, Phys. Rev. Lett. **129**, 160502 (2022).

[46] T. Wong, Spatial search by continuous-time quantum walk with multiple marked vertices, Quantum Inf. Proc. **15** (2016).

[47] A. Kurt, M. A. C. Rossi, and J. Piilo, Efficient quantum transport in a multi-site system combining classical noise and quantum baths, New J. Phys. **22**, 013028 (2020).

[48] D. Abbott, P. C. W. Davies, A. K. Pati, and R. Penrose, *Quantum Aspects of Life* (World Scientific, London, UK, 2008).