Localization from Incomplete Noisy Distance Measurements

Adel Javanmard · Andrea Montanari

Received: 13 July 2011 / Accepted: 4 May 2012 / Published online: 6 July 2012
© SFoCM 2012

Abstract We consider the problem of positioning a cloud of points in the Euclidean space $\mathbb{R}^d$, using noisy measurements of a subset of pairwise distances. This task has applications in various areas, such as sensor network localization and reconstruction of protein conformations from NMR measurements. It is also closely related to dimensionality reduction problems and manifold learning, where the goal is to learn the underlying global geometry of a data set using local (or partial) metric information. Here we propose a reconstruction algorithm based on semidefinite programming. For a random geometric graph model and uniformly bounded noise, we provide a precise characterization of the algorithm’s performance: in the noiseless case, we find a radius $r_0$ beyond which the algorithm reconstructs the exact positions (up to rigid transformations). In the presence of noise, we obtain upper and lower bounds on the reconstruction error that match up to a factor that depends only on the dimension $d$, and the average degree of the nodes in the graph.

Keywords Rigidity theory · Global rigidity · Stress matrix · Graph realization · Network localization · Manifold learning · Semidefinite programming
1 Introduction

1.1 Problem Statement

Given a set of \( n \) nodes in \( \mathbb{R}^d \), the localization problem requires one to reconstruct the positions of the nodes from a set of pairwise measurements \( \tilde{d}_{ij} \) for \((i, j) \in E \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\}\). Thus an instance of the problem is given by the graph \( G = (V, E) \), \( V = \{1, \ldots, n\} \), and the vector of distance measurements \( \tilde{d}_{ij} \) associated to the edges of this graph.

In this paper we consider the random geometric graph model \( G(n, r) = (V, E) \), whereby the \( n \) nodes in \( V \) are independent and uniformly random in the \( d \)-dimensional hypercube \([-0.5, 0.5]^d\), and \( E \subseteq V \times V \) is a set of edges that connect the nodes which are close to each other. More specifically, we let \((i, j) \in E \) if and only if \( d_{ij} = \|x_i - x_j\| \leq r \). For each edge \((i, j) \in E\), \( \tilde{d}_{ij} \) denotes the measured distance between nodes \( i \) and \( j \). Letting \( z_{ij} \equiv \tilde{d}_{ij}^2 - d_{ij}^2 \) the measurement error, we will study a “worst-case model,” in which the errors \( \{z_{ij}\}_{(i, j) \in E} \) are arbitrary but uniformly bounded \( |z_{ij}| \leq \Delta \). We propose an algorithm for this problem based on semidefinite programming and provide a rigorous analysis of its performance, focusing in particular on its robustness properties.

Notice that the positions of the nodes can only be determined up to rigid transformations (a combination of rotation, reflection, and translation) of the nodes, because the inter-point distances are invariant to rigid transformations. For future use, we introduce a formal definition of rigid transformation. Let \( X \in \mathbb{R}^{n \times d} \) be the matrix whose \( i \)th row, \( x_i^T \in \mathbb{R}^d \), is the coordinate of node \( i \). Further, let \( O(d) \) denote the orthogonal group of \( d \times d \) matrices. A set of positions \( Y \in \mathbb{R}^{n \times d} \) is a rigid transform of \( X \), if there exists a \( d \)-dimensional shift vector \( s \in \mathbb{R}^d \) and an orthogonal matrix \( O \in O(d) \) such that

\[
Y = XO + us^T. \tag{1}
\]

Throughout \( u \in \mathbb{R}^n \) is the all-ones vector. Therefore, \( Y \) is obtained by first rotating (and/or reflecting) nodes in position \( X \) by matrix \( O \) and then shifting by \( s \). Also, two position matrices \( X \) and \( Y \) are called equivalent up to rigid transformations, if there exists \( O \in O(d) \) and a shift \( s \in \mathbb{R}^d \) such that \( Y = XO + us^T \). We use the following metric, similar to the one defined in [16], to evaluate the distance between the original position matrix \( X \in \mathbb{R}^{n \times d} \) and the estimation \( \hat{X} \in \mathbb{R}^{n \times d} \). Let \( L = I - (1/n)uu^T \) be the centering matrix. Note that \( L \) is an \( n \times n \) symmetric matrix of rank \( n - 1 \) which eliminates the contribution of the translation, in the sense that \( LX = L(X + us^T) \) for all \( s \in \mathbb{R}^d \). Furthermore, \( LXX^TL \) is invariant under rigid transformations and \( LXX^L = L\hat{X}\hat{X}^TL \) implies that \( X \) and \( \hat{X} \) are equal up to rigid transformations. The metric is defined as

\[
d(X, \hat{X}) \equiv \frac{1}{n^2} \left\| LXX^TL - L\hat{X}\hat{X}^TL \right\|_1. \tag{2}
\]
This is a measure of the average reconstruction error per point, when \( X \) and \( \hat{X} \) are aligned optimally. To get a better intuition about this metric, consider the case in which all the entries of \( LXX^TL - L\hat{X}\hat{X}^TL \) are roughly of the same order. Then

\[
d(X, \hat{X}) \approx d_2(X, \hat{X}) = \frac{1}{n} \left\| LXX^TL - L\hat{X}\hat{X}^TL \right\|_F.
\]

Denote by \( Y = \hat{X} - X \) the estimation error, and assume without loss of generality that both \( X \) and \( \hat{X} \) are centered. Then for small \( Y \), we have

\[
d_2(X, \hat{X}) = \frac{1}{n} \left\| XY^T + YY^T \right\|_F \approx \frac{1}{n} \left\| XY^T + YX^T \right\|_F \geq \frac{C}{\sqrt{n}} \| Y \|_F = C \left\{ \frac{1}{n} \sum_{i=1}^{n} \| \hat{X}_i - x_i \|^2 \right\}^{1/2},
\]

where the bound (a) holds with high probability for a suitable constant \( C \), if \( X \) is distributed according to our model.\(^1\)

Remark Clearly, connectivity of \( G \) is a necessary assumption for the localization problem to be solvable. It is a well-known result that the graph \( G(n, r) \) is connected with high probability (w.h.p.) if \( K_d r^d > (\log n + c_n)/n \), where \( K_d \) is the volume of the \( d \)-dimensional unit ball and \( c_n \to \infty \) [18]. Vice versa, the graph is disconnected with positive probability if \( K_d r^d \leq (\log n + C)/n \) for some constant \( C \). Hence, we focus on the regime where \( r \geq \alpha (\log n / n)^{1/d} \) for some constant \( \alpha \). We further notice that, under the random geometric graph model, the configuration of the points is almost surely generic, in the sense that the coordinates do not satisfy any nonzero polynomial equation with integer coefficients.

1.2 Algorithm and Main Results

Algorithm 1 uses semidefinite programming (SDP) to solve the localization problem. Here \( M_{ij} = e_i e_j^T \in \mathbb{R}^{n \times n} \), where \( e_i \in \mathbb{R}^n \) is the vector with +1 at the \( i \)th position, −1 at the \( j \)th position, and zero everywhere else. Also, \( \langle A, B \rangle \equiv \text{Tr}(A^T B) \). Note that, with a slight abuse of notation, the solution of the SDP problem in the first step is denoted by \( Q \).

Let \( Q_0 := XX^T \) be the Gram matrix of the node positions, namely \( Q_{0,ij} = x_i \cdot x_j \). A key observation is that \( Q_0 \) is a low-rank matrix—\( \text{rank}(Q_0) \leq d \)—and obeys the constraints of the SDP problem. By minimizing \( \text{Tr}(Q) \) in the first step, we promote low-rank solutions \( Q \) (since \( \text{Tr}(Q) \) is the sum of the eigenvalues of \( Q \)). Alternatively, this minimization can be interpreted as setting the center of gravity of \( \{x_1, \ldots, x_n\} \) to coincide with the origin, thus removing the degeneracy due to translational invariance.

In step 2, the algorithm computes the eigen-decomposition of \( Q \) and retains the \( d \) largest eigenvalues. This is equivalent to computing the best rank-\( d \) approximation

\(^1\)Estimates of this type will be repeatedly proved in what follows.
Algorithm 1 SDP-based Algorithm for Localization

**Input:** dimension $d$, distance measurements $\tilde{d}_{ij}$ for $(i,j) \in E$, bound on the measurement noise $\Delta$

**Output:** estimated coordinates in $\mathbb{R}^d$

1: Solve the following SDP problem:

$$\begin{align*}
\text{minimize} & \quad \text{Tr}(Q) \\
\text{s.t.} & \quad \left| \langle M_{ij}, Q \rangle - \tilde{d}_{ij} \right|^2 \leq \Delta, \quad (i,j) \in E \\
                & \quad Q \succeq 0.
\end{align*}$$

2: Compute the best rank-$d$ approximation $U_d \Sigma_d U_d^T$ of $Q$

3: Return $\hat{X} = U_d \Sigma_d^{1/2}$.

of $Q$ in Frobenius norm. The center of gravity of the reconstructed points remains at the origin after this operation.

Our main result characterizes the robustness properties of the SDP-based algorithm. Hereafter “with high probability (w.h.p.)” means with probability converging to 1 as $n \to \infty$ for $d$ fixed.

**Theorem 1.1** Let $\{x_1, \ldots, x_n\}$ be $n$ nodes distributed uniformly at random in the hypercube $[-0.5, 0.5]^d$. Further, assume connectivity radius $r \geq \alpha (\log n / n)^{1/d}$, with $\alpha \geq 10\sqrt{d}$. Then w.h.p., the error distance between the estimate $\hat{X}$ returned by the SDP-based algorithm and the correct coordinate matrix $X$ is upper bounded as

$$d(X, \hat{X}) \leq C_1 (nr^d)^{5} \frac{\Delta}{r^4}. \quad (3)$$

Conversely, w.h.p., there exist adversarial measurement errors $\{z_{ij}\}_{(i,j) \in E}$ such that

$$d(X, \hat{X}) \geq C_2 \min \left\{ \frac{\Delta}{r^4}, 1 \right\}. \quad (4)$$

Here, $C_1$ and $C_2$ denote universal constants that depend only on $d$.

The proof of this theorem relies on several technical results of independent interest. First, we will prove a general deterministic error estimate in terms of the condition number of the stress matrix of the graph $G$; see Theorem 5.1. Next we will use probabilistic arguments to control the stress matrix of random geometric graphs (Theorem 5.2). Finally, we will prove several estimates on the rigidity matrix of $G$, cf. in particular Theorem 6.1. The necessary background in rigidity theory is summarized in Sect. 2.1.

1.3 Related Work

The localization problem and its variants have attracted significant interest over the past years due to their applications in numerous areas, e.g., sensor network localization [6], nuclear magnetic resonance (NMR) spectroscopy [14], and manifold learning [19, 23]. Of particular interest to our work are the algorithms proposed for the
localization problem [6, 16, 21, 24]. In general, few performance guarantees have been proved for these algorithms, in particular in the presence of noise.

The existing algorithms can be categorized into two groups. The first group consists of algorithms that try first to estimate the missing distances and then use multidimensional scaling (MDS) to find the positions from the reconstructed distance matrix [10, 16]. MDS-MAP [10] and ISOMAP [23] are two well-known examples of this class where the missing entries of the distance matrix are approximated by computing the shortest paths between all pairs of nodes. The algorithms in the second group formulate the localization problem as a non-convex optimization problem and then use different relaxation schemes to solve it. An example of this type is relaxation to an SDP [1, 6, 22, 24, 25]. A crucial assumption in these works is the existence of some anchors among the nodes whose exact positions are known. The SDP is then used to efficiently check whether the graph is uniquely \( d \)-localizable and to find its unique realization.

Maximum variance unfolding (MVU) is an SDP-based algorithm with a very similar “flavor” as ours [24]. MVU is an approach to solving dimensionality reduction problems using local metric information and is based on the following simple interpretation. Assume \( n \) points lying on a low-dimensional manifold in a high-dimensional ambient space. In order to find a low-dimensional representation of this data set, the algorithm attempts to somehow unfold the underlying manifold. To this end, MVU pulls the points apart in the ambient space, maximizing the total sum of their pairwise distances, while respecting the local information. However, to the best of our knowledge, no performance guarantee has been proved for the MVU algorithm.

Given the large number of applications and computational methods developed in this broad area, the present paper is in many respects a first step. While we focus on a specific model and a relatively simple algorithm, we expect that the techniques developed here will be applicable to a broader setting, and to a number of algorithms in the same class.

1.4 Organization of the Paper

The remainder of this paper is organized as follows. Section 2 is a brief review of some notions in rigidity theory and some properties of \( G(n, r) \) which will be useful in this paper. In Sect. 3, we discuss the implications of Theorem 1.1 in different applications. The proof of Theorem 1.1 (upper bound) is given in Sect. 4. Sections 5 and 6 contain the proof of two important lemmas used in proving Theorem 1.1. Several technical steps are discussed in the Appendices. Finally, we prove Theorem 1.1 (lower bound) in Sect. 7.

For the reader’s convenience, an overview of the symbols used throughout this paper is given in Table 1 in Appendix N.

2 Preliminaries

2.1 Rigidity Theory

Rigidity theory studies whether a given partial set of pairwise distances \( d_{ij} = \|x_i - x_j\| \) between a finite set of nodes in \( \mathbb{R}^d \) uniquely determines the coordinates
of the points up to rigid transformations. This section is a very brief overview of definitions and results in rigidity theory which will be useful in this paper. We refer the interested reader to [2, 13] for a thorough discussion.

A framework \( GX \) in \( \mathbb{R}^d \) is an undirected graph \( G = (V, E) \) along with a configuration \( X \in \mathbb{R}^{n \times d} \) which assigns a point \( x_i \in \mathbb{R}^d \) to each vertex \( i \) of the graph. The edges of \( G \) correspond to the distance constraints. In the following, we discuss two important notions, the rigidity matrix and the stress matrix. As mentioned above, a crucial part of the proof of Theorem 1.1 consists in establishing some properties of the stress matrix and of the rigidity matrix of the random geometric graph \( G(n, r) \).

**Rigidity Matrix** Consider a motion of the framework with \( x_i(t) \) being the position vector of point \( i \) at time \( t \). Any smooth motion that instantaneously preserves the distance \( d_{ij} \) must satisfy

\[
\frac{d}{dt} \| x_i(t) - x_j(t) \|^2 = 0 \quad \forall (i, j) \in E,
\]

where \( \dot{x}_i \) is the velocity of the \( i \)th point. Given a framework \( GX \in \mathbb{R}^d \), as a solution \( \dot{X} = [\dot{x}_1^T \dot{x}_2^T \cdots \dot{x}_n^T]^T \), with \( \dot{x}_i \in \mathbb{R}^d \), for the linear system of equations (5) is called an infinitesimal motion of the framework \( GX \). This linear system of equations consists of \( |E| \) equations in \( dn \) unknowns and can be written in the matrix form \( R_G(X)\dot{X} = 0 \), where \( R_G(X) \) is called the \(|E| \times dn \) rigidity matrix of \( GX \).

It is easy to see that for every antisymmetric matrix \( A \in \mathbb{R}^{d \times d} \) and for every vector \( b \in \mathbb{R}^d \), \( \dot{x}_i = Ax_i + b \) is an infinitesimal motion. Notice that these motions are the derivative of rigid transformations. (A corresponds to orthogonal transformations and \( b \) corresponds to translations.) Further, these motions span a \( d(d+1)/2 \) dimensional subspace of \( \mathbb{R}^{dn} \), accounting for \( d(d-1)/2 \) degrees of freedom for orthogonal transformations (corresponding to the choice of \( A \), and \( d \) degrees of freedom for translations (corresponding to the choice of \( b \)). Hence, \( \dim \ker(R_G(X)) \geq d(d+1)/2 \). A framework is said to be infinitesimally rigid if \( \dim \ker(R_G(X)) = d(d+1)/2 \).

**Stress Matrix** A stress for a framework \( GX \) is an assignment of scalars \( \omega_{ij} \) to the edges such that for each \( i \in V \),

\[
\sum_{j:(i,j) \in E} \omega_{ij}(x_i - x_j) = \left( \sum_{j:(i,j) \in E} \omega_{ij} \right) x_i - \sum_{j:(i,j) \in E} \omega_{ij} x_j = 0.
\]

A stress vector can be rearranged into an \( n \times n \) symmetric matrix \( \Omega \), known as the stress matrix, such that for \( i \neq j \), the \((i, j)\) entry of \( \Omega \) is \( \Omega_{ij} = -\omega_{ij} \), and the diagonal entries for \((i, i)\) are \( \Omega_{ii} = \sum_{j:j \neq i} \omega_{ij} \). Since all the coordinate vectors of the configuration as well as the all-ones vector are in the null space of \( \Omega \), the rank of the stress matrix for generic configurations is at most \( n - d - 1 \).

There is an important relation between stress matrices of a framework and the notion of global rigidity. A framework \( GX \) is said to be globally rigid in \( \mathbb{R}^d \) if all frameworks in \( \mathbb{R}^d \) with the same set of edge lengths are congruent to \( GX \), i.e. are a rigid transformation of \( GX \). Further, a framework \( GX \) is generically globally rigid in
If $G_X$ is globally rigid at all generic configurations $X$. (Recall that a configuration of points is called generic if the coordinates of the points do not satisfy any nonzero polynomial equation with integer coefficients.)

The connection between global rigidity and stress matrices is demonstrated in the following two results proved in [9] and [13].

**Theorem 2.1** (Connelly, 2005) If $X$ is a generic configuration in $\mathbb{R}^d$ with a stress matrix $\Omega$ of rank $n - d - 1$, then $G_X$ is globally rigid in $\mathbb{R}^d$.

**Theorem 2.2** (Gortler, Healy, Thurston, 2010) Suppose that $X$ is a generic configuration in $\mathbb{R}^d$, such that $G_X$ is globally rigid in $\mathbb{R}^d$. Then either $G_X$ is a simplex or it has a stress matrix $\Omega$ with rank $n - d - 1$.

Among other results in this paper, we construct a special stress matrix $\Omega$ for the random geometric graph $G(n,r)$. We also provide an upper bound and a lower bound on the maximum and the minimum nonzero singular values of this stress matrix. These bounds are used in proving Theorem 1.1.

### 2.2 Some Properties of $G(n,r)$

In this section, we study some of the basic properties of $G(n,r)$ which will be used several times throughout the paper. Our first remark provides probabilistic bounds on the number of nodes contained in a region $\mathcal{R} \subseteq [-0.5, 0.5]^d$.

**Remark 2.1** (Sampling Lemma) Let $\mathcal{R}$ be a measurable subset of the hypercube $[-0.5, 0.5]^d$, and let $V(\mathcal{R})$ denote its volume. Assume $n$ nodes are deployed uniformly at random in $[-0.5, 0.5]^d$, and let $n(\mathcal{R})$ be the number of nodes in region $\mathcal{R}$. Then,

$$n(\mathcal{R}) \in nV(\mathcal{R}) + \left[-\sqrt{2cnV(\mathcal{R}) \log n}, \sqrt{2cnV(\mathcal{R}) \log n}\right],$$

with probability at least $1 - 2/n^c$.

The proof is immediate and deferred to Appendix A.

In the graph $G(n,r)$, every node is connected to all the nodes within its $r$-neighborhood. Using Remark 2.1 for the $r$-neighborhood of each node, and the fact that $r \geq 10\sqrt{d(\log n/n)^{1/d}}$, we obtain the following corollary after applying a union bound over all the $r$-neighborhoods of the nodes.

**Corollary 2.1** In the graph $G(n,r)$, with $r \geq 10\sqrt{d(\log n/n)^{1/d}}$, the degrees of all nodes are in the interval $[(1/2)K_d n r^d, (3/2)K_d n r^d]$, with high probability. Here, $K_d$ is the volume of the $d$-dimensional unit ball.

Next, we discuss some properties of the spectrum of $G(n,r)$.

Recall that the Laplacian $\mathcal{L}$ of the graph $G$ is the symmetric matrix indexed by the vertices $V$, such that $\mathcal{L}_{ij} = -1$ if $(i, j) \in E$, $\mathcal{L}_{ii} = \text{degree}(i)$ and $\mathcal{L}_{ij} = 0$ otherwise. The all-ones vector $u \in \mathbb{R}^n$ is an eigenvector of $\mathcal{L}(G)$ with eigenvalue 0.
Further, the multiplicity of eigenvalue 0 in the spectrum of $L(G)$ is equal to the number of connected components in graph $G$. Let us stress that our definition of $L(G)$ has opposite sign with respect to the one adopted in part of the computer science literature. In particular, with the present definition, $L(G)$ is a positive semidefinite matrix.

It is useful to recall a basic estimate on the Laplacian of random geometric graphs.

**Remark 2.2** Let $L_n$ denote the normalized Laplacian of the random geometric graph $G(n,r)$, defined as $L_n = D^{-1/2}L D^{-1/2}$, where $D$ is the diagonal matrix with the node degrees on the diagonal. Then, w.h.p., $\lambda_2(L_n)$, the second smallest eigenvalue of $L_n$, is at least $Cr^2$ ([7, 18]). Also, using the result of [8] (Theorem 4) and Corollary 2.1, we have $\lambda_2(L) \geq C(nrd)r^2$, for some constant $C = C(d)$.

### 2.3 Notation

For a vector $v \in \mathbb{R}^n$ and a subset $T \subseteq \{1, \ldots, n\}$, $v_T \in \mathbb{R}^T$ is the restriction of $v$ to indices in $T$. We use the notation $\langle v_1, \ldots, v_n \rangle$ to represent the subspace spanned by vectors $v_i$, $1 \leq i \leq n$. The orthogonal projections onto subspaces $V$ and $V^\perp$ are respectively denoted by $P_V$ and $P_V^\perp$. The identity matrix, in any dimension, is denoted by $I$.

Further, $e_i$ always refers to the $i$th standard basis element, e.g., $e_1 = (1, 0, \ldots, 0)$. Throughout this paper, $u \in \mathbb{R}^n$ is the all-ones vector and $C$ is a constant depending only on the dimension $d$, whose value may change from case to case.

For a matrix $A$, we denote its operator norm by $\|A\|_2$, its Frobenius norm by $\|A\|_F$, its nuclear norm by $\|A\|_*$, and its $\ell_1$-norm by $\|A\|_1$. ($\|A\|_*$ is simply the sum of the singular values of $A$ and $\|A\|_1 = \sum_{ij} |A_{ij}|$.) We also use $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ to respectively denote the maximum and the minimum nonzero singular values of $A$.

For a graph $G$, we denote by $V(G)$ the set of its vertices and we use $E(G)$ to denote the set of edges in $G$. Following the convention adopted above, the Laplacian of $G$ is represented by $L(G)$.

Finally, we denote by $x(i) \in \mathbb{R}^n$, $i \in \{1, \ldots, d\}$ the $i$th column of the positions matrix $X$. In other words $x(i)$ is the vector containing the $i$th coordinate of points $x_1, \ldots, x_n$.

Throughout the proof we shall adopt the convention of using the notation $X$, $\{x_j\}_{j \in [n]}$, and $\{x(i)\}_{i \in [d]}$ to denote the centered positions. In other words $X = LX'$, where the rows of $X'$ are i.i.d. uniform in $[-0.5, 0.5]^d$.

### 3 Discussion

In this section, we make some remarks about Theorem 1.1 and its implications.

**Tightness of the Bounds** The upper and the lower bounds in Theorem 1.1 match up to the factor $C(nr^d)^5$. Note that $nr^d$ is the average degree of the nodes in $G$ (up to a constant) and when the range $r$ is of the same order as the connectivity threshold, i.e., $r = O((\log n/n)^{1/d})$, it is logarithmic in $n$. Furthermore, we believe that this factor is an artifact of our analysis. The numerical experiments in Sect. 8.

© Springer
also support the idea that the performance of the SDP-based algorithm, evaluated by 
\( d(X, \hat{X}) \), scales as \( C\Delta/r^4 \) for some constant \( C \). In addition, the theorem states the bounds for \( r \geq \alpha(\log n/n)^{1/d} \), with \( \alpha \geq 10\sqrt{d} \). However, numerical experiments in Sect. 8 show that the bounds hold for much smaller \( \alpha \), namely \( \alpha \geq 3 \) for \( d = 2, 4 \). Finally, one sees immediately that, under the worst-case model for the measurement errors, no algorithm can perform better than \( C\Delta/r^2 \). More specifically, for any algorithm \( d(X, \hat{X}) \geq C\Delta/r^2 \), for some constant \( C \). The reason is that letting \( \hat{d}_{ij} = (1 + \Delta/r^2)d_{ij} \), no algorithm can differentiate between \( X \) and its scaled version \( \hat{X} = \sqrt{1+\Delta/r^2}X \). Also \( d(X, \hat{X}) = (\Delta/r^2)(1/n^2)\|LXX^TL\|_1 \geq C\Delta/r^2 \), w.h.p. and for some constant \( C \) that depends on the dimension \( d \).

**Global Rigidity of \( G(n, r) \)** As a special case of Theorem 1.1 we can consider the problem of reconstructing the point positions from exact measurements. The case of exact measurements was also studied recently in [20] following a different approach. This corresponds to setting \( \Delta = 0 \). The underlying question is whether the point positions \( \{x_i\}_{i \in V} \) can be efficiently determined (up to a rigid motion) by the set of distances \( \{d_{ij}\}_{(i, j) \in E} \). If this is the case, then, in particular, the random graph \( G(n, r) \) is globally rigid.

Since the right-hand side of our error bound Eq. (3) vanishes for \( \Delta = 0 \), we immediately obtain the following.

**Corollary 3.1** Let \( \{x_1, \ldots, x_n\} \) be \( n \) nodes distributed uniformly at random in the hypercube \([-0.5, 0.5]^d \). If \( r \geq 10\sqrt{d}(\log n/n)^{1/d} \), and the distance measurements are exact, then w.h.p., the SDP-based algorithm recovers the exact positions (up to rigid transformations). In particular, the random geometric graph \( G(n, r) \) is w.h.p. globally rigid if \( r \geq 10\sqrt{d}(\log n/n)^{1/d} \).

In [3], the authors prove a similar result on the global rigidity of \( G(n, r) \). Namely, they show that if \( n \) points are drawn from a Poisson process in \([0, 1]^2 \), then the random geometric graph \( G(n, r) \) is globally rigid w.h.p. when \( r \) is of the order \( \sqrt{\log n/n} \).

As already mentioned above, the graph \( G(n, r) \) is disconnected w.h.p. if \( r \leq K_d^{-1/d}(\log n + C)/n^{1/d} \) for some constant \( C \). Hence, our result establishes the following **rigidity phase transition** phenomenon: There exist dimension-dependent constants \( C_1(d), C_2(d) \) such that a random geometric graph \( G(n, r) \) is w.h.p. not globally rigid if \( r \leq C_1(d)(\log n/n)^{1/d} \), and w.h.p. globally rigid if \( r \geq C_2(d)(\log n/n)^{1/d} \). Applying Stirling’s formula, it is easy to see that the above arguments yield \( C_1(d) \geq C_{1, *} \sqrt{d} \) and \( C_2(d) \leq C_{2, *} \sqrt{d} \) for some numerical (dimension-independent) constants \( C_{1, *}, C_{2, *} \).

It is natural to conjecture that the rigidity phase transition is sharp.

**Conjecture 1** Let \( G(n, r_n) \) be a random geometric graph with \( n \) nodes, and range \( r_n \), in \( d \) dimensions. Then there exists a constant \( C_*(d) \) such that, for any \( \varepsilon > 0 \), the following happens. If \( r_n \leq (C_*(d) - \varepsilon)(\log n/n)^{1/d} \), then \( G(n, r_n) \) is w.h.p. not globally rigid. If \( r_n \geq (C_*(d) + \varepsilon)(\log n/n)^{1/d} \), then \( G(n, r_n) \) is w.h.p. globally rigid.
Sensor Network Localization  Research in this area aims at developing algorithms and systems to determine the positions of the nodes of a sensor network exploiting inexpensive distributed measurements. Energy and hardware constraints rule out the use of global positioning systems, and several proposed systems exploit pairwise distance measurements between the sensors [15, 17]. These techniques have acquired new industrial interest due to their relevance to indoor positioning. In this context, global positioning systems are not a method of choice because of their limited accuracy in indoor environments.

Semidefinite programming (SDP) methods for sensor network localization have been developed starting with [6]. It is common to study and evaluate different techniques within the random geometric graph model, but no performance guarantees have been proven for advanced (SDP-based) algorithms, with inaccurate measurements. We shall therefore consider $n$ sensors placed uniformly at random in the unit hypercube, with ambient dimension either $d = 2$ or $d = 3$ depending on the specific application. The connectivity range $r$ is dictated by various factors: power limitations; interference between nearby nodes; loss of accuracy with distance.

The measurement error $z_{ij}$ depends on the method used to measure the distance between nodes $i$ and $j$. We will limit ourselves to measurement errors due to noise (as opposed—for instance—to malicious behavior of the nodes) and discuss two common techniques for measuring distances between wireless devices: received signal strength indicator (RSSI) and time difference of arrival (TDOA). The RSSI technique measures the ratio of the power present in a received radio signal ($P_r$) and a reference transmitted power ($P_s$). The ratio $P_r / P_s$ is inversely proportional to the square of the distance between the receiver and the transmitter. Hence, RSSI can be used to estimate the distance. It is reasonable to assume that the dominant error is in the measurement of the received power, and that it is proportional to the transmitted power. We thus assume that there is an error $\varepsilon P_s$ in measuring the received power $P_r$, i.e., $\tilde{P}_r = P_r + \varepsilon P_s$, where $\tilde{P}_r$ denotes the measured received power. Then, the measured distance is given by

$$\tilde{d}_{ij}^2 \propto \frac{P_s}{\tilde{P}_r} = \frac{P_s}{P_r} \left(1 + \frac{P_s}{P_r} \varepsilon \right)^{-1} \approx \frac{P_s}{P_r} \left(1 - \frac{P_r}{P_r} \varepsilon \right) \propto d_{ij}^2 \left(1 - C d_{ij}^2 \varepsilon \right). \quad (7)$$

Therefore the overall error $|z_{ij}| \propto d_{ij}^4 \varepsilon$ and its magnitude is $\Delta \propto r^4 \varepsilon$. Applying Theorem 1.1, we obtain an average error per node of order

$$d(X, \hat{X}) \leq C'_1 \left(nr^d \right)^{\frac{5}{2}} \varepsilon.$$  

In other words, the positioning accuracy is linear in the measurement accuracy, with a proportionality constant that is polynomial in the average node degree. Remarkably, the best accuracy is obtained by using the smallest average degree, i.e., the smallest measurement radius that is compatible with connectivity.

The TDOA technique uses the time difference between the receipt of two different signals with different velocities, for instance, ultrasound and radio signals. The time difference is proportional to the distance between the receiver and the transmitter, and, given the velocity of the signals, the distance can be estimated from the time difference. Now, assume that there is a relative error $\varepsilon$ in measuring this time
difference (this might be related to inaccuracies in ultrasound speed). We thus have \( \tilde{t}_{ij} = t_{ij}(1 + \varepsilon) \), where \( \tilde{t}_{ij} \) is the measured time while \( t_{ij} \) is the “ideal” time difference. This leads to an error in estimating \( d_{ij} \) which is proportional to \( d_{ij}\varepsilon \). Therefore, \( |z_{ij}| \propto d_{ij}^2 \varepsilon \) and \( \Delta \propto r^2 \varepsilon \). Again applying Theorem 1.1, we obtain an average error per node of order

\[
\Delta \leq C(nrd)^{5/2} \varepsilon \frac{r}{r_0^2}.
\]

In other words, the reconstruction error decreases with the measurement radius, which suggests a somewhat different network design for such a system.

Note that the above error bounds are proved under an adversarial error model (see below). It would be useful to complement them with similar analyses carried out for other, more realistic, models.

**Manifold Learning**

Manifold learning deals with finite data sets of points in ambient space \( \mathbb{R}^N \) which are assumed to lie on a smooth submanifold \( M^d \) of dimension \( d < N \). The task is to recover \( M \) given only the data points. Here, we discuss the implications of Theorem 1.1 for applications of SDP methods to manifold learning.

It is typically assumed that the manifold \( M^d \) is isometrically equivalent to a region in \( \mathbb{R}^d \). For simplicity we shall assume that this region is convex (see [12] for a discussion of this point). With little loss of generality we can identify the region with the unit hypercube \( [-0.5, 0.5]^d \). A typical manifold learning algorithm ([23] and [24]) estimates the geodesic distances between a subset of pairs of data points \( d_M(y_i, y_j), y_i \in \mathbb{R}^N, \) and then tries to find a low-dimensional embedding (i.e., positions \( x_i \in \mathbb{R}^d \)) that reproduces these distances.

The unknown geodesic distance between nearby data points \( y_i \) and \( y_j \), denoted by \( d_M(y_i, y_j) \), can be estimated by their Euclidean distance in \( \mathbb{R}^n \). Therefore, the manifold learning problem reduces mathematically to the localization problem whereby the distance “measurements” are \( \tilde{d}_{ij} = \|y_i - y_j\|_{\mathbb{R}^N} \), while the actual distances are \( d_{ij} = d_M(y_i, y_j) \). The accuracy of these estimates depends on the curvature of the manifold \( M \). Let \( r_0 = r_0(M) \) be the minimum radius of curvature defined by

\[
\frac{1}{r_0} = \max_{\gamma, t} \{\|\dot{\gamma}(t)\|\},
\]

where \( \gamma \) varies over all unit-speed geodesics in \( M \) and \( t \) is in the domain of \( \gamma \). For instance, a Euclidean sphere of radius \( r_0 \) has minimum radius of curvature equal to \( r_0 \).

As shown in [5, Lemma 3], \( (1 - d_{ij}^2/24r_0^2)d_{ij} \leq \tilde{d}_{ij} \leq d_{ij} \). Therefore, \( |z_{ij}| \propto d_{ij}^3/r_0^2 \), and \( \Delta \propto r^4/r_0^2 \). Theorem 1.1 supports the claim that the estimation error \( d(X, \hat{X}) \) is bounded by \( C(nrd)^{5/2}/r_0^2 \).

As mentioned several times, this paper focuses on a particularly simple SDP relaxation and noise model. This opens the way to a number of interesting directions.

1. **Stochastic Noise Models.** A somewhat complementary direction to the one taken here would be to assume that the distance measurements are \( \tilde{d}_{ij} = d_{ij}^2 + z_{ij} \) with \( \{z_{ij}\} \) a collection of independent zero-mean random variables. This would be a good model, for instance, for errors in RSSI measurements.
Another interesting case would be the one in which a small subset of measurements is grossly incorrect (e.g., due to node malfunctioning, obstacles, etc.).

2. Tighter Convex Relaxations. The relaxation considered here is particularly simple, and can be improved in several ways. For instance, in manifold learning it is useful to maximize the embedding variance $\text{Tr}(Q)$ under the constraint $Qu = 0$ [24].

Also, for any pair $(i, j) \notin E$ it is possible to add a constraint of the form $(M_{ij}, Q) \leq \tilde{d}_{ij}^2$, where $\tilde{d}_{ij}$ is an upper bound on the distance obtained by computing the shortest path between $i$ and $j$ in $G$.

3. More General Geometric Problems. This paper analyzes the problem of reconstructing the geometry of a cloud of points from incomplete and inaccurate measurements of the points’ local geometry. From this point of view, a number of interesting extensions can be explored. For instance, instead of distances, it might be possible to measure angles between edges in the graph $G$ (in sensor networks, angles of arrival might be available [15, 17]).

4 Proof of Theorem 1.1 (Upper Bound)

Let $V = \langle u, x^{(1)}, \ldots, x^{(d)} \rangle$, and for any matrix $S \in \mathbb{R}^{n \times n}$, define

$$
\tilde{S} = P_V SP_V + P_V SP_V^\perp + P_V^\perp SP_V,
$$

$$
S^\perp = P_V^\perp SP_V^\perp.
$$

(8)

Thus $S = \tilde{S} + S^\perp$. Also, denote by $R$ the difference between the optimum solution $Q$ and the actual Gram matrix $Q_0$, i.e., $R = Q - Q_0$. The proof of Theorem 1.1 is based on the following key lemmas that bound $R^\perp$ and $\tilde{R}$ separately.

**Lemma 4.1** There exists a numerical constant $C = C(d)$ such that, w.h.p.,

$$
\|R^\perp\|_* \leq C \frac{n}{r^4} (nr^d)^5 \Delta.
$$

(9)

**Lemma 4.2** There exists a numerical constant $C = C(d)$ such that, w.h.p.,

$$
\|\tilde{R}\|_1 \leq C \frac{n^2}{r^4} (nr^d)^5 \Delta.
$$

(10)

We defer the proof of Lemmas 4.1 and 4.2 to the next sections.

**Proof of Theorem 1.1** Let $Q = \sum_{i=1}^n \sigma_i u_i u_i^T$, where $\|u_i\| = 1$, $u_i^T u_j = 0$ for $i \neq j$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Let $P_d(Q) = \sum_{i=1}^d \sigma_i u_i u_i^T$ be the best rank-$d$ approximation of $Q$ in Frobenius norm (step 2 in the algorithm). Recall that $Q u = 0$, because $Q$ minimizes $\text{Tr}(Q)$. Consequently, $P_d(Q) u = 0$ and $P_d(Q) = L P_d(Q) L$. Further, by our assumption $Q_0 u = 0$ and thus $Q_0 = L Q_0 L$. Using the triangle inequality,

$$
\|L P_d(Q) L - L Q_0 L\|_1 = \|P_d(Q) - Q_0\|_1
\leq \|P_d(Q) - \tilde{Q}\|_1 + \|\tilde{Q} - Q_0\|_1.
$$

(11)
Observe that $\tilde{Q} = Q_0 + \tilde{R}$ and $Q^\perp = R^\perp$. Since $P_d(Q) - \tilde{Q}$ has rank at most $3d$, it follows that $\|P_d(Q) - \tilde{Q}\|_1 \leq n\|P_d(Q) - \tilde{Q}\|_F \leq \sqrt{3dn}\|P_d(Q) - \tilde{Q}\|_2$ (for any matrix $A$, $\|A\|_F^2 \leq \text{rank}(A)\|A\|_2^2$). By the triangle inequality, we have

$$\|P_d(Q) - \tilde{Q}\|_2 \leq \|P_d(Q) - Q\|_2 + \|Q - \tilde{Q}\|_2. \tag{12}$$

Note that $\|P_d(Q) - Q\|_2 = \sigma_{d+1}$. Recall the variational principle for the eigenvalues:

$$\sigma_q = \min_{H, \dim(H) = n-d+1} \max_{y \in H, \|y\| = 1} y^T Q y.$$

Taking $H = \langle x(1), \ldots, x(d) \rangle^\perp$, for any $y \in H$, $y^T Q y = y^T P_\perp V Q P_\perp V y = y^T Q^\perp y$, where we used the fact $Qu = 0$ in the first equality. Therefore, $\sigma_{d+1} \leq \max_{\|y\| = 1} y^T R^\perp y = \|R^\perp\|_2$. It follows from Eqs. (11) and (12) that

$$\|LP_d(Q) L - LQ_0 L\|_1 \leq 2\sqrt{3dn}\|R^\perp\|_2 + \|\tilde{R}\|_1.$$

Using Lemmas 4.1 and 4.2, we obtain

$$d(X, \hat{X}) = \frac{1}{n^2} \|LP_d(Q) L - LQ_0 L\|_1 \leq C(nr^d)^{s\Delta / r^4},$$

which proves the claimed upper bound on the error.

The lower bound is proved in Sect. 7. \hfill \Box

5 Proof of Lemma 4.1

The proof is based on the following three steps: (i) Upper bound $\|R^\perp\|_\ast$ in terms of $\sigma_{\min}(\Omega)$ and $\sigma_{\max}(\Omega)$, where $\Omega$ is an arbitrary positive semidefinite (PSD) stress matrix of rank $n - d - 1$ for the framework; (ii) construct a particular PSD stress matrix $\Omega$ of rank $n - d - 1$ for the framework; (iii) upper bound $\sigma_{\max}(\Omega)$ and lower bound $\sigma_{\min}(\Omega)$.

**Theorem 5.1** Let $\Omega$ be an arbitrary PSD stress matrix for the framework $G_X$ such that $\text{rank}(\Omega) = n - d - 1$. Then,

$$\|R^\perp\|_\ast \leq 2\frac{\sigma_{\max}(\Omega)}{\sigma_{\min}(\Omega)} |E| \Delta. \tag{13}$$

**Proof** Note that $R^\perp = Q^\perp = P_\perp V Q P_\perp V \geq 0$. Write $R^\perp = \sum_{i=1}^{n-d-1} \lambda_i u_i u_i^T$, where $\|u_i\| = 1$, $u_i^T u_j = 0$ for $i \neq j$, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-d-1} \geq 0$. Therefore,

$$\langle \Omega, R^\perp \rangle = \left( \Omega, \sum_{i=1}^{n-d-1} \lambda_i u_i u_i^T \right) = \sum_{i=1}^{n-d-1} \lambda_i u_i^T \Omega u_i \geq \sigma_{\min}(\Omega) \|R^\perp\|_\ast. \tag{14}$$
Here, we used the fact that $u_i \in V^\perp = \text{Ker}^\perp(\Omega)$. Note that $\sigma_{\text{min}}(\Omega) > 0$, since $\Omega \succeq 0$.

Now, we need to upper bound the quantity $\langle \Omega, R^\perp \rangle$. Since $\Omega u = 0$, the stress matrix $\Omega = [\omega_{ij}]$ can be written as $\Omega = \sum_{(i,j) \in E} \omega_{ij} M_{ij}$. Define $\omega_{\text{max}} = \max_{i \neq j} |\omega_{ij}|$. Then,

$$\langle \Omega, R^\perp \rangle \leq \sum_{(i,j) \in E} \omega_{\text{max}} \left| \langle M_{ij}, Q \rangle - Q_0 \right|$$

$$\leq \sum_{(i,j) \in E} \omega_{\text{max}} \left( \left| \langle M_{ij}, Q \rangle - 2d_{ij} \right| + \sum_{z_{ij}} \left| d_{ij} - d_{ij} \right| \right)$$

$$\leq 2\omega_{\text{max}} |E| \Delta,$$

where (a) follows from the fact that $\langle PV, \Omega \rangle = 0$. Since $\Omega \succeq 0$, we have $\omega_{ij}^2 \leq \omega_{ii} \omega_{jj} = (e_i^T \Omega e_i)(e_j^T \Omega e_j) \leq \sigma_{\text{max}}^2(\Omega)$, for $1 \leq i, j \leq n$. Hence, $\omega_{\text{max}} \leq \sigma_{\text{max}}(\Omega)$. Combining Eqs. (14) and (15), we get the desired result. \hfill \Box

The next step is constructing a PSD stress matrix of rank $n - d - 1$. For each node $i \in V(G)$ define $C_i = \{ j \in V(G) : d_{ij} \leq r/2 \}$. Note that the nodes in each $C_i$ form a clique in $G$. In addition, let $S_i$ be the following set of cliques:

$$S_i := \bigcup_{k \in C_i} C_i \setminus k \cup \{ C_i \}.$$  

Therefore, $S_i$ is a set of $|C_i| + 1$ cliques. For the graph $G$, we define $\text{cliq}(G) := S_1 \cup \cdots \cup S_n$. The next proposition establishes a simple property of cliques $C_i$. Its proof is immediate and deferred to Appendix B.

**Proposition 5.1** If $r \geq 4c \sqrt{d} (\log n/n)^{1/d}$ with $c > 1$, the following is true w.h.p. For any two nodes $i$ and $j$, such that $||x_i - x_j|| \leq r/2$, $|C_i \cap C_j| \geq d + 1$.

Now we are ready to construct a special stress matrix $\Omega$ of $G_X$. Define the $|Q_k| \times |Q_k|$ matrix $\Omega_k$ as follows:

$$\Omega_k = \sum_{Q_k \in \text{cliq}(G)} \hat{\Omega}_k.$$

Let $\hat{\Omega}_k$ be the $n \times n$ matrix obtained from $\Omega_k$ by padding it with zeros. Define

$$\Omega = \sum_{Q_k \in \text{cliq}(G)} \hat{\Omega}_k.$$

The proof of the next statement is again immediate and is given in Appendix C.
Proposition 5.2 The matrix $\Omega$ defined above is a PSD stress matrix for the framework $G_X$. Further, almost surely, $\text{rank}(\Omega) = n - d - 1$.

The final step is to upper bound $\sigma_{\text{max}}(\Omega)$ and lower bound $\sigma_{\text{min}}(\Omega)$.

Claim 5.1 There exists a constant $C = C(d)$ such that, w.h.p.,

$$\sigma_{\text{max}}(\Omega) \leq C(nr^d)^2.$$  

Proof For any vector $v \in \mathbb{R}^n$,

$$v^T \Omega v = \sum_{Q_k \in \text{cliq}(G)} v^T \hat{\Omega}_k v = \sum_{Q_k \in \text{cliq}(G)} \|\hat{\Omega}_k v\|^2 = \sum_{Q_k \in \text{cliq}(G)} \|P_{u_{Q_k}} x_{u_{Q_k}}^{(1)} ... x_{u_{Q_k}}^{(d)} v_{Q_k}\|^2 \leq \sum_{Q_k \in \text{cliq}(G)} \|v_{Q_k}\|^2 = \sum_{j=1}^n v_j^2 \sum_{k: j \in Q_k} 1 = \sum_{j=1}^n \left( \sum_{i \in C_j} |C_i| \right) v_j^2 \leq (Cnr^d \|v\|^2).$$

The last inequality follows from the fact that, w.h.p., $|C_j| \leq Cnr^d$ for all $j$ and some constant $C$ (see Corollary 2.1). \hfill \Box

We now pass to lower bounding $\sigma_{\text{min}}(\Omega)$.

Theorem 5.2 There exists a constant $C = C(d)$ such that, w.h.p., $\Omega \perp \succeq C(nr^d)^{-3} \times r^2 L^\perp$ (see Eq. (8)).

The proof is given in Sect. 5.1. We are finally in position to prove Lemma 4.1.

Proof of Lemma 4.1 Following Theorem 5.2 and Remark 2.2, we obtain $\sigma_{\text{min}}(\Omega) \geq C(nr^d)^{-2}r^4$. Also, by Corollary 2.1, w.h.p., the node degrees in $G$ are bounded by $3/2 K_d nr^d$. Hence, w.h.p., $|E| \leq 3/4n^2 K_d r^d$. Using the bounds on $\sigma_{\text{max}}(\Omega)$, $\sigma_{\text{min}}(\Omega)$, and $|E|$ in Theorem 5.1 yields the thesis. \hfill \Box

5.1 Proof of Theorem 5.2

Before proceeding with the proof, we note that the authors in [4] propose a heuristic argument showing $\Omega v \approx L^2 v$ for smoothly varying vectors $v$. Since $\sigma_{\text{min}}(L) \geq C(nr^d)r^2$ (see Remark 2.2), this heuristic supports the claim of the theorem.

In the following, we first establish some claims and definitions which will be used in the proof.

Claim 5.2 There exists a constant $C = C(d)$ such that, w.h.p.,

$$L \leq C \sum_{k=1}^n P_{u_{C_k}}^\perp.$$
The argument is closely related to the Markov chain comparison technique [11]. The proof is given in Appendix D.

The next claim provides a concentration result about the number of nodes in the cliques $\mathcal{C}_i$. Its proof is immediate and deferred to Appendix E.

**Claim 5.3** For every node $i \in V(G)$, define $\tilde{\mathcal{C}}_i = \{ j \in V(G) : d_{ij} \leq \frac{1}{2}(\frac{1}{2} + \frac{1}{100}) \}$. There exists an integer number $m$ such that the following is true w.h.p.:

$$|\tilde{\mathcal{C}}_i| \leq m \leq |\mathcal{C}_i|, \quad \forall i \in V(G).$$

Now, for any node $i$, let $i_1, \ldots, i_m$ denote the $m$ nearest neighbors of that node. Using Claim 5.3, $\tilde{\mathcal{C}}_i \subseteq \{i_1, \ldots, i_m\} \subseteq \mathcal{C}_i$. Define the set $\tilde{S}_i$ as follows:

$$\tilde{S}_i = \{\mathcal{C}_i, \mathcal{C}_i \setminus i_1, \ldots, \mathcal{C}_i \setminus i_m\}.$$

Therefore, $\tilde{S}_i$ is a set of $(m + 1)$ cliques. Let $\text{cliq}^*(G) = \tilde{S}_1 \cup \cdots \cup \tilde{S}_n$. Note that $\text{cliq}^*(G) \subseteq \text{cliq}(G)$. Construct the graph $G^*$ in the following way. For every element in $\text{cliq}^*(G)$, there is a corresponding vertex in $G^*$. (Thus, $|V(G^*)| = n(m + 1)$.) Also, for any two nodes $i$ and $j$, such that $\|x_i - x_j\| \leq r/2$, every vertex in $V(G^*)$ corresponding to an element in $\tilde{S}_i$ is connected to every vertex in $V(G^*)$ corresponding to an element in $\tilde{S}_j$.

Our next claim establishes some properties of the graph $G^*$. For its proof, we refer to Appendix F.

**Claim 5.4** With high probability, the graph $G^*$ has the following properties.

(i) The degree of each node is bounded by $C(nr^d)^2$, for some constant $C = C(d)$.

(ii) Let $L^*$ denote the Laplacian of $G^*$. Then $\sigma_{\text{min}}(L^*) \geq C(nr^d)^2 r^2$, for some constant $C$.

Now, we are in position to prove Theorem 5.2.

**Proof of Theorem 5.2** Let $v \in V^\perp$ be an arbitrary vector. For every clique $Q_i \in \text{cliq}(G)$, decompose $v$ locally as $v_{Q_i} = \sum_{\ell=1}^d \tilde{P}^{(\ell)}_{i \perp} \tilde{x}^{(\ell)}_{Q_i} + \gamma_i u_{Q_i} + w^{(i)}$, where $\tilde{x}^{(\ell)}_{Q_i} = P^{\perp}_{u_{Q_i}} x^{(\ell)}_{Q_i}$ and $w^{(i)} \in \langle x^{(1)}_{Q_i}, \ldots, x^{(d)}_{Q_i}, u_{Q_i} \rangle^\perp$. Hence,

$$v^T \Omega v = \sum_{Q_i \in \text{cliq}(G)} \|w^{(i)}\|^2.$$

Note that $v_{Q_i} \cap Q_j$ has two representations; one is obtained by restricting $v_{Q_i}$ to indices in $Q_j$, and the other is obtained by restricting $v_{Q_j}$ to indices in $Q_i$. From these two representations, we get

$$w^{(i)}_{Q_i \cap Q_j} - w^{(j)}_{Q_i \cap Q_j} = \sum_{\ell=1}^d (\beta^{(\ell)}_j - \beta^{(\ell)}_i) \tilde{x}^{(\ell)}_{Q_i \cap Q_j} + \tilde{y}_{i,j} u_{Q_i \cap Q_j}.$$

(16)
Here, \( \tilde{x}^{(\ell)}_{Q_i \cap Q_j} = P_{u_{Q_i \cap Q_j}} x^{(\ell)}_{Q_i \cap Q_j} \). The value of \( \gamma_{i,j} \) does not matter for our argument; however, it can be given explicitly.

Note that \( \{C_1, \ldots, C_n\} \subseteq \text{cliq}^*(G) \subseteq \text{cliq}(G) \). Invoking Claim 5.2,

\[
v^T \mathcal{L} v \leq C \sum_{k=1}^n \| P_{\perp u_{C_k}} v_{C_k} \|^2 \\
\leq C \sum_{Q_i \in \text{cliq}^*(G)} \| P_{u_{Q_i}} v_{Q_i} \|^2 \\
= C \sum_{Q_i \in \text{cliq}^*(G)} \left\| \sum_{\ell=1}^d \beta_i^{(\ell)} \tilde{x}^{(\ell)}_{Q_i} + w(i) \right\|^2 \\
= C \left( \sum_{Q_i \in \text{cliq}^*(G)} \left\| \sum_{\ell=1}^d \beta_i^{(\ell)} \tilde{x}^{(\ell)}_{Q_i} \right\|^2 + \sum_{Q_i \in \text{cliq}^*(G)} \| w(i) \|^2 \right) \\
\leq C \left( d \sum_{Q_i \in \text{cliq}^*(G)} \sum_{\ell=1}^d \| \beta_i^{(\ell)} \tilde{x}^{(\ell)}_{Q_i} \|^2 + \sum_{Q_i \in \text{cliq}^*(G)} \| w(i) \|^2 \right).
\]

Hence, we only need to show that

\[
\sum_{Q_i \in \text{cliq}(G)} \| w(i) \|^2 \geq C (nr^d)^{-3} r^2 \sum_{Q_i \in \text{cliq}^*(G)} \sum_{\ell=1}^d \| \beta_i^{(\ell)} \tilde{x}^{(\ell)}_{Q_i} \|^2, \tag{17}
\]

for some constant \( C = C(d) \).

In the following we adopt the convention that for \( j \in V(G^*) \), \( Q_j \) is the corresponding clique in \( \text{cliq}^*(G) \). We have

\[
\sum_{Q_i \in \text{cliq}(G)} \| w(i) \|^2 \geq \sum_{Q_i \in \text{cliq}^*(G)} \| w(i) \|^2 = \sum_{i \in V(G^*)} \| w(i) \|^2 \\
\overset{(a)}{\geq} C (nr^d)^{-2} \sum_{(i,j) \in E(G^*)} (\| w(i) \|^2 + \| w(j) \|^2) \\
\overset{(b)}{\geq} C (nr^d)^{-2} \sum_{(i,j) \in E(G^*)} \left\| w_{Q_i \cap Q_j} - w_{Q_i \cap Q_j} \right\|^2 \\
\overset{(c)}{\geq} C (nr^d)^{-1} r^2 \sum_{(i,j) \in E(G^*)} \sum_{\ell=1}^d (\beta_j^{(\ell)} - \beta_i^{(\ell)}) \tilde{x}^{(\ell)}_{Q_i \cap Q_j} \right\|^2 \\
\overset{(d)}{\geq} C (nr^d)^{-1} r^2 \sum_{(i,j) \in E(G^*)} \sum_{\ell=1}^d (\beta_j^{(\ell)} - \beta_i^{(\ell)})^2. \tag{18}
\]
Here, (a) follows from the fact that the degrees of the nodes in $G^*$ are bounded by $C(nr^d)^2$ (Claim 5.4, part (i)); (b) follows from Eq. (16), and (c) follows from Claim 5.5, whose proof is deferred to Appendix G.

Claim 5.5 There exists a constant $C = C(d)$ such that, for any set of values $\{\beta^{(\ell)}_i\}$, the following holds w.h.p.:

$$\left\| \sum_{\ell=1}^{d} (\beta^{(\ell)}_j - \beta^{(\ell)}_i) \tilde{x}^{(\ell)}_{Q_i \cap Q_j} \right\|^2 \geq C(nr^d)r^2 \sum_{\ell=1}^{d} (\beta^{(\ell)}_j - \beta^{(\ell)}_i)^2, \quad \forall (i, j) \in E(G^*).$$

Also note that $\|\tilde{x}^{(\ell)}_{Q_i}\|^2 \leq |Q_i|r^2$ and, w.h.p., $|Q_i| \leq C(nr^d)$, for all $i \in V(G^*)$. Therefore, using Eq. (18), in order to prove (17) it suffices to show that

$$\sum_{\ell=1}^{d} \sum_{(i,j) \in E(G^*)} (\beta^{(\ell)}_j - \beta^{(\ell)}_i)^2 \geq C(nr^d)^r \sum_{\ell=1}^{d} \sum_{i \in V(G^*)} (\beta^{(\ell)}_i)^2.$$

Define $\beta^{(\ell)}_i = (\beta^{(\ell)}_i)_{i \in V(G^*)}$. Observe that

$$\sum_{(i,j) \in E(G^*)} (\beta^{(\ell)}_j - \beta^{(\ell)}_i)^2 = (\beta^{(\ell)})^T L^* \beta^{(\ell)} \geq \sigma_{\text{min}}(L^*) \|P_u \perp \beta^{(\ell)}\|^2.$$

Using Claim 5.4 (part (ii)) we obtain

$$\sum_{(i,j) \in E(G^*)} (\beta^{(\ell)}_j - \beta^{(\ell)}_i)^2 \geq C(nr^d)^r \sum_{\ell=1}^{d} \|P_u \perp \beta^{(\ell)}\|^2.$$

The proof is completed by the following claim, whose proof is given in Appendix H.

Claim 5.6 There exists a constant $C = C(d)$ such that the following holds w.h.p. Consider an arbitrary vector $v \in V^\perp$ with local decompositions $v_{Q_i} = \sum_{\ell=1}^{d} \beta^{(\ell)}_i \tilde{x}^{(\ell)}_{Q_i} + \gamma_i u_{Q_i} + w^{(i)}$. Then,

$$\sum_{\ell=1}^{d} \|P_u \perp \beta^{(\ell)}\|^2 \geq C(nr^d)^{-3} \sum_{\ell=1}^{d} \|\beta^{(\ell)}\|^2.$$

6 Proof of Lemma 4.2

Recall that $\tilde{R} = P_V R P_V + P_V R P_V^\perp + P_V^\perp R P_V$, and $V = \langle x^{(1)}, \ldots, x^{(d)}, u \rangle$. Therefore, there exist a matrix $Y \in \mathbb{R}^{n \times d}$ and a vector $a \in \mathbb{R}^n$ such that $\tilde{R} = XY^T + YX^T + au^T + au^T$. We can further assume that $Y^T u = 0$. Otherwise, define $\tilde{Y} = \tilde{X}^T + au^T$. The proof is completed by the following claim, whose proof is given in Appendix H.

Claim 5.6 There exists a constant $C = C(d)$ such that the following holds w.h.p. Consider an arbitrary vector $v \in V^\perp$ with local decompositions $v_{Q_i} = \sum_{\ell=1}^{d} \beta^{(\ell)}_i \tilde{x}^{(\ell)}_{Q_i} + \gamma_i u_{Q_i} + w^{(i)}$. Then,

$$\sum_{\ell=1}^{d} \|P_u \perp \beta^{(\ell)}\|^2 \geq C(nr^d)^{-3} \sum_{\ell=1}^{d} \|\beta^{(\ell)}\|^2.$$

The proof is completed by the following claim, whose proof is given in Appendix H.
\[ Y - u(u^T Y/\|u\|^2) \] and \( \tilde{a} = a + X(Y^T u/\|u\|^2) \). Then \( \tilde{R} = X\tilde{Y}^T + \tilde{Y}X^T + u\tilde{a}^T + \tilde{a}u^T \), and \( \tilde{Y}^T u = 0 \).

Also note that \( u^T Qu = u^T \tilde{R} u = 2(a^T u)\|u\|^2 \). Hence, \( a^T u = 0 \), since \( Qu = 0 \). In addition, \( Qu = \tilde{R} u = a\|u\|^2 \), which implies that \( a = 0 \). Therefore, \( \tilde{R} = X\tilde{Y}^T + YX^T \) where \( Y^T u = 0 \). Denote by \( y_i^T \in \mathbb{R}^d \), \( i \in [n] \), the \( i \)th row of the matrix \( Y \).

Define the operator \( R_{G,X} : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^E \) as \( R_{G,X}(Y) = RG(X)Y \), where \( Y = [y_1^T, \ldots, y_n^T]^T \) and \( RG(X) \) is the rigidity matrix of framework \( GX \). Observe that \( \|R_{G,X}(Y)\|_1 = \sum_{(l,k) \in E(G)} |\langle x_l - x_k, y_l - y_k \rangle| \).

The following theorem compares the operators \( R_{G,X} \) and \( R_{K_n,X} \), where \( G = G(n,r) \) and \( K_n \) is the complete graph with \( n \) vertices. This theorem is the key ingredient in the proof of Lemma 4.2.

**Theorem 6.1** There exists a constant \( C = C(d) \), such that, w.h.p.,

\[ \|R_{K_n,X}(Y)\|_1 \leq Cr^{-d-2}\|R_{G,X}(Y)\|_1, \quad \text{for all} \quad Y \in \mathbb{R}^{n \times d}. \]

The proof of Theorem 6.1 is discussed in the next subsection. The following statement provides an upper bound on \( \|\tilde{R}\|_1 \). Its proof is immediate and is given in Appendix I.

**Proposition 6.1** Given \( \tilde{R} = XY^T + YX^T \), with \( Y^T u = 0 \), we have

\[ \|\tilde{R}\|_1 \leq 5\|R_{K_n,X}(Y)\|_1. \]

Now we have in place all we need to prove Lemma 4.2.

**Proof of Lemma 4.2** Define the operator \( A_G : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^E \) as \( A_G(S) = (\langle M_{ij}, S \rangle)_{(i,j) \in E} \). By our assumptions,

\[ |\langle M_{ij}, \tilde{R} \rangle + \langle M_{ij}, R^\perp \rangle| = |\langle M_{ij}, Q \rangle - \langle M_{ij}, Q_0 \rangle| \leq |\langle M_{ij}, Q \rangle - \tilde{a}_{ij}^2 + |\tilde{a}_{ij}^2 - \langle M_{ij}, Q_0 \rangle| \leq 2\Delta. \]

Therefore, \( \|A_G(\tilde{R})\|_1 \leq 2|E|\Delta + \|A_G(R^\perp)\|_1 \). Write the Laplacian matrix \( \mathcal{L} \) as \( \mathcal{L} = \sum_{(i,j) \in E} M_{ij} \). Then, \( \langle \mathcal{L}, R^\perp \rangle = \sum_{(i,j) \in E} \langle M_{ij}, R^\perp \rangle = \|A_G(R^\perp)\|_1 \). Here, we used the fact that \( \langle M_{ij}, R^\perp \rangle \geq 0 \), since \( M_{ij} \succeq 0 \) and \( R^\perp \succeq 0 \). Hence, \( \|A_G(\tilde{R})\|_1 \leq 2|E|\Delta + \langle \mathcal{L}, R^\perp \rangle \). Due to Theorem 5.2, Eq. (15), and Claim 5.1,

\[ \langle \mathcal{L}, R^\perp \rangle \leq C(nr^d)^3 r^{-2}\|\Omega, R^\perp \| \leq C(nr^d)^{6} \frac{n}{r^2} \Delta, \]

whence we obtain

\[ \|A_G(\tilde{R})\|_1 \leq C(nr^d)^{6} \frac{n}{r^2} \Delta. \]
Fig. 1 An illustration of a chain $G_{uv}$

The last step is to write $\|A_G(\tilde{R})\|_1$ more explicitly. Notice that

$$\|A_G(\tilde{R})\|_1 = \sum_{(l,k) \in E} |\langle M_{lk}, XY^T + YX^T \rangle|$$

$$= 2 \sum_{(l,k) \in E} |\langle x_l - x_k, y_l - y_k \rangle| = 2 \|\mathcal{R}_{G,X}(Y)\|_1.$$

Invoking Theorem 6.1 and Proposition 6.1, we have

$$\|\tilde{R}\|_1 \leq C r^{-d-2} \|\mathcal{R}_{G,X}(Y)\|_1$$

$$= C r^{-d-2} \|A_G(\tilde{R})\|_1 \leq C (nr^d)^{\frac{5n^2}{r^4}} \Delta. \quad \square$$

6.1 Proof of Theorem 6.1

We begin with some definitions and the initial setup.

**Definition 1** The $d$-dimensional hypercube $M_d$ is the simple graph whose vertices are the $d$-tuples with entries in $\{0, 1\}$ and whose edges are the pairs of $d$-tuples that differ in exactly one position. Also, we use $M_d^{(2)}$ to denote the graph with the same set of vertices as $M_d$, whose edges are the pairs of $d$-tuples that differ in at most two positions.

**Definition 2** An isomorphism of graphs $G$ and $H$ is a bijection between the vertex sets of $G$ and $H$, say $\phi : V(G) \rightarrow V(H)$, such that any two vertices $u$ and $v$ of $G$ are adjacent in $G$ if and only if $\phi(u)$ and $\phi(v)$ are adjacent in $H$. The graphs $G$ and $H$ are called isomorphic, denoted by $G \simeq H$, if an isomorphism exists between $G$ and $H$.

**Chains and Force Flows** A chain $G_{ij}$ between nodes $i$ and $j$ is a sequence of subgraphs $H_1, H_2, \ldots, H_k$ of $G$ such that $H_p \simeq M_d^{(2)}$ for $1 \leq p \leq k$, $H_p \cap H_{p+1} \simeq M_d^{(2)}$ for $1 \leq p \leq k-1$, and $H_p \cap H_{p+2}$ is empty for $1 \leq p \leq k-2$. Further, $i$ (resp. $j$) is connected to all vertices in $V(H_1) \setminus V(H_2)$ (resp. $V(H_k) \setminus V(H_{k-1})$). See Fig. 1 for an illustration of a chain in the case $d = 2$.

A force flow $\gamma$ is a collection of chains $\{G_{ij}\}_{1 \leq i \neq j \leq n}$ for all $\binom{n}{2}$ node pairs. Let $\Gamma$ be the collection of all possible $\gamma$. Consider the probability distribution induced on $\Gamma$ by selecting the chains between all node pairs in the following manner. Chains are chosen independently for different node pairs. Consider a particular node pair $(i, j)$. Let $\ell = \|x_i - x_j\|$ and $a = (x_i - x_j)/\|x_i - x_j\|$. Define $\tilde{r} = \frac{3r}{4\sqrt{2}}$, and choose
nonnegative numbers \( m \in \mathbb{Z} \) and \( \eta \in \mathbb{R} \) such that \( \ell = m\tilde{r} + \eta \) and \( \eta < \tilde{r} \). Consider the following set of points on the line segment between \( x_i \) and \( x_j \):

\[
\xi_k = x_i + \eta \frac{k}{2} + (k - 1)\tilde{r}a, \quad \text{for } 1 \leq k \leq m + 1.
\]

Construct the sequence of hypercubes in the direction of \( a \), with centers at \((\xi_k + \xi_{k+1})/2\), and side length \( \tilde{r} \). (See Fig. 2 for an illustration.) Denote the set of vertices in this construction by \( \{z_k\} \). Now, partition the space \([-0.5, 0.5]^d\) into hypercubes (bins) of side length \( \frac{r}{8\sqrt{d}} \). From the proof of Proposition 5.1, w.h.p., every bin contains at least one of the nodes \( \{x_k\}_{k \in [n]} \). For every vertex \( z_k \), choose a node \( x_k \) uniformly at random among the nodes in the bin that contains \( z_k \). Hence, \( \|x_k - z_k\| \leq \frac{r}{8} \) and

\[
\|x_l - x_k\| \leq \|x_l - z_l\| + \|z_l - z_k\| + \|z_k - x_k\|
\leq \frac{r}{4} + \|z_l - z_k\|, \quad \forall l, k.
\]

By wigging points \( \{z_k\} \) to nodes \( \{x_k\} \), we obtain a perturbation of the sequence of hypercubes; call it \( G_{ij} \). It is easy to see that \( G_{ij} \) is a chain between nodes \( i \) and \( j \).

Under the preceding setup, we claim the following two lemmas.

**Lemma 6.1** Under the probability distribution on \( \Gamma \) as described above, the expected number of chains containing a particular edge is upper bounded by \( Cr^{-d-1} \), w.h.p., where \( C = C(d) \) is a constant.

The proof is discussed in Appendix J.

**Lemma 6.2** Let \( G_{ij} \) be the chain between nodes \( i \) and \( j \) as described above. There exists a constant \( C = C(d) \) such that

\[
\left| \langle x_l - x_j, y_i - y_j \rangle \right| \\
\leq Cr^{-1} \sum_{(l,k) \in E(G_{ij})} \left| \langle x_l - x_k, y_l - y_k \rangle \right|, \quad \forall 1 \leq i, j \leq n.
\]
The proof is deferred to Sect. 6.1.1. Now, we are in position to prove Theorem 6.1.

Proof of Theorem 6.1 Consider a force flow $\gamma = \{G_{ij}\}_{1 \leq i, j \leq n}$. Using Lemma 6.2, we have

$$\sum_{i,j} |\langle x_i - x_j, y_i - y_j \rangle| \leq Cr^{-1} \sum_{i,j} \sum_{(l,k) \in E(G_{ij})} |\langle x_l - x_k, y_l - y_k \rangle|$$

$$\leq Cr^{-1} \sum_{(l,k) \in E(G)} \left( \sum_{G_{ij}: (l,k) \in E(G_{ij})} 1 \right) |\langle x_l - x_k, y_l - y_k \rangle|$$

$$= Cr^{-1} \sum_{(l,k) \in E(G)} b(\gamma, (l,k)) |\langle x_l - x_k, y_l - y_k \rangle|, \quad (19)$$

where $b(\gamma, (l, k))$ denotes the number of chains passing through edge $(l, k)$. Notice that, in Eq. (19), $b(\gamma, (l, k))$ is the only term that depends on the force flow $\gamma$. Hence, $b(\gamma, (l, k))$ can be replaced by its expectation under a probability distribution on $\Gamma$. According to Lemma 6.1, under the described distribution on $\Gamma$, the average number of chains containing any particular edge is upper bounded by $Cr^{-d-1}$, w.h.p. Therefore,

$$\sum_{i,j} |\langle x_i - x_j, y_i - y_j \rangle| \leq Cr^{-d-2} \sum_{(l,k) \in E(G)} |\langle x_l - x_k, y_l - y_k \rangle|.$$ 

Equivalently, $\|\mathcal{R}_{K_n,X}(Y)\|_1 \leq Cr^{-d-2}\|\mathcal{R}_{G,X}(Y)\|_1$, w.h.p. \qed

6.1.1 Proof of Lemma 6.2

Proof Assume that $|V(G_{ij})| = m + 1$. Relabel the vertices in the chain $G_{ij}$ such that the nodes $i$ and $j$ have labels 0 and $m$ respectively, and all the other nodes are labeled in $\{1, \ldots, m - 1\}$. Since both sides of the desired inequality are invariant to translations, without loss of generality we assume that $x_0 = y_0 = 0$. For a fixed vector $y_m$ consider the following optimization problem:

$$\Theta = \min_{y_1, \ldots, y_{m-1} \in \mathbb{R}^d} \sum_{(l,k) \in E(G_{ij})} |\langle x_l - x_k, y_l - y_k \rangle|.$$ 

To each edge $(l, k) \in E(G_{ij})$, assign a number $\lambda_{lk}$. (Note that $\lambda_{lk} = \lambda_{kl}$.) For any assignment with $\max_{(l,k) \in E(G_{ij})} |\lambda_{lk}| \leq 1$, we have
\[ \Theta \geq \min_{y_1, \ldots, y_{m-1} \in \mathbb{R}^d} \sum_{(l,k) \in E(G_{ij})} \lambda_{lk} (x_l - x_k, y_l - y_k) \]

\[ = \min_{y_1, \ldots, y_{m-1} \in \mathbb{R}^d} \sum_{l \in G_{ij}} \sum_{k \in \partial l \setminus \{0\}} \lambda_{lk} (y_l, x_l - x_k) \]

\[ = \min_{y_1, \ldots, y_{m-1} \in \mathbb{R}^d} \sum_{l \in G_{ij}} \left( y_l, \sum_{k \in \partial l \setminus \{0\}} \lambda_{lk} (x_l - x_k) \right), \]

where \( \partial l \) denotes the set of adjacent vertices to \( l \) in \( G_{ij} \). Therefore,

\[ \Theta \geq \max_{l \in G_{ij}} \min_{y_1, \ldots, y_{m-1} \in \mathbb{R}^d} \sum_{k \in \partial l \setminus \{0\}} \lambda_{lk} (x_l - x_k) \right), \]

Note that the numbers \( \lambda_{lk} \) that maximize the right-hand side should satisfy

\[ \frac{1}{\lambda_{lk}} \sum_{k \in \partial l} \lambda_{lk} (x_l - x_k) = 0, \quad \forall l \neq 0, m. \]

It is convenient to generalize the constraints in Eq. (21). Consider the following linear system of equations with unknown variables \( \lambda_{lk} \):

\[ \sum_{l \in V(G_{ij})} \sum_{k \in \partial l} \lambda_{lk} (x_l - x_k) = \sum_{k \in \partial 0} \lambda_{0k} (x_0 - x_k) + x_m \]

\[ \sum_{l \in V(G_{ij})} \sum_{k \in \partial l} \lambda_{lk} (x_l - x_k) = \sum_{(l,k) \in E(G_{ij})} \lambda_{lk} (x_l - x_k) + \lambda_{kl} (x_k - x_l) = 0. \]

Hence, \( \sum_{k \in \partial 0} \lambda_{0k} (x_0 - x_k) = -x_m. \)

It is convenient to generalize the constraints in Eq. (21). Consider the following linear system of equations with unknown variables \( \lambda_{lk} :\)

\[ \sum_{k \in \partial l} \lambda_{lk} (x_l - x_k) = u_l, \quad \text{for } l = 0, \ldots, m. \]

Writing Eq. (22) in terms of the rigidity matrix of \( G_{ij} \), and using the characterization of its null space as discussed in Sect. 2.1, it follows that Eq. (22) has a solution if and
only if
\[ \sum_{i=0}^{m} u_i = 0, \quad \sum_{i=0}^{m} u_i^T Ax_i = 0, \quad (23) \]

where \( A \in \mathbb{R}^{d \times d} \) is an arbitrary antisymmetric matrix.

**A Mechanical Interpretation**

For any pair \((l,k) \in E(G_{ij})\), assume a spring with spring constant \( \lambda_{lk} \) between nodes \( l \) and \( k \). Then, by Eq. (22), \( u_l \) will be the force imposed on node \( l \). The first constraint in Eq. (23) states that the net force on \( G_{ij} \) is zero (force equilibrium), while the second condition states that the net torque is zero (torque equilibrium).

Indeed, \( \sum_{i=0}^{m} u_i^T Au_i = (A, \sum_{i=0}^{m} u_i x_i^T) = 0 \), for every antisymmetric matrix \( A \) if and only if \( \sum_{i=0}^{m} u_i x_i^T \) is a symmetric matrix. Therefore,
\[
\sum_{i=0}^{m} u_i \wedge x_i = \sum_{i=0}^{m} \left( \sum_{\ell=1}^{d} u_i^{(\ell)} e_\ell \right) \wedge \left( \sum_{k=1}^{d} x_i^{(k)} e_k \right) = \sum_{\ell,k} \sum_{i=0}^{m} \left( u_i^{(\ell)} x_i^{(k)} - x_i^{(\ell)} u_i^{(k)} \right) (e_\ell \wedge e_k) = 0.
\]

With this interpretation in mind, we propose a two-part procedure to find the spring constants \( \lambda_{lk} \) that obey the constraints in (21).

**Part (i)** For simplicity, we focus here on the special case \( d = 2 \). The general argument proceeds along the same lines and is deferred to Appendix K.

Consider the chain \( G_{ij} \) between nodes \( i \) and \( j \); cf. Fig. 1. For every \( 1 \leq p \leq k \), let \( \mathcal{F}_p \) denote the common side of \( H_p \) and \( H_{p+1} \). Without loss of generality, assume that \( V(\mathcal{F}_p) = \{1, 2\} \), and \( x_m \) is in the direction of \( e_1 \). Find the forces \( f_1, f_2 \) such that
\[
f_1 + f_2 = x_m, \quad f_1 \wedge x_1 + f_2 \wedge x_2 = 0,
\]
\[
\|f_1\|^2 + \|f_2\|^2 \leq C \|x_m\|^2.
\]

To this end, we solve the following optimization problem:
\[
\begin{align*}
\text{minimize} \quad & \frac{1}{2} \left( \|f_1\|^2 + \|f_2\|^2 \right) \\
\text{subject to} \quad & f_1 + f_2 = x_m, \quad f_1 \wedge x_1 + f_2 \wedge x_2 = 0.
\end{align*}
\]

(25)

It is easy to see that the solutions of (25) are given by
\[
\begin{align*}
f_1 &= \frac{1}{2} x_m + \frac{1}{2} \gamma A (x_1 - x_2) \\
f_2 &= \frac{1}{2} x_m - \frac{1}{2} \gamma A (x_1 - x_2)
\end{align*}
\]
\[
\gamma = -\frac{1}{\|x_1 - x_2\|^2} x_m^T A (x_1 + x_2), \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.
\]
Now, we should show that the forces $f_1$ and $f_2$ satisfy the constraint $\|f_1\|^2 + \|f_2\|^2 \leq C\|x_m\|^2$, for some constant $C$. Clearly, it suffices to prove $\|\gamma(x_1 - x_2)\| \leq C\|x_m\|$.

Observe that

$$\frac{\|\gamma(x_1 - x_2)\|}{\|x_m\|} = \frac{1}{\|x_1 - x_2\|} \left| \frac{x_m^T}{\|x_m\|} A(x_1 + x_2) \right|$$

$$= \frac{1}{\|x_1 - x_2\|} \left| e_1^T A(x_1 + x_2) \right|$$

$$= \frac{1}{\|x_1 - x_2\|} \left| e_2^T (x_1 + x_2) \right|.$$ 

From the construction of chain $G_{ij}$, we have

$$\left| e_2^T (x_1 + x_2) \right| \leq \frac{r}{4}, \quad \|x_1 - x_2\| \geq \frac{r}{4},$$

which shows that $\|\gamma(x_1 - x_2)\| \leq \|x_m\|$.

**Part (ii)** For each $H_p$ consider the following set of forces:

$$u_i = \begin{cases} 
 f_i & \text{if } i \in V(F_p), \\
 -f_i & \text{if } i \in V(F_{p-1}).
\end{cases} \quad \text{(26)}$$

Also, let $u_0 = -x_m$ and $u_m = x_m$ (see Fig. 3).

Notice that $\sum_{i \in V(H_p)} u_i = 0$, $\sum_{i \in V(H_p)} u_i \wedge x_i = 0$, and thus by the discussion prior to Eq. (23), there exist values $\lambda^{(H_p)}_{lk}$ such that

$$\sum_{k:(l,k) \in E(H_p)} \lambda^{(H_p)}_{lk} (x_l - x_k) = u_l, \quad \forall l \in V(H_p).$$

Writing this in terms of $R^{(H_p)}$, the rigidity matrix of $H_p$, we have

$$\left( R^{(H_p)} \right)^T \Lambda^{(H_p)} = u, \quad \text{(27)}$$

where the vector $\Lambda^{(H_p)} = [\lambda^{(H_p)}_{lk}]$ has size $|E(H_p)| = d(d + 1)2^{d-2}$, and the vector $u = [u_l]$ has size $d \times |V(H_p)| = d2^d$. Among the solutions of Eq. (27), choose the one that is orthogonal to the null space of $(R^{(H_p)})^T$. Therefore,
\[ \sigma_{\min}(R(H_p)) \| \lambda(H_p) \|_\infty \leq \sigma_{\min}(R(H_p)) \| \lambda(H_p) \|_2 \leq \| u \| \leq C \| x_m \|. \]

From the construction of the chains, \( H_p \) is a perturbation of the \( d \)-dimensional hypercube with side length \( \tilde{r} = \frac{3r}{4\sqrt{2}} \) (each vertex wiggles by at most \( \frac{r}{8} \)). Using the fact that \( \sigma_{\min}(\cdot) \) is a Lipschitz continuous function of its argument, we get that \( \sigma_{\min}(R(H_p)) \geq Cr \), for some constant \( C = C(d) \). Also, \( \| x_m \| \leq 1 \). Hence, \( \| \lambda(H_p) \|_\infty \leq Cr^{-1} \).

Now define
\[ \lambda_{lk} = \sum_{H_p:(l,k) \in E(H_p)} \lambda_{lk}^{(H_p)}, \forall (l,k) \in E(G_{ij}). \] (28)

We claim that the values \( \lambda_{lk} \) satisfy the constraints in (21).

First, note that for every node \( l \),
\[ \sum_{k \in \partial l} \lambda_{lk} (x_l - x_k) = \sum_{k \in \partial l} \left( \sum_{H_p:(l,k) \in E(H_p)} \lambda_{lk}^{(H_p)} \right) (x_l - x_k) = \sum_{H_p : l \in V(H_p)} \left( \sum_{k : (l,k) \in E(H_p)} \lambda_{lk}^{(H_p)} (x_l - x_k) \right) = \sum_{H_p : l \in V(H_p)} u_l. \] (29)

For nodes \( l \notin \{0, m\} \), there are two \( H_p \) containing \( l \). In one of them \( u_l = f_l \), and in the other \( u_l = -f_l \). Hence, the forces \( u_l \) cancel each other in Eq. (29) and the sum is zero. At nodes 0 and \( m \), this sum is equal to \( -x_m \) and \( x_m \) respectively.

Second, since each edge participates in at most two \( H_p \), it follows from Eq. (28) that \( |\lambda_{lk}| \leq Cr^{-1} \).  

\[ \square \]

7 Proof of Theorem 1.1 (Lower Bound)

**Proof** Consider the “bending” map \( T : [-0.5, 0.5]^d \to \mathbb{R}^{d+1} \), defined as
\[ T(t_1, t_2, \ldots, t_d) = \left( R \sin \frac{t_1}{R}, t_2, \ldots, t_d, R \left( 1 - \cos \frac{t_1}{R} \right) \right). \]

This map bends the hypercube in the \( d + 1 \) dimensional space. Here, \( R \) is the curvature radius of the embedding (for instance, \( R \gg 1 \) corresponds to slightly bending the hypercube; see Fig. 4).

Now for a given \( \Delta \), let \( R = \max\{1, r^2 \Delta^{-1/2}\} \) and give the distances \( \tilde{d}_{ij} = \| T(x_i) - T(x_j) \| \) as the input distance measurements to the algorithm. First we show that these adversarial measurements satisfy the noise constraint \( \| \tilde{d}_{ij}^2 - d_{ij}^2 \| \leq \Delta \):
Fig. 4  Bending map $T$, with $R = 2$, and $d = 2$

Fig. 5  An illustration of subspaces $U$ and $W$

\[
d_{ij}^2 - \tilde{d}_{ij}^2 = (x_i^{(1)} - x_j^{(1)})^2 - R^2 \left[ \sin \left( \frac{x_i^{(1)}}{R} \right) - \sin \left( \frac{x_j^{(1)}}{R} \right) \right]^2
\]
\[
- R^2 \left[ \cos \left( \frac{x_i^{(1)}}{R} \right) - \cos \left( \frac{x_j^{(1)}}{R} \right) \right]^2
\]
\[
= (x_i^{(1)} - x_j^{(1)})^2 - R^2 \left[ 2 - 2 \cos \left( \frac{x_i^{(1)} - x_j^{(1)}}{R} \right) \right]
\]
\[
\leq \frac{(x_i^{(1)} - x_j^{(1)})^4}{2R^2} \leq \frac{r^4}{2R^2} \leq \Delta.
\]

Also, $\tilde{d}_{ij} \leq d_{ij}$. Therefore, $|z_{ij}| = |\tilde{d}_{ij}^2 - d_{ij}^2| \leq \Delta$.

The crucial point is that the SDP in the first step of the algorithm is oblivious of dimension $d$. Therefore, given the measurements $\tilde{d}_{ij}$ as the input, the SDP will return the Gram matrix $Q$ of the positions $\tilde{x}_i = LT(x_i)$, i.e., $Q_{ij} = \tilde{x}_i \cdot \tilde{x}_j$. Denote by $\{u_1, \ldots, u_d\}$ the eigenvectors of $Q$ corresponding to the $d$ largest eigenvalues. Next, the algorithm projects the positions $\{\tilde{x}_i\}_{i \in [n]}$ onto the space $U = \langle u_1, \ldots, u_d \rangle$ and returns them as the estimated positions in $\mathbb{R}^d$. Hence,
\[ d(X, \hat{X}) = \frac{1}{n^2} \| XX^T - PU \tilde{X} \tilde{X}^T PU \|_1. \]

Let \( W = \langle e_1, e_2, \ldots, e_d \rangle \) (see Fig. 5). Then,
\[
\begin{align*}
\| X - \hat{X} \|_1 & \geq \frac{1}{n^2} \| XX^T - \tilde{X} P_W \tilde{X}^T \|_1 - \frac{1}{n^2} \| \tilde{X} P_W \tilde{X}^T - PU \tilde{X} \tilde{X}^T PU \|_1. 
\end{align*}
\]

(30)

We bound each term on the right-hand side separately. For the first term,
\[
\frac{1}{n^2} \| XX^T - \tilde{X} P_W \tilde{X}^T \|_1 
\]
\[
\begin{align*}
&= \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \left| x_i^{(1)} x_j^{(1)} - R^2 \sin \left( \frac{x_i^{(1)}}{R} \right) \sin \left( \frac{x_j^{(1)}}{R} \right) \right| \\
&\overset{(a)}{=} \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \left| x_i^{(1)} x_j^{(1)} - R^2 \left( \frac{x_i^{(1)}}{R} - \frac{(x_i^{(1)})^3}{3!R^3} + \frac{\xi_i^5}{5!R^5} \right) \right. \\
&\quad \times \left. \left( \frac{x_j^{(1)}}{R} - \frac{(x_j^{(1)})^3}{3!R^3} + \frac{\xi_j^5}{5!R^5} \right) \right| \\
&\overset{(b)}{=} C \left( \frac{R}{n} \right)^2 \sum_{1 \leq i, j \leq n} \left| \frac{x_i^{(1)}}{R} \right| \left( \frac{x_j^{(1)}}{R} \right)^3 + \frac{1}{3!} \left( \frac{x_j^{(1)}}{R} \right)^3 \left( \frac{x_i^{(1)}}{R} \right)^3 \\
&\geq \frac{C}{(nR)^2} \left( \sum_{1 \leq i \leq n} |x_i^{(1)}| \right) \left( \sum_{1 \leq j \leq n} |x_i^{(1)}| \right)^3 \geq \frac{C}{R^2}, 
\end{align*}
\]

(31)

where \((a)\) follows from Taylor’s theorem, and \((b)\) follows from \(|\xi_i/R| \leq |x_i/R| \leq 1/2\).

The next proposition provides an upper bound for the second term on the right-hand side of Eq. (30).

**Proposition 7.1** The following is true:
\[
\frac{1}{n^2} \| \tilde{X} P_W \tilde{X}^T - PU \tilde{X} \tilde{X}^T PU \|_1 \to 0 \text{ a.s.}, \quad \text{as } n \to \infty.
\]

The proof of this proposition is provided in the next section.

Using the bounds given by Proposition 7.1 and Eq. (31), we obtain that, w.h.p.,
\[
d(X, \hat{X}) \geq \frac{C_1}{R^2} \geq C \min \left\{ 1, \frac{\Delta}{r^4} \right\}.
\]

The result follows.

\[ \square \]

7.1 Proof of Proposition 7.1

We first establish the following remarks.
Remark 7.1 Let $a, b \in \mathbb{R}^m$ be two unitary vectors. Then,
\[
\|aa^T - bb^T\|_2 = \sqrt{1 - (a^T b)^2}.
\]

For the proof, we refer to Appendix L.

Remark 7.2 Assume $A$ and $\tilde{A}$ are $p \times p$ matrices. Let $\{\lambda_i\}$ be the eigenvalues of $A$ such that $\lambda_1 \geq \cdots \geq \lambda_{p-1} > \lambda_p$. Also, let $v$ and $\tilde{v}$ respectively denote the eigenvectors of $A$ and $\tilde{A}$ corresponding to their smallest eigenvalues. Then,
\[
1 - (v^T \tilde{v})^2 \leq \frac{4\|A - \tilde{A}\|_2}{\lambda_{p-1} - \lambda_p}.
\]

The proof is deferred to Appendix M.

Proof of Proposition 7.1 Let $\tilde{X} = \sum_{i=1}^{d+1} \sigma_i u_i^T \hat{w}_i^T$ be the singular value decomposition of $\tilde{X}$, where $\|u_i\| = \|\hat{w}_i\| = 1$, $u_i \in \mathbb{R}^n$, $\hat{w}_i \in \mathbb{R}^{d+1}$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{d+1}$. Notice that $\hat{P}_U \tilde{X} = \sum_{i=1}^d \sigma_i u_i \hat{w}_i^T = \left( \sum_{i=1}^d \sigma_i u_i \hat{w}_i^T \right) = \hat{X} \hat{P}_W$, where $\hat{W} = (\hat{w}_1, \ldots, \hat{w}_d)$, and $\hat{P}_W \in \mathbb{R}^{(d+1) \times (d+1)}$. Hence, $\hat{P}_U \tilde{X} \hat{X}^T \hat{P}_U = \hat{X} \hat{P}_W \hat{X}^T$. Define $M = \hat{P}_W - P_W$. Then, we have
\[
\frac{1}{n^2} \|P_U \tilde{X} - P_U \hat{X} \hat{X}^T P_U\|_1 = \frac{1}{n^2} \|\tilde{X} M \hat{X}^T\|_1 = \frac{1}{n^2} \sum_{1 \leq i, j \leq n} |\hat{x}_i^T M \hat{x}_j| \leq \frac{1}{n^2} \|M\|_2 \sum_{1 \leq i, j \leq n} \|\hat{x}_i\| \|\hat{x}_j\| \leq \|M\|_2. \tag{32}
\]

Now, we need to bound $\|M\|_2$. We have
\[
M = \hat{P}_W - P_W = (I - P_{\hat{w}_{d+1}}) - (I - P_{\hat{e}_{d+1}}) = P_{\hat{e}_{d+1}} - P_{\hat{w}_{d+1}}.
\]
Using Remark 7.1, we obtain
\[
\|M\|_2 = \left\| \hat{e}_{d+1} \hat{e}_{d+1}^T - \hat{w}_{d+1} \hat{w}_{d+1}^T \right\|_2 = \sqrt{1 - (\hat{e}_{d+1}^T \hat{w}_{d+1})^2}.
\]

Let $Z_i = \tilde{x}_i \hat{x}_i^T \in \mathbb{R}^{(d+1) \times (d+1)}$, $\tilde{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$, and $Z = \mathbb{E}(Z_i)$. Notice that $\tilde{Z} = \frac{1}{n} \tilde{X} \hat{X} = \frac{1}{n} \sum_{i=1}^{d+1} \sigma_i^2 \hat{w}_i \hat{w}_i^T$. Therefore, $\hat{w}_{d+1}$ is the eigenvector of $\tilde{Z}$ corresponding to its smallest eigenvalue. In addition, $Z$ is a diagonal matrix (with $Z_{(d+1),(d+1)}$ the smallest diagonal entry). Hence, $\hat{e}_{d+1}$ is its eigenvector corresponding to the smallest eigenvalue, $Z_{(d+1),(d+1)}$.

By applying Remark 7.2, we have
\[ \|M\|_2 \leq \sqrt{1 - \left( e^{T}_{d+1} \hat{w}_{d+1} \right)^2} \leq \sqrt{\frac{4\|Z - \hat{Z}\|_2}{\lambda_d - \lambda_{d+1}}}, \]  

where \( \lambda_d > \lambda_{d+1} \) are the two smallest eigenvalues of \( Z \). Let \( t \) be a random variable, uniformly distributed in \([-0.5, 0.5]\). Then,

\[
\lambda_d = \mathbb{E} \left[ R^2 \sin^2 \left( \frac{t}{R} \right) \right] \quad \text{and} \quad \lambda_{d+1} = \mathbb{E} \left[ R^2 \left( 1 - \cos^2 \left( \frac{t}{R} \right) \right) \right].
\]

Hence, \( \lambda_d - \lambda_{d+1} = R^2 (-1/R - \sin(1/R) + 4 \sin(1/2R)) \geq 0.07 \), since \( R \geq 1 \).

Also, note that \( \{Z_i\}_{1 \leq i \leq n} \) is a sequence of i.i.d. random matrices with dimension \((d + 1)\) and \( \|Z\|_\infty = \|\mathbb{E}(Z_i)\|_\infty < \infty \). By the law of large numbers, \( \hat{Z} \to Z \), almost surely. Now, since the operator norm is a continuous function, we have \( \|Z - \hat{Z}\|_2 \to 0 \), almost surely. The result follows directly from Eqs. (32) and (33). \( \square \)

### 8 Numerical Experiments

Theorem 1.1 considers a worst-case model for the measurement noise in which the errors \( \{z_{ij}\}_{(i,j) \in E} \) are arbitrary but uniformly bounded as \( |z_{ij}| \leq \Delta \). The proof of the lower bound (cf. Sect. 7) introduces errors \( \{z_{ij}\}_{(i,j) \in E} \) defined based on a bending map, \( T \). This set of errors results in the claimed lower bound. For clarity, we denote this set of errors by \( \{z_{ij}^T\} \). In this section, we consider a mixture model for the measurement errors. For given parameters \( \Delta \) and \( \varepsilon \), we let

\[
z_{ij} \sim \varepsilon \gamma_{\Delta/2} + (1 - \varepsilon)\delta_{z_{ij}}, \tag{34}
\]

where \( \gamma_\sigma(x) = 1/(\sqrt{2\pi}\sigma)e^{-x^2/2\sigma^2} \) is the density function of the normal distribution with mean zero and variance \( \sigma^2 \). The goal of the numerical experiments is to show the dependency of the algorithm performance on each of the parameters \( n, r, \) and \( \Delta \).

We consider the following configurations. For each configuration we run the SDP-based algorithm and evaluate \( d(X, \hat{X}) \). The error bars in the figures correspond to 10 realizations of that configuration. Throughout the measurement errors are defined according to (34) with \( \varepsilon = 0.1 \).

1. Fix \( \Delta = 0.005 \) and \( d \in \{2, 4\} \). Let \( r = 3(\log n/n)^{1/d} \), with \( n \in \{100, 120, 140, \ldots, 300\} \). Figure 6 summarizes the results. According to the plot, \( d(X, \hat{X}) \propto n^2 \) for \( d = 2 \) and \( d(X, \hat{X}) \propto n \) for \( d = 4 \).
2. Fix \( \Delta = 0.005 \), \( d = 2 \), and \( n = 150 \). Let \( r \in \{0.5, 0.55, 0.6, \ldots, 0.8\} \). The results are shown in Fig. 7. As we see, \( d(X, \hat{X}) \) is fairly proportional to \( r^{-4} \).
3. Fix \( n = 150 \), \( r = 0.6 \), and \( d = 2 \). Let \( \Delta \in \{0.005, 0.01, 0.015, 0.02, 0.025\} \). Figure 8 showcases the results. The performance deteriorates linearly with respect to \( \Delta \).

**Acknowledgements**

Adel Javanmard is supported by the Caroline and Fabian Pease Stanford Graduate Fellowship. This work was partially supported by NSF CAREER award CCF-0743978, NSF grant DMS-0806211, and AFOSR grant FA9550-10-1-0360. The authors thank the anonymous reviewers for their insightful comments.
Fig. 6 Performance results for $\Delta = 0.005$, $d = 2, 4$, and $r = 3(\log n/n)^{1/d}$. The plot shows $\log d(X, \hat{X})$ vs. $\log n$ for a set of values of $n$. The solid line and the dashed line respectively correspond to $d(X, \hat{X}) \propto n^2$ and $d(X, \hat{X}) \propto n$ and are plotted as reference.

Fig. 7 Performance results for $\Delta = 0.005$, $d = 2$, and $n = 150$. The plot shows $\log d(X, \hat{X})$ vs. $\log r$ for a set of values of $r$. The solid line corresponds to $d(X, \hat{X}) \propto r^{-4}$ and is plotted as reference.

Fig. 8 Performance results for $n = 150$, $r = 0.6$, and $d = 2$. The plot shows $\log d(X, \hat{X})$ vs. $\log \Delta$ for a set of values of $\Delta$. The solid line corresponds to $d(X, \hat{X}) \propto \Delta$ and is plotted as reference.
Appendix A: Proof of Remark 2.1

For $1 \leq j \leq n$, let random variable $z_j$ be 1 if node $j$ is in region $\mathcal{R}$ and 0 otherwise. The variables $\{z_j\}$ are i.i.d. Bernoulli with probability $V(\mathcal{R})$ of success. Also, $n(\mathcal{R}) = \sum_{j=1}^{n} z_j$.

By application of the Chernoff bound we obtain

$$
P\left( \left| \sum_{j=1}^{n} z_j - nV(\mathcal{R}) \right| \geq \delta nV(\mathcal{R}) \right) \geq 2 \exp\left( -\frac{\delta^2 nV(\mathcal{R})}{2} \right).$$

Choosing $\delta = \sqrt{\frac{2\log n}{nV(\mathcal{R})}}$, the right-hand side becomes $2 \exp(-c \log n) = 2/n^c$. Therefore, with probability at least $1 - 2/n^c$,

$$n(\mathcal{R}) \in nV(\mathcal{R}) + \left[ -\sqrt{2cnV(\mathcal{R}) \log n}, \sqrt{2cnV(\mathcal{R}) \log n} \right].$$

(35)

Appendix B: Proof of Proposition 5.1

We apply the bin-covering technique. Cover the space $[-0.5, 0.5]^d$ with a set of nonoverlapping hypercubes (bins) whose side lengths are $\delta$. Thus, there is a total of $m = \lceil 1/\delta \rceil^d$ bins, each of volume $\delta^d$. In formula, bin $(j_1, \ldots, j_d)$ is the hypercube $[(j_1 - 1)\delta, j_1\delta) \times \cdots \times [(j_d - 1)\delta, j_d\delta)$, for $j_k \in \{1, \ldots, \lceil 1/\delta \rceil\}$ and $k \in \{1, \ldots, d\}$. Denote the set of bins by $\{B_k\}_{1 \leq k \leq m}$. Assume $n$ nodes are deployed uniformly at random in $[-0.5, 0.5]^d$. We claim that if $\delta \geq (c \log n/n)^{1/d}$, where $c > 1$, then w.h.p., every bin contains at least $d + 1$ nodes.

Fix $k$ and let random variable $\xi_l$ be 1 if node $l$ is in bin $B_k$ and 0 otherwise. The variables $\{\xi_l\}_{1 \leq l \leq n}$ are i.i.d. Bernoulli with probability $1/m$ of success. Also $\xi = \sum_{l=1}^{n} \xi_l$ is the number of nodes in bin $B_k$. By the Markov inequality, $P(\xi \leq d) \leq \mathbb{E}\{Z^{\xi-d}\}$, for any $0 \leq Z \leq 1$. Choosing $Z = md/n$, we have

$$P(\xi \leq d) \leq \mathbb{E}\{Z^{\xi-d}\} = Z^{-d} \prod_{l=1}^{n} \mathbb{E}\{Z^{\xi_l}\}$$

$$= Z^{-d} \left( \frac{1}{m}Z + 1 - \frac{1}{m} \right)^n = \left( \frac{n}{md} \right)^d \left( 1 + \frac{d}{n} - \frac{1}{m} \right)^n$$

$$\leq \left( \frac{ne}{md} \right)^d e^{-n/m} = \left( \frac{ne \delta^d}{d} \right)^d e^{-n \delta^d} \leq \left( \frac{c \log n}{d} \right)^d n^{-c}. $$

By applying a union bound over all the $m$ bins, we get the desired result.

Now take $\delta = r/(4\sqrt{d})$. Given that $r \geq 4c\sqrt{d}(\log n/n)^{1/d}$, for some $c > 1$, every bin contains at least $d + 1$ nodes, w.h.p. Note that for any two nodes $x_i, x_j \in [-0.5, 0.5]^d$ with $\|x_i - x_j\| \leq r/2$, the point $(x_i + x_j)/2$ (the midpoint of the line segment between $x_i$ and $x_j$) is contained in one of the bins, say $B_k$. For any point $s$ in this bin,

$$\|s - x_i\| \leq \left\| s - \frac{x_i + x_j}{2} \right\| + \left\| \frac{x_i + x_j}{2} - x_i \right\| \leq \frac{r}{4} + \frac{r}{4} = \frac{r}{2}. $$
Similarly, \( \| s - x_j \| \leq r/2 \). Since \( s \in B_k \) was arbitrary, \( C_i \cap C_j \) contains all the nodes in \( B_k \). This implies the thesis, since \( B_k \) contains at least \( d + 1 \) nodes.

**Appendix C: Proof of Proposition 5.2**

Let \( m_k = | Q_k | \) and define the matrix \( R_k \) as follows:

\[
R_k = \left[ x(1)_{Q_k} | \cdots | x(d)_{Q_k} | u_{Q_k} \right]^T \in \mathbb{R}^{(d+1) \times m_k}.
\]

Compute an orthonormal basis \( w_{k,1}, \ldots, w_{k,m_k-d-1} \in \mathbb{R}^{m_k} \) for the null space of \( R_k \). Then

\[
\Omega_k = P_{\perp \langle u_{Q_k}, x(1)_{Q_k}, \ldots, x(d)_{Q_k} \rangle} = \sum_{l=1}^{m_k-d-1} w_{k,l} w_{k,l}^T.
\]

Let \( \hat{w}_{k,l} \in \mathbb{R}^n \) be the vector obtained from \( w_{k,l} \) by padding it with zeros. Then, \( \hat{\Omega}_k = \sum_{l=1}^{m_k-d-1} \hat{w}_{k,l} \hat{w}_{k,l}^T \). In addition, the \((i, j)\) entry of \( \hat{\Omega}_k \) is nonzero only if \( i, j \in Q_k \).

Any two nodes in \( Q_k \) are connected in \( G \) (recall that \( Q_k \) is a clique of \( G \)). Hence, \( \hat{\Omega}_k \) is zero outside \( E \). Since \( \Omega = \sum_{Q_k \in \text{cliq}(G)} \hat{\Omega}_k \), the matrix \( \Omega \) is also zero outside \( E \).

Notice that for any \( v \in \langle x(1), \ldots, x(d), u \rangle \),

\[
\Omega v = \left( \sum_{Q_k \in \text{cliq}(G)} \hat{\Omega}_k \right) v = \sum_{Q_k \in \text{cliq}(G)} \Omega_k v_{Q_k} = 0.
\]

So far we have proved that \( \Omega \) is a stress matrix for the framework. Clearly, \( \Omega \succeq 0 \), since \( \hat{\Omega}_k \succeq 0 \) for all \( k \). We only need to show that \( \text{rank}(\Omega) = n - d - 1 \). Since \( \text{Ker}(\Omega) \supseteq \langle x(1), \ldots, x(d), u \rangle \), we have \( \text{rank}(\Omega) \leq n - d - 1 \). Define

\[
\tilde{\Omega} = \sum_{Q_k \in \{ C_1, \ldots, C_n \}} \hat{\Omega}_k.
\]

Since \( \Omega \succeq \tilde{\Omega} \succeq 0 \), it suffices to show that \( \text{rank}(\tilde{\Omega}) \geq n - d - 1 \). For an arbitrary vector \( v \in \text{Ker}(\tilde{\Omega}) \),

\[
v^T \tilde{\Omega} v = \sum_{i=1}^{n} \left\| P_{\perp \langle u_{C_i}, x(1)_{C_i}, \ldots, x(d)_{C_i} \rangle} v_{C_i} \right\|^2 = 0,
\]

which implies that \( v_{C_i} \in \langle u_{C_i}, x(1)_{C_i}, \ldots, x(d)_{C_i} \rangle \). Hence, the vector \( v_{C_i} \) can be written as

\[
v_{C_i} = \sum_{\ell=1}^{d} \beta_i^{(\ell)} x_{C_i}^{(\ell)} + \beta_i^{(d+1)} u_{C_i},
\]
or some scalars $\beta_i^{(\ell)}$. Note that for any two nodes $i$ and $j$, the vector $v_{C_i \cap C_j}$ has the following two representations:

$$v_{C_i \cap C_j} = \sum_{\ell=1}^{d} \beta_i^{(\ell)} x_{C_i \cap C_j}^{(\ell)} + \beta_i^{(d+1)} u_{C_i \cap C_j} = \sum_{\ell=1}^{d} \beta_j^{(\ell)} x_{C_i \cap C_j}^{(\ell)} + \beta_j^{(d+1)} u_{C_i \cap C_j}.$$

Therefore,

$$\sum_{\ell=1}^{d} (\beta_i^{(\ell)} - \beta_j^{(\ell)}) x_{C_i \cap C_j}^{(\ell)} + (\beta_i^{(d+1)} - \beta_j^{(d+1)}) u_{C_i \cap C_j} = 0. \quad (36)$$

According to Proposition 5.1, w.h.p., for any two nodes $i$ and $j$ with $\|x_i - x_j\| \leq r/2$, we have $|C_i \cap C_j| \geq d + 1$. Thus, the vectors $x_{C_i \cap C_j}^{(\ell)}$, $u_{C_i \cap C_j}$, $1 \leq \ell \leq d$ are linearly independent, since the configuration is generic. More specifically, let $Y$ be the matrix with $d + 1$ columns $\{x_{C_i \cap C_j}^{(\ell)}\}_{\ell=1}^{d}$, $u_{C_i \cap C_j}$. Then, $\det(Y^T Y)$ is a nonzero polynomial in the coordinates $x_k^{(\ell)}$, $k \in C_i \cap C_j$ with integer coefficients. Since the configuration of the points is generic, $\det(Y^T Y) \neq 0$, yielding the linear independence of the columns of $Y$. Consequently, Eq. (36) implies that $\beta_i^{(\ell)} = \beta_j^{(\ell)}$ for any two adjacent nodes in $G(n, r/2)$. Given that $r > 10^{1/d} (\log n/n)^{1/d}$, the graph $G(n, r/2)$ is connected w.h.p., and thus the coefficients $\beta_i^{(\ell)}$ are the same for all $i$. Dropping the subscript $(i)$, we obtain

$$v = \sum_{\ell=1}^{d} \beta^{(\ell)} x^{(\ell)} + \beta^{(d+1)} u,$$

proving that $\text{Ker}(\tilde{\Omega}) \subseteq \langle u, x^{(1)}, \ldots, x^{(d)} \rangle$, and thus $\text{rank}(\tilde{\Omega}) \geq n - d - 1$.

**Appendix D: Proof of Claim 5.2**

Let $\tilde{G} = (V, \tilde{E})$, where $\tilde{E} = \{(i, j) : d_{ij} \leq r/2\}$. The Laplacian of $\tilde{G}$ is denoted by $\tilde{\mathcal{L}}$.

We first show that, for some constant $C$,

$$\tilde{\mathcal{L}} \leq C \sum_{k=1}^{n} P_{u_{C_k}}. \quad (37)$$

Note that

$$\sum_{k=1}^{n} P_{u_{C_k}} \leq \sum_{k=1}^{n} \left( I - \frac{1}{|C_k|} u_{C_k} u_{C_k}^T \right) = \sum_{k=1}^{n} \frac{1}{|C_k|} \left( \sum_{i,j \in C_k} M_{ij} \right) \leq \sum_{(i,j) \in \tilde{E}} \left( \sum_{k : (i,j) \in C_k} \frac{1}{|C_k|} \right) M_{ij} = \sum_{(i,j) \in \tilde{E}} \left( \sum_{k \in C_i \cap C_j} \frac{1}{|C_k|} \right) M_{ij}.$$
The inequality follows from the fact that $M_{ij} \succeq 0$, \( \forall i, j \). By applying Remark 2.1, we have $|C_k| \leq C_1(nrd)$ and $|C_i \cap C_j| \geq C_2nrd$, for some constants $C_1$ and $C_2$ (depending on $d$) and $\forall k, i, j$. Therefore,

$$\sum_{k=1}^{n} P_{\nabla C_k} \geq \sum_{(i,j) \in \tilde{E}} \frac{C_2}{C_1}M_{ij} = \frac{C_2}{C_1} \tilde{L}.$$ 

Next we prove that, for some constant $C$,

$$\mathcal{L} \preceq C \tilde{\mathcal{L}}.$$ (38)

To this end, we use the Markov chain comparison technique.

A path between two nodes $i$ and $j$, denoted by $\gamma_{ij}$, is a sequence of nodes $(i, v_1, \ldots, v_t, j)$, such that the consecutive pairs are connected in $\tilde{G}$. Let $\gamma = (\gamma_{ij})_{(i,j) \in E}$ denote a collection of paths for all pairs connected in $G$, and let $\Gamma$ be the collection of all possible $\gamma$. Consider the probability distribution induced on $\Gamma$ by choosing paths between all connected pairs in $G$ in the following way.

Cover the space $[-0.5, 0.5]^d$ with bins of side length $r/(4\sqrt{d})$ (similar to the proof of Proposition 5.1. As discussed there, w.h.p., every bin contains at least one node). Paths are selected independently for different node pairs. Consider a particular path $\gamma_{ij}$ as follows. If $i$ and $j$ are in the same bin or in the neighboring bins, then $\gamma_{ij} = (i, j)$. Otherwise, consider all bins intersecting the line joining $i$ and $j$. From each of these bins, choose a node $v_k$ uniformly at random. Then the path $\gamma_{ij}$ is $(i, v_1, \ldots, j)$.

In the following, we compute the average number of paths passing through each edge in $\tilde{E}$. The total number of paths is $|E| = \Theta(n^2r^d)$. Also, since any connected pair in $G$ are within distance $r$ of each other and the side length of the bins is $O(r)$, there are $O(1)$ bins intersecting a straight line joining a pair $(i, j) \in E$. Consequently, each path contains $O(1)$ edges. The total number of bins is $\Theta(r^{-d})$. Hence, by symmetry, the number of paths passing through each bin is $\Theta(n^2r^d)$. Consider a particular bin $B$ and the paths passing through it. All these paths are equally likely to choose any of the nodes in $B$. Therefore, the average number of paths containing a particular node in $B$, say $i$, is $\Theta(n^2r^d/nr^d) = \Theta(nr^d)$. In addition, the average number of edges between $i$ and the neighboring bins is $\Theta(nr^d)$. Due to symmetry, the average number of paths containing an edge incident on $i$ is $\Theta(1)$. Since this is true for all nodes $i$, the average number of paths containing an edge is $\Theta(1)$.

Now, let $v \in \mathbb{R}^n$ be an arbitrary vector. For a directed edge $e \in \tilde{E}$ from $i \rightarrow j$, define $v(e) = v_i - v_j$. Also, let $|\gamma_{ij}|$ denote the length of the path $\gamma_{ij}$. We have

$$v^T \mathcal{L} v = \sum_{(i,j) \in E} (v_i - v_j)^2 = \sum_{(i,j) \in E} \left( \sum_{e \in \gamma_{ij}} v(e) \right)^2 \leq \sum_{(i,j) \in E} |\gamma_{ij}| \sum_{e \in \gamma_{ij}} v(e)^2 \leq \sum_{e \in \tilde{E}} v(e)^2 \sum_{\gamma_{ij} \ni e} |\gamma_{ij}| \leq \gamma^* \sum_{e \in \tilde{E}} v(e)^2 b(\gamma, e).$$ (39)
where $\gamma_*$ is the maximum path length and $b(\gamma, e)$ denotes the number of paths passing through $e$ under $\gamma = (\gamma_{ij})$. The first inequality follows from the Cauchy-Schwarz inequality. Since all paths have length $O(1)$, we have $\gamma_* = O(1)$. Also, note that in Eq. (39), $b(\gamma, e)$ is the only term that depends on the paths. Therefore, we can replace $b(\gamma, e)$ with its expectation under the distribution on $\Gamma$, i.e., $b(e) = \sum_{\gamma \in \Gamma} P(\gamma) b(\gamma, e)$. We proved above that the average number of paths passing through an edge is $\Theta(1)$. Hence, $\max_{e \in E} b(\gamma, e) = \Theta(1)$. Using these bounds in Eq. (39), we obtain
\begin{equation}
 v^T \mathcal{L} v \leq C \sum_{e \in \tilde{E}} v(e)^2 = C v^T \tilde{\mathcal{L}} v,
\end{equation}
for some constant $C$ and all vectors $v \in \mathbb{R}^n$. Combining Eqs. (37) and (40) implies the thesis.

**Appendix E: Proof of Claim 5.3**

In Remark 2.1, let region $\mathcal{R}$ be the $r/2$-neighborhood of node $i$, and take $c = 2$. Then, with probability at least $1 - 2/n^2$,
\begin{equation}
 |C_i| \in np_d + [-\sqrt{4n p_d \log n}, \sqrt{4n p_d \log n}],
\end{equation}
where $p_d = K_d(r/2)^d$.

Similarly, with probability at least $1 - 2/n^2$,
\begin{equation}
 |	ilde{C}_i| \in n\tilde{p}_d + [-\sqrt{4n \tilde{p}_d \log n}, \sqrt{4n \tilde{p}_d \log n}],
\end{equation}
where $\tilde{p}_d = K_d(r^2)\left(\frac{1}{2} + \frac{1}{100}\right)^d$. By applying a union bound over all $1 \leq i \leq n$, Eqs. (41) and (42) hold for any $i$, with probability at least $1 - 4/n$. Given that $r > 10\sqrt{d(\log n/n)^{1/d}}$, the result follows after some algebraic manipulations.

**Appendix F: Proof of Claim 5.4**

**Part (i)** Let $\tilde{G} = (V, \tilde{E})$, where $\tilde{E} = \{(i, j) : d_{ij} \leq r/2\}$. Also, let $A_{\tilde{G}}$ and $A_{G^*}$ respectively denote the adjacency matrices of the graphs $\tilde{G}$ and $G^*$. Therefore, $A_{\tilde{G}} \in \mathbb{R}^{n \times n}$ and $A_{G^*} \in \mathbb{R}^{N \times N}$, where $N = |V(G^*)| = n(m + 1)$. From the definition of $G^*$, we have
\begin{equation}
 A_{G^*} = A_{\tilde{G}} \otimes B, \quad B = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{(m+1) \times (m+1)},
\end{equation}
where $\otimes$ stands for the Kronecker product. Hence,
\begin{equation}
 \max_{i \in V(G^*)} \deg_{G^*}(i) = (m + 1) \max_{i \in V(\tilde{G})} \deg_{\tilde{G}}(i).
\end{equation}
Since the degree of the nodes in $\tilde{G}$ are bounded by $C(nr^d)$ for some constant $C$, and $m \leq C(nr^d)$ (by definition of $m$ in Claim 5.3), we have that the degree of the nodes in $G^*$ is bounded by $C(nr^d)^2$, for some constant $C$.

**Part (ii)** Let $D_{\tilde{G}} \in \mathbb{R}^{n \times n}$ be the diagonal matrix with the degrees of the nodes in $\tilde{G}$ on its diagonal. Define $D_{G^*} \in \mathbb{R}^{N \times N}$ analogously. From Eq. (43), it is easy to see that
\[
(D_{\tilde{G}}^{-1/2}A_{\tilde{G}}D_{\tilde{G}}^{-1/2}) \otimes \left(\frac{1}{m+1} B\right) = D_{G^*}^{-1/2}A_{G^*}D_{G^*}^{-1/2}.
\]
Now for any two matrices $A$ and $B$, the eigenvalues of $A \otimes B$ are all products of eigenvalues of $A$ and $B$. The matrix $1/(m+1)B$ has eigenvalues 0, with multiplicity $m$, and 1, with multiplicity one. Thereby,
\[
\sigma_{\min}(I - D_{G^*}^{-1/2}A_{G^*}D_{G^*}^{-1/2}) \geq \min\{\sigma_{\min}(I - D_{\tilde{G}}^{-1/2}A_{\tilde{G}}D_{\tilde{G}}^{-1/2}), 1\} \geq Cr^2,
\]
where the last step follows from Remark 2.2. Due to the result of [8] (Theorem 4), we obtain
\[
\sigma_{\min}(\mathcal{L}_{G^*}) \geq d_{\min,G^*}\sigma_{\min}(\mathcal{L}_{n,G^*}),
\]
where $d_{\min,G^*}$ denotes the minimum degree of the nodes in $G^*$, and $\mathcal{L}_{n,G^*} = I - D_{G^*}^{-1/2}A_{G^*}D_{G^*}^{-1/2}$ is the normalized Laplacian of $G^*$. Since $d_{\min,G^*} = (m+1)d_{\min,\tilde{G}} \geq C(nr^d)^2$, for some constant $C$, we obtain
\[
\sigma_{\min}(\mathcal{L}_{G^*}) \geq C(nr^d)^2 r^2,
\]
for some constant $C$.

**Appendix G: Proof of Claim 5.5**

Fix a pair $(i, j) \in E(G^*)$. Let $m_{ij} = |Q_i \cap Q_j|$, and without loss of generality assume that the nodes in $Q_i \cap Q_j$ are labeled with $\{1, \ldots, m_{ij}\}$. Let $z^{(\ell)} = \tilde{x}^{(\ell)}_{Q_i \cap Q_j}$, for $1 \leq \ell \leq d$, and let $z_k = (z_k^{(1)}, \ldots, z_k^{(d)})$, for $1 \leq k \leq m_{ij}$. Define the matrix $M^{(i,j)} \in \mathbb{R}^{d \times d}$ as $M^{(i,j)}_{\ell,\ell'} = \langle z^{(\ell)}, z^{(\ell')} \rangle$, for $1 \leq \ell', \ell \leq d$. Finally, let $\beta_{ij} = (\beta_{ij}^{(1)} - \tilde{\beta}_i^{(1)}, \ldots, \beta_{ij}^{(d)} - \tilde{\beta}_i^{(d)}) \in \mathbb{R}^d$. Then,
\[
\left\| \sum_{\ell=1}^{d} (\beta_{ij}^{(\ell)} - \tilde{\beta}_i^{(\ell)}) \tilde{x}^{(\ell)}_{Q_i \cap Q_j} \right\|^2 = \beta_{ij}^T M^{(i,j)} \beta_{ij} \geq \sigma_{\min}(M^{(i,j)}) \|\beta_{ij}\|^2. \tag{44}
\]

In the following, we lower bound $\sigma_{\min}(M^{(i,j)})$. Notice that
\[
M^{(i,j)} = \sum_{k=1}^{m_{ij}} z_k z_k^T = \sum_{k=1}^{m_{ij}} \left\{z_k z_k^T - \mathbb{E}(z_k z_k^T) \right\} + \sum_{k=1}^{m_{ij}} \mathbb{E}(z_k z_k^T). \tag{45}
\]
We first lower bound the quantity $\sigma_{\text{min}}(\sum_{k=1}^{m_{ij}} E(\hat{z}_k z_k^T))$. Let $S \in \mathbb{R}^{d \times d}$ be an orthogonal matrix that aligns the line segment between $x_i$ and $x_j$ with $e_1$. Now, let $\hat{z}_k = S z_k$ for $1 \leq k \leq m_{ij}$. Then,

$$\sum_{k=1}^{m_{ij}} E(\hat{z}_k z_k^T) = \sum_{k=1}^{m_{ij}} S^T E(\hat{z}_k \hat{z}_k^T) S.$$

The matrix $E(\hat{z}_k \hat{z}_k^T)$ is the same for all $1 \leq k \leq m_{ij}$. Further, it is a diagonal matrix whose diagonal entries are bounded from below by $C_1 r^2$, for some constant $C_1$. Therefore, $\sigma_{\text{min}}(\sum_{k=1}^{m_{ij}} E(\hat{z}_k \hat{z}_k^T)) \geq m_{ij} C_1 r^2$. Consequently,

$$\sigma_{\text{min}}(\sum_{k=1}^{m_{ij}} E(\hat{z}_k z_k^T)) \geq m_{ij} C_1 r^2. \quad (46)$$

Let $Z^{(k)} = z_k z_k^T - E(z_k z_k^T)$, for $1 \leq k \leq m_{ij}$. Next, we upper bound the quantity $\sigma_{\text{max}}(\sum_{k=1}^{m_{ij}} Z^{(k)})$. Note that for any matrix $A \in \mathbb{R}^{d \times d}$,

$$\sigma_{\text{max}}(A) = \max_{\|x\| = \|y\| = 1} x^T A y \leq \max_{\|x\| = \|y\| = 1} \sum_{1 \leq p, q \leq d} |A_{pq}| |x_p y_q|$$

$$\leq \max_{1 \leq p, q \leq d} |A_{pq}| \cdot \max_{\|x\| = 1} \left( \sum_{p=1}^{d} |x_p| \right) \cdot \max_{\|y\| = 1} \left( \sum_{q=1}^{d} |y_q| \right) \leq d \max_{1 \leq p, q \leq d} |A_{pq}|.$$

Taking $A = \sum_{k=1}^{m_{ij}} Z^{(k)}$, we have

$$\Pr\left( \sigma_{\text{max}}\left( \sum_{k=1}^{m_{ij}} Z^{(k)} \right) > \epsilon \right) \leq \Pr\left( \max_{1 \leq p, q \leq d} \left| \sum_{k=1}^{m_{ij}} Z^{(k)}_{pq} \right| > \frac{\epsilon}{d} \right)$$

$$\leq d^2 \max_{1 \leq p, q \leq d} \Pr\left( \left| \sum_{k=1}^{m_{ij}} Z^{(k)}_{pq} \right| > \frac{\epsilon}{d} \right), \quad (47)$$

where the last inequality follows from the union bound. Take $\epsilon = C_1 m_{ij} r^2/2$. Note that $\{Z^{(k)}_{pq}\}_{1 \leq k \leq m_{ij}}$ is a sequence of independent random variables with $E(Z^{(k)}_{pq}) = 0$, and $|Z^{(k)}_{pq}| \leq r^2/4$, for $1 \leq k \leq m_{ij}$. Applying Hoeffding’s inequality,

$$\Pr\left( \sum_{k=1}^{m_{ij}} Z^{(k)}_{pq} > \frac{C_1 m_{ij} r^2}{2d} \right) \leq 2 \exp\left( -\frac{2C_1^2 m_{ij}}{d^2} \right) \leq 2n^{-3}. \quad (48)$$

Combining Eqs. (47) and (48), we obtain

$$\Pr\left( \sigma_{\text{max}}(\sum_{k=1}^{m_{ij}} Z^{(k)}) > \frac{C_1 m_{ij} r^2}{2} \right) \leq 2d^2 n^{-3}. \quad (49)$$
Using Eqs. (45), (46), and (49), we have

\[ \sigma_{\min}(M^{(ij)}) \geq \sigma_{\min}\left(\sum_{k=1}^{m_{ij}} \mathbb{E}(z_k z_k^T)\right) - \sigma_{\max}\left(\sum_{k=1}^{m_{ij}} Z^{(k)}\right) \geq \frac{C_1 m_{ij} r^2}{2}, \]

with probability at least \(1 - 2d^2 n^{-3}\). Applying a union bound over all pairs \( (i, j) \in E(G^*) \), we obtain that, w.h.p., \( \sigma_{\min}(M^{(ij)}) \geq C_1 m_{ij} r^2 / 2 \geq C(nr^d)r^2 \), for all \( (i, j) \in E(G^*) \). Invoking Eq. (44),

\[ \left\| \sum_{\ell=1}^{d} (\beta^{(\ell)}_j - \beta^{(\ell)}_i) \tilde{x}^{(\ell)}_{Q_i \cap Q_j} \right\|^2 \geq C(nr^d)r^2 \|\beta_{ij}\|^2 = C(nr^d)r^2 \sum_{\ell=1}^{d} (\beta^{(\ell)}_j - \beta^{(\ell)}_i)^2. \]

**Appendix H: Proof of Claim 5.6**

**Proof** Let \( N = |V(G^*)| = n(m + 1) \). Define \( \bar{\beta}^{(\ell)} = (1/N) \sum_{i=1}^{N} \beta^{(\ell)}_i \) and let \( \tilde{v} = v - \sum_{\ell=1}^{d} \bar{\beta}^{(\ell)} x^{(\ell)} \). Then, the vector \( \tilde{v} \) has the following local decompositions.

\[ \tilde{v}_{Q_i} = \sum_{\ell=1}^{d} (\beta^{(\ell)}_i - \bar{\beta}^{(\ell)}) \tilde{x}^{(\ell)}_{Q_i} + \tilde{\gamma}_i u_{Q_i} + w^{(i)}, \]

where \( \tilde{\gamma}_i = \gamma_i - \sum_{\ell=1}^{d} \bar{\beta}^{(\ell)} \frac{1}{|Q_i|} \langle \tilde{x}^{(\ell)}_{Q_i}, u_{Q_i} \rangle \). For convenience, we establish the following definitions.

\( M \in \mathbb{R}^{d \times d} \) is a matrix with \( M_{\ell, \ell'} = \langle x^{(\ell)}, x^{(\ell')} \rangle \). Also, for any \( 1 \leq i \leq N \), define the matrix \( M^{(i)} \in \mathbb{R}^{d \times d} \) as \( M^{(i)}_{\ell, \ell'} = \langle \tilde{x}^{(\ell)}_{Q_i}, \tilde{x}^{(\ell')}_{Q_i} \rangle \). Let \( \hat{\beta}^{(\ell)}_i := \beta^{(\ell)}_i - \bar{\beta}^{(\ell)} \) and \( \eta_{i}^{(\ell)} = \sum_{\ell'} M^{(i)}_{\ell, \ell'} \hat{\beta}^{(\ell')} \). Finally, for any \( 1 \leq \ell \leq d \), define the matrix \( B^{(\ell)} \in \mathbb{R}^{N \times n} \) as follows:

\[ B^{(\ell)}_{i, j} = \begin{cases} \tilde{x}^{(\ell)}_{Q_i, j} & \text{if } j \in Q_i, \\ 0 & \text{if } j \notin Q_i. \end{cases} \]

Now, note that \( \langle \tilde{v}_{Q_i}, \tilde{x}^{(\ell)}_{Q_i} \rangle = \sum_{\ell'=1}^{d} M^{(i)}_{\ell, \ell'} \hat{\beta}^{(\ell')} = \eta_{i}^{(\ell)} \). Writing it in matrix form, we have \( B^{(\ell)} \tilde{v} = \eta^{(\ell)} \).

Our first lemma provides a lower bound for \( \sigma_{\min}(B^{(\ell)}) \). For its proof, we refer to Sect. H.1.

**Lemma H.1** Let \( \tilde{G} = (V, \tilde{E}) \), where \( \tilde{E} = \{(i, j) : d_{ij} \leq r/2\} \) and denote by \( \tilde{\mathcal{L}} \) the Laplacian of \( \tilde{G} \). Then, there exists a constant \( C = C(d) \), such that, w.h.p.,

\[ B^{(\ell)} (B^{(\ell)})^T \geq C(nr^d)^{-1} r^2 \tilde{\mathcal{L}}, \quad \forall 1 \leq \ell \leq d. \]

The next lemma establishes some properties of the spectra of the matrices \( M \) and \( M^{(i)} \). Its proof is deferred to Sect. H.2.
Lemma H.2 There exist constants $C_1$ and $C_2$ such that, w.h.p.,

$$
\sigma_{\min}(M) \geq C_1 n, \quad \sigma_{\max}(M^{(i)}) \leq C_2 (nr^d)r^2, \quad \forall \ 1 \leq i \leq N.
$$

Now, we are in position to prove Claim 5.6. Using Lemma H.1 and since $\langle \tilde{v}, u \rangle = 0$,

$$
\| \eta^{(\ell)} \|^2 \geq \sigma_{\min}(B^{(\ell)}(B^{(\ell)})^T) \| \tilde{v} \|^2 \geq C(nr^d)^{-1}r^2 \sigma_{\min}(\tilde{L}) \geq Cr^4 \| \tilde{v} \|^2,
$$

for some constant $C$. The last inequality follows from the lower bound on $\sigma_{\min}(\tilde{L})$ provided by Remark 2.2. Moreover,

$$
\left[ \sum_{\ell'=1}^{d} M_{\ell, \ell'} \tilde{\beta}^{(\ell')} \right]^2 = \langle \tilde{v}, x^{(\ell)} \rangle^2 \leq \| \tilde{v} \|^2 \| x^{(\ell)} \|^2 \leq Cr^{-4} \| \eta^{(\ell)} \|^2 \| x^{(\ell)} \|^2.
$$

Summing both sides over $\ell$ and using $\| x^{(\ell)} \|^2 \leq Cn$, we obtain

$$
\sum_{\ell=1}^{d} \left[ \sum_{\ell'=1}^{d} M_{\ell, \ell'} \tilde{\beta}^{(\ell')} \right]^2 \leq C(nr^{-4}) \sum_{\ell=1}^{d} \| \eta^{(\ell)} \|^2.
$$

Equivalently,

$$
\sum_{\ell=1}^{d} \langle M_{\ell, \ldots}, \tilde{\beta} \rangle^2 \leq C(nr^{-4}) \sum_{\ell=1}^{d} \sum_{i=1}^{N} \langle M^{(i)}_{\ell, \ldots}, \hat{\beta}_i \rangle^2.
$$

Here, $\tilde{\beta} = (\tilde{\beta}^{(1)}, \ldots, \tilde{\beta}^{(d)}) \in \mathbb{R}^d$ and $\hat{\beta}_i = (\hat{\beta}_i^{(1)}, \ldots, \hat{\beta}_i^{(d)}) \in \mathbb{R}^d$. Writing this in matrix form,

$$
\| M \tilde{\beta} \|^2 \leq C(nr^{-4}) \sum_{i=1}^{N} \| M^{(i)} \hat{\beta}_i \|^2.
$$

Therefore,

$$
\sigma_{\min}^2(M) \| \tilde{\beta} \|^2 \leq C(nr^{-4}) \left[ \max_{1 \leq i \leq N} \sigma_{\max}^2(M^{(i)}) \right] \sum_{i=1}^{N} \| \hat{\beta}_i \|^2.
$$

Using the bounds on $\sigma_{\min}(M)$ and $\sigma_{\max}(M^{(i)})$ provided in Lemma H.2, we obtain

$$
\| \tilde{\beta} \|^2 \leq \frac{C}{n} (nr^d)^2 \sum_{i=1}^{N} \| \hat{\beta}_i \|^2. \quad (50)
$$

Now, note that

$$
\| \tilde{\beta} \|^2 = \sum_{\ell=1}^{d} (\tilde{\beta}^{(\ell)})^2 = \sum_{\ell=1}^{d} \left( \sum_{i=1}^{N} \beta_i^{(\ell)} \right)^2 = \frac{1}{N} \sum_{\ell=1}^{d} \| P_{\ell} \beta^{(\ell)} \|^2, \quad (51)
$$

Springer
\[
\sum_{i=1}^{N} \| \hat{\beta}_i \|^2 = \sum_{\ell=1}^{d} \left( \sum_{i=1}^{N} (\beta^{(\ell)}_i - \bar{\beta}^{(\ell)})^2 \right) = \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2. \tag{52}
\]

Consequently,
\[
\sum_{\ell=1}^{d} \left\| \beta^{(\ell)} \right\|^2 = \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2 + \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2,
\]
\[
\overset{(a)}{\equiv} N \| \hat{\beta} \|^2 + \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2,
\]
\[
\overset{(b)}{\leq} \frac{CN}{n} (nr^d)^2 \sum_{i=1}^{N} \| \hat{\beta}_i \|^2 + \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2
\]
\[
\overset{(c)}{=} \left( 1 + \frac{CN}{n} (nr^d)^2 \right) \sum_{\ell=1}^{d} \left\| P_u \beta^{(\ell)} \right\|^2
\]
\[
\leq C (nr^d)^3 \sum_{\ell=1}^{d} \left\| P_u \left( \beta^{(\ell)} \right) \right\|^2.
\]

Here, (a) follows from Eq. (51), (b) follows from Eq. (50), and (c) follows from Eq. (52). The result follows. \qed

### H.1 Proof of Lemma H.1

Recall that \(e_{ij} \in \mathbb{R}^n\) is the vector with +1 at the \(i\)th position, −1 at the \(j\)th position, and zero everywhere else. For any two nodes \(i\) and \(j\) with \(\|x_i - x_j\| \leq r/2\), choose a node \(k \in \tilde{C}_i \cap \tilde{C}_j\) uniformly at random and consider the cliques \(Q_1 = C_k\), \(Q_2 = C_k \setminus i\), and \(Q_3 = C_k \setminus j\). Define \(S_{ij} = \{Q_1, Q_2, Q_3\}\). Note that \(S_{ij} \subset \text{cliq}^*(G)\).

Let \(a_1, a_2,\) and \(a_3\) respectively denote the center of mass of the points in cliques \(Q_1, Q_2,\) and \(Q_3\). Find scalars \(\xi_1^{(ij)}, \xi_2^{(ij)},\) and \(\xi_3^{(ij)}\), such that
\[
\begin{aligned}
\xi_1^{(ij)} + \xi_2^{(ij)} + \xi_3^{(ij)} &= 0, \\
\xi_1^{(ij)} a_1^{(\ell)} + \xi_2^{(ij)} a_2^{(\ell)} + \xi_3^{(ij)} a_3^{(\ell)} &= 0, \\
\xi_1^{(ij)} (x_i^{(\ell)} - a_1^{(\ell)}) + \xi_3^{(ij)} (x_i^{(\ell)} - a_3^{(\ell)}) &= 1.
\end{aligned}
\tag{53}
\]

Note that the space of the solutions of this linear system of equations is invariant to translation of the points. Hence, without loss of generality, assume that \(\sum_{\ell \in Q_1 \setminus i \neq j} x_\ell = 0\). Also, let \(m = |C_k|\). Then, it is easy to see that
\[
a_1 = \frac{x_i + x_j}{m}, \quad a_2 = \frac{x_j}{m}, \quad a_3 = \frac{x_i}{m},
\]

\(\mathbb{E}\) Springer
and the solution of equations (53) is given by
\[
\begin{align*}
\xi_1^{(ij)} &= \frac{x_j^{(\ell)} - x_i^{(\ell)}}{x_j^{(\ell)} (x_i^{(\ell)} - \frac{x_j^{(\ell)}}{m})}, \\
\xi_2^{(ij)} &= -\frac{x_j^{(\ell)}}{x_j^{(\ell)} (x_i^{(\ell)} - \frac{x_j^{(\ell)}}{m})}, \\
\xi_3^{(ij)} &= \frac{x_i^{(\ell)}}{x_j^{(\ell)} (x_i^{(\ell)} - \frac{x_j^{(\ell)}}{m})}.
\end{align*}
\]

First, observe that
\[
\begin{align*}
\xi_1^{(ij)} (x_i^{(\ell)} - a_1^{(\ell)}) + \xi_2^{(ij)} (x_i^{(\ell)} - a_3^{(\ell)}) &= 1, \\
\xi_1^{(ij)} (x_j^{(\ell)} - a_1^{(\ell)}) + \xi_2^{(ij)} (x_j^{(\ell)} - a_2^{(\ell)}) &= -1, \\
\text{For } t \in C_k \text{ and } t \neq i, j: \\
\xi_1^{(ij)} (x_t^{(\ell)} - a_1^{(\ell)}) + \xi_2^{(ij)} (x_t^{(\ell)} - a_2^{(\ell)}) + \xi_3^{(ij)} (x_t^{(\ell)} - a_3^{(\ell)}) &= 0.
\end{align*}
\]

Therefore,
\[
\xi_1^{(ij)} x_{Q_1,t}^{(\ell)} + \xi_2^{(ij)} x_{Q_2,t}^{(\ell)} + \xi_3^{(ij)} x_{Q_3,t}^{(\ell)} = \begin{cases} 
1 & \text{if } t = i, \\
-1 & \text{if } t = j, \\
0 & \text{if } t \in C_k, t \neq i, j.
\end{cases}
\tag{54}
\]

Let \( \xi^{(ij)} \in \mathbb{R}^N \) be the vector with \( \xi_1^{(ij)}, \xi_2^{(ij)}, \) and \( \xi_3^{(ij)} \) at the positions corresponding to the cliques \( Q_1, Q_2, Q_3 \) and zero everywhere else. Then, Eq. (54) gives \( (B^{(\ell)})^T \xi^{(ij)} = e_{ij} \).

Second, note that \( \|\xi^{(ij)}\|^2 = (\xi_1^{(ij)})^2 + (\xi_2^{(ij)})^2 + (\xi_3^{(ij)})^2 \leq C r^2 \), for some constant \( C \).

Now, we are in position to prove Lemma H.1.

For any vector \( z \in \mathbb{R}^n \), we have
\[
(\mathcal{L}z, z) = \sum_{(i,j) \in \mathcal{E}} (e_{ij}, z)^2
\]
\[
= \sum_{(i,j) \in \mathcal{E}} (\xi^{(ij)}, B^{(\ell)}z)^2 = \sum_{(i,j) \in \mathcal{E}} \left( \sum_{Q_t \in S_{ij}} (\xi_t^{(ij)}, B^{(\ell)}z) \right)^2
\]
\[
\leq \sum_{(i,j) \in \mathcal{E}} \left( \sum_{Q_t \in S_{ij}} (\xi_t^{(ij)})^2 \right) \left( \sum_{Q_t \in S_{ij}} (B^{(\ell)}z)^2 \right)
\]
\[
\leq \max_{(i,j) \in \mathcal{E}} \|\xi^{(ij)}\|^2 \sum_{Q_t \in S_{ij}} (B^{(\ell)}z)^2 \left( \sum_{s_j \supseteq Q_t} 1 \right)
\]
\[
\leq \frac{C}{r^2} (n_r^d) \|B^{(\ell)}z\|^2.
\]
Hence, $B(\ell)(B(\ell))^T \succeq C(nr^d)^{-1}r^2 \tilde{C}$.

H.2 Proof of Lemma H.2

First, we prove that $\sigma_{\min}(M) \geq Cn$, for some constant $C$.

By definition, $M = \sum_{i=1}^{n} x_ix_i^T$. Let $Z_i = x_ix_i^T \in \mathbb{R}^{d \times d}$, and $\bar{Z} = 1/n \sum_{i=1}^{n} Z_i$. Note that $\{Z_i\}_{1 \leq i \leq n}$ is a sequence of i.i.d. random matrices with $Z = \mathbb{E}(Z_i) = 1/12I_d \times d$. By the law of large numbers we have $\bar{Z} \to Z$, almost surely. In addition, since $\sigma_{\max}(.)$ is a continuous function of its argument, we obtain $\sigma_{\max}(\bar{Z} - Z) \to 0$, almost surely. Therefore,

$$\sigma_{\min}(M) = n\sigma_{\min}(\bar{Z}) \geq n(\sigma_{\min}(Z) - \sigma_{\max}(\bar{Z} - Z))$$

$$= n\left(\frac{1}{12} - \sigma_{\max}(\bar{Z} - Z)\right),$$

whence we obtain $\sigma_{\min}(M) \geq n/12$, w.h.p.

Now we prove the second part of the claim.

Let $m_i = |\mathcal{Q}_i|$, for $1 \leq i \leq N$. Since $M^{(i)} \succeq 0$, we have

$$\sigma_{\max}(M^{(i)}) \leq \text{Tr}(M^{(i)}) = \sum_{\ell=1}^{d} \|\tilde{x}_{\ell}^{(i)}\|^2 \leq Cm_ir^2.$$

With high probability, $m_i \leq C(nr^d)$, $\forall 1 \leq i \leq N$, and for some constant $C$. Hence,

$$\max_{1 \leq i \leq N} \sigma_{\max}(M^{(i)}) \leq C(nr^d)r^2,$$

w.h.p. The result follows.

Appendix I: Proof of Proposition 6.1

Proof Recall that $\tilde{R} = XY^T + YX^T$ with $X, Y \in \mathbb{R}^{n \times d}$ and $Y^Tu = 0$. By the triangle inequality, we have

$$|\langle x_i - x_j, y_i - y_j \rangle| \geq |\langle x_i, y_j \rangle + \langle x_j, y_i \rangle| - |\langle x_i, y_i \rangle| - |\langle x_j, y_j \rangle|$$

$$= |\tilde{R}_{ij}| - \frac{|\tilde{R}_{ii}|}{2} - \frac{|\tilde{R}_{jj}|}{2}.$$

Therefore,

$$\sum_{i,j} |\langle x_i - x_j, y_i - y_j \rangle| \geq \sum_{i,j} |\tilde{R}_{ij}| - n \sum_{i} |\tilde{R}_{ii}|. \quad (55)$$

Again, by the triangle inequality,
\[ \sum_{ij} |\langle x_i - x_j, y_i - y_j \rangle| \geq \sum_i n \langle x_i, y_i \rangle + \sum_j \langle x_j, y_j \rangle - \langle x_i, \sum_j y_j \rangle - \langle \sum_j x_j, y_i \rangle \]

\[ = n \sum_i \langle x_i, y_i \rangle + \frac{1}{n} \sum_j \langle x_j, y_j \rangle, \quad (56) \]

where the last equality follows from \( Y^T u = 0 \) and \( X^T u = 0 \).

**Remark 17.1** For any \( n \) real values \( \xi_1, \ldots, \xi_n \), we have

\[ \sum_i |\xi_i + \bar{\xi}| \geq \frac{1}{2} \sum_i |\xi_i|, \]

where \( \bar{\xi} = \frac{1}{n} \sum_i \xi_i \).

**Proof of Remark 17.1** Without loss of generality, we assume \( \bar{\xi} \geq 0 \). Then,

\[ \sum_i |\xi_i + \bar{\xi}| \geq \sum_{i: \xi_i \geq 0} \xi_i \geq \frac{1}{2} \left( \sum_{i: \xi_i \geq 0} \xi_i - \sum_{i: \xi_i < 0} \xi_i \right) = \frac{1}{2} \sum_i |\xi_i|, \]

where the second inequality follows from \( \sum_i \xi_i = n \bar{\xi} \geq 0 \). \( \square \)

Using Remark 17.1 with \( \xi_i = \langle x_i, y_i \rangle \), Eq. (56) yields

\[ \sum_{ij} |\langle x_i - x_j, y_i - y_j \rangle| \geq \frac{n}{2} \sum_i |\langle x_i, y_i \rangle| = \frac{n}{4} \sum_i |\tilde{R}_{ii}|. \quad (57) \]

Combining Eqs. (55) and (57), we obtain

\[ \| R_{Kn} X(Y) \|_1 = \sum_{ij} |\langle x_i - x_j, y_i - y_j \rangle| \geq \frac{1}{5} \| \tilde{R} \|_1. \quad (58) \]

which proves the desired result. \( \square \)

**Appendix J: Proof of Lemma 6.1**

We will compute the average number of chains passing through a particular edge in the order notation. Notice that the total number of chains is \( \Theta(n^2) \) since there are \( \binom{n}{2} \) node pairs. Each chain has \( O(1/r) \) vertices and thus intersects \( O(1/r) \) bins. The total number of bins is \( \Theta(1/r^d) \). Hence, by symmetry, the number of chains intersecting each bin is \( \Theta(n^2 r^{d-1}) \). Consider a particular bin \( B \), and the chains intersecting it. Such chains are equally likely to select any of the nodes in \( B \). Since the expected number of nodes in \( B \) is \( \Theta(n r^d) \), the average number of chains containing a particular node, say \( i \), in \( B \), is \( \Theta(n^2 r^{d-1}/nr^d) = \Theta(nr^{-1}) \). Now consider node \( i \) and one of
its neighbors in the chain, say \( j \). Denote by \( B^\ast \) the bin containing node \( j \). The number of edges between \( i \) and \( B^\ast \) is \( \Theta(nr^d) \). Hence, by symmetry, the average number of chains containing an edge incident on \( i \) will be \( \Theta(nr^{-1}/nr^d) = \Theta(r^{-d-1}). \) This is true for all nodes. Therefore, the average number of chains containing any particular edge is \( O(r^{-d-1}) \). In other words, on average, no edge belongs to more than \( O(r^{-d-1}) \) chains.

Appendix K: The Two-Part Procedure for General \( d \)

In the proof of Lemma 6.2, we stated a two-part procedure to find the values \( \{ \lambda_{lk} \}_{(l,k) \in E(G_{ij})} \) that satisfy Eq. (21). Part (i) of the procedure was demonstrated for the special case \( d = 2 \). Here, we discuss this part for general \( d \).

Let \( G_{ij} = \{ i \} \cup \{ j \} \cup H_1 \cup \cdots \cup H_k \) be the chain between nodes \( i \) and \( j \). Let \( F_p = H_p \cap H_{p+1} \). Without loss of generality, assume \( V(F_p) = \{ 1, 2, \ldots, q \} \), where \( q = 2^{d-1} \). The goal is to find a set of forces, namely \( f_1, \ldots, f_q \), such that

\[
\sum_{i=1}^{q} f_i = x_m, \quad \sum_{i=1}^{q} f_i \wedge x_i = 0, \quad \sum_{i=1}^{q} \| f_i \|^2 \leq C \| x_m \|^2.
\]

(59)

It is more convenient to write this problem in matrix form. Let \( X = [x_1, x_2, \ldots, x_q] \in \mathbb{R}^{d \times q} \) and \( \Phi = [f_1, f_2, \ldots, f_q] \in \mathbb{R}^{d \times q} \). Then, the problem can be recast as finding a matrix \( \Phi \in \mathbb{R}^{d \times d} \) such that

\[
\Phi u = x_m, \quad X \Phi^T = \Phi X^T, \quad \| \Phi \|_F^2 \leq C \| x_m \|^2.
\]

(60)

Define \( \tilde{X} = X(I - 1/q uu^T) \), where \( I \in \mathbb{R}^{q \times q} \) is the identity matrix and \( u \in \mathbb{R}^q \) is the all-ones vector. Let

\[
\Phi = \frac{1}{q} x_m u^T + \left( \frac{1}{q} X u x_m^T + S \right)(\tilde{X} \tilde{X}^T)^{-1} \tilde{X},
\]

(61)

where \( S \in \mathbb{R}^{d \times d} \) is an arbitrary symmetric matrix. Observe that

\[
\Phi u = x_m, \quad X \Phi^T = \Phi X^T.
\]

(62)

Now, we only need to find a symmetric matrix \( S \in \mathbb{R}^{d \times d} \) such that the matrix \( \Phi \) given by Eq. (61) satisfies \( \| \Phi \|_F \leq C \| x_m \| \). Without loss of generality, assume that the vector \( x_m \) is in the direction \( e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^{d} \). Let \( x_c = \frac{1}{q} X u \) be the center of the nodes \( \{ x_i \}_{i=1}^{q} \), and let \( x_c = (x_c^{(1)}, \ldots, x_c^{(d)}) \). Take \( S = -\| x_m \| x_c^{(1)} e_1 e_1^T \). From the construction of the chain \( G_{ij} \), the nodes \( \{ x_i \}_{i=1}^{q} \) are obtained by wiggling the vertices of a hypercube aligned in the direction \( x_m/\| x_m \| = e_1 \), and with side length...
\( \tilde{r} = 3r/4\sqrt{2} \) (each node wiggles by at most \( \frac{r}{8} \)). Therefore, \( x_c \) is almost aligned with \( e_1 \), and has small components in the other directions. Formally, \( |x_c^{(i)}| \leq \frac{r}{8} \), for \( 2 \leq i \leq d \). Therefore,

\[
\frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T + S = \left( \sum_{i=1}^{d} x_c^{(i)} e_i \right) \cdot \left( \| \mathbf{x}_m \| e_1 \right) - \| \mathbf{x}_m \| x_c^{(1)} e_1 e_1^T \\
= \sum_{i=2}^{d} \| \mathbf{x}_m \| x_c^{(i)} e_i e_1^T .
\]

Hence, \( \frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T + S \in \mathbb{R}^{d \times d} \) has entries bounded by \( \frac{r}{8} \| \mathbf{x}_m \| \). In the following we show that there exists a constant \( C = C(d) \) such that all entries of \( (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} \) are bounded by \( C/r \). Once we show this, it follows that

\[
\| \mathbf{\Phi} \|_F \leq \left\| \frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T \right\|_F + \left\| \left( \frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T + S \right) (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} \right\|_F \leq C \| \mathbf{x}_m \| ,
\]

for some constant \( C = C(d) \). Therefore,

\[
\| \mathbf{\Phi} \|_F \leq \left\| \frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T \right\|_F + \left\| \left( \frac{1}{q} \mathbf{X} \mathbf{u} \mathbf{x}_m^T + S \right) (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} \right\|_F \leq C \| \mathbf{x}_m \| ,
\]

for some constant \( C \).

We are now left with the task of showing that all entries of \( (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} \) are bounded by \( C/r \), for some constant \( C \).

The nodes \( x_i \) were obtained by wiggling the vertices of a hypercube of side length \( \tilde{r} = 3r/4\sqrt{2} \) (each node wiggles by at most \( r/8 \)). Let \( \{ z_i \}_{i=1}^{q} \) denote the vertices of this hypercube, and thus \( \| x_i - z_i \| \leq \frac{r}{8} \). Define

\[
Z = \frac{1}{\tilde{r}} [z_1, \ldots, z_q], \quad \delta Z = \frac{1}{\tilde{r}} \tilde{X} - Z.
\]

Then, \( \tilde{X} \tilde{X}^T = \tilde{r}^2 (Z + \delta Z)(Z + \delta Z)^T = \tilde{r}^2 (ZZ^T + \tilde{Z}) \), where \( \tilde{Z} = Z(\delta Z)^T + (\delta Z)Z^T + (\delta Z)(\delta Z)^T \). Consequently,

\[
(\tilde{X} \tilde{X}^T)^{-1} \tilde{X} = \frac{1}{\tilde{r}} (ZZ^T + \tilde{Z})^{-1} (Z + \delta Z).
\]

Now notice that the columns of \( Z \) represent the vertices of a unit \((d - 1)\)-dimensional hypercube. Also, the norm of each column of \( \delta Z \) is bounded by \( \frac{r}{8\tilde{r}} < \frac{1}{4} \). Therefore, \( \sigma_{\min}(ZZ^T + \tilde{Z}) \geq C \), for some constant \( C = C(d) \). Hence, for every \( 1 \leq i \leq q \),

\[
\| (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} e_i \| \leq \frac{1}{\tilde{r}} \sigma_{\min}^{-1}(ZZ^T + \tilde{Z}) \| (Z + \delta Z)e_i \| \leq \frac{C}{r},
\]

for some constant \( C \). Therefore, all entries of \( (\tilde{X} \tilde{X}^T)^{-1} \tilde{X} \) are bounded by \( C/r \).
Appendix L: Proof of Remark 7.1

Let \( \theta \) be the angle between \( a \) and \( b \) and define \( a_\perp = \frac{b - \cos(\theta) a}{\|b - \cos(\theta) a\|} \). Therefore, \( b = \cos(\theta) a + \sin(\theta) a_\perp \). In the basis \( (a, a_\perp) \), we have

\[
aa^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad bb^T = \begin{bmatrix} \cos^2(\theta) & \sin(\theta) \cos(\theta) \\ \sin(\theta) \cos(\theta) & \sin^2(\theta) \end{bmatrix}.
\]

Therefore,

\[
\|aa^T - bb^T\|_2 = \left\| \begin{bmatrix} \sin^2(\theta) & -\sin(\theta) \cos(\theta) \\ -\sin(\theta) \cos(\theta) & -\sin^2(\theta) \end{bmatrix} \right\|_2 = |\sin(\theta)| = \sqrt{1 - (a^T b)^2}.
\]

Appendix M: Proof of Remark 7.2

**Proof** Let \( \{\tilde{\lambda}_i\} \) be the eigenvalues of \( \tilde{A} \) such that \( \tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_p \). Notice that

\[
\|A - \tilde{A}\|_2 \geq v^T (\tilde{A} - A)v \\
\geq \tilde{\lambda}_p (v^T \tilde{v})^2 + \tilde{\lambda}_{p-1} \|P_{v_\perp}(v)\|^2 - \lambda_p \\
= \tilde{\lambda}_p (v^T \tilde{v})^2 + \tilde{\lambda}_{p-1} (1 - (v^T \tilde{v})^2) - \lambda_p \\
= (\tilde{\lambda}_p - \tilde{\lambda}_{p-1})(v^T \tilde{v})^2 + \tilde{\lambda}_{p-1} - \lambda_p.
\]

Therefore,

\[
(v^T \tilde{v})^2 \geq \frac{\tilde{\lambda}_{p-1} - \lambda_p - \|A - \tilde{A}\|_2}{\tilde{\lambda}_{p-1} - \lambda_p}.
\]

Furthermore, due to Weyl’s inequality, \( |\tilde{\lambda}_i - \lambda_i| \leq \|A - \tilde{A}\|_2 \). Therefore,

\[
(v^T \tilde{v})^2 \geq \frac{\lambda_{p-1} - \lambda_p - 2\|A - \tilde{A}\|_2}{\lambda_{p-1} - \lambda_p + 2\|A - \tilde{A}\|_2},
\]

which implies the thesis after some algebraic manipulations. \( \square \)
Appendix N: Table of Symbols

Table 1  Table of symbols

| Symbol | Description |
|--------|-------------|
| n      | Number of nodes |
| d      | Dimension (the nodes are scattered in \([-0.5, 0.5]^d\)) |
| L ∈ R^{n×n} | \(I - \frac{1}{n} uu^T\), where \(I\) is the identity matrix and \(u\) is the all-ones vector |
| x_i ∈ R^d | Coordinate of node \(i\), for \(1 ≤ i ≤ n\) |
| x(ℓ) ∈ R^n | The vector containing the \(\ell\)th coordinate of the nodes, for \(1 ≤ \ell ≤ d\) |
| X ∈ R^{n×d} | The (original) position matrix |
| \(\hat{X}\) | Estimated position matrix |
| Q ∈ R^{n×n} | Solution of SDP in the first step of the algorithm |
| Q_0 ∈ R^{n×n} | Gram matrix of the node (original) positions, namely \(Q_0 = XX^T\) |
| Subspace \(V\) | The subspace spanned by vectors \(x(1), \ldots, x(d), u\) |
| \(R \in R^{n×n}\) | \(P_V R P_V + P_V R P_V^\perp + P_V^\perp R P_V\) |
| \(R^\perp \in R^{n×n}\) | \(P_V^\perp R P_V^\perp\) |
| \(C_i\) | \(\{j ∈ V(G) : d_{ij} ≤ r/2\}\) (the nodes in \(C_i\) form a clique in \(G(n, r)\)) |
| \(\tilde{C}_i\) | \(\{j ∈ V(G) : d_{ij} ≤ r/2(1/2 + 1/100)\}\) |
| \(S_i\) | \(\{C_i, C_i \setminus k\}_{k ∈ C_i}\) |
| \(\text{cliq}(G)\) | \(\{S_1 ∪ \ldots ∪ S_n\}\) |
| \(V^\perp\) | The subspace of vectors \(x(1), \ldots, x(d), u\) orthogonal to the all-ones vector \(u\) |
| \(\beta(\ell)_i\) | Coefficients in local decomposition of an arbitrary (fixed) vector \(v \in V^\perp\) |
| \(\bar{\beta}(\ell)_i\) | Average of numbers \(\beta(\ell)_1, \ldots, \beta(\ell)_N\) |
| \(\tilde{\beta}(\ell)_i\) | \(\beta(\ell)_i - \bar{\beta}(\ell)\) |
| \(\hat{\beta}(\ell)_i\) | \(\beta(\ell)_i^{(1)}, \ldots, \beta(\ell)_i^{(d)}\), for \(i = 1,\ldots, N\) |
References

1. A.Y. Alfakih, A. Khandani, H. Wolkowicz, Solving Euclidean distance matrix completion problems via semidefinite programming, *Comput. Optim. Appl.* **12**, 13–30 (1999).

2. L. Asimow, B. Roth, The rigidity of graphs, *Trans. Am. Math. Soc.* **245**, 279–289 (1978).

3. J. Aspnes, T. Eren, D.K. Goldberg, A.S. Morse, W. Whiteley, Y.R. Yang, B.D.O. Anderson, P.N. Belhumeur, A theory of network localization, *IEEE Trans. Mob. Comput.* **5**(12), 1663–1678 (2006).

4. M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Comput.* **15**, 1373–1396 (2002).

5. M. Bernstein, V. de Silva, J. Langford, J. Tenenbaum, *Graph approximations to geodesics on embedded manifolds*. Technical Report, Stanford University, Stanford, CA, 2000.

6. P. Biswas, Y. Ye, Semidefinite programming for ad hoc wireless sensor network localization, in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks, IPSN ’04* (ACM, New York, 2004), pp. 46–54.

7. S.P. Boyd, A. Ghosh, B. Prabhakar, D. Shah, Mixing times for random walks on geometric random graphs, in *Proceedings of the 7th Workshop on Algorithm Engineering and Experiments and the 2nd Workshop on Analytic Algorithmics and Combinatorics, ALENEX/ANALCO 2005*, Vancouver, BC, Canada, 22 January 2005 (SIAM, Philadelphia, 2005), pp. 240–249.

8. S. Butler, *Eigenvalues and structures of graphs*. Ph.D. thesis, University of California, San Diego, CA, 2008.

9. R. Connelly, Generic global rigidity, *Discrete Comput. Geom.* **33**, 549–563 (2005).

10. T. Cox, M. Cox, *Multidimensional Scaling*, Monographs on Statistics and Applied Probability, vol. 88 (Chapman & Hall, London, 2001).

11. P. Diaconis, L. Saloff-Coste, Comparison theorems for reversible Markov chains, *Ann. Appl. Probab.* **3**(3), 696–730 (1993).

12. D.L. Donoho, C. Grimes, Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data, *Proc. Natl. Acad. Sci. USA* **100**(10), 5591–5596 (2003).

13. S.J. Gortler, A.D. Healy, D.P. Thurston, Characterizing generic global rigidity, *Am. J. Math.* **132**, 897–939 (2010).

14. F. Lu, S.J. Wright, G. Wahba, Framework for kernel regularization with application to protein clustering, *Proc. Natl. Acad. Sci. USA* **102**(35), 12332–12337 (2005).

15. G. Mao, B. Fidan, B.D.O. Anderson, Wireless sensor network localization techniques, *Comput. Netw. ISDN Syst.* **51**, 2529–2553 (2007).

16. S. Oh, A. Karbasi, A. Montanari, Sensor network localization from local connectivity: performance analysis for the MDS-MAP algorithm, in *IEEE Information Theory Workshop 2010* (ITW 2010) (2010).

17. N. Patwari, J.N. Ash, S. Kyperountas, R. Moses, N. Correal, Locating the nodes: cooperative localization in wireless sensor networks, *IEEE Signal Process. Mag.* **22**, 54–69 (2005).

18. M. Penrose, *Random Geometric Graphs* (Oxford University Press, Oxford, 2003).

19. L.K. Saul, V. Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, *J. Mach. Learn. Res.* **4**, 119–155 (2003).

20. D. Shamsi, Y. Ye, N. Taheri, On sensor network localization using SDP relaxation. arXiv:1010.2262 (2010).

21. A. Singer, A remark on global positioning from local distances, *Proc. Natl. Acad. Sci. USA* **105**(28), 9507–9511 (2008).

22. A.M.-C. So, Y. Ye, Theory of semidefinite programming for sensor network localization, in *Symposium on Discrete Algorithms* (2005), pp. 405–414.

23. J.B. Tenenbaum, V. Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, *Science* **290**(5500), 2319–2323 (2000).

24. K.Q. Weinberger, L.K. Saul, An introduction to nonlinear dimensionality reduction by maximum variance unfolding, in *Proceedings of the 21st National Conference on Artificial Intelligence*, 2 (AAAI Press, Menlo Park, 2006), pp. 1683–1686.

25. Z. Zhu, A.M.-C. So, Y. Ye, Universal rigidity: towards accurate and efficient localization of wireless networks, in *IEEE International Conference on Computer Communications* (2010), pp. 2312–2320.