Navigability is a robust property*

Dimitris Achlioptas[1] ** and Paris Siminelakis[2] ***

[1] Department of Computer Science
University of California, Santa Cruz
optas@cs.ucsc.edu
[2] Department of Electrical Engineering
Stanford University
psimin@stanford.edu

Abstract. The Small World phenomenon has inspired researchers across a number of fields. A breakthrough in its understanding was made by Kleinberg who introduced Rank Based Augmentation (RBA): add to each vertex independently an arc to a random destination, selected from a carefully crafted probability distribution. Kleinberg proved that RBA makes many networks *navigable*, i.e., it allows greedy routing to successfully deliver messages between any two vertices in a polylogarithmic number of steps. Our goal in this work is to prove that navigability is an inherent, robust property of many random networks. Our framework assigns a cost to each edge and considers the uniform measure over all graphs on n vertices that satisfy a total budget constraint. We show that when the cost function is sufficiently correlated with the underlying geometry of the vertices and for a wide range of budgets, the overwhelming majority of all feasible graphs with the given budget are navigable. We provide a new set of geometric conditions that generalize Kleinberg's set systems as well as a unified analysis of navigability.

* A preliminary version of this work was presented at the 12th Workshop on Algorithms and Models for the Web Graph (WAW 15), Eindhoven, Netherlands, December 2015.

** Research supported by NSF Grant CCF-1514128, an Alfred P. Sloan Fellowship, the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARISTEIA II.

*** Supported in part by an Onassis Foundation Scholarship.

1 Introduction

The Small World phenomenon (SW), popularly known as Six Degrees of Separation [1], refers to the empirical fact that one can connect any two people in the world through short chains of personal acquaintances. Sociologists, mathematicians and more recently computer scientists have undertaken efforts to formalize this phenomenon, provide plausible explanations (mechanisms) of its emergence and finally exploit it in the design of networks and protocols. In light of these efforts the SW phenomenon has been at center stage of the developments in the study of mathematical models of real world networks [2, 3] and of decentralized routing protocols.

The early interpretation of the SW was in terms of the *diameter* of the network, that is, the largest distance between any two individuals. A network (graph) is said to be a "small world", if it's diameter is at most logarithmic in the number of vertices. This is motivated by the fact that any network with constant degree has to have diameter at least logarithmic. Many mathematical models (typically involving some amount of randomness) have been shown to satisfy this requirement and underlying proofs of such results lies *expansion*, i.e. that there is non-trivial density of edges between sets of vertices.

The first concrete empirical evidence that gave credence to Small World phenomenon came from Sociologist Stanley Milgram in one of his famous experiments [4]. In this experiment participants were given a letter addressed to a certain person specified by his name, location and occupation. They were asked to forward the letter (along with instructions) to an acquaintance of theirs that is most likely to know the target. A large fraction of the letters sent did arrive to the target and on average each letter had to go through 5 or 6 people before it reached its target. Since then, more modern forms of the same experiment have been repeated [5] and in our era we have the phenomenon of 4 degrees of separation [6]. However, it was not until Kleinberg's work that Milgram's experiment and the Small World phenomenon in general was fully appreciated and given its modern interpretation.

Jon Kleinberg realized that Milgram's experiment indicated not only that short chains of acquaintances exist, but also that they can be found in a decentralized manner using only some global information of the target (occupation and location). In his groundbreaking work [7, 8], Kleinberg formulated mathematically the property of finding short-paths in a decentralized manner as *navigability*. Since then, the concept of navigability has also found applications in the design of peer-to-peer networks [9, 10], data-structures [11, 12] and search algorithms [13–15].

In this paper we revisit this most recent algorithmic interpretation of the SW phenomenon, that of Navigability. We propose a unified and simplified framework that separates the geometric and probabilistic requirements for navigability and offers insights on its emergence in the real world.

1.1 Navigability and Rank Based Augmentation

Navigability is a property of networks that allows efficient decentralized communication. Key to decentralization is shared knowledge in the form of geometry. In Milgram's experiment individuals were embedded in a occupation-location space and distance was defined based on this information. Here, more generally, we will assume shared knowledge of a (distance) function on pairs of vertices (not necessarily satisfying the triangle inequality) that can guide the search through an otherwise unstructured network.

Definition 1. *A* geometry (V, d) *consists of a set of vertices* V *and a distance function* $d : V \times V \to \mathbb{R}_+$*, where* $d(x, y) \geq 0$*,* $d(x, y) = 0$ *iff* $x = y$*, and* $d(x, y) = d(y, x)$*, i.e., the function* d *is a semi-metric.*

Given a graph $G(V, E)$ on a geometry (V, d), a *decentralized search algorithm* is any algorithm that given a target vertex t and current vertex v selects the next edge $\{v, u\} \in E$ to cross by only considering the distance $d(u, t)$ of each neighbor u of v to the target t. Navigability concerns the performance of decentralized algorithms in finding paths in a graph guided by a distance function.

Definition 2. *A graph* $G(V, E)$ *on geometry* (V, d) *is* d-navigable *if there exists a decentralized search algorithm that given any two vertices* $s, t \in V$ *will find a path from* s *to* t *of length* $O\left(\mathrm{poly}(\log n)\right)$.

In the above definition, the requirement that paths have polylogarithmic length essentially means that decentralized search incurs cost at most polynomial larger than if we had full information, i.e., we could actually compute the shortest path between any pair of vertices.

In his original work on navigability [7], Kleinberg showed that if G is the two dimensional grid (with d being the ℓ_1 distance on the grid) then adding a single random edge independently to each $v \in V$ according to some distribution results in a navigable graph. The distribution for selecting the other endpoint u of each added edge is crucial. Indeed, if it can only depend on $d(v, u)$ and distinct vertices are augmented independently, Kleinberg showed that there is a *unique* suitable distribution, the one in which the probability is proportional to $d(v, u)^{-2}$ (and, more generally, $d(v, u)^{-r}$ for r-dimensional lattices).

The underlying principle behind Kleinberg's augmentation scheme has by now become known as *Rank Based Augmentation* (RBA) [8, 16].

Definition 3. *Given a geometry* (V, d) *and two vertices* $v, u \in V$*, let* $\mathrm{rank}_v(u)$ *be the number of vertices that have distance at most* $d(u, v)$ *from* v*. Rank Based Augmentation (RBA) refers to a probability distribution* μ *over edges such that the probability of an edge* (v, u) *being included is proportional to the rank of* u *with respect to* v*:*

$$\mathbb{P}_\mu((v, u) \in E) \propto \frac{1}{\mathrm{rank}_v(u)} \tag{1}$$

The intuition behind RBA is that navigability is attained because the added edges provide connectivity *across all distance scales* and crucially *in all possible directions*, i.e., target vertices. The first property is not hard to achieve. For instance, if all distances to v are unique, then $rank_v$ actually induces a permutation of the vertices. As such, if we pick any constant $\gamma > 1$ and partition vertices depending whether their rank is between $[\gamma^{k-1}, \gamma^k)$ for $k \geq 1$, we see that for each "scale" k the expected number of edges to vertices in that scale is proportional to $1/\log n$. However, even though we have a non-trivially probability of finding an edge to cross the right distance scale, we still need the edge to bring us closer to the target by a constant factor if we are ever going to reach the destination in a poly-logarithmic number of steps. For that to happen, we roughly need *a constant fraction* of the vertices in that scale to *reduce the distance* to the target by a constant factor. Both conditions are effortlessly true in regular lattices of finite dimension as there are a constant number (though exponential in the dimension) of possible "directions" and the rank function grows only polynomially with the distance of the vertex.

An important facet of Kleinberg's work is that navigability is achieved by only adding a constant number of edges per node. Observe, that trivially the complete graph is navigable, so the question of navigability is of special interest when the graph is sparse. From a practical perspective, the amount of links that a node has translates into the number of bits that each node has to store in order to be able to run the decentralized search algorithm [12].

In subsequent work [8], Kleinberg generalized his results beyond lattices and showed that the geometric conditions needed for RBA to render a network navigable are satisfied by the geometries induced by a family of *set-systems*, whose definition (Section 6) corresponds roughly to an abstraction of the above considerations. To make this intuition more concrete consider a regular lattice and for all $k \geq 0$ define balls of radius γ^k centered at each vertex, the resulting collection of sets of vertices would form a valid set-system. The motivation for studying set-systems has also a conceptual importance as in the real world people belong in groups and typically the distance between two individuals can be defined as the size of the smallest group that both belong to. RBA in this context expresses the idea that the probability of two people knowing each other is inversely proportional to their distance in this *group space*. Besides, its conceptual appeal it is also known [12] that set-systems encompass geometries defined by metric spaces of bounded doubling dimension. Thus, set-systems are a very general and conceptually appealing abstraction of the geometric requirements of navigability and along with RBA go far in providing a solid mathematical explanation of the Small world phenomenon.

The remarkable success of RBA in conferring navigability rests crucially on its *perfect adaptation* to the underlying geometry. This adaptation, though, requires not only all vertices to behave identically and independently, but also a very specific, indeed unique, functional form for the probability distribution of edge formation. This exact fine tuning renders RBA unnatural, severely undermining its plausibility.

1.2 Our contribution

Our goal in this paper is to demonstrate that navigability is in fact a robust property of networks that does not require independence, coordination or fine tuning, but rather arises naturally under the right geometric and "economical" conditions. We achieve this by first presenting a set of sufficient conditions for navigability (unifying previous approaches) and then showing that they can be easily satisfied under natural models of generating the geometry and the graph. Roughly speaking, we isolate three ingredients that suffice for navigability on a geometry (V, d):

- *Geometric requirements:* some degree of *coherence* of the semi-metric d (similar to Kleinberg's set systems).
- *Local progress:* a *substrate* of connections between nearby points in V compatible with the semi-metric, making it impossible to get stuck locally.
- *Probabilistic requirements:* sufficient edge density across all distance scales and "directions" (target vertices).

The first two ingredients are generalizations of existing work and, as we will see, fully compatible with RBA. The third ingredient is also motivated by the RBA viewpoint, but we will prove that it can be achieved in far more-light handed, and thus natural, manner than RBA. Moreover, in the course of doing so, we will give RBA a very natural *economic* interpretation, as the distribution on edges arising when the cost of each edge is the *cost of indexing* among neighbours at the same distance scale.

As mentioned, at the foundation of navigability lies shared knowledge in the form of geometry. At the same time, geometry imposes *global* constraints on the set of feasible networks. Most obviously, in a physical network where edges (wire, roads) correspond to a resource (copper, concrete) there is typically an upper bound on how much can be invested to create the network. More generally, cost may represent a number of different notions (e.g., class membership) that distinguish between edges.

We will formalize the above intuition by (i) allowing edges to have costs given by an arbitrary function c on the edges, and (ii) taking as input an upper bound on the *total* cost of feasible graphs, i.e., a budget B. For instance, the cost of each edge may express the propensity (low cost) or reluctance (high cost) of two individuals interacting. In that case, an upper bound on the total cost expresses that feasible social interaction graphs are selected to not cause too much discomfort to the participating individuals.

Geometry, either of physical or of "concept" space, is an extremely natural backdrop for network formation that brings along both notions of cost and budget. In general, we expect that cost will correlate with geometry and that the budget, for any given cost function, will be such that the average degree of the network will be small (a property of nearly all real networks). Within these highly generic considerations, given a geometry, a cost function, and a budget we would like to study the set of all graphs satisfying the budget constraint, i.e., the set of all *feasible* graphs, and answer the following question: is it the case that the *overwhelming majority* is navigable?

This viewpoint departs from previous work where the aim was to provide a network creation *mechanism* that would lead to navigable graphs. Our viewpoint is motivated by the fact that, in reality, navigability is almost never an explicit goal of the network formation process yet, at the same time, navigability appears to be prevalent in a wide variety of settings. To demonstrate the power of our approach at this point we give only an informal a flavor of our results and postpone the precise statements for the main part of the paper.

Theorem 1 (Informal). *Given a random set of n points on the square $[0, \sqrt{n}]^2$ endowed with some local connectivity (e.g. edges between nearest neighbours), define the cost of an edge to be the logarithm of ℓ_∞-distance between the points. If one picks a graph uniformly at random among all graphs with total cost at most $n \cdot \mathrm{poly}(\log n)$, then with high probability the graph will be navigable.*

Though the precise quantitative dependences are omitted in this theorem, the theorem captures the main point that we want to make with this work, namely, that navigability is a *robust property*, in the sense that it emerges almost inevitably under very different ways of defining the graph as long as the geometry permits it (e.g. set systems, random points) and there is some local connectivity (substrate).

1.3 Related work

The Small World phenomenon and Navigability are by now well studied topics. The review by Kleinberg [17] provides an excellent introduction and covers almost all of the earlier results up to 2006. Here, we would like to highlight three major questions that were left open and the work that has been made towards their resolution the past years.

Robustness of Navigbability. Kleinberg's work identified a specific graph augmentation mechanism (RBA) that renders networks navigable, thus providing a plausible explanation of the Small World Phenomenon. However, he provided no indication about how such a distribution might come about in the real world or whether networks even approximately exhibit such behavior. Early empirical studies [16, 18] of real world social networks have shown that although there are deviations from the idealized behavior that Kleinberg assumed (RBA), the distribution of edges scales proportionally with $\mathrm{rank}_u(v)^{-\beta}$, albeit with $\beta < 1$. This was perplexing as Kleinberg had shown that $\beta = 1$ is the only exponent for which the network is navigable. Previously researchers attempted to reconcile this by suggesting that the observed deviations from the exponent are due to finite size (n finite) effects [19]. In this paper, we show that such a discrepancy can be to attributed to the network having super-constant average degree (see proof of Theorem 5), that underlies Kleinberg's results.

Evolution of Navigability. Another intriguing question that was inspired by Kleinberg's work was to exhibit a network creation mechanism that might lead to a distribution similar to RBA. The first work to provide an indication about how that might happen was carried out by Clauset and Moore [19]. They considered a network rewiring process, that samples source-target pairs x, y according to some distribution $Q(x, y)$ and performs greedy routing for at most T steps. If the target has not been reached until time T then one of the out-going links of x are rewired to point to y. They showed empirically that this process seems to converge to RBA but did not provide any rigorous evidence of that fact. Later, Sandberg [20] proved that a very similar process does converge to a unique stationary distribution over links but again only provided experimental evidence that the networks created by this mechanism are navigable. The first rigorous proof that a mechanism is able to reproduce Kleinberg's distribution came from Chaintreau et al. [21]. The authors considered a setting where nodes are embedded in a d-dimensional regular lattice and each node independently picks an outgoing link by having the end of the link performing a random walk with some time varying restart probability starting at the source. When the restart probability is approximately harmonic with the number of steps taken the stationary distribution of the link approximately matches RBA. A different approach to providing an explanation for the SW phenomenon was put forward by Gulyás et al. [22]. Instead of assuming that nodes are embedded in a euclidean space as previous works, they consider that nodes are embedded in the hyperbolic plane. The important fact about hyperbolic plane is that space expands exponentially with the distance away from the center. This behavior is very similar to what happens on trees, and indeed hyperbolic space and trees have very close connections [23]. The authors exploit this phenomenon and under a game-theoretic framework, where each node tries to connect to the closer/smallest set of neighbors that would allow them to send messages to any node, show semi-rigorously that the graphs at equilibrium are navigable.

Searchability of arbitrary Networks. In achieving Navigability, an underlying geometry is indispensable. Kleinberg originally considered the 2-dimensional grid and showed that indeed navigability is achievable through RBA. A natural question that was raised by Duchon et al. [24] is whether any any graph $G(V, E_0)$ could become navigable after being augmented randomly with long range edges. They proved that as ufficient condition for a graph to become navigable through RBA is to have some bounded growth. In the same direction, other authors have been looking at other general sufficient conditions on the underlying graph that enable navigability through augmentation. Fraigniaud [25] showed that this is possible for bounded-treewidth graphs, and Abraham et al. [26] showed it, further, for minor-free graphs. The work of Slivkins [12] proved that augmentation always works if the doubling dimension of the graph is $O(\log \log n)$ and Fraigniaud et al.[27] proved that this is actually best possible. Since [27] research in this topic has turned to proving upper bounds for the performance of decentralized routing algorithms for arbitrary graphs. In that direction, Peleg first proved an $O(\sqrt{n})$ upper bound which was consequently improved to $O(n^{1/3})$ (up to poly-logarithmic factors) by Fraigniaud [28]. The best upper bound to date is $O(2^{(\log n)^{1/2+o(1)}})$ due to the work of Fraigniaud and Giakkoupis [29] almost matching a lower bound of $\Omega(2^{\sqrt{\log n}})$ for "monotone" decentralized algorithms by Fraigniaud et al. [27].

2 Our Results

Notation. Throughout the paper the set of vertices V is considered to be fixed and large, i.e., $n := |V|$ is finite but large. Any asymptotic notation, e.g. $f(|V|) = O(g(|V|))$ should be interpreted as comparing two functions of $|V|$ (eq. n) and only means that there are some constants independent of $|V|$ such that the corresponding inequalities hold, e.g. $f(|V|) \leq Cg(|V|)$. In particular, all the statements in this paper apply to a single geometry or graph rather than a sequence. Lastly, to make the presentation more readable we will often say that a property $\mathcal{A}$ holds *with high probability (w.h.p)* to indicate that $\mathbb{P}(\mathcal{A}) \geq 1 - o(1)$.

2.1 Geometric requirements and a unifying framework for RBA

We start by introducing the geometric requirements for navigability through the notion of *coherence*[3], that comes with an associated *scale factor* $\gamma > 1$ and a resolution $\kappa \geq 1$.

[3] Note that coherence here is unrealted to the corresponding notion for matrices used in Compressed Sensing.

Specifically, given a geometry (V, d) we will refer to the vertices whose distance from a given vertex $v \in V$ lie in the interval $(\gamma^{k-1}, \gamma^k]$ as the vertices in the k-th (distance) γ-scale from v and denote their number as $P_k(v)$. Additionally for any two vertices $v \neq t \in V$ we will use k_{vt} to denote the integer k such that $d(v,t) \in (\gamma^{k-1}, \gamma^k]$. For a fixed $\lambda < 1$ and any target vertex $t \neq v$, we will say that a vertex u is t-helpful to v if $d(v,u) \leq \gamma^{k_{vt}}$ (u is within the same or lower γ-scale as t from v), and $d(u,t) < \lambda d(v,t)$ (reduces the distance to t by a constant). We denote the set of t-helpful nodes of v by $D_\lambda(v,t)$.

Definition 4. *Fix $\gamma > 1$ where $K = \lceil \log_\gamma |V| \rceil$. A geometry (V,d) is (γ, κ)-*coherent *for $\kappa \in [K]$ if:*

(H1) Bounded Growth: $\exists A > 1, \alpha \in (0,1)$ *such that*

$$P_k(v) \in \gamma^k[\alpha, A], \textit{ for all } v \in V \textit{ and } k \in [\kappa, K].$$

(H2) Isotropy: $\exists \phi > 0, \lambda \in (0,1)$ *such that*

$$|D_\lambda(v,t)| \geq \phi\gamma^{k_{vt}}, \textit{ for all } v \neq t \in V \textit{ such that } k_{vt} \geq \kappa.$$

The two conditions above endow the, otherwise arbitrary, semi-metric d with sufficient regularity and consistency to guide the search. Although our definition of coherence is far more general, in order to convey intuition about the two conditions, think for a moment of V as a set of points in Euclidean space. The first condition guarantees that there are no "holes" when one looks at an appropriate resolution $\kappa \geq 1$, as the variance in the density of points is bounded in every distance scale. In particular, it implies that the largest "distance" is proportional to $\gamma^k \sim |V|$. The second condition guarantees that around any vertex v the density of points does not change much depending on the direction (target vertex t) and distance scale. A concrete example is presented in Section 2.4. Besides these two conditions, we make *no further* assumptions on d and, in particular, we do *not* assume the triangle inequality. For convenience, we will usually omit κ and simply call a geometry γ-coherent when $\kappa = 1$.

Coherent geometries allow us to provide a unified treatment of navigability since they encompass finite-dimensional lattices, hierarchical models, any vertex transitive graph with bounded doubling dimension and more generally as we show Kleinberg's set systems.

Theorem 2. *Every set system satisfying the conditions of [8] is a γ-coherent geometry for some explicit $\gamma > 1$.*

Our second requirement is to assume the existence of a substrate, that implies that greedy routing will not get trivially stuck, i.e., that we can always move towards the target even incrementally.

Substrate. *A set of edges E_0 forms a* substrate *for a geometry (V,d), if for every $(s,t) \in V \times V$ with $s \neq t$, there is at least one vertex v such that $\{s,v\} \in E_0$ and $d(v,t) \leq d(s,t) - 1$. If there are multiple such vertices, we distinguish one arbitrarily and call it the* local t-connection *of s. A path starting from s and ending to t using only local t-connections is called a local (s,t)-path.*

In the graph augmentation setting this was given by the fact that the initial set of edges formed a known connected graph, while in Kleinberg's work on set systems it was circumvented by making the vertex degrees $\Theta(\log^2 n)$, so that the probability of ever being stuck at a vertex is polynomially small. We chose to use the notion of a substrate to encompass the graph augmentation setting but also generalize it since the semi-metric d is only *locally consistent* with the substrate. We show that those two requirements are sufficient for RBA to create a navigable graph.

Theorem 3. *Let (V, d) be any (γ, κ)-coherent geometry and let E_0 be any substrate for it. If E_d is the (random) set of edges obtained by applying RBA to (V, d), then the graph $G(V, E_0 \cup E_d)$ is d-navigable w.h.p.*

Theorem 3 subsumes and unifies a number of previous positive results on RBA-induced navigability. Our main contribution, though, lies in showing that given a substrate and coherence, navigability can emerge without any coordination or independence, merely from the alignment of cost and geometry.

2.2 Navigability from organic growth

As mentioned earlier, the success of RBA stems from the fact that the edge-creation mechanism is *perfectly* adapted to the underlying geometry so as to induce navigability. In contrast, we will not specify any edge-creation mechanism, but rather consider the set of *all* graphs feasible with a given budget. Our requirement is merely that the cost function is *informed* by the geometry, in the following sense.

γ-consistency. *Given a γ-coherent geometry (V, d), a cost function $c : V \times V \to \mathbb{R}$ is* γ-consistent *if c takes the same value c_k for every edge $\{u, v\}$ such that $d(u, v) \in (\gamma^{k-1}, \gamma^k]$.*

In other words, γ-consistency means that the partition of edges according to cost is a coarsening of the partition of the edges by γ-scale. Note that beyond γ-consistence we do not impose *any* constraint on the values $\{c_k\}$, not even a rudimentary one such as being increasing in k. In fact, even the γ-consistency requirement can be weakened significantly, as long as some correlation between the two partitions remains, but it is technically much simpler to assume γ-consistency as it greatly simplifies the exposition. One can think of consistency as limited sensitivity with respect to distance. As an example, it means that making friends with the people next door might be more likely than making friends with other people on the same floor, and that making friends with people on the same floor is more likely than making friends with people in a different floor, but it does not really matter which floor.

Cost-geometries. *We say that $\Gamma = \Gamma(V, d, c)$ is a* coherent cost-geometry *if there exists $\gamma > 1$ such that (V, d) is a γ-coherent geometry and c is γ-consistent cost function.*

We are now in a position to state the set of feasible graphs that we consider.

Random graphs of bounded cost. *Given a coherent cost-geometry $\Gamma(V, d, c)$ and a real number $B \geq 0$, let $G_\Gamma(B) = \{E \subseteq V \times V : \frac{1}{n} \sum_{e \in E} c(e) \leq B\}$, i.e., $G_\Gamma(B)$ is the set of all graphs (edge sets) on V with total cost at most Bn. A uniformly random element of $G_\Gamma(B)$ will be denoted as $E_\Gamma = E_\Gamma(B)$.*

Obtaining bounds on the probability that a uniformly random element out of $G_\Gamma(B)$ is navigable, is an intuitive and technically enabling way to obtain bounds on the fraction of feasible graphs that are navigable. Our main result is the following general theorem.

Theorem 4. *For every coherent cost-geometry $\Gamma(V, d, c)$ with substrate E_0, there exist numbers $B^{\pm}$ such that if E_Γ is a uniformly random element of $G_\Gamma(B)$ then:*

– *For all $B \leq B^+$, w.h.p. $|E_\Gamma| = O(n \cdot \mathrm{poly}(\log n))$.* *(Sparsity)*
– *For all $B \geq B^-$, w.h.p. the graph $G(V, E_0 \cup E_\Gamma)$ is d-navigable.* *(Navigability)*

In the proving this theorem, the fact that we were able to get a close-form expression for the probabilities of each edge, exponentially decreasing in its cost (Lemma 4), was instrumental. Note that Theorem 4 shows that navigability arises eventually, i.e., for all $B \geq B^-$, without *any* further assumptions on the cost function or geometry. The caveat, if we think of B as increasing from 0, is that by the time there are enough edges across all distance scales, i.e., $B \geq B^-$, the total number of edges may be much greater than linear. This is because for an arbitrary cost structure $\{c_k\}$, by the time the "slowest growing" distance scale has the required number of edges, the other scales may be replete with edges (due to the exponential dependence), possibly many more than $\Omega(n/\mathrm{poly}\log n)$ that are required in order for greedy decentralized search to have probability of crossing a distance scale at least inverse poly-logarithmic. This is reflected in the ordering between B^- and B^+ that determines whether the sparsity and navigability regimes are overlapping. In particular, we would like $B^- \leq B^+$ and, ideally, the ratio $R = B^+/B^- > 0$ to be large. Whether this is the case or not depends precisely on the degree of adaptation of the cost-structure to the geometry, as we discuss next.

2.3 Navigability as a reflection of the cost of indexing

Recall that for every vertex v in a γ-coherent geometry and for every distance scale $k \in [K]$, the number of vertices whose distance from v is in the k-th γ-distance scale is $P_k(v) = \Theta(\gamma^k)$. Let $p_k := \frac{1}{|V|}\sum_{v \in V} P_k(v)$ be the average number of vertices at distance scale k from a random vertex. A coherent-cost geometry is parametrized by the numbers $\{p_k\}$ and the values of the cost function $\{c_k\}$.

We will now exhibit a class of cost functions that (i) have an intuitive interpretation as the average *cost of indexing*, (ii) achieve a ratio $R = B^+/B^- > 0$ that *grows* with n, i.e., a very wide range of budgets for which we have both navigability and sparsity, and (iii) recover RBA as a special case corresponding to a particular budget choice. To motivate the cost of indexing consider a vertex v that needs to forward a message to a neighbor u at the k-th distance scale. To do so, v needs to distinguish u among all other $P_k(v)$ vertices in the k-th distance scale, i.e., v needs to be able to *index* into that scale. Storing the unique ID of u among the other members of its equivalence class (in the eyes of v) has a cost of $\Theta(\log_2 P_k(v)) = \Theta(\log p_k) = \Theta(k)$ bits. Motivated by this we consider cost functions where for some $\beta > 0$,

$$c_k^* = \frac{1}{\beta}\log p_k \quad . \tag{2}$$

We also assume that c_k^* is non-decreasing. This assumption is not needed and only changes some absolute constants but deals with some tedious technical issues.

Theorem 5. *For any coherent cost-geometry $\Gamma(V, d, c^*)$, there exist $B^{\pm}$ such that :*

(a) $B^+/B^- = \omega(\text{poly}\log n)$.
(b) For all $B \in [B^-,\ B^+]$, w.h.p. $|E_\Gamma(B)| = O(n\ \ \text{poly}\log n))$ and the graph $G\left(V, E_0 \cup E_\Gamma(B)\right)$ is d-navigable.
(c) There exists $B_\beta \in [B^-,\ B^+]$ such that $\mathbb{P}\left[(u,v) \in E_\Gamma(B_\beta)\right] = \Theta\left(\frac{1}{\text{rank}_u(v)}\right)$ for all edges $u, v \in V$ such that $k_{uv} \geq \kappa$.

This result shows that Navigability and sparsity are both compatible for a large range of values of B and hence that Navigability is indeed a *robust property* of networks. Since, our results are rather abstract and general, we show here how our theorems apply in a simple setting where both the geometry and the graph itself is generated randomly.

2.4 Application: Random Cube Model

Let $\mathcal{X}_n^t = [0, R_n]^t$ be the t-dimensional cube of length $R_n = n^{1/t}$ and $\mathcal{B}(\mathcal{X}_n^t)$ be its Borel σ-algebra. A probability measure $\mu : \mathcal{B}(\mathcal{X}_n^t) \to [0,1]$ with density $\frac{d\mu}{d\lambda}(x) = p(x)$ satisfies the *Random Cube Model* $C(n, t, \Delta)$ iff $\max\limits_{x,y \in \mathcal{X}_n^t} \frac{p(x)}{p(y)} \leq \Delta$.

Theorem 6. *Let V be a set of points sampled from a probability measure satisfying the Random cube model $C(n, t, \Delta)$, then the semi-metric $d(u,v) := \|u - v\|_\infty^t$ defines with high probability a coherent geometry at resolution $\kappa = \log\log n$ with parameters $(\gamma, \alpha, A, \phi) = (2^D, \frac{1-\epsilon}{\Delta} 2^{-t}(1 - 2^{-t}), \Delta(1+\epsilon), \frac{1-\epsilon}{\Delta} 2^{-2t})$.*

Proof. We first show the bounded growth properties of coherent geometries for scales $k \geq \kappa$. We need to get for each $u \in V$ upper and lower bounds on the number of vertices v that $d(u,v) \in [\gamma^{k-1}, \gamma^k]$ or equivalently (by definition of the semi-metric) $\|u - v\|_\infty \in [2^{k-1}, 2^k]$. Letting $B_k(u) := \sum_{\ell=1}^{k} P_k(u)$ we have that for integer $r \geq 0$:

$$P_{\kappa+r}(u) = B_{\kappa+r}(u) - B_{\kappa+r-1}(u)$$

To control this quantity, we partition the cube $\mathcal{X}_n^t = [0, n^{1/t}]^t$ into $N_0 = \frac{n}{c\log n}$ boxes of size $r_0 = (c\log n)^{1/t}$ for some constant $c > 0$ to be selected later. Using concentration of measure We first obtain bounds on N_b the number of points in each box $b \in [N_0]$.

Lemma 1. *Let $\epsilon > 0, \delta > 0$ and $c > 3(1+\delta)\Delta\epsilon^{-2}$, with probability at least $1 - n^{-\delta}$ we have that for all $b \in [N_0]$, $N_b \in \left[\frac{1-\epsilon}{\Delta}, (1+\epsilon)\Delta\right] c\log n$.*

Proof. By standard Chernoff bounds for i.i.d Bernouli random variables, the number of points N_b that fall into any box $b \in [N_0]$ satisfies

$$\mathbb{P}_\mu\left(N_b \geq (1+\epsilon)\Delta c\log n\right) \leq \exp\left(-\frac{\epsilon^2}{3} c\log n\right)$$

$$\mathbb{P}_\mu\left(N_b \leq (1-\epsilon)\frac{c\log n}{\Delta}\right) \leq \exp\left(-\frac{\epsilon^2}{3\Delta}c\log n\right)$$

Thus, we have that for all $b \in [N_0]$:

$$\mathbb{P}_\mu\left(\bigcap_{b\in[N_0]}\left(N_b \in \left[\frac{1-\epsilon}{\Delta},(1+\epsilon)\Delta\right]c\log n\right)\right) \geq 1-2\exp\left(-\frac{\epsilon^2}{3}\frac{c\log n}{\Delta}+\log N_0\right)$$

For $c(\epsilon,\delta,\Delta) = 2^{\lceil \log_2(3(1+\delta)\Delta)\epsilon^{-2}\rceil}$ we have that the event holds with probability at least $n^{-\delta}$. □

Next, we obtain upper and lower bounds on the number of boxes that are within the specified radius. To obtain a lower bound imagine a node that is located at a vertex of the cube $\mathcal{X}_n^t$. For this node there are $\frac{1}{2^t}2^{tr}$ boxes within ℓ_∞ distance $2^{\kappa+r}$. Thus, we have that for any vertex there are at least $\frac{1}{2^t}2^{tr} - \frac{1}{2^t}2^{t(r-1)}$ boxes. By Lemma 1 each box has at least $\frac{1-\epsilon}{\Delta}\gamma^\kappa$ vertices. This gives us the following lower bound:

$$P_{\kappa+r}(u) \geq \frac{1-\epsilon}{\Delta}2^{-t}(1-2^{-t})2^{rt}\gamma^\kappa = \left(\frac{1-\epsilon}{\Delta}2^{-t}(1-2^{-t})\right)\gamma^{(\kappa+r)}$$

To obtain an upper bound we follow the same strategy but identify a vertex in the center of the cube. For such a vertex, there are at most $2^{tr}(1-2^{-t})$ boxes of length r_0 fully contained in the specified interval of distances. Again by Lemma 1 this gives the following upper bound

$$P_{\kappa+r}(u) \leq (1+\epsilon)\Delta 2^{rt}\gamma^\kappa = [(1+\epsilon)\Delta]\,\gamma^{(\kappa+r)}$$

Finally, to prove the coherence property of the geometry, we count again the number of boxes that are both contained in the range $[2^{\kappa+r-1}, 2^{\kappa+r}]$ from u and $[0, 2^{\kappa+r-1}]$ from v. The number of such boxes is at least $2^{-t}2^{t(r-1)}$ and thus we get that the number of good vertices is

$$|D_\lambda(u,v)| \geq 2^{-2t}\frac{1-\epsilon}{\Delta}2^{tr}\gamma^\kappa = \left(2^{-2t}\frac{1-\epsilon}{\Delta}\right)\gamma^{\kappa+r}$$

Since, the event described in Lemma 1 holds with high probability we know that the geometry is coherent with the same probability. □

Having proven that the geometry defined by the random cube model is coherent at an appropriate scale, constructing a graph that is navigable becomes an easy matter. Assuming the existence of an appropriate substrate, for instance one defined by connecting k-nearest neighbours with edges, we consider two ways of defining the graph:

1. *Distance based augmentation:* we may add a constant number of shortcuts with probability proportional to $\frac{1}{1+\|u-v\|_\infty^t}$. This is very similar to RBA augmentation, in fact Lemma 1 shows that indeed $\|u-v\|_\infty^t$ is within constant factors of $\mathrm{rank}_u(v)$ and by Theorem 3 the resulting graph will be navigable. Alternatively, if we assume that edges are added independently of each other, as we will see we may invoke Lemma 2 to get the same conclusion.

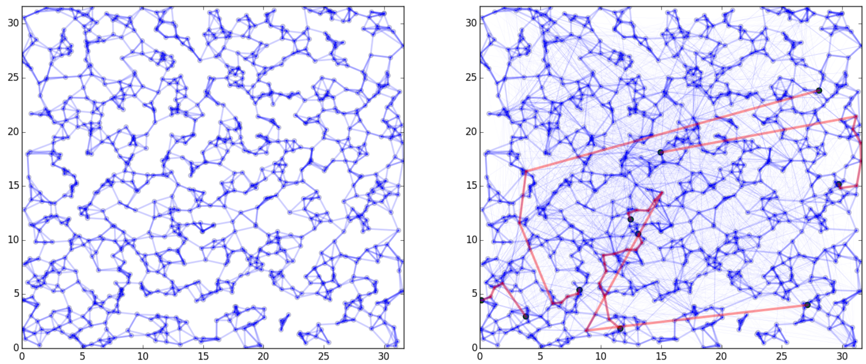

Fig. 1. Left: Geometry and substrate (4-nearest neighbors) by sampling points $n = 1000$ from the density $p(x_1, x_2) = 1 + 0.5 \cdot \sin(\frac{6\pi x_1}{\sqrt{n}}) \cdot \sin(\frac{12\pi x_2}{\sqrt{n}})$ for which $\Delta = 3$. Right: resulting graph after distance based augmentation, where we also include a number of sample searches.

2. *Random Graphs of Bounded Cost:* a different way to define the graph would be to set for each edge (u, v) a cost $c_{uv} := \lceil \log_\gamma \|u - v\|_\infty \rceil$ and pick a graph uniformly at random subject to the total budget constraint of $B = \Omega(n \mathrm{poly}(\log n))$. It is easy to see that in this case the conditions of Theorem 5 are satisfied and the resulting graph would be navigable.

The ease with which we can show that the above construction produces navigable graphs is indicative of the generality and power of our results. To summarize, our work identifies three components for navigability:

(i) *Geometry:* we require bounded (in reality at most poly-logarithmic in n) fluctuations of the density of points above some scale κ (resolution). Mathematically, we capture this requirement with the notion of coherent geometries and have showed that natural candidates as set systems and random geometries satisfy our definition.
(ii) *Augmentation:* any probability measure over edges that is uniformly rich, i.e., it recovers up to poly-logarithmic factors RBA, is sufficient, even if edges are not independent. We give a concrete example of Random Graphs of Bounded cost that is a minimal naturalistic model of graph generation.
(iii) *Local Connectivity:* allows to effectively care for distances from a resolution and above, as well as, deals with the trivial issue of getting stuck locally.

2.5 Outline of the rest of the paper

This concludes the presentation of our results and the rest of the paper is devoted to providing their proofs. In Section 3, we present a concise framework that allows one to prove that a graph on a coherent geometry is navigable . Then in Sections 4 and 5, we show respectively how one can analyze a random graph of bounded cost using a coupling with an explicitly constructed product measure and under what conditions such a graph is navigable. Section 6, then proves that Kleinberg's set systems are coherent geometries. Finally, in Section 7, we include for completeness a proof that classical RBA makes graphs on coherent geometries navigable.

3 Navigability via reducibility and uniform richness

In this section we present structural results about navigability on coherent geometries that allow us to reduce navigability to a "richness" property of the probability measure on the non-substrate edges. We first define a sufficient deterministic property for navigability.

Reducibility. *If $G(V, E)$ is a graph on a coherent geometry (V, d) with substrate $E_0 \subseteq E$, we will say that $(s, t) \in V \times V$ is p-*reducible *if there is $C > 0$ such that among the first $C(\log |V|)^p$ vertices of the local (s, t)-path there is at least one vertex u such that $(u, v) \in E$ and $d(v, t) \le \lambda d(s, t)$. If every pair $(s, t) \in V \times V$ is p-reducible, we will say that G is p-*reducible.

Reducibility expresses that as we move along the local path we never have to wait too long in order to encounter an edge that reduces the remaining distance by a constant factor. The motivation for introducing reducibility is that it allows us to separate the construction of the random graph from the analysis of the algorithm. Reducibility implies navigability in a straightforward manner.

Proposition 1. *If G is p-reducible, then greedy routing on G takes $O(\log^{1+p} n)$ steps.*

Proof. Given any arbitrary pair of vertices (s, t) with distance at most n, the reducibility property of G guarantees us that after at most $C \log^p n$ steps we will obtain a new pair (s', t) with distance reduced by a constant factor. Since, the new pair is also p-reducible, we can repeat the process until we reduce the distance again by a constant. After at most $\log_{1/\lambda} n$ iterations we will reach the target. Since, the pairs were arbitrary, this holds for all pairs and thus the graph is navigable in 1+$C(\log n)^{1+p}$ steps. □

Reducibility is easiest to establish for random graphs whose edges are included independently, for concreteness we provide the following definition.

Product measure. *Given a set of vertices V with $|V| = n$, let $\mathcal{G}_n$ denote the set of all $2^{\binom{n}{2}}$ possible graphs (edge-sets) on n vertices. A* product measure *on $\mathcal{G}_n$ is specified succinctly by a symmetric matrix $\mathbf{Q} \in [0, 1]^{n \times n}$ of probabilities where $Q_{ii} = 0$ for $i \in [n]$. We denote by $G(n, \mathbf{Q})$ the distribution over $\mathcal{G}_n$ in which possible edge $\{i, j\}$ is included independently with probability $Q_{ij} = Q_{ji}$.*

We next introduce the probabilistic requirement that suffices for reducibility.

Uniform richness. *Let (V, d) be a γ-coherent geometry with parameter $\alpha \in (0, 1)$ (see H1). For $\theta \ge 1$, a product measure $G(n, \mathbf{Q})$ is θ-*uniformly rich *for (V, d) if there is a constant $M > 0$ such that for every $k \ge k_\theta$, for every pair (i, j) with $d(i, j) \in (\gamma^{k-1}, \gamma^k]$ we have:*

$$Q_{ij} \ge \frac{1}{M \log^\theta n} \frac{1}{\gamma^k}$$

where $k_\theta := \frac{\theta \log\log n - \log\alpha}{\log\gamma}$.

The number k_θ simply denotes the distance scale that would take $O(\log^\theta n)$ "slow" steps to cross, and is used to impose density requirements only for non-trivial distance scales as opposed to all scales. As we show next, uniform richness is a sufficient condition for reducibility on coherent geometries.

Lemma 2. *If* (V, d) *is a* γ*-coherent geometry with substrate* E_0 *and* E_q *is sampled from a* θ*-uniformly rich product measure* $G(n, \mathbf{Q})$*, then* $G(V, E_0 \cup E_q)$ *is* $(\theta+1)$*-reducible with probability at least* $1 - n^{-5}$.

Proof. To prove that the graph is $(\theta+1)$-reducible we will (i) prove that the event B_{st} that any fixed source-destination pair (s,t) is not $(\theta+1)$-reducible has very small probability under $G(n, \mathbf{Q})$, and (ii) use union-bound to argue that the probability that any pair is not $(\theta+1)$-reducible is small as well. To simplify the proof, we first distinguish between pairs (s,t) where within the first $C\log^{\theta+1}(n)$ steps of the t-local path there is a vertex with distance smaller than $d(s,t)$ by a constant factor $\lambda < 1$ and where there is no such vertex. Pairs (s,t) that belong in the first case, are $(\theta+1)$-reducible with probability 1. Hence, we only need to focus on the latter case, where all vertices on the first $C\log^{(\theta+1)}(n)$ steps are within the same distance scale $k_{st} := \lceil \log_\gamma d(s,t) \rceil$ as s from t. We will refer to k_{st} as k to ease the notation. For each such vertex v on the t-local path, property (H2) of coherent geometries tells us that there are at least $\phi\gamma^k$ candidate edges that would reduce the distance from t by a constant factor $\lambda < 1$. The probability Q_{vz} of each such good edge (v,z) is lower bounded by $\frac{1}{M\log^{\theta+1}(n)}\frac{1}{\gamma^k}$, since the measure $G(n, \mathbf{Q})$ is θ-uniformly rich. Let $T(s,t)$ be the set of all such good edges. We can write the probability of the event B_{st} as:

$$\mathbb{P}_{\mathbf{Q}}(B_{st}) = \prod_{e \in T(s,t)} (1 - Q_e) \leq \left(1 - \frac{1}{M(\log n)^{\theta+1}\gamma^k}\right)^{|T(s,t)|} \leq e^{-\frac{C\log^{\theta+1}(n)\phi\gamma^k}{M\log^\theta(n)\gamma^k}} \leq n^{-\frac{C\phi}{M}}$$

where we used that $|T(s,t)| \geq C(\log n)^{\theta+1} \cdot \phi\gamma^k$ due to (H2) and the definition of reducibility. For any $\ell > 0$ and $C \geq (2+\ell)\frac{M}{\phi}$ we get that $\mathbb{P}(B_{st}) \leq n^{-(2+\ell)}$. To finish the proof, we perform a Union Bound over all possible sets (s,t). Let B be the even that the graph $G(V, E_0 \cup E_d)$ is not $(\theta+1)$-reducible, then:

$$\mathbb{P}_{\mathbf{Q}}(B) = \mathbb{P}_{\mathbf{Q}}(\bigcup B_{st}) \leq \sum_{st} \mathbb{P}_{\mathbf{Q}}(B_{st}) \leq n^2 n^{-(2+\ell)} = n^{-\ell}$$

for any $\ell > 0$. Taking $\ell = 5$ we see that the graph $G(V, E_0 \cup E_d)$ is d-navigable with the desired probability. □

Deriving navigability from uniform richness may strike the reader as odd, given that a central goal of our work is to show that independence assumptions are *not* needed for navigability. There is no cause for alarm: we we will never *assume* uniform richness. Instead, we will prove that under certain conditions, the (random) set of edges of a typical element of the set of all graphs feasible within a certain budget *dominates* a θ-uniformly rich product measure. Our capacity to do so is enabled by a very recent general theorem we developed in [30] which asserts that if a family of graphs $S \subseteq \mathcal{G}_n$ is sufficiently symmetric, then the uniform measure on S can be well-approximated by a product measure on the $\binom{n}{2}$ edges. We discuss this next.

4 Analyzing Random Graphs of Bounded Cost

A classic result of random graph theory is that to study monotone properties of graphs with n vertices and m edges it suffices to study $G(n,p)$ random graphs, i.e., graphs generated by including each edge independently of all other with probability $p = p(m) = m/\binom{n}{2}$. The reason for this is that the uniform measure on graphs with exactly m edges is *sandwiched* by the $G(n, p(m))$ product measure, in the following sense.

Sandwichability. *The uniform measure $U(S)$ on an arbitrary set of graphs $S \subseteq \mathcal{G}_n$ is (ϵ, δ)-*sandwichable *if there exists a $n \times n$ symmetric matrix $\mathbf{Q}$ such that the two distributions $G^{\pm} \sim G(n, (1 \pm \epsilon)\mathbf{Q})$, and the distribution $G \sim U(S)$ can be coupled so that $G^- \subseteq G \subseteq G^+$ with probability at least $1 - \delta$.*

When S is the set of all graphs with exactly m edges we have $\mathbf{Q}_{ij} = p(m)$ for all non-diagonal entries. To make a sandwich, i.e., simultaneously generate G^-, G, G^+, one generates $\binom{n}{2}$ i.i.d. uniformly distributed real numbers in $[0, 1]$, one for each potential edge. The graph G^- contains all edges whose r.v. is less than $(1 - \epsilon)p$, the graph G contains the edges corresponding to the m smallest r.v.'s, while G^+ contains all edges whose r.v. is less than $(1+\epsilon)p$. As long as the m-th smallest r.v. is in $((1 - \epsilon)p, (1 + \epsilon)p)$ we have $G^- \subseteq G \subseteq G^+$.

The set of all graphs with m edges is highly symmetric: its characteristic function is invariant under every permutation of the input $x \in \{0, 1\}^{\binom{n}{2}}$; it only cares about $|x|$. When considering graphs with bounded total cost, symmetry comes from the fact that edges with the same cost are interchangeable. Thus, if the number of distinct cost-classes is not too big we can hope for a product measure approximation (indeed, the set of all graphs with m edges can be seen as the case where there is only one cost class, unit cost, and the total budget is m). As discussed earlier, navigability requires some degree of structure in the underlying geometry in the form of coherence. Our requirement that the cost function is consistent with the (coherent) geometry, giving rise to a coherent cost-geometry, is what will give us enough symmetry to apply the main theorem of [30] and derive the following approximation.

In all of the following, $\Gamma(V, d, c)$ is an arbitrary coherent cost-geometry and $K = \lceil \log_\gamma |V| \rceil$. As before, we denote by c_k the cost of an edge of scale k and by p_k the average number of neighbors at distance scale k from a random vertex in V. For a given budget $B \geq 0$, let $\lambda(B) = g^{-1}(B) \geq 0$, where

$$g(\lambda) := \sum_{k=1}^{K} c_k \frac{p_k}{1 + \exp(\lambda c_k)} \quad .$$

Intuitively, $\lambda(B)$ will control the drop in likelihood of costlier edges as a function of the budget B (mathematically, $\lambda(B)$ is a Lagrange multiplier, physically, it is an inverse temperature).

Theorem 7. *For every coherent cost-geometry Γ, there exists a constant $B_0(\Gamma) > 0$ such that for every $B \geq B_0(\Gamma)$ the uniform measure on $G_\Gamma(B)$ is (ϵ, δ)-sandwichable by the product measure $G(n, \mathbf{Q}^*(B))$ in which each edge of cost c_k has probability*

$$\mathbf{Q}^*_{ij}(B) = \frac{1}{1 + \exp(\lambda_\Gamma(B) c_k)} \quad , \tag{3}$$

where $(\epsilon, \delta) = \left(\sqrt{\frac{24}{\log n}},\ 2n^{-5K}\right)$.

4.1 Proof of Theorem 7

We start with some definitions that will allow us to state the main theorem of [30]. A set of graphs $S \subseteq \mathcal{G}_n$ is symmetric with respect to a partition $\mathcal{P}$ of the set of all possible $\binom{n}{2}$ edges, if the characteristic function of S depends only on the number of edges from each part of $\mathcal{P}$ but not on which edges.

Edge profile. *Given a partition $\mathcal{P} = (\mathcal{P}_1, \ldots, \mathcal{P}_K)$ of the set of all possible $\binom{n}{2}$ edges, for a set of edges $E \in \mathcal{G}_n$ and for each $k \in [K]$, let $m_k(E)$ denote the number of edges in E from $\mathcal{P}_k$. The* edge profile *of E is $\mathbf{m}(E) := (m_1(E), \ldots, m_K(E))$.*

We denote the image of a symmetric set S under the edge-profile as $\mathbf{m}(S)$. As before let $P_k := |\mathcal{P}_k| = \frac{1}{2}\sum_{u \in V} P_k(u)$ be the total number of edges in part k of partition $\mathcal{P}$.

Edge profile entropy. *Given an edge profile $\mathbf{v} = (v_1, \ldots, v_k)$ the entropy of $\mathbf{v}$ is* $\textsc{Ent}(\mathbf{v}) = \sum_{k=1}^{K} \log \binom{P_k}{v_k}$.

The edge-profile entropy is used to express the number of graphs with a particular edge profile $\mathbf{v}$ as $\exp(\textsc{Ent}(\mathbf{v}))$. Given any symmetric set $S \subseteq \mathcal{G}_n$, the probability of observing an edge profile $\mathbf{v}$ when sampling an element uniformly at random from S is then given by $\mathbf{P}_S(\mathbf{v}) = \frac{1}{|S|} e^{\textsc{Ent}(\mathbf{v})}$. Thus, in order to analyze the distribution of a random edge-profile, and consequently of a random element of $G_c(n, B)$, we are going to exploit analytic properties of the entropy on the set of feasible edge profiles $\mathbf{m}(S)$.

Convexity. *Let $\mathrm{Conv}(A)$ denote the convex hull of a set A. Say that a $\mathcal{P}$-symmetric set $S \subseteq \mathcal{G}_n$ is convex iff the convex hull of $\mathbf{m}(S)$ contains no new integer points, i.e., if* $\mathrm{Conv}(\mathbf{m}(S)) \cap \mathbb{N}^k = \mathbf{m}(S)$.

Entropic optimizer. *Given a symmetric set S, let $\mathbf{m}^* = \mathbf{m}^*(S) \in \mathbb{R}^k$ be the unique solution to*

$$\max_{\mathbf{v} \in \mathrm{Conv}(\mathbf{m}(S))} -\sum_{k=1}^{K} \left[v_k \log\left(\frac{v_k}{P_k}\right) + (P_k - v_k) \log\left(\frac{P_k - v_k}{P_k}\right) \right] . \tag{4}$$

Given the maximizer $\mathbf{m}^*(S)$, the matrix $\mathbf{Q}^* = \mathbf{Q}^*(S)$ is given by letting for all $k \in [K]$ the probability of an edge $e \in \mathcal{P}_k$ be $Q^*_e := m^*_k / P_k$. To state the theorem, we need the following parameters that quantify the concentration of the uniform measure around its mode.

Thickness and condition number. *Given a partition $\mathcal{P}$ and a $\mathcal{P}$-symmetric set S, we define*

$$\text{Thickness:} \qquad \mu = \mu(S) = \min_{k\in[K]} \min\{m_k^*, P_k - m_k^*\} \tag{5}$$

$$\text{Condition number:} \qquad \tau = \tau(S) = \frac{5K\log n}{\mu(S)} \tag{6}$$

We now state the main theorem employed in the proof.

Theorem 8 ([30]). *Let $\mathcal{P}$ be an edge-partition and let S be ay $\mathcal{P}$-symmetric convex set. For every $1 > \epsilon > \sqrt{12\tau(S)}$, the uniform measure over S is (ϵ, δ)-sandwichable with $\delta = 2\exp\left[-\mu(S)\left(\frac{\epsilon^2}{12} - \tau(S)\right)\right]$.*

In our setting, S is the set $G_\Gamma(B) := \{E \subset V \times V : \frac{1}{n}\sum_{e\in E} c_e \leq B\}$ of graphs with bounded average cost and $\mathcal{P}$ is the partition induced by the coherent cost function c. The set $\mathbf{m}(S)$ is then given by $\mathbf{m}(S) = \{\mathbf{v} \in \mathbb{N}^k : \frac{1}{n}\sum_{k=1}^K c_k v_k \leq B\}$. Hence, it is easy to see that $G_\Gamma(B)$ is convex and symmetric, according to the previous definition, for all values of B. To prove Theorem 7, we need to find:

(i) an analytic expression for the vector $\mathbf{m}^*$ as a function of B
(ii) the range of values of B for which applying Theorem 8 gives high probability bounds.

Finding the Entropic Optimal Edge Profile We start by introducing a slight reparametrization in terms of the average-degree profile. For an edge set E, define the vector $\mathbf{a}(E) := \mathbf{m}(E)/n$, where as before $\mathbf{m}$ is the edge-profile. In the same spirit, let $p_k := P_k/n$ denote the average number of edges in part (scale) k. Using this parametrization and by explicitly writing $\text{Conv}(\mathbf{a}(S))$, we can equivalently express the optimization problem (4) as:

$$\begin{aligned} \max_{\mathbf{a}} \ H(\mathbf{a}) = \ & -\sum_{k=1}^K [(p_k - a_k)\log(p_k - a_k) + a_k \log(a_k)] \\ \text{subject to} \quad & \sum_{k=1}^K a_k c_k \leq B \\ & 0 \leq a_k \leq p_k, \qquad \forall k \in [K] \ . \end{aligned}$$

We will refer to the above optimization problem as (Λ) and to its solution as $\mathbf{a}^* = \mathbf{a}^*(B)$. Towards obtaining an analytic expression for $\mathbf{a}^*$, we first show that no coordinate $k \in [K]$ lies on the natural boundary $\{0, p_k\}$.

Lemma 3. *The optimal profile* $\mathbf{a}^* \in \mathcal{D}(B) := \{\mathbf{a} \in (0, p_1)\times\ldots\times(0, p_K) : \sum_k^K c_k a_k \leq B\}$.

Proof. We prove the lemma by contradiction. We show that if $\mathbf{a}^*$ is a solution of (Λ) such that $\mathbf{a}^* \notin \mathcal{D}$, then there is an $\hat{\mathbf{a}}^* \in \mathcal{D}$ for which objective function takes a higher value. Specifically, there exists $\epsilon > 0$ small enough such that there are indices[4] $1 \leq i, j \leq K$ for which $a_i^* = 0$ and $a_j^* > \delta(\epsilon)$, where $\delta(\epsilon) = \epsilon\, c_i/c_j$. Define $\hat{\mathbf{a}}^*(\epsilon) = (a_1^*, \ldots, a_i^* + \epsilon, \ldots, a_j^* - \delta(\epsilon), \ldots, a_K^*)$. If $h(\epsilon) = H(\hat{\mathbf{a}}^*) - H(\mathbf{a}^*)$ is the difference in the objective function between the assumed optimal $\mathbf{a}^*$ and the perturbation $\hat{\mathbf{a}}^*$, then

$$h'(\epsilon) = -\log(\epsilon) + \log(p_i - a_i - \epsilon) + \frac{c_i}{c_j} \left(\log(a_j - \delta(\epsilon)) - \log(p_j - a_j + \delta(\epsilon))\right) \ .$$

Observe, that $\lim_{\epsilon \to 0} h^{'}(\epsilon) = +\infty$, since we have assumed that $a_j^* > 0$. This shows that every maximizer satisfies $\mathbf{a}^* > 0$. The same argument establishes that $a_k^* < p_k$ for all $k \in [K]$. Combining the two statements we get that any maximizer belongs in $\mathcal{D}$. □

As a consequence, since they are inactive at the optimum, we can omit separable inequalities from the formulation. Further, define $\bar{B} := \frac{1}{2} \sum_{k=1}^{K} p_k c_k$ the average cost of the solution to the unconstrained version of (Λ), i.e., where $\bar{a}_k := p_k/2$. If $B > \bar{B}$ then the absolute maximum entropic point $\bar{\mathbf{a}}$ is still in $\mathcal{D}(B)$ and thus the solution will be always $a_k^* = \bar{a}_k$ for every such B.

Lemma 4. *There is a unique function $\lambda(B)$ that is one-to-one for all $0 \leq B \leq \bar{B}$ and $\lambda(B) = 0$ for all $B \geq \bar{B}$, such that the unique solution of (Λ) is given by:*

$$a_k^*(B) = \frac{p_k}{1 + \exp\left[\lambda(B) \cdot c_k\right]}, \ \forall k \in [K] \ . \tag{7}$$

Proof. Uniqueness of the solution follows easily from convexity of the domain and concavity of the objective function. Further, by Lemma 3, we can reduce the optimization problem (Λ) to the following:

$$\max_{\mathbf{a}} - \sum_{k=1}^{K} \left[(p_k - a_k)\log(p_k - a_k) + a_k \log(a_k)\right]$$
$$\text{subject to} \quad \sum_{k=1}^{K} a_k c_k \leq B \ .$$

To obtain an analytical solution, we form the Lagrangian of the reduced problem

$$L(\mathbf{a}, \lambda) = - \sum_{k=1}^{K} \left[(p_k - a_k)\log(p_k - a_k) + a_k \log(a_k)\right] + \lambda \left(B - \sum_{k=1}^{K} a_k c_k\right) \ .$$

with the additional constraint that $\lambda \geq 0$. The Karush-Kuhn-Tacker conditions read

$$\frac{\partial L}{\partial a_k} = 0 \Longleftrightarrow \log\left(\frac{a_k}{p_k - a_k}\right) = -\lambda c_k \tag{8}$$

$$\frac{\partial L}{\partial \lambda} = 0 \Longleftrightarrow \sum_{k=1}^{K} a_k c_k = B \ . \tag{9}$$

[4] For any nontrivial values of B such an index can always be found.

Solving the first equation for $a_k(\lambda)$ we get

$$a_k^* = \frac{p_k}{1+\exp(\lambda c_k)} \quad ,$$

Substituting this expression in (9), we get the following function of λ:

$$g(\lambda) = \sum_{k=1}^{K} c_k \cdot \frac{p_k}{1+\exp(\lambda c_k)} \tag{10}$$

and the second constraint can now be written as $g(\lambda) = B$. The domain of g is the set of non-negative numbers on which g is continuous and infinitely differentiable. Under positive costs $\{c_k\}$, it is easy to see that $g'(\lambda) < 0$ for all $B < \bar{B}$, hence, g is strictly decreasing in the interval $[0, \infty)$ and $g(0) = \bar{B}$. Thus, $g : [0, \infty) \to [0, \bar{B}]$ is 1-to-1 and thus invertible. This means that every budget in $[0, \bar{B}]$ is feasible and that for each such budget there is a unique $\lambda(B) := g^{-1}(B)$. For $B \geq \bar{B}$, $\lambda(B) = 0$. Therefore, we conclude that the maximizer is always unique for any feasible B and implicitly given by $g(\lambda) = B$. □

Thickness $\mu(B)$ of $G_{\Gamma}(B)$ and Sandwiching Our next step is to use the analytical solution to the optimization problem to instantiate the thickness parameter μ defined in (5). Using (7), we can write:

$$\mu(B) = \min_{k\in[K]} m_k^* = n \cdot \min_{k\in[K]} \frac{p_k}{1+\exp\left[\lambda(B) c_k\right]} \tag{11}$$

where we have used the facts that that $a_k^* = m_k^*/n$ and $a_k^*(B) \leq 1/2 \Rightarrow m_k^* \leq P_k - m_k^*$. To get a more convenient expression, since $0 < c_k < \infty$ we can write the cost as $c_k = \frac{1}{\beta_k}\log(p_k)$ where $0 < \beta_k < \infty$ when $p_k \geq 1$. Thus, approximately[5] for large p_k (eq. k) we have $\mu(B) \approx n \cdot \min_{k\in[K]}\left[p_k^{1-\lambda(B)/\beta_k}\right]$. Theorem 8, gives strong (non-constant) probability bounds as long as $\tau(B) \ll 1$. For our purposes we are going to consider that the maximum $\tau(B)$ (respectively minimum B) that we allow is $\tau_0 = \log^{-1}(n)$ (respectively B_0). Substituting the above expression for $\mu(B)$ in (6), we get that the condition $\tau \leq \tau_0$ can be rewritten as $\lambda(B) \leq \lambda_0$, where

$$\lambda_0 = \lambda_0(\{p_k\},\{\beta_k\}) := \min_{k\in[K]}\left[\log\left(\frac{n\log p_k}{5K\log^2(n)}\right)\frac{\beta_k}{\log p_k}\right] \quad . \tag{12}$$

Using the function $g(\lambda)$ defined in (10), we can express this constraint as $B \geq B_0(\Gamma) := g(\lambda_0)$.

To conclude the proof we see that $\mu(B) \geq 5K\log^2(n)$ and $\tau(B) \leq \frac{1}{\log(n)}$, for all $B \geq B_0(\Gamma)$. Applying Theorem 8, for $\epsilon_0 = \sqrt{\frac{24}{\log n}}$ that is greater than $\sqrt{12\tau_0}$, we get that $\delta \leq 2\exp\left[\mu(B)\left(\frac{\epsilon_0^2}{12} - \tau(B)\right)\right]$. The proof is concluded by substituting the bounds in the last expression.

[5] When the approximation does not hold it means that $\mu(B) = \Omega(n)$ which trivially satisfies all the requirements we need for "sandwiching" and navigability.

5 Navigability for Random Graphs of Bounded Cost Model

Having established the connection of Random Graphs of Bounded cost with an explicit product measure (Theorem 7), in order to prove navigability we simply need to show that for a range of values of B, the product measure defined through (3) is θ-*uniformly rich* for some $\theta > 0$.

Proposition 2. *Let* $\lambda_\theta(\{p_k\},\{c_k\}) := \min_{k_\theta \le k \le K}\left[\frac{\log p_k}{c_k}\left(1+\frac{\theta \log\log n}{\log p_k}\right)\right]$. *Let* $B_\theta^- := \max\{B_0(\Gamma), g(\lambda_\theta)\}$. *For all* $B \ge B_\theta^-$*, the product measure* $G(n, \mathbf{Q}^*(B))$ *is* θ*-uniformly rich.*

Proof. This follows easily by the definition of λ_θ and the monotonicity of $\lambda(B) = g^{-1}(B)$ with respect to B. In particular, for any pair (i,j) of distance scale $k \ge k_\theta$ we have

$$Q^*_{ij}(B) = [1+\exp(c_k\lambda(B))]^{-1} \ge \left[p_k \log^\theta(n)\right]^{-1} \ge \frac{1}{A\log^\theta(n)\gamma^k} \quad ,$$

where the last inequality follows from (H1). □

Proposition 3. *Let* $\Lambda_\theta(\{p_k\},\{c_k\}) := \max_{k_\theta \le k \le K}\left[\frac{\log p_k}{c_k}\left(1-\frac{\theta \log\log n}{\log p_k}\right)\right]$ *and* $B_\theta^+ := g(\Lambda_\theta)$. *For all* $B \le B_\theta^+$*, the product measure* $G(n, \mathbf{Q}^*(B))$ *has* $O(n\log^{\theta+1} n)$ *edges with probability at least* $1-n^{-5}$.

Proof. For all $B \le B_\theta^+$, by definition of Λ_θ we have that for all $k \ge k_\theta$:

$$Q^*_{ij}(B) = [1+\exp(c_k\lambda(B))]^{-1} \le \left[p_k \log^{-\theta}(n)\right]^{-1} \quad .$$

Thus, the *expected* number of edges $n \cdot \sum_{k=1}^K p_k[1+\exp(\lambda(B)c_k)]^{-1}$ is upper bounded by

$$n\cdot\left[Ak_\theta p_{k_\theta} + (K-k_\theta)\max_{k\ge k_\theta} p_k \frac{\log^\theta n}{p_k}\right] = n\cdot O\left(\log\log(n)\log^\theta n + \log(n)\log^\theta n\right) \quad ,$$

since $k_\theta = O(\log\log n)$, $p_{k_\theta} = O(\log^\theta n)$ by (H1), and $K = O(\log n)$. Expressing the number of edges as a sum of independent Bernoulli random variables and applying standard Chernoff bounds [31] we get the required conclusion. □

Proof of Theorem 4. For any $B \ge B_0(\Gamma)$, consider two random elements generated according to $E^\pm \sim G(n, (1\pm\epsilon)\mathbf{Q}^*(B))$ and let W be the event that $E^- \subseteq E_\Gamma \subseteq E^+$. Theorem 7 implies that for $\epsilon = \sqrt{24/\log(n)}$ the probability of W is at least $1-n^{-5K}$. Further, for any constant $p>0$ and for an arbitrary set of edges E let $N_p(E)$ denote the event that that the graph $G(V, E_0 \cup E))$ is not p-reducible and let $N_d(E)$ be the event that the same graph is not d-navigable. Since p-reducibility is a monotone increasing property with respect to edge inclusion and since $N_d \subseteq N_p$ by Proposition 1, we get

$$\mathbb{P}(N_d(E_\Gamma)) = \mathbb{P}(N_d(E_\Gamma) \cap W) + \mathbb{P}(N_d(E_\Gamma) \cap \overline{W}) \tag{13}$$
$$\leq \mathbb{P}(N_p(E_\Gamma)|W) + \mathbb{P}(\overline{W}) \tag{14}$$
$$\leq \mathbb{P}_{\mathbf{Q}^*}(N_p(E^-)) + 2n^{-5K} \tag{15}$$
$$\leq n^{-5} + 2n^{-5K} \ , \tag{16}$$

where we used the law of total probability in the first equality, Bayes Theorem in the second inequality, Theorem 7 and monotonicity of reducibility in the third. The last inequality follows from Lemma 2 and Proposition 2. This proves part (a) of the theorem. To prove part (b) we follow the same method but for the event $\{|E_\Gamma| = \omega(n\text{polylog}n)\}$ and exploit that, conditional on W occurring, $E_\Gamma \subseteq E^+$. Using Proposition 3 and Theorem 7 we get the required conclusion. □

5.1 Recovering RBA: Proof of Theorem 5

Let us write $c_k^* = \frac{1}{\beta} \log p_k$ and let $B_\beta := g(\beta)$. For any edge (u, v) of scale $k \geq \kappa$, we have

$$Q_{uv}^*(B_\beta) = \frac{1}{1 + \exp(\lambda(B_\beta)c_k^*)} = \frac{1}{1 + \exp(\beta \frac{\log p_k}{\beta})} = \frac{1}{1 + p_k} \ .$$

Now, by property (H1) we know that for any vertex u and every vertex v within distance scale $k \geq \kappa$ from u, $\text{rank}_u(v) \leq \sum_{t=1}^{k} P_t(u) \leq \frac{A\gamma}{\gamma-1}\gamma^k$ and $\text{rank}_u(v) \geq P_{k-1}(u) \geq \frac{a}{\gamma}\gamma^k$. This relation enables as to provide a connection with RBA, proving Part (c):

$$\left(\frac{a}{2A\gamma}\right)\frac{1}{\text{rank}_u(v)} \leq \frac{1}{2A\gamma^k} \leq Q_{uv}^*(B_\beta) \leq \frac{1}{a\gamma^k} \leq \left(\frac{A\gamma}{a(\gamma-1)}\right)\frac{1}{\text{rank}_u(v)} \ . \tag{17}$$

To further highlight the correspondence between RBA and Random Graphs of Bounded Cost when the cost is proportional to the cost of indexing, observe that for every distance scale $k \in [K]$, the per vertex average number, $a_k^*(B_\beta)$, of edges of scale k is

$$a_k^*(B_\beta) = \frac{p_k}{1 + p_k} = \Theta(1) \ .$$

Thus, we recover RBA's property of (approximately) one long-distance edge per vertex per distance scale.

To prove the first part of the theorem we see that when $c_k^* = \frac{1}{\beta} \log p_k$, the quantities λ_θ and Λ_θ in Propositions 2 and 3, respectively, equal[6]

$$\lambda_\theta^* = \beta\left(1 + \theta\frac{\log\log n}{\log p_K}\right) \tag{18}$$
$$\Lambda_\theta^* = \beta\left(1 - \theta\frac{\log\log n}{\log p_K}\right) \ . \tag{19}$$

[6] This is the only place we use monotonicity of c_k^* and this is only to not burden the reader with the extraneous constants that would arise.

By property (H1) we know that $\log p_K = \Theta(\log n)$. Define $B^+ = g(\Lambda_\theta^*)$ and $B^- = g(\lambda_\theta^*)$ and let $a_k^*(B_\beta) = \frac{p_k}{1+\exp(\lambda(B_\beta)c_k^*)}$ be the average number of edges of scale k per vertex. Then for every $B^- \leq B \leq B^+$ or equivalently for $\Lambda_\theta^* \leq \lambda(B) \leq \lambda_\theta^*$, we have that for some $C > 0$,

$$\Omega\left(\left[\log n^{-\frac{C\theta}{\log n}}\right]^k\right) = a_k^*(B) = O\left(\left[\log n^{\frac{C\theta}{\log n}}\right]^k\right) \quad .$$

where the far left and far right hand sides are achieved for the extremal exponents. Using the facts that $\log p_K = \Theta(\log n)$, that $\lambda(B^+) = \Lambda_\theta^*$ and consequently $a_K^*(B^+) = \Omega\left(\left[\log n^{\frac{C\theta}{\log n}}\right]^k\right)$, we get

$$B^+ = \frac{1}{\beta}\sum_{k=1}^{K} a_k^*(B^+)\log p_k \geq \frac{1}{\beta}a_K^*(B^+)\log p_K = \Omega(\log(n)^{1+C'\theta})$$

where the constant $C' = CK/\log n = \Theta(1)$. Further, $B^- \leq B_\beta = \sum_{k=1}^{K} c_k^* \frac{p_k}{1+p_k} = \Theta(\log^2(n))$. Hence, $B^+/B^- = \Omega(\text{poly}(\log n))$ as $\theta > 0$ can be as large a constant as we want.

Finally, the second part is a direct corollary of Theorem 4 since the numbers $B^\pm$ defined above are those in Theorem 4 for the cost function $c_k^* \propto \log p_k$. Hence, for all $B \in [B^-, B^+]$ a random element $E_\Gamma(B)$ of $G_\Gamma(B)$ is navigable and has poly-logarithmic degree with high probability.

6 Proof of Theorem 2

We begin by recalling the definitions of set-systems from [8].

Definition 5 (Set System). *Let V be a finite set of vertices and let $\Sigma = \{S_1, \ldots, S_m\}$ be a collection of subsets of V. If a set S contains a vertex t we will say that S is* t-bound. *Fix $0 < \lambda < 1$ and $\beta > 1$. We say that Σ is a (λ, β)-set system if all the following hold:*

(K1) $V \in \Sigma$.
(K2) If $|S| > 1$, then for every $t \in S$, there is a t-bound $S' \subset S$ of size $|S'| \geq \min\{\lambda|S|, |S| - 1\}$.
(K3) If $S_L(v)$ is the union of sets that contain v and have size at most $L \geq 2$, then $|S_L(v)| \leq \beta L$.

Given a set system Σ on a set of vertices V, we define the distance (semi-metric) between two vertices.

Definition 6. *For any two vertices $u, v \in V$, their* distance *in Σ, denoted by $d_\Sigma(u, v)$, is the size of the smallest set in Σ containing both vertices minus 1, i.e. $d_\Sigma(u, v) = \min_{S\in\Sigma}\{|S| - 1 : u, v \in S\}$.*

The goal of this section is to show that the geometry (V, d_Σ) is coherent for any (λ, β)-set system, i.e., prove that for a suitable $\gamma > 1$ the semi-metric d_Σ satisfies properties (H1) and (H2) in the definition of Coherence in Section 3. Towards that direction, the main hurdle is obtaining for all $v \in V$ upper and lower bounds on $P_k(v)$, the number of vertices at distance in $(\gamma^{k-1}, \gamma^k]$ from v. The basic observation that guides the proof is that for all v and $k \geq 1$

$$P_k(v) = |B_k(v)| - |B_{k-1}(v)| \tag{20}$$

where $B_k(v)$ is the set of all vertices having distance from v at most γ^k. This representation is very convenient because the properties of set systems are directly related to $|B_k(v)|$. In particular, if we get good upper and lower bound for $|B_k(v)|$ then we can obtain upper and lower bounds for $P_k(v)$ and prove $(H1)$, which comprises the main challenge.

Obtaining the upper bound is trivial, since it is directly given by (K3). However, the lower bound on $B_k(v)$ requires more thought as it needs to be tight enough so that when substituting both bounds in (4) (in order to obtain a lower bound on $P_k(v)$) the difference is strictly positive. It turns out that the last property depends on the particular values of the parameters λ, β. We show that it is always possible to select $\gamma = \gamma(\beta, \lambda) > 1$ such that the last property holds. The main observation that will provide a lower bound on $|B_k(v)|$ is that the existence of a set S with size in $(\gamma^{k-1}, \gamma^k]$ implies that $|B_k(v)| \geq |S|$ for all $v \in S$. This is because all vertices in S have distance at most $|S| - 1$ from v. Thus, what remains is to show the existence of such set S for all $v \in V$ and k. To that end, we need the following axillary lemma that was implicitly stated and used in Kleinberg's original work [8].

Proposition 4 (Shrinkage). *For every $S \in \Sigma$ with $|S| \geq 1/(\lambda - \lambda^2)$ and for every $t \in S$, there exists a t-bound set $S' \in \Sigma$ with $\lambda^2|S| \leq |S'| \leq \lambda|S|$.*

Proof. Given any set $S \in \Sigma$ and $t \in S$, we start with S and invoke (K2) iteratively until we reach t, producing a sequence $S = S_1 \supset S_2 \cdots \supset S_k = \{t\}$ of t-bound subsets of S. Since $|S| > \lambda|S|$, there is a largest index i such that $|S_i| > \lambda|S|$ and $|S_i| \geq 2$, since $\lambda|S| \geq \frac{\lambda}{\lambda - \lambda^2} > 1$. Applying (K2) to S_i yields a t-bound set S_{i+1} of size at least $\min\{\lambda|S_i|, |S_i| - 1\}$ and less than or equal to $\lambda|S|$ (by maximality of i). Trivially, $\lambda|S_i| > \lambda^2|S|$ so we only need to show that $|S_i| - 1 \geq \lambda^2|S|$, for which a sufficient condition is that $\lambda|S| - \lambda^2|S| \geq 1$. □

This lemma will be used to show that for all vertices v one can start from the set V, that belongs in Σ by (K1), and inductively apply Proposition 4 to deduce the existence of sets S containing v at all scales. More specifically, given a (λ, β)-set system Σ, let M be the smallest integer such that $\lambda^{-2M} \geq |V|$. We partition the range of possible set-sizes in Σ as $\mathcal{I} = (I_1, \ldots, I_M)$ by letting $I_k = (\lambda^{-2(k-1)}, \lambda^{-2k}]$, for $k \in [M]$. The partition $\mathcal{I}$ implicitly partitions all pairs of vertices into groups, such that all pairs in a group have roughly the same distance in Σ, i.e., up to a factor of λ^2. We show that for every vertex and for every interval of the partition, there is a set with size in that interval that contains the vertex.

Proposition 5. *For every $t \in V$, for every $k \in [M]$, there exists a t-bound set $S \in \Sigma$ with $|S| \in I_k$.*

Proof. Assume, for the sake of contradiction, that there exists a vertex t for which the proposition does not hold. Let $k_0 \in [M]$ be the largest integer such that there is no t-bound set $S' \in \Sigma$ with $|S'| \in I_{k_0}$. If we start with V (which is trivially t-bound) and invoke (K2) iteratively until we reach t, we get a sequence $V = S_1 \supset S_2 \cdots \supset S_k = \{t\}$ of t-bound sets. Let i_{k_0} be the largest index $i \leq k$ such that $|S_i| \in I_{k_0+1}$ (by maximality of k_0). The maximality of k_0 implies $|S_{i_{k_0}+1}| \in I_{k_0-1}$. But invoking Proposition 4 for $S_{i_{k_0}}$ implies $|S_{i_{k_0}+1}| \in I_{k_0}$, a contradiction. □

Treating $\mathcal{I}$ as a distance scale, our next goal is to obtain for each vertex t, upper and lower bounds on the number of vertices that lie at each distance-scale from t. To achieve this we need to consider a coarser partition of the set sizes than $\mathcal{I}$. To do that it will be beneficial to use a partition built out of blocks of $\mathcal{I}$, thus allowing us to utilize Proposition 5, proven for $\mathcal{I}$. In particular, the existence of a t-bound set of each size will be the basis for obtaining lower bounds on the number of vertices at each new distance scale from t.

We let $r = r(\beta, \lambda) \geq 2$ denote the smallest integer such that $\lambda^{-2(r-1)} > \beta$ and consider the partition that results by grouping together every r consecutive intervals of $\mathcal{I}$. That is, for $\gamma(\beta, \lambda) = \lambda^{-2r(\beta,\lambda)}$, we define the partition $\mathcal{A} = \mathcal{A}(\gamma)$ consisting of the intervals $A_k = (\gamma^{k-1}, \gamma^k],\ k \in [K]$, where K is the smallest integer such that $\gamma^K \geq |V| - 1$. Having defined $\mathcal{A}$, we now let $P_k(v)$ denote the number of vertices whose distance from v lies in the set A_k and we let $P_k = \frac{1}{2}\sum_{v \in V} P_k(v)$ denote the total number of pairs of vertices whose distance lies in A_k.

Lemma 5 (Bounded Growth). *Let $\alpha = (\lambda^2 - \beta/\gamma) > 0$ and $A = (\beta - \lambda^2/\gamma)$. For all $k \in [K]$ and $v \in V$,*

$$\alpha \cdot \gamma^k \leq P_k(v) \leq A \cdot \gamma^k \ .$$

Proof. First observe that $\mathcal{A}$ is a coarsening of $\mathcal{I}$ since $\gamma = \lambda^{-2r}$ and $r \geq 2$ is an integer. Next, let $B_k(v) = \sum_{i \leq k} P_k(v)$ be the number vertices in V whose distance from v lies in $A_1 \cup \cdots \cup A_k$, i.e., is no more than γ^k. Condition (K3) asserts that $B_k(v) \leq \beta\gamma^k$. On the other hand, by Proposition 5, we know that for any $v \in V$ there is a v-bound set $S \in I_{rk} \subset A_k$. Since, all vertices in S have distance at most $|S| \leq \gamma^k$ from v, we get that $B_k(v) \geq |S| \geq \lambda^{-2(rk-1)} = \gamma^k\lambda^2$. Therefore, for all $k \in [K]$,

$$\lambda^2\gamma^k \leq B_k(v) \leq \beta\gamma^k \ . \tag{21}$$

Using the representation (20) and invoking (21), we get

$$\lambda^2\gamma^k - \beta\gamma^{k-1} \leq P_k(v) \leq \beta\gamma^k - \lambda^2\gamma^{k-1}$$

which is equivalent to the claimed statement. The fact $\alpha > 0$ is implied by our choice of γ. □

Thus we have shown property (H1). Proceeding further, we need to show that the semi-metric d_Σ satisfies also the isotropy property (Section 3), i.e. that the size of the set

$D_\lambda(s,t) = \{v \in V : d(s,v) \leq \gamma^{k_{st}}$ and $d(v,t) \leq \lambda d(s,t)\}$ is proportional to $\gamma^{k_{st}}$, where k_{st} is the scale of $d(s,t)$. To do that we are going to show something stronger. Given any two vertices $s \neq t \in V$, consider a $S_{st} \in \Sigma$ of minimal size such that both $s,t \in S$. Then for all $k \leq k_{st}$ define the following set $G_k(s,t) = \{v \in S_{st} : d(s,v) \in A_k$ and $d(v,t) \leq \lambda|S|\}$ of vertices in S_{st} whose distance from s lies in the interval A_k (scale k) and whose distance from t is no more than $\lambda|S_{st}|$.

Lemma 6 (Isotropy). *For every $s \neq t \in V$ with $|S_{st}| \geq 1/(\lambda - \lambda^2)$, we have that*

$$|G_{k_{st}}(s,t) \cup G_{k_{st}-1}(s,t)| \geq \left(\frac{\alpha}{\gamma}\right)\gamma^{k_{st}} \quad .$$

Proof. Proposition 4 implies that there is a t-bound set $S' \in \Sigma$ with $\lambda^2|S_{st}| \leq |S'| \leq \lambda|S_{st}|$. Thus, a λ^2 fraction of the vertices in S_{st} have distance from t at least a factor λ less that $|S_{st}|$. Having established an abundance of "good" vertices in S_{st}, we are left to show that a constant fraction of them are in the top two distance scales $k_{st}, k_{st}-1$ from s (recall that $|S_{st}| \in A_{k_{st}}$). We start by noting that $Z = \sum_{i \leq k} |G_i(s,t)| \geq |S'|$, as the sum must count the vertices in S'. Since $S_{st} \in |A_{k_{st}}|$ and $|S'| \geq \lambda^2|S_{st}|$, we get $Z \geq \lambda^2\gamma^{k_{st}-1}$. On the other hand, the good vertices in the bottom $k_{st}-2$ distance scales from s are a subset of all vertices containing s at those distance scales, a quantity bounded by (K3) as $\sum_{i \leq k-2} |G_i(s,t)| \leq \beta\gamma^{k_{st}-2}$. Therefore, $|G_{k_{st}}(s,t) \cup G_{k_{st}-1}(s,t)| \geq \lambda^2\gamma^{k_{st}-1} - \beta\gamma^{k_{st}-2}$. □

Proof of Theorem 2. In order to prove that the set system defines a coherent geometry, we need to show that properties $(H1)$ and $(H2)$ hold for some $\gamma > 1$. Our two lemmas achieve exactly that. The first property follows from Lemma 5 and the second property follows from Lemma 6 since $G_{k_{st}}(s,t) \cup G_{k_{st}-1}(s,t) \subset D_\lambda(s,t)$. □

7 Proof of Theorem 3

Recall that in RBA, for each vertex $u \in V$, a single link is added from u to a random vertex v with probability given by

$$P_{\text{RBA}}(u,v) = \frac{1}{Z}\frac{1}{\text{rank}_u(v)}. \tag{22}$$

Here we show that the Kleinberg's original proof can be applied with ease when instead of the semi-metric induced by set-system, we have a semi-metric corresponding to a coherent geometry. There are basically two steps. We first upper bound the normalizing constant Z and then lower bound the probability that for a given pair (s,t) we find an edge in the first $C\log^2(n)$ steps of a path along the substrate that reduces the distance to t by a constant factor.

Proposition 6 (Bounded Growth). *For a coherent geometry (V,d), $\exists C < \infty$ such that $Z \leq C\log(n)$.*

Proof. For a given vertex u, we divide vertices depending on their distance scale $k \in \{0, \ldots, \log_\gamma(n)\}$ from u. For $k \geq 0$, we know from property (H1) that there are at most $A\gamma^k$ such vertices. Further, we also know that $|B_{k-1}(u)| = \sum_{i=0}^{k-1} P_k(u) \geq a\frac{\gamma^k - 1}{\gamma - 1}$. Using these two facts we have:

$$
\begin{aligned}
Z = \sum_{v \in V} P_\alpha(u, v) &\leq \frac{A}{a} + \sum_{k=1}^{\log(n)} P_k(u) \frac{1}{|B_{k-1}(u)|} \\
&\leq \frac{A}{a} + \frac{A}{a} \sum_{k=1}^{\log(n)} \gamma^k \frac{\gamma - 1}{\gamma^k - 1} \leq \frac{A}{a} \left(1 + \gamma \log_\gamma(n)\right)
\end{aligned}
$$

□

Finally, to complete the proof, we are going to employ once again reducibility.

Proof of Theorem 3. Fix any two vertices s, t, the probability of finding a long-range edge at s reducing the distance by a constant factor is at least:

$$
\frac{|D(s,t)|}{Z} \frac{1}{P_k(s)} \geq \frac{1}{C \log n} \frac{\phi \gamma^k}{A \gamma^k} = \frac{\phi}{AC} \frac{1}{\log n}
$$

Thus, the probability of the event B_{st} that no such edge exists after $C' \log^2(n)$ trials is at most:

$$
\mathbb{P}(B_{st}) \leq \left(1 - \frac{\phi}{AC} \frac{1}{\log n}\right)^{C' \log^2(n)} \leq e^{-\frac{\phi C'}{AC} \log n} \leq n^{-\frac{\phi}{AC} C'}
$$

For C' large enough and a union bound over the $\Theta(n^2)$ possible pairs of vertices, we get that if E_d is the random set of edges added through RBA and E_0 is a substrate for the coherent geometry (V, d), then the graph $G(V, E_0 \cup E_d)$ is d-navigable with high probability. □

References

1. D. J. Watts, *Six degrees: The science of a connected age*. WW Norton, 2004.
2. R. van der Hofstad, “Random Graphs and Complex Networks”. Volume 1. *Cambridge Series in Statistical and Probabilistic Mathematics*, Cmapridge university press, 2017.
3. M. Newman, “Networks: an introduction,” Oxford university press, 2010.
4. S. Milgram, “The small world problem,” *Psychology today*, vol. 2, no. 1, pp. 60–67, 1967.
5. S. Goel, R. Muhamad, and D. Watts, “Social Search in ”Small-World” Experiments,” in *Proceedings of the 18th International Conference on World Wide Web*. ACM, 2009.
6. L. Backstorm, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, “Four degrees of separation,” in *Proceedings of the 4th Annual ACM Web Science Conference*, pp. 33–42, ACM, 2012.
7. J. Kleinberg, “The small-world phenomenon: An algorithmic perspective,” in *Proceedings of the 32nd annual ACM symposium on Theory of computing*, pp. 163–170, ACM, 2000.
8. J. Kleinberg, “Small-world phenomena and the dynamics of information,” *Advances in neural information processing systems*, vol. 1, pp. 431–438, 2002.

9. I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, “Freenet: A distributed anonymous information storage and retrieval system,” in *Designing Privacy Enhancing Technologies*, pp. 46–66, Springer, 2001.
10. H. Zhang, A. Goel, and R. Govindan, “Using the small-world model to improve freenet performance,” *Computer Networks*, vol. 46, no. 4, pp. 555–574, 2004.
11. J. Aspnes, Z. Diamadi, and G. Shah, “Fault-tolerant routing in peer-to-peer systems,” in *Proceedings of the twenty-first annual symposium on Principles of distributed computing*, pp. 223–232, ACM, 2002.
12. A. Slivkins, “Distance estimation and object location via rings of neighbors,” *Distributed Computing*, vol. 19, no. 4, pp. 313–333, 2007.
13. G. S. Manku, M. Naor, and U. Wieder, “Know thy neighbor’s neighbor: the power of lookahead in randomized p2p networks,” in *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pp. 54–63, ACM, 2004.
14. J. Zeng, W.-J. Hsu, and J. Wang, “Near optimal routing in a small-world network with augmented local awareness,” in *Parallel and Distributed Processing and Applications*, pp. 503–513, Springer, 2005.
15. A. Karbasi, S. Ioannidis, and L. Massoulie, “From small-world networks to comparison-based search,” *Information Theory, IEEE Transactions on*, vol. 61, no. 6, pp. 3056–3074, 2015.
16. D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins, “Geographic routing in social networks,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 33, pp. 11623–11628, 2005.
17. J. Kleinberg, “Complex networks and decentralized search algorithms,” in *Proceedings of the International Congress of Mathematicians (ICM)*, vol. 3, pp. 1019–1044, 2006.
18. L. Adamic and E. Adar, “How to search a social network,” in *Social networks*, vol. 27, no. 3, pp. 187–203, Elsevier, 2005.
19. A. Clauset and C. Moore, “How do networks become navigable?,”, *arXiv preprint* cond-mat/0309415, 2003.
20. O. Sandberg, “Neighbor selection and hitting probability in small-world graphs,” *Ann. Appl. Probab.*, vol. 18, pp. 1771–1793, 10 2008.
21. A. Chaintreau, P. Fraigniaud, and E. Lebhar, “Networks become navigable as nodes move and forget,” in *Automata, Languages and Programming*, pp. 133–144, Springer, 2008.
22. A. Gulys, J.J Bíró, A. Kőrösi, G. Rétvári, and D. Krioukov, “Navigable networks as Nash equilibria of navigation games,” *Nature communications 6*, 2015.
23. W. Chen, W. Fang, G. Hu, and M.W. Mahoney, “On the hyperbolicity of small-world and treelike random graphs,” in *Internet Mathematics*, vol. 9, no. 4, pp. 434–491, Taylor & Francis, 2013.
24. P. Duchon, N. Hanusse, E. Lebhar, and N. Schabanel, “Could any graph be turned into a small-world?,” *Theor. Comput. Sci.*, vol. 355, no. 1, pp. 96–103, 2006.
25. P. Fraigniaud, “Greedy routing in tree-decomposed graphs,” in *Algorithms–ESA 2005*, pp. 791–802, Springer, 2005.
26. I. Abraham, C. Gavoille, and D. Malkhi, “Compact routing for graphs excluding a fixed minor,” in *Distributed Computing*, pp. 442–456, Springer, 2005.
27. P. Fraigniaud, E. Lebhar, and Z. Lotker, “A lower bound for network navigability,” *SIAM J. Discrete Math.*, vol. 24, no. 1, pp. 72–81, 2010.
28. P. Fraigniaud, C. Gavoille, A. Kosowski, E. Lebhar, and Z. Lotker, “Universal augmentation schemes for network navigability,” *Theoretical Computer Science*, vol. 410, no. 21, pp. 1970–1981, 2009.
29. P. Fraigniaud and G. Giakkoupis, “On the searchability of small-world networks with arbitrary underlying structure,” in *Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 389–398, ACM, 2010.

30. D. Achlioptas and P. Siminelakis, "Symmetric graph properties have independent edges," in *Automata, Languages, and Programming*, pp. 467–478, Springer, 2015.
31. N. Alon and J. Spencer, *The Probabilistic Method.* John Wiley, 1992.