Exact Camera Location Recovery by Least Unsquared Deviations*

Gilad Lerman[1], Yunpeng Shi[1], and Teng Zhang[2]

[1]School of Mathematics, University of Minnesota
[2]Department of Mathematics, University of Central Florida
{lerman, shixx517}@umn.edu, Teng.Zhang@ucf.edu

arXiv:1709.09683v2 [cs.CV] 3 Oct 2017

Abstract

We establish exact recovery for the Least Unsquared Deviations (LUD) algorithm of Özyesil and Singer. More precisely, we show that for sufficiently many cameras with given corrupted pairwise directions, where both camera locations and pairwise directions are generated by a special probabilistic model, the LUD algorithm exactly recovers the camera locations with high probability. A similar exact recovery guarantee was established for the ShapeFit algorithm by Hand, Lee and Voroninski, but with typically less corruption.

1 Introduction

The Structure from Motion (SfM) problem asks to recover the 3D structure of an object from its 2D images. These images are taken by many cameras at different orientations and locations. In order to recover the underlying structure, both the orientations and locations of the cameras need to be estimated [19].

The common procedure is to first estimate the relative orientations between pairs of cameras from the corresponding essential matrices and then use them to obtain the pairwise directions between cameras [12]. The global orientations up to an arbitrary rotation can be concluded via synchronization from the pairwise orientations [1, 6, 10, 13, 15, 18]. The locations can be derived from the pairwise directions [1, 2, 8, 9, 10, 11, 16, 17, 18, 22, 23].

This paper mathematically addresses the latter subproblem of estimating global camera locations when given corrupted pairwise directions with missing values. In doing so, it follows the corruption model and the mathematical problem of Hand, Lee and Voroninski (HLV) [11], which are described next.

The HLV model: Assume n camera locations $\{\boldsymbol{t}_i^*\}_{i=1}^n \subseteq \mathbb{R}^3$ i.i.d. sampled from $N(\mathbf{0},\mathbf{I})$. Let $G(n,p)$ be the Erdös-Rényi graph of n vertices $V=\{\boldsymbol{t}_i^*\}_{i=1}^n$ with probability of connection p. That is, denoting $[n]=\{1,2,...,n\}$, an edge with index $ij\in[n]\times[n]$ is independently drawn between $\boldsymbol{t}_i^*$ and $\boldsymbol{t}_j^*$ with probability p. Let E denote the set of indices of drawn edges and WLOG assume that if $ij\in E$, then $i<j$, so ji is not repeated. Note that the set of indices of missing edges is $[n]\times[n]\setminus E$. Assume further that $G(V,E)$ is parallel rigid. Namely, all locations $\{\boldsymbol{t}_i^*\}_{i=1}^n$ can be uniquely recovered from the true pairwise directions $\{\boldsymbol{\gamma}_{ij}^*\}_{ij\in E}$, where

$$\boldsymbol{\gamma}_{ij}^*=\frac{\boldsymbol{t}_i^*-\boldsymbol{t}_j^*}{\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|} \tag{1}$$

and $\|\cdot\|$ denotes the Euclidean norm. This assumption is necessary for the well-posedness of the recovery problem. For each edge with index $ij\in E$, a possibly corrupted pairwise direction vector $\boldsymbol{\gamma}_{ij}\in S^2$ is assigned. More precisely, E is partitioned into sets of "good" and "bad" indices, E_g and E_b respectively, and the direction vectors are obtained in each set as follows: If $ij\in E_g$, then $\boldsymbol{\gamma}_{ij}=\boldsymbol{\gamma}_{ij}^*$; otherwise, $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E_b}$ are arbitrarily assigned in S^2. The level of corruption of the HLV model is quantified by $\epsilon_b=\frac{1}{n}$(maximal degree of E_b). Note that $|E_b|<\frac{1}{2}\epsilon_b n^2$, where $|E_b|$ denotes the number of elements in E_b. The parameters of the HLV model are n, p and ϵ_b.

*This work was supported by NSF award DMS-14-18386.

The HLV problem and its solutions: Given data sampled from the HLV model and assuming a bound on the corruption parameter ϵ_b, the exact recovery problem is to reconstruct, up to ambiguous translation and scale, $\{\boldsymbol{t}_i^*\}_{i=1}^n$ from $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$. Hand, Lee and Voroninski addressed this problem while assuming $\epsilon_b = O(p^5/\log^3 n)$ and using their ShapeFit algorithm [11]. Here we address this problem with the weaker assumption $\epsilon_b = O(p^{7/3}/\log^{9/2} n)$, while using the LUD algorithm [17].

1.1 Previous Works

In the past two decades, a variety of algorithms have been proposed for estimating global camera locations from corrupted pairwise directions [19]. The earliest methods use least squares optimization [1, 2, 9] and often result in collapsed solutions. That is, they tend to wrongly estimate many camera locations near the origin. Constrained Least Squares (CLS) [22, 23] utilizes a least squares formulation with an additional constraint to avoid collapsed solutions. Another least squares solver with anti-collapse constraint is semidefinite relaxation (SDR) [18]. Its constraint is non-convex and makes it hard to solve even after convex relaxation. Other solvers include the L_∞ method [16] and the Lie-Algebraic averaging method [10]. However, all the above methods are sensitive to outliers.

Recently, Özyesil and Singer proposed the Least Unsquared Deviation (LUD) algorithm [17] and numerically demonstrated its robustness to outliers and noise. Given the pairwise directions $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$, the LUD algorithm estimates the camera locations $\{\boldsymbol{t}_i^*\}_{i=1}^n$ by $\{\hat{\boldsymbol{t}}_i\}_{i=1}^n \subset \mathbb{R}^3$, which solve the following constrained optimization problem with the additional parameters $\{\hat{\alpha}_{ij}\}_{ij\in E} \subset \mathbb{R}$:

$$(\{\hat{\boldsymbol{t}}_i\}_{i=1}^n, \{\hat{\alpha}_{ij}\}_{ij\in E}) = \underset{\substack{\{\boldsymbol{t}_i\}_{i=1}^n \subset \mathbb{R}^3 \\ \{\alpha_{ij}\}_{ij\in E} \subset \mathbb{R}}}{\operatorname{argmin}} \sum_{ij\in E} \|\boldsymbol{t}_i - \boldsymbol{t}_j - \alpha_{ij}\boldsymbol{\gamma}_{ij}\| \text{ s.t. } \alpha_{ij} \geq 1 \text{ and } \sum_i \boldsymbol{t}_i = \boldsymbol{0}. \tag{2}$$

This formulation is very similar to that of CLS, but uses least absolute deviations instead of least squares in order to gain robustness to outliers. Numerical results in [17] demonstrate that LUD can exactly recover the original locations even when some pairwise directions are maliciously corrupted.

Following Özyesil and Singer, Hand, Lee and Voroninski [11] proposed the ShapeFit algorithm as a theoretically guaranteed solver. Given the pairwise directions $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$, the ShapeFit algorithm estimates the locations $\{\boldsymbol{t}_i^*\}_{i=1}^n$ by solving the following convex optimization problem:

$$\min_{\{\boldsymbol{t}_i\}_{i=1}^n \subset \mathbb{R}^3} \sum_{ij\in E} \|P_{\boldsymbol{\gamma}_{ij}^\perp}(\boldsymbol{t}_i - \boldsymbol{t}_j)\| \text{ s.t. } \sum_{ij\in E} \langle \boldsymbol{t}_i - \boldsymbol{t}_j, \boldsymbol{\gamma}_{ij}\rangle = 1 \text{ and } \sum_{i=1}^n \boldsymbol{t}_i = \boldsymbol{0},$$

where $P_{\boldsymbol{\gamma}_{ij}^\perp}$ denotes the orthogonal projection onto the orthogonal complement of $\boldsymbol{\gamma}_{ij}$. Empirically, for low levels of noise and corruption, ShapeFit is more accurate than LUD. However, for high levels of corruption and noise, LUD is more accurate and stable. Figure 1 demonstrates the empirical behavior of ShapeFit and LUD for synthetic data. We remark that in this case of synthetic data, stability can be measured as the magnitude of the rate of change of accuracy with respect to corruption or noise. Figures 1 and 2 of Goldstein et al. [8] demonstrate similar behavior, but emphasize exact recovery at lower corruption levels, where ShapeFit often outperforms LUD. Practical results are demonstrated in [8, 21] and seem to indicate similar behavior. Most notably, LUD is more stable, where stability for real data sets is demonstrated by consistent performance of different simulations for the same data set as well as consistent performance among different data sets. We remark that [8] presents an accelerated version of ShapeFit and [21] presents a novel heuristic for estimating the fundamental matrices, which directly relies on LUD.

The mathematical problem discussed in this paper is an example of a convex recovery problem. Other such problems include, for example, recovering sparse signals, low-dimensional signals and underlying subspaces. There seem to be two different kinds of theoretical guarantees for convex recovery problems. Guarantees of the first kind construct dual certificates [3, 4, 5]. Guarantees of the second kind show that the underlying object is the minimizer of the convex objective function, and it is sufficient to show this in a small local neighborhood [7, 14, 20, 24, 25]. The latter guarantees often require geometric methods. It is evident from page 10 of [11] that the guarantees of ShapeFit are of the second kind. Nevertheless, the graph-theoretic approach of [11] is completely innovative and enlightening. In particular, it clarifies the effect of vertex perturbation on edge deformation.

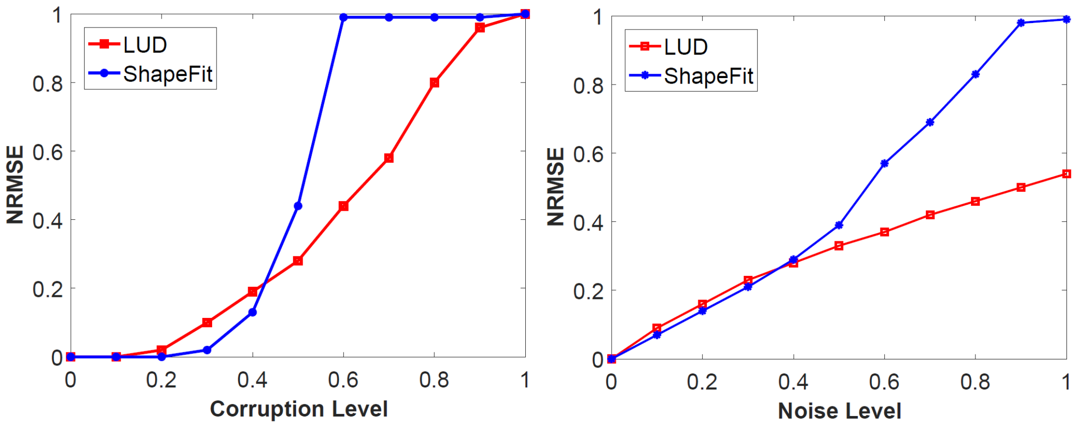

Figure 1: Empirical performance of LUD and ShapeFit under corruption and noise for synthetic data. Left: Data is generated by the HLV model with $n=50$ and $p=0.5$. The corruption level is measured by $|E_b|/|E|$ instead of ϵ_b and takes values in [0,1]. Right: The ground truth is generated by the HLV model with $n=50$, $p=0.5$ and $E_b=\emptyset$. For each $ij\in E$, $\gamma_{ij}=(\gamma^*_{ij}+\sigma\boldsymbol{v}_{ij})/\|\gamma^*_{ij}+\sigma\boldsymbol{v}_{ij}\|$, where $\boldsymbol{v}_{ij}$ is uniformly distributed on S^2 and $0\le\sigma\le1$ is the noise level. In both figures the performance is measured by the normalized root mean squared error (NRMSE): $\text{NRMSE}^2=\sum_{i=1}^n\|\kappa^*\hat{\boldsymbol{t}}_i-\boldsymbol{t}^*_i\|^2/\sum_{i=1}^n\|\boldsymbol{t}^*_i\|^2$, where $\kappa^*=\operatorname{argmin}_{\kappa\in\mathbb{R}}\sum_{ij\in E}\|\kappa\hat{\boldsymbol{t}}_i-\boldsymbol{t}^*_i\|^2$.

1.2 This Work

This paper proves exact recovery of LUD under the HLV model up to ambiguous scale and translation. More precisely, it establishes the following theorem.

Theorem 1. *There exist absolute constants n_0, C_0 and C_1 such that for $n>n_0$ and for $\{\boldsymbol{t}^*_i\}_{i=1}^n\subseteq\mathbb{R}^3$, $E\subseteq[n]\times[n]$ and $\{\gamma_{ij}\}_{ij\in E}\subseteq\mathbb{R}^3$ generated by the HLV model with parameters n, p and ϵ_b satisfying $C_0n^{-1/3}\log^{1/3}n\le p\le1$ and $\epsilon_b\le C_1p^{7/3}/\log^{9/2}n$, LUD recovers $\{\boldsymbol{t}^*_i\}_{i=1}^n$ up to translation and scale with probability $1-1/n^4$.*

To the best of our knowledge this theorem is the first exact recovery result for LUD under a corrupted model. Theorem 2 of Hand, Lee and Voroninski [11] provides exact recovery for ShapeFit under the same model. Both theorems restrict the minimal value of p and the maximal degree of corruption ϵ_b. In Theorem 2 of [11], $p=\Omega(n^{-1/5}\log^{3/5}n)$, whereas in Theorem 1, $p=\Omega(n^{-1/3}\log^{1/3}n)$. Therefore, our setting allows exact recovery for sparser graphs. For example, if $p\approx n^{-\alpha}$, our theorem allows $0<\alpha<1/3$, whereas [11] allows $0<\alpha<1/5$. More importantly, Theorem 1 tolerates more corruption. Indeed, the higher the upper bound on ϵ_b, the higher the corruption that the algorithm can tolerate. Theorem 2 of [11] requires a bound of order $O(p^5/\log^3n)$ and Theorem 1 requires a bound of order $O(p^{7/3}/\log^{9/2}n)$. Therefore in sparse settings where $p\ll1$, e.g., $p\approx n^{-\alpha}$, Theorem 1 guarantees recovery with more corruption than Theorem 2 of [11]. Nevertheless, our analysis borrows various ideas from the work of Hand, Lee and Voroninski [11]. In fact, we find it interesting to show that their innovative and nontrivial ideas are not limited to a specific objective function, but can be extended to another one.

The main ideas of the proof of Theorem 1 are discussed in Section 2, while additional technical details are left to other sections. The novelties of this work are emphasized in Section 2.5.

2 Proof of Theorem 1

The outline of the proof is as follows. Section 2.1 reformulates the LUD problem. Section 2.2 uses the new formulation to define the "good-long-dominance condition" and states that under this condition LUD exactly recovers $\{\boldsymbol{t}^*_i\}_{i=1}^n$. Section 2.3 defines the "good-shape condition" and claims that it implies the good-long-dominance condition. Section 2.4 shows that under the HLV model the good-shape condition is satisfied with high probability and thus concludes the proof of the theorem. At last, Section 2.5 discusses the novelties in our proof. Details of proofs of the main results of this section are left to Sections 3-5 and the Appendix.

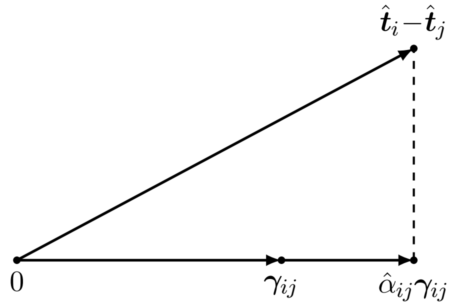

Figure 2: Demonstration of the choice of $\hat{\alpha}_{ij}$ when $\kappa > 1$ ($P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j) = \kappa\boldsymbol{\gamma}_{ij}$). By definition, $\hat{\alpha}_{ij} = \|P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j)\|$.

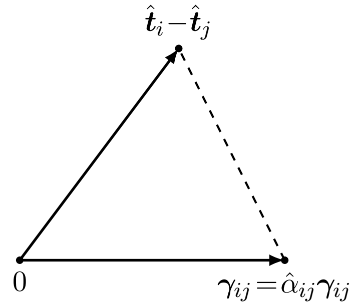

Figure 3: Demonstration of the choice of $\hat{\alpha}_{ij}$ when $\kappa \le 1$ ($P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j) = \kappa\boldsymbol{\gamma}_{ij}$). Since $\kappa \le 1$ and $\hat{\alpha}_{ij} \ge 1$, $\hat{\alpha}_{ij} = 1$.

Throughout the rest of the paper we assume that $|E_b| > 0$. This assumption is sufficient for concluding the proof. Indeed, Proposition 1 of [17] implies that when $E_b = \emptyset$, LUD recovers the true solution $\{t_i^*\}_{i=1}^n$ up to translation and scale.

2.1 Reformulation of the Problem

In order to reformulate the LUD optimization problem, we use the following observation: If $\{\hat{\boldsymbol{t}}_i\}_{i=1}^n$ and $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ are known, then for each $ij \in E$

$$\hat{\alpha}_{ij} = \begin{cases} \|P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j)\|, & \text{if } P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j) = \kappa\boldsymbol{\gamma}_{ij} \text{ for } \kappa > 1; \\ 1, & \text{if } P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j) = \kappa\boldsymbol{\gamma}_{ij} \text{ for } \kappa \le 1, \end{cases}$$

where $P_{\boldsymbol{\gamma}_{ij}}$ denotes the orthogonal projection onto $\boldsymbol{\gamma}_{ij}$. Indeed, following (2) and the demonstration in Figure 2, if $P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j) = \kappa\boldsymbol{\gamma}_{ij}$ with $\kappa > 1$, then

$$\hat{\alpha}_{ij} \equiv \underset{\substack{\{\alpha_{ij}\}_{ij\in E}: \\ \alpha_{ij}\ge 1, \forall ij\in E}}{\operatorname{argmin}} \sum_{ij\in E} \|\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j - \alpha_{ij}\boldsymbol{\gamma}_{ij}\| = \|P_{\boldsymbol{\gamma}_{ij}}(\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j)\|.$$

Otherwise, as demonstrated in Figure 3, $\hat{\alpha}_{ij} = 1$. Plugging the above values of $\{\hat{\alpha}_{ij}\}_{ij\in E}$ into (2), we obtain the equivalent LUD formulation:

$$\{\hat{\boldsymbol{t}}_i\}_{i=1}^n = \underset{\{\boldsymbol{t}_i\}_{i=1}^n \subset \mathbb{R}^3}{\operatorname{argmin}} \sum_{ij\in E} f_{ij}(\boldsymbol{t}_i, \boldsymbol{t}_j) \text{ subject to } \sum_{i=1}^n \boldsymbol{t}_i = \boldsymbol{0}, \tag{3}$$

where

$$f_{ij}(\boldsymbol{t}_i, \boldsymbol{t}_j) = \begin{cases} \|P_{\boldsymbol{\gamma}_{ij}^\perp}(\boldsymbol{t}_i - \boldsymbol{t}_j)\|, & \text{if } P_{\boldsymbol{\gamma}_{ij}}(\boldsymbol{t}_i - \boldsymbol{t}_j) = \kappa\boldsymbol{\gamma}_{ij} \text{ for } \kappa > 1; \\ \|\boldsymbol{t}_i - \boldsymbol{t}_j - \boldsymbol{\gamma}_{ij}\|, & \text{if } P_{\boldsymbol{\gamma}_{ij}}(\boldsymbol{t}_i - \boldsymbol{t}_j) = \kappa\boldsymbol{\gamma}_{ij} \text{ for } \kappa \le 1. \end{cases} \tag{4}$$

Our analysis requires formulating an oracle problem that determines the particular shift and scale found by LUD. That is, we assume we know the ground truth solution $\{\boldsymbol{t}_i^*\}_{i=1}^n$ and we ask for the scale c^* and shift $\boldsymbol{t}_s$ such that $\{c^*\boldsymbol{t}_i^* + \boldsymbol{t}_s\}_{i=1}^n$ minimizes the LUD problem. This oracle problem is formulated as follows:

$$(c^*, \boldsymbol{t}_s) = \underset{c\in\mathbb{R}, \boldsymbol{t}\in\mathbb{R}^3}{\operatorname{argmin}} \sum_{ij\in E} f_{ij}(\boldsymbol{t}_i, \boldsymbol{t}_j) \text{ subject to } \sum_{i=1}^n \boldsymbol{t}_i = \boldsymbol{0} \text{ and } \boldsymbol{t}_i = c\boldsymbol{t}_i^* + \boldsymbol{t}. \tag{5}$$

We later show in Appendix A that c^* is unique with overwhelming probability under our assumption that $E_b \ne \emptyset$. The uniqueness of $\boldsymbol{t}_s$ follows from the LUD constraint $\sum_i \boldsymbol{t}_i = \boldsymbol{0}$. We will prove Theorem 1 by showing that $\hat{\boldsymbol{t}}_i = c^*\boldsymbol{t}_i^* + \boldsymbol{t}_s$ for all $i \in [n]$.

2.2 Exact Recovery under the Good-Long-Dominance Condition

We establish the recovery of the ground truth locations $\{\boldsymbol{t}_i^*\}_{i=1}^n$ by LUD up to translation and scale under a geometric condition, which we refer to as the good-long-dominance condition. The set of good and long edges, E_{gl}, and its complement are defined by

$$E_{\text{gl}} = \{ij \in E_g \mid \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| > 1/c^*\} \text{ and } E_{\text{gl}}^c = E \setminus E_{\text{gl}}. \tag{6}$$

We say that $E_{\rm gl}$ and $G(V,E)$ satisfy the good-long-dominance condition if for any perturbation vectors $\{\boldsymbol{\epsilon}_i\}_{i=1}^n \in \mathbb{R}^3$ such that $\sum_{i=1}^n \boldsymbol{\epsilon}_i = \mathbf{0}$ and $\sum_{i=1}^n \langle \boldsymbol{\epsilon}_i, \boldsymbol{t}_i^* \rangle = 0$,

$$\sum_{ij\in E_{\rm gl}} \|P_{\gamma_{ij}^{*\perp}}(\boldsymbol{\epsilon}_i - \boldsymbol{\epsilon}_j)\| \geq \sum_{ij\in E_{\rm gl}^c} \|\boldsymbol{\epsilon}_i - \boldsymbol{\epsilon}_j\|. \tag{7}$$

In order to clarify this condition, we assume that the variables $\{\boldsymbol{t}_i\}_{i=1}^n$ are perturbed by $\{\boldsymbol{\epsilon}_i\}_{1=1}^n$ respectively from the ground truth $\{c^*\boldsymbol{t}_i^* + \boldsymbol{t}_s\}_{i=1}^n$. As explained later in (16), the change in the objective function of (3), when restricted to the sum over $E_{\rm gl}$, is the LHS of (7). Furthermore, as explained later in (17), the change in the objective function of (3), when restricted to $E_{\rm gl}^c$, is bounded above by the RHS of (7). The condition thus shows that the change in the objective function due to the good and long edges dominates the change due to all other edges.

At last, we formulate the following theorem, which is proved in Section 3.

Theorem 2. *If $V = \{\boldsymbol{t}_i^*\}_{i=1}^n$, $E \subseteq [n]\times[n]$ and $\{\gamma_{ij}\}_{ij\in E}$ are generated by the HLV model and E_{gl} and $G(V,E)$ satisfy the good-long-dominance condition, then LUD exactly recovers the ground truth solution up to translation and scale. That is, the solution of (3) has the form $\hat{\boldsymbol{t}}_i = c^*\boldsymbol{t}_i^* + \boldsymbol{t}_s$ for $i\in[n]$, where c^* and $\boldsymbol{t}_s$ solve (5).*

2.3 Exact Recovery under the Good-Shape Condition

We show that the good-long-dominance condition is satisfied when the graph E has certain properties. We first review the definitions of the following two properties suggested in [11]: a p-typical graph and c-well distributed vertices.

Definition 2.1. *A graph $G(V,E)$ is p-typical if it satisfies the following propositions:*
1. G is connected.
2. Each vertex of G has degree between $\frac{1}{2}np$ and $2np$.
3. Each pair of vertices has codegree between $\frac{1}{2}np^2$ and $2np^2$, where the codegree of a pair of vertices ij is defined as $|\{k\in[n]: ik, jk\in E\}|$.

Definition 2.2. *Let $G = G(V,E)$ be a graph with vertices $V = \{\boldsymbol{t}_i\}_{i=1}^n \subseteq \mathbb{R}^3$. For $\boldsymbol{x}$, $\boldsymbol{y}\in\mathbb{R}^3$, $c>0$ and $A\subseteq V$, we say that A is c-well-distributed with respect to $(\boldsymbol{x},\boldsymbol{y})$ if the following holds for any $\boldsymbol{h}\in\mathbb{R}^3$:*

$$\frac{1}{|A|}\sum_{t\in A} \|P_{\mathrm{Span}\{\boldsymbol{t}-\boldsymbol{x},\boldsymbol{t}-\boldsymbol{y}\}^\perp}(\boldsymbol{h})\| \geq c\|P_{(\boldsymbol{x}-\boldsymbol{y})^\perp}(\boldsymbol{h})\|.$$

We say that V is c-well-distributed along G if for all distinct $1\leq i,j\leq n$, the set $S_{ij} = \{\boldsymbol{t}_k\in V : ik, jk\in E(G)\}$ is c-well-distributed with respect to $(\boldsymbol{t}_i,\boldsymbol{t}_j)$.

Let K_n denote the complete graph with n given vertices and $E(K_n)$ denote the set of edges of K_n. When saying that V is c-well-distributed along K_n, we assume that V has n vertices and K_n is the complete graph with these vertices.

Using the above notation and definitions, we formulate a geometric condition on $E_{\rm gl}$ and $G(V,E)$ that guarantees exact recovery by LUD.

Definition 2.3 (Good-Shape Condition)**.** *Let p, β, ϵ_0, $c_1\in(0,1]$, $c_0\geq 1$ and let $V = \{\boldsymbol{t}_i^*\}_{i=1}^n$, $E\subseteq[n]\times[n]$ and $\{\gamma_{ij}\}_{ij\in E}$ be generated by the HLV model with E_{gl} and E_{gl}^c defined above. We say that E_{gl} and $G(V,E)$ satisfy the good-shape condition with the parameters p, β, ϵ_0, c_1, c_0, if the following hold:*
1. G is p-typical.
2. For any distinct $ij\in E(K_n)$, there exists at least $n-\epsilon_1 n$ indices $k\neq i,j$ such that $1-\langle\gamma_{ij}^*,\gamma_{ik}^*\rangle\geq\beta^2$ and $1-\langle\gamma_{ij}^*,\gamma_{jk}^*\rangle\geq\beta^2$.
3. For any distinct $ij\in E(K_n)$, $\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\leq c_0\mu$, where

$$\mu = \frac{1}{|E(K_n)|}\sum_{ij\in E(K_n)} \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\|. \tag{8}$$

4. The maximal degree of E_{gl}^c is bounded by $\epsilon_0 n$.
5. V is c_1-well distributed along G and along K_n.
6. For any distinct i, j ,$k\in[n]$, $\boldsymbol{t}_i^*$, $\boldsymbol{t}_j^*$ and $\boldsymbol{t}_k^*\in V$ are not collinear.

At last, we claim that under the HLV model the good-shape condition with certain restriction on its parameters implies exact recovery. The proof verifies that the good-long-dominance condition holds and then applies Theorem 2.

Theorem 3. *If $V=\{\boldsymbol{t}_i^*\}_{i=1}^n$, $E\subseteq[n]\times[n]$ and $\{\gamma_{ij}\}_{ij\in E}$ are generated by the HLV model, E_{gl} and $G(V,E)$ satisfy the good-shape condition with respect to the parameters p, β, ϵ_0, ϵ_1, c_1,, c_0 and if*

$$\epsilon_0<\min\left\{\frac{\beta c_1 p}{2^{22}c_0^3},\frac{\beta c_1^2 p}{2^{19}c_0},\frac{c_1 p^2}{16}\right\} \text{ and } \epsilon_1\leq\min\left(\frac{1}{144c_0},\frac{1}{96}\right), \tag{9}$$

then the solution $\{\hat{\boldsymbol{t}}_i\}_{i=1}^n$ of (3) has the form $\hat{\boldsymbol{t}}_i=c^*\boldsymbol{t}_i^*+\boldsymbol{t}_s$ for $i\in[n]$, where c^* and $\boldsymbol{t}_s$ solve (5).

2.4 Conclusion of Theorem 1

We verify that under the HLV model the good-shape condition holds with parameters satisfying (9) and with high probability. Combining this observation with Theorem 3 results in Theorem 1.

We assume the conditions of Theorem 1 and set the following parameters

$$\beta=\frac{p}{2^{18}\log n},\ c_1=\frac{c}{\sqrt{\log n}},\ \epsilon_1=\frac{p}{192c_0} \text{ and } c_0=64\sqrt{\log n},$$

where c is a constant used in lemma 18 of [11] and is also the same as the constant q, which is clarified in the proof of Lemma 17 of [11]. The second inequality of (9) is clearly satisfied with these parameters. We note that establishing the first inequality of (9) requires establishing the inequality $\epsilon_0\leq c'p^2/\log^3 n$, where c' linearly depends on c. That is, it requires establishing $\epsilon_0=O(p^2/\log^3 n)$.

We note that Lemma 12 of [11] and the assumption of Theorem 1 that $p=\Omega(\sqrt[3]{\log n/n})$[1] imply property 1 of Definition 2.3 with probability larger than $1-O(n^{-5})$. Lemma 18 of [11] and the assumption of Theorem 1 that $p=\Omega(\sqrt[3]{\log n/n})$ imply properties 2, 3 and 5 of Definition 2.3 with probability $1-O(n^{-5})$ and with the above choice of parameters. Furthermore, property 6 of Definition 2.3 holds almost surely since the vertices are generated by i.i.d. Gaussian distributions.

Property 4 of Definition 2.3 is about ϵ_0, whereas [11] considered instead ϵ_b. The following theorem bounds with high probability ϵ_0 by a function of ϵ_b. Combining this theorem with the assumption of Theorem 1 that $\epsilon_b=O(p^{7/3}/\log^{9/2}n)$ implies that property 4 of Definition 2.3 holds with probability $1-O(n^{-5})$ and with $\epsilon_0=O(p^2/\log^3 n)$. We recall from the discussion above that this requirement on ϵ_0 is consistent with satisfying the first inequality of (9).

Theorem 4. *If $V=\{\boldsymbol{t}_i^*\}_{i=1}^n$, $E\subseteq[n]\times[n]$ and $\{\gamma_{ij}\}_{ij\in E}$ are generated by the HLV model with $p=\Omega(\sqrt[3]{\log n/n})$ and $\epsilon_b=O(p^{7/3}/\log^{9/2}n)$, then*

$$\epsilon_0=O\left(\max\left\{p^2/\log^4 n,(p^{1/4}\log^{3/8}n)\cdot\epsilon_b^{3/4}\right\}\right) \quad w.p.\ 1-O(n^{-5}). \tag{10}$$

We have shown that all properties of the good-shape condition hold with probability $1-O(n^{-5})$, which can be written as $1-n^{-4}$ for sufficiently large n. This concludes the proof of Theorem 1.

We remark that the bound on ϵ_b in Theorem 1 is chosen so that (10) and the first inequality of (9) hold. Note that the lower bound on p in Theorem 1 is the one required by Theorem 4. With this bound and the assumption on ϵ_b, we include the non-trivial case where $n\epsilon_b\to\infty$ as $n\to\infty$.

2.5 Novelties of This Paper

This work uses ideas and techniques of [11], but considers LUD instead of ShapeFit and guarantees a stronger rate of corruption. Here we highlight the main technical differences between the two works and emphasize the novel arguments for handling these differences in the current work.

Reformulation: The objective function of ShapeFit depends only on $\{\boldsymbol{t}_i\}_{i=1}^n$, while the objective function of LUD has the additional variables $\{\alpha_{ij}\}_{ij\in E}$, which introduce more degrees of freedom. To handle this issue, we reformulated the LUD problem in (3) as a new convex optimization problem with objective function depending only on $\{\boldsymbol{t}_i\}_{i=1}^n$. We also needed to introduce the oracle problem (5) that provided the scale and shift of LUD with respect to the ground truth.

[1]Recall that for a, $b\in\mathbb{R}$, the notation $a=\Omega(b)$ is equivalent with $b=O(a)$.

Adaptation to the new formulation: The reformulated objective function for LUD is different than that of ShapeFit only in the case where $P_{\gamma_{ij}}(\hat{\boldsymbol{t}}_i-\hat{\boldsymbol{t}}_j)=\kappa\gamma_{ij}$ and $\kappa\leq 1$. We note that for $ij\in E_{\text{gl}}$, $P_{\gamma_{ij}}(\hat{\boldsymbol{t}}_i-\hat{\boldsymbol{t}}_j)=\kappa\gamma_{ij}$ for $\kappa>1$. Therefore, for $ij\in E_{\text{gl}}$ the objective functions of ShapeFit and LUD coincide. Our analysis thus tries to follow that of [11], while replacing E_g and E_b in [11] with E_{gl} and E^c_{gl} respectively. Some modifications in the analysis of [11] are needed, in particular, the two mentioned below.

More faithful constraint on perturbation: Both works introduce constraints on the perturbed solutions $\{c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i\}_{i=1}^n$. Even though c^* is not defined in [11], it can be defined as the constant satisfying $\sum_{ij\in E}\langle c^*\boldsymbol{t}_i^*-c^*\boldsymbol{t}_j^*,\gamma_{ij}\rangle=1$, where the ground truth $\{\boldsymbol{t}_i^*\}_{i=1}^n$ is denoted by $\{\boldsymbol{t}_i^0\}_{i=1}^n$ in [11]. Hand, Lee and Voroninski [11] require that

$$\sum_{ij\in E}\langle\boldsymbol{\epsilon}_i-\boldsymbol{\epsilon}_j,\gamma_{ij}\rangle=0 \tag{11}$$

so that any perturbed solution $\{\tilde{\boldsymbol{t}}_i\}_{i=1}^n$, where $\tilde{\boldsymbol{t}}_i=c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i$ for all $i\in[n]$, satisfy

$$\sum_{ij\in E}\langle\tilde{\boldsymbol{t}}_i-\tilde{\boldsymbol{t}}_j,\gamma_{ij}\rangle=1.$$

The perturbation constraint of our work appears in the formulation of the good-long-dominance condition. That is, the perturbation vectors $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$ need to satisfy $\sum_{i=1}^n\langle\boldsymbol{\epsilon}_i,\boldsymbol{t}_i^*\rangle=0$ and $\sum_{i=1}^n\boldsymbol{\epsilon}_i=0$. This requirement implies that

$$\sum_{ij\in E(K_n)}\langle\boldsymbol{\epsilon}_i-\boldsymbol{\epsilon}_j,\boldsymbol{t}_{ij}^*\rangle=0. \tag{12}$$

We note that the perturbation constraint in (12) replaces $\boldsymbol{\gamma}_{ij}$ and E in (11) with $\boldsymbol{t}_{ij}^*=\boldsymbol{t}_i^*-\boldsymbol{t}_j^*$ and $E(K_n)$ respectively. Any perturbed solution $\{\tilde{\boldsymbol{t}}_i\}_{i=1}^n$ thus needs to satisfy

$$\sum_{ij\in E(K_n)}\langle\tilde{\boldsymbol{t}}_i-\tilde{\boldsymbol{t}}_j,\boldsymbol{t}_{ij}^*\rangle=\sum_{ij\in E(K_n)}\langle c^*\boldsymbol{t}_i^*-c^*\boldsymbol{t}_j^*,\boldsymbol{t}_{ij}^*\rangle=c^*\sum_{ij\in E(K_n)}\|\boldsymbol{t}_{ij}^*\|^2. \tag{13}$$

We believe that our perturbation constraint is more faithful to the underlying structure of the problem. First of all, it uses the correct directions $\boldsymbol{t}_{ij}^*$ instead of the corrupted ones $\boldsymbol{\gamma}_{ij}$. More importantly, it uses $\boldsymbol{t}_{ij}^*$ for any pair of locations, even if they are not connected by an edge. The latter property results in improved estimates in comparison to [11]. For example, our lower bound in (37) is tighter than the one in [11, page 13], which is multiplied by $2p^2$ and suffers when $p\ll 1$.

Effective way of controlling $\boldsymbol{\epsilon}_0$:A deterministic upper bound on ϵ_b was obtained in pages 24 and 26 of [11], where ϵ_b is denoted in [11] by ϵ_0. A direct analogous bound on the maximal degree of E^c_{gl}, ϵ_0, depends on the unknown scale c^* and is thus not appealing. The proof of Theorem 4 shows that with high probability $1/c^*$ concentrates around a function of ϵ_b, n and p and consequently ϵ_0 can also be controlled with high probability by a function of ϵ_b, n and p, as stated in Theorem 4. The proof of this theorem is delicate and does not follow ideas of [11].

3 Proof of Theorem 2

We recall that c^* is uniquely defined with overwhelming probability and we assume WLOG that $\boldsymbol{t}_s=\boldsymbol{0}$, or equivalently $\sum_{i=1}^n\boldsymbol{t}_i^*=\boldsymbol{0}$. Indeed, the statement of Theorem 2, in particular, the good-long-dominance condition, is independent of any shift of the locations $\{\boldsymbol{t}_i^*\}_{i=1}^n$.

Since the objective function in (3) is convex, in order to prove that $\{c^*\boldsymbol{t}_i^*\}_{i=1}^n$ solves (3), it is sufficient to prove that for any sufficiently small perturbations $\{\boldsymbol{\epsilon}_i\}_{i=1}^n\in\mathbb{R}^3$,

$$\sum_{ij\in E}f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)\geq\sum_{ij\in E}f_{ij}(c^*\boldsymbol{t}_i^*,c^*\boldsymbol{t}_j^*). \tag{14}$$

For any $i\in[n]$, $\boldsymbol{\epsilon}_i$ can be decomposed as $\boldsymbol{\epsilon}_i=\boldsymbol{\epsilon}_i^{\parallel}+\boldsymbol{\epsilon}_i^{\perp}$, where $\boldsymbol{\epsilon}_i^{\parallel}=\kappa\boldsymbol{t}_i^*$ for some $\kappa\in\mathbb{R}$, where κ is independent of i, and $\sum_{i=1}^n\langle\boldsymbol{\epsilon}_i^{\perp},\boldsymbol{t}_i^*\rangle=0$. To clarify this, we stack $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$, $\{\boldsymbol{\epsilon}_i^{\parallel}\}_{i=1}^n$, $\{\boldsymbol{\epsilon}_i^{\perp}\}_{i=1}^n$, $\{\boldsymbol{t}_i^*\}_{i=1}^n$ as rows of matrices $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}^{\parallel}$, $\boldsymbol{\Sigma}^{\perp}$ and $\boldsymbol{T}^*$ respectively so that $\boldsymbol{\Sigma}^{\parallel}=\kappa\boldsymbol{T}^*$, $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}^{\parallel}+\boldsymbol{\Sigma}^{\perp}$ and $\langle\boldsymbol{\Sigma}^{\parallel},\boldsymbol{\Sigma}^{\perp}\rangle=\text{tr}(\boldsymbol{\Sigma}^{\parallel\boldsymbol{T}}\boldsymbol{\Sigma}^{\perp})=0$. Furthermore, the assumption $\boldsymbol{t}_s=\boldsymbol{0}$ implies that $\sum_{i=1}^n\boldsymbol{\epsilon}_i^{\perp}=\sum_{i=1}^n\boldsymbol{\epsilon}_i=\boldsymbol{0}$. Therefore, the perturbations $\{\boldsymbol{\epsilon}_i^{\perp}\}_{i=1}^n$ satisfy the required assumptions on the perturbations used in the good-long-dominance condition.

Letting $c'=c^*+\kappa$, the relation $\boldsymbol{\epsilon}_i=\kappa\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i^{\perp}$ implies that

$$c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i=c'\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i^{\perp}\quad\text{for all } i\in[n]. \tag{15}$$

Since $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$ have sufficiently small norms, we may assume that c' is sufficiently close to c^*.

Next, we obtain useful estimates in two complimentary cases.

Case A: $\boldsymbol{ij\in E_{\mathrm{gl}}}$. In this case, $\boldsymbol{\gamma}_{ij}=(\boldsymbol{t}_i^*-\boldsymbol{t}_j^*)/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|=\boldsymbol{\gamma}_{ij}^*$ and $\|P_{\boldsymbol{\gamma}_{ij}}(c^*(\boldsymbol{t}_i^*-\boldsymbol{t}_j^*))\|>1$. Combining the latter inequality, the fact that the perturbations are arbitrarily small and the proximity of c' to c^* result in $\|P_{\boldsymbol{\gamma}_{ij}}(c'(\boldsymbol{t}_i^*-\boldsymbol{t}_j^*)+\boldsymbol{\epsilon}_i^\perp-\boldsymbol{\epsilon}_j^\perp)\|>1$. Applying (15), then the latter inequality and (4), and at last the assumption $ij\in E_{\mathrm{gl}}$ concludes that

$$f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)=f_{ij}(c'\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i^\perp,c'\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j^\perp)=\|P_{\boldsymbol{\gamma}_{ij}^\perp}(c'(\boldsymbol{t}_i^*-\boldsymbol{t}_j^*)+\boldsymbol{\epsilon}_i^\perp-\boldsymbol{\epsilon}_j^\perp)\|=\|P_{\boldsymbol{\gamma}_{ij}^\perp}(\boldsymbol{\epsilon}_i^\perp-\boldsymbol{\epsilon}_j^\perp)\|.$$

This equation and the observation $f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)=0$ imply the inequality:

$$\sum_{ij\in E_{\mathrm{gl}}}\left(f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)-f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)\right)=\sum_{ij\in E_{\mathrm{gl}}}\|P_{\boldsymbol{\gamma}_{ij}^\perp}(\boldsymbol{\epsilon}_i^\perp-\boldsymbol{\epsilon}_j^\perp)\|. \tag{16}$$

Case B: $\boldsymbol{ij\in E_{\mathrm{gl}}^c}$. Following the demonstration in Figures 2 and 3, we note that $f_{ij}(\boldsymbol{t}_i,\boldsymbol{t}_j)$ is the distance between the following two convex sets: $\{\alpha\boldsymbol{\gamma}_{ij}:\alpha\geq 1\}$ and the singleton $\{\boldsymbol{t}_i-\boldsymbol{t}_j\}$. Application of (15) and then the triangle inequality for a distance between convex sets of $\mathbb{R}^3$ results in

$$|f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)-f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)|=|f_{ij}(c'\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i^\perp,c'\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j^\perp)-f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)|\leq\|\boldsymbol{\epsilon}_i^\perp-\boldsymbol{\epsilon}_j^\perp\|. \tag{17}$$

At last, we combine the above estimates with the good-long-dominance condition to verify (14). We first apply (16), then the good-long-dominance condition of (7) with $\{\boldsymbol{\epsilon}_i^\perp\}_{i=1}^n$ that satisfy its necessary requirements, and at last (17), and consequently conclude that

$$\sum_{ij\in E_{\mathrm{gl}}}\left(f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)-f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)\right)\geq\sum_{ij\in E_{\mathrm{gl}}^c}\left(f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*)-f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)\right).$$

By rearranging terms, this equation becomes

$$\sum_{ij\in E}f_{ij}(c^*\boldsymbol{t}_i^*+\boldsymbol{\epsilon}_i,c^*\boldsymbol{t}_j^*+\boldsymbol{\epsilon}_j)\geq\sum_{ij\in E}f_{ij}(c'\boldsymbol{t}_i^*,c'\boldsymbol{t}_j^*).$$

By the definition of c^* in (5) and the assumption $\boldsymbol{t}_s=\boldsymbol{0}$, this equation implies (14) and thus concludes the proof.

4 Proof of Theorem 3

We show that under the assumptions of Theorem 3, the good-shape condition implies the good-long-dominance condition and consequently Theorem 3 follows from Theorem 2. Section 4.1 reviews notation and auxiliary lemmas, which were borrowed from [11]. Section 4.2 presents the details of the proof.

While the outline of the proof in this section resembles the outline of the proof of Theorem 4 of [11], there are some nontrivial modifications. A main difference between the proofs appears in the perturbation constraints stated earlier in (11) and (12).

4.1 Preliminaries

We first review some notation that we mainly borrowed from [11]. We denote $\boldsymbol{t}_{ij}^*=:\boldsymbol{t}_i^*-\boldsymbol{t}_j^*$ and for $\{\boldsymbol{\epsilon}_i\}_{i=1}^n\subseteq\mathbb{R}^3$, we define $\eta_{ij}=\|P_{\boldsymbol{\gamma}_{ij}^{*\perp}}(\boldsymbol{\epsilon}_i-\boldsymbol{\epsilon}_j)\|$ and $\delta_{ij}\|\boldsymbol{t}_{ij}^*\|=\langle\boldsymbol{\epsilon}_i-\boldsymbol{\epsilon}_j,\boldsymbol{\gamma}_{ij}^*\rangle$. The function $\eta:E(K_n)\times E(K_n)\to\mathbb{R}$ of [11] is defined as

$$\eta(ij,kl)=\sum_{\substack{m,n\in\{i,j,k,l\}\\ m<n}}\eta_{mn}. \tag{18}$$

That is, if ij and kl do not have common elements, then $\eta(ij,kl)=\eta_{ij}+\eta_{kl}+\eta_{ik}+\eta_{il}+\eta_{jk}+\eta_{jl}$. If they have one common element, e.g., $i=k$, then $\eta(ij,kl)=\eta_{ij}+\eta_{il}+\eta_{jl}$. We modify the definition of E_g' in [11] and define $E'(K_n)$ as follows:

$$E'(K_n)=\{ij\in E(K_n):\|\boldsymbol{t}_{ij}^*\|\geq\frac{1}{2}\mu\}, \tag{19}$$

where μ was defined in equation (8). Let $B(ij)$ denote the set of all $kl\in E(K_n)$ for which there exist distinct $a,b,c\in\{i,j,k,l\}$ satisfying $\{a,b\}\neq\{i,j\}$ and $\sqrt{1-\langle\boldsymbol{\gamma}_{ab}^*,\boldsymbol{\gamma}_{bc}^*\rangle}<\beta$.

The following lemmas are from [11]. We remark that Lemma 2 was formulated in [11] for $E'=E_g$ as a matter of convenience, however, its formulation below still hold.

Lemma 1 (Lemma 3 of [11] with $\alpha=1$)**.** *Let $V_4^*=\{\boldsymbol{t}_i^*\}_{i=1}^4\subset\mathbb{R}^3$ be a set of 4 distinct vertices, K_4 be the complete graph with the set of vertices V_4^* and let $\{\boldsymbol{\epsilon}_i\}_{i=1}^4\subset\mathbb{R}^3$ be perturbation vectors. Then*

$$\eta(12,34)\geq\frac{\beta_0}{4}\|\boldsymbol{t}_{12}^*\||\delta_{12}-\delta_{34}|,\quad \textit{where}\quad \beta_0=\min_{\substack{\{i,j,k\}\in[4]\\ \{j,k\}\neq\{1,2\}}}\sqrt{1-\langle\boldsymbol{\gamma}_{ij}^*,\boldsymbol{\gamma}_{ik}^*\rangle}. \tag{20}$$

Lemma 2 (Lemmas 5 and 6 of [11])**.** *Let $G(V,E)$ be p-typical and c_1-well-distributed graph with n vertices for $0<p$, $c_1\leq 1$ and let E' be a subset of E, where the maximal degree of its complement, E'^c, is bounded by $\epsilon' n$. If $\epsilon'\leq c_1p^2/8$, then*

$$\sum_{ij\in E'}\eta_{ij}\geq\frac{c_1p^2}{8\epsilon'}\sum_{ij\in E'^c}\eta_{ij}\ \textit{and}\ \sum_{ij\in E'}\eta_{ij}\geq\frac{c_1p}{16}\sum_{ij\in E(K_n)}\eta_{ij}. \tag{21}$$

Since K_n is 1-typical, the next corollary follows from the first inequality of Lemma 2.

Corollary 1. *Let K_n be c_1-well-distributed and let E' be a subset of $E(K_n)$, where the maximal degree of its complement, E'^c, is bounded by $\epsilon' n$. If $\epsilon'\leq c_1/8$, then*

$$\sum_{ij\in E'}\eta_{ij}\geq\frac{c_1}{8\epsilon'}\sum_{ij\in E'^c}\eta_{ij}. \tag{22}$$

Lemma 4 (Lemma 14 of [11])**.** *For any $ij\in E(K_n)$,*

$$|B(ij)|\leq 6\epsilon_1 n^2, \tag{23}$$

where ϵ_1 is the constant specified in property 2 of Definition 2.3.

4.2 Details of Proof

In order to verify the good-long-dominance condition of (7), it is sufficient to prove that

$$\sum_{ij\in E_{\mathrm{gl}}}\eta_{ij}\geq 2\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|. \tag{24}$$

Indeed, since $\epsilon_0\leq c_1p^2/16$ we can apply the first inequality of Lemma 2 and obtain that $\sum_{ij\in E_{\mathrm{gl}}}\eta_{ij}> 2\sum_{ij\in E_{\mathrm{gl}}^c}\eta_{ij}$. The combination of the latter inequality with (24) and the following triangle inequality: $\|\boldsymbol{\epsilon}_i-\boldsymbol{\epsilon}_j\|\leq|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|+\eta_{ij}$ yields (7).

Following [11], we prove (24) by considering three complementary cases, which depend on the parameter $\bar{\delta}=\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|/\sum_{ij\in E_{\mathrm{gl}}^c}\|\boldsymbol{t}_{ij}^*\|$.

Case 1: $\boldsymbol{\bar{\delta}=0}$ or $\boldsymbol{E_{\mathrm{gl}}^c=\emptyset}$. Since either $E_{\mathrm{gl}}^c=\emptyset$ or $\delta_{ij}=0$ for all $ij\in E_{\mathrm{gl}}^c$, the RHS of (24) is 0.

Case 2: $\boldsymbol{\bar{\delta}\neq 0}$, $\boldsymbol{E_{\mathrm{gl}}^c\neq\emptyset}$ and $\boldsymbol{\sum_{ij\in E'(K_n)}|\delta_{ij}|<\bar{\delta}|E'(K_n)|/8}$. First, we obtain a lower bound on $|E'(K_n)|/|E(K_n)|$. The definition of $E'(K_n)$ and then the definition of μ in (8) result in

$$\sum_{ij\in E(K_n)\setminus E'(K_n)}\|\boldsymbol{t}_{ij}^*\|<\frac{1}{2}\mu|E(K_n)|=\frac{1}{2}\sum_{ij\in E(K_n)}\|\boldsymbol{t}_{ij}^*\|.$$

Consequently,

$$\sum_{ij\in E'(K_n)}\|\boldsymbol{t}_{ij}^*\|\geq\frac{1}{2}\sum_{ij\in E(K_n)}\|\boldsymbol{t}_{ij}^*\|=\frac{1}{2}\mu|E(K_n)|. \tag{25}$$

Using assumption 3 of the good-shape condition (Definition 2.3) and then (25), we obtain that

$$c_0\mu|E'(K_n)|\geq\sum_{ij\in E'(K_n)}\|\boldsymbol{t}_{ij}^*\|\geq\frac{1}{2}\mu|E(K_n)|$$

and consequently

$$|E'(K_n)|\geq\frac{1}{2c_0}|E(K_n)|. \tag{26}$$

We change the definition of L_b in [11] to $L=\{ij\in E_{\mathrm{gl}}^c:|\delta_{ij}|\geq\frac{1}{2}\bar{\delta}\}$ and derive the following inequality, which is analogous to (14) of [11]:

$$\sum_{ij\in L}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|=\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|-\sum_{ij\in E_{\mathrm{gl}}^c\setminus L}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|\geq\frac{1}{2}\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|. \tag{27}$$

We modify the definition of F_g in [11] to $F'(K_n)=\{ij\in E'(K_n):|\delta_{ij}|<\frac{1}{4}\bar{\delta}\}$ and following [11], while using the last assumption of this case (case 2), we obtain that

$$\frac{1}{8}\bar{\delta}|E'(K_n)|>\sum_{ij\in E'(K_n)}|\delta_{ij}|\geq\sum_{ij\in E'(K_n)\setminus F'(K_n)}|\delta_{ij}|\geq\frac{1}{4}\bar{\delta}|E'(K_n)\setminus F'(K_n)|.$$

We thus conclude that $|F'(K_n)|>\frac{1}{2}|E'(K_n)|$. Combining this inequality with (26) we conclude that for $n\geq 3$,

$$|F'(K_n)|>\frac{1}{4c_0}|E(K_n)|=\frac{n(n-1)}{8c_0}\geq\frac{n^2}{12c_0}. \tag{28}$$

By Lemma 4, $|B(ij)|\leq 6\epsilon_1 n^2$ for all $ij\in E(K_n)$. Combining this with (28), we obtain that for $\epsilon_1\leq\frac{1}{144c_0}$,

$$|F'(K_n)\setminus B(ij)|\geq\frac{n^2}{12c_0}-6\epsilon_1 n^2\geq\frac{n^2}{24c_0}. \tag{29}$$

The rest of the proof uses the above inequalities to obtain a lower bound on the LHS of (24) and a similar upper bound on the RHS of (24). To get the lower bound, we first note that the second inequality of Lemma 2 implies that

$$\sum_{ij\in E_{\mathrm{gl}}}\eta_{ij}\geq\frac{c_1 p}{16}\sum_{ij\in E(K_n)}\eta_{ij}. \tag{30}$$

We thus need to find a lower bound for the RHS of (30).

By applying assumption 4 of the good-shape condition (Definition 2.3) and following the combinatorial argument establishing case 1 in the proof of Theorem 4 in [11], but replacing E_b and E_g with E_{gl}^c and $E(K_n)$ respectively, we obtain that

$$\sum_{ij\in E_{\mathrm{gl}}^c}\sum_{\substack{kl\in E(K_n)\\ kl\neq ij}}\eta(ij,kl)\leq\sum_{ij\in E_{\mathrm{gl}}^c}3n^2\eta_{ij}+\sum_{ij\in E(K_n)}18\epsilon_0 n^2\eta_{ij}.$$

We recall that $\epsilon_0\leq c_1p^2/8\leq c_1/8$ and thus Corollary 1 implies that

$$\sum_{ij\in E(K_n)}\eta_{ij}\geq\sum_{ij\in E(K_n)\setminus E_{\mathrm{gl}}^c}\eta_{ij}\geq\frac{c_1}{8\epsilon_0}\sum_{ij\in E_{\mathrm{gl}}^c}\eta_{ij}.$$

The above two equations yield

$$\sum_{ij\in E_{\mathrm{gl}}^c}\sum_{\substack{kl\in E(K_n)\\ kl\neq ij}}\eta(ij,kl)\leq\frac{42\epsilon_0}{c_1}n^2\sum_{ij\in E(K_n)}\eta_{ij}. \tag{31}$$

The combination of (30) and (31) results in the following lower bound on the LHS of (24)

$$\sum_{ij\in E_{\mathrm{gl}}}\eta_{ij}\geq\frac{c_1 p}{16}\sum_{ij\in E(K_n)}\eta_{ij}\geq\frac{c_1^2 p}{3\cdot 2^8\epsilon_0 n^2}\sum_{ij\in E_{\mathrm{gl}}^c}\sum_{\substack{kl\in E(K_n)\\ kl\neq ij}}\eta(ij,kl). \tag{32}$$

In order to upper bound the RHS of (24) we first apply Lemma 1, which implies that for $ij\in L$ and $kl\in F'(K_n)$

$$\eta(ij,kl)\geq\frac{\beta}{4}|\delta_{kl}-\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|.$$

For $ij\in L$, $|\delta_{ij}|>\frac{1}{2}\bar{\delta}$ and for $kl\in F'(K_n)$, $|\delta_{kl}|<\frac{1}{4}\bar{\delta}$. Consequently, $|\delta_{kl}|<|\delta_{ij}|/2$ and

$$\eta(ij,kl)\geq\frac{\beta}{4}\big||\delta_{kl}|-|\delta_{ij}|\big|\|\boldsymbol{t}_{ij}^*\|\geq\frac{\beta}{8}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|. \tag{33}$$

Applying first the inclusions $L\subseteq E_{\mathrm{gl}}^c$ and $F'(K_n)\subseteq E(K_n)$, then (33), next (29) and at last (27), we obtain that

$$\sum_{ij\in E_{\mathrm{gl}}^c}\sum_{\substack{kl\in E(K_n)\\ kl\neq ij}}\eta(ij,kl)\geq\sum_{ij\in L}\sum_{\substack{kl\in F'(K_n)\\ kl\neq ij}}\eta(ij,kl)$$

$$\geq\sum_{ij\in L}|F'(K_n)\setminus B(ij)|\cdot\frac{\beta}{8}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|\geq\frac{\beta}{8}\cdot\frac{n^2}{24c_0}\sum_{ij\in L}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|\geq\frac{\beta}{16}\cdot\frac{n^2}{24c_0}\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|.$$

This equation implies the following upper bound for the RHS of (24):

$$2\sum_{ij\in E_{\mathrm{gl}}^c}|\delta_{ij}|\|\boldsymbol{t}_{ij}^*\|\leq\frac{3\cdot 2^8 c_0}{\beta n^2}\sum_{ij\in E_{\mathrm{gl}}^c}\sum_{\substack{kl\in E(K_n)\\ kl\neq ij}}\eta(ij,kl). \tag{34}$$

Note that (9) implies that the RHS of (34) is less than the LHS of (32). This observation concludes (24) and consequently the proof of the current case.

Case 3: $\bar{\boldsymbol{\delta}}\neq 0$, $\boldsymbol{E^c_{\mathrm{gl}}}\neq\emptyset$ and $\sum_{\boldsymbol{ij\in E'(K_n)}}|\boldsymbol{\delta_{ij}}|\geq\bar{\boldsymbol{\delta}}|\boldsymbol{E'(K_n)}|/8$. Similarly to case 2, in order to prove (24), we obtain a lower bound for the LHS of (24) and a similar upper bound for the RHS of (24).

Following [11], we define $E_+=\{ij\in E(K_n):\delta_{ij}\geq 0\}$ and $E_-=\{ij\in E(K_n):\delta_{ij}<0\}$. Using this notation, we rewrite the perturbation constraint of (12) as

$$\sum_{ij\in E_+}\delta_{ij}\|\boldsymbol{t}^*_{ij}\|^2+\sum_{ij\in E_-}\delta_{ij}\|\boldsymbol{t}^*_{ij}\|^2=0$$

and conclude that

$$\sum_{ij\in E_+}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2=\sum_{ij\in E_-}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2=\frac{1}{2}\sum_{ij\in E(K_n)}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2. \tag{35}$$

Next, we upper bound the RHS of (24) by a constant times the term $\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl)$. We first lower bound the latter term by following [11] and applying Lemma 1 as follows

$$\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl)\geq\sum_{ij\in E_-}\sum_{kl\in E_+\setminus B(ij)}\frac{\beta}{4}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|\geq\frac{\beta}{4}(|E_+|-|B(ij)|)\sum_{ij\in E_-}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|.$$

The successive application of property 3 of the good-shape condition, (35), the inclusion $E'(K_n)\subseteq E(K_n)$, the definition of $E'(K_n)$ together with the assumption $\sum_{ij\in E'(K_n)}|\delta_{ij}|\geq\frac{1}{8}\bar{\delta}|E'(K_n)|$ and (26) results in

$$\begin{aligned}\sum_{ij\in E_-}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|&\geq\frac{1}{c_0\mu}\sum_{ij\in E_-}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2=\frac{1}{2c_0\mu}\sum_{ij\in E(K_n)}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2\\&\geq\frac{1}{2c_0\mu}\sum_{ij\in E'(K_n)}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|^2\geq\frac{1}{2c_0\mu}\cdot\frac{1}{4}\mu^2\cdot\frac{1}{8}\bar{\delta}|E'(K_n)|\geq\frac{\mu\bar{\delta}n^2}{512c_0^2}.\end{aligned} \tag{36}$$

Assuming $|E_+|\geq|E(K_n)|/2$ and combining (36), the latter assumption, the fact that $|E(K_n)|=n(n-1)/2\geq n^2/4$ for $n\geq 2$, and the assumption $\epsilon_1\leq 1/96$, gives

$$\frac{\beta}{4}(|E_+|-|B(ij)|)\sum_{ij\in E_-}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|\geq\frac{\beta\mu\bar{\delta}n^2}{2048c_0^2}\Big(\frac{1}{2}|E(K_n)|-6\epsilon_1 n^2\Big)\geq\frac{\beta\mu\bar{\delta}n^4}{2^{15}c_0^2}.$$

Consequently,

$$\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl)\geq\frac{\beta\mu\bar{\delta}n^4}{2^{15}c_0^2}. \tag{37}$$

Assuming on the contrary that $|E_-|\geq|E(K_n)|/2$ and following the same arguments, while switching between E_+ and E_-, also yield (37).

We conclude with the following upper bound on the RHS of (24) by first applying the definition of $\bar{\delta}$, then condition 3 of Definition 2.3, then condition 4 of Definition 2.3, and at last (37):

$$\sum_{ij\in E^c_{\mathrm{gl}}}|\delta_{ij}|\|\boldsymbol{t}^*_{ij}\|=\bar{\delta}\sum_{ij\in E^c_{\mathrm{gl}}}\|\boldsymbol{t}^*_{ij}\|\leq\bar{\delta}c_0\mu|E^c_{\mathrm{gl}}|\leq\bar{\delta}c_0\mu\epsilon_0 n^2\leq\frac{2^{15}c_0^3\epsilon_0}{\beta n^2}\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl). \tag{38}$$

In order to obtain a lower bound on the LHS of (24), we use the following result from [11, page 25], which is obtained by counting the number of elements in the sum of η's:

$$\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl)\leq 3n^2\sum_{ij\in E(K_n)}\eta_{ij}. \tag{39}$$

We remark that although we modified the definition of E_+ and E_-, this result still holds. We conclude a lower bound on the LHS of (24) by applying the second inequality of Lemma 2 and then (39) as follows:

$$\sum_{ij\in E_{\mathrm{gl}}}\eta_{ij}\geq\frac{c_1 p}{16}\sum_{ij\in E(K_n)}\eta_{ij}\geq\frac{c_1 p}{48n^2}\sum_{ij\in E_-}\sum_{kl\in E_+}\eta(ij,kl). \tag{40}$$

The combination of (38), (40) and the assumption $\frac{\beta c_1 p}{2^{21}c_0^3\epsilon_0}>2$ verifies (24).

5 Proof of Theorem 4

Note that $E_{\rm gl}^c \subseteq E_b \cup E_s$, where $E_s = \{ij \in E : \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < 1/c^*\}$ is the set of short edges. Therefore, to conclude the theorem it is enough to estimate the maximal degree of E_s. Our estimate uses the following notation: I denotes the indicator function, the neighborhood $N(\boldsymbol{t}_i^*)$ of $\boldsymbol{t}_i^* \in V$ includes all indices $j \in [n]$ such that $ij \in E$, and for a, $b \in \mathbb{R}$, $a \lesssim b$ if and only if $b = \Omega(a)$. We will prove that for any fixed $\boldsymbol{t}_i^* \in V$

$$\sum_{j \in N(\boldsymbol{t}_i^*)} I\left(\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}\right) \lesssim \max\left\{\frac{np^2}{\log^4 n}, p^{\frac{1}{4}} \epsilon_b^{\frac{3}{4}} n \log^{\frac{3}{8}} n\right\} \text{ w.p. } 1 - O(n^{-6}). \tag{41}$$

Taking a union bound yields

$$\frac{\text{Maximal degree of } E_s}{n} \lesssim \max\left\{\frac{p^2}{\log^4 n}, p^{\frac{1}{4}} \epsilon_b^{\frac{3}{4}} \log^{\frac{3}{8}} n\right\} \text{ w.p. } 1 - O(n^{-5})$$

and this implies (10) and thus concludes the proof of the theorem.

We derive (41) by using the following function of c^*, which is defined with respect to a Gaussian random variable $\boldsymbol{x} \sim N(\mathbf{0}, \mathbf{I})$ with pdf Φ:

$$g(c^*) = \Pr\left(\left\{\|\boldsymbol{x}\| < \frac{1}{c^*}\right\}\right) = \int_{B(\mathbf{0}, \frac{1}{c^*})} \Phi(\boldsymbol{t}) d\boldsymbol{t}. \tag{42}$$

We note that for fixed $\boldsymbol{t}_i^* \in V$,

$$\Pr(\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < 1/c^*) = \int_{B(\boldsymbol{t}_i^*, \frac{1}{c^*})} \Phi(\boldsymbol{t}) d\boldsymbol{t} \leq \int_{B(0, \frac{1}{c^*})} \Phi(\boldsymbol{t}) d\boldsymbol{t} = \Pr(\|\boldsymbol{t}_j^*\| < 1/c^*) = g(c^*).$$

Furthermore, $I(ij \in E$ and $\|\boldsymbol{t}_i^* - \boldsymbol{t}_i^*\| < 1/c^*)$ is a Bernoulli random variable $B(1, \mu)$ with $\mu \leq pg(c^*)$. This observation and Chernoff bound can be used to conclude (41). It is easily done in Section 5.1 when $g(c^*) \lesssim 1/\sqrt{n}$, while only using the first term in the RHS of (41). The other case, where $g(c^*) \gtrsim 1/\sqrt{n}$, is more complicated and verified in Section 5.2 and uses the second term in the RHS of (41).

5.1 Proof for the case where $g(c^*) \lesssim 1/\sqrt{n}$.

In order to verify (41), we use the following version of Chernoff bound for Bernoulli random variables: If $X_1, X_2, \cdots, X_n \sim B(1, \mu)$ i.i.d., then

$$\Pr\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu > \delta\mu\right) < \exp(-\delta n\mu/3) \text{ for any } \delta \geq 1. \tag{43}$$

We apply this inequality to

$$X_{ij} = I(ij \in E \text{ and } \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < 1/c^*), \text{ where } i \in [n] \text{ is fixed and } j \in [n] \backslash \{i\}. \tag{44}$$

As we explained above, $X_{ij} \sim B(1, \mu)$, where $\mu \geq pg(c^*)$ and thus with probability $1 - \exp(-\Omega(\delta npg(c^*)))$

$$\sum_{j \in N(\boldsymbol{t}_i^*)} I\left(\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}\right) = \sum_{j \in [n] \backslash \{i\}} X_{ij} \lesssim (\delta + 1) npg(c^*) \approx \delta npg(c^*).$$

Taking $\delta = p/(\log^4 n g(c^*))$ results in

$$\sum_{j \in N(\boldsymbol{t}_i^*)} I\left(\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}\right) \lesssim \frac{np^2}{\log^4 n} \text{ w.p. } 1 - e^{-\Omega\left(\frac{np^2}{\log^4 n}\right)}. \tag{45}$$

Note that the assumptions $g(c^*) \lesssim n^{-1/2}$ and $p \gtrsim \sqrt[3]{\log n / n}$ guarantee that our choice of δ satisfies the constraint $\delta \geq 1$ in (43). Indeed, $\delta = p/(\log^4 n g(c^*)) = \Omega(n^{1/6}/\log^{11/3} n) > 1$ for n sufficiently large. Also, the assumption $p \gtrsim \sqrt[3]{\log n / n}$ implies that $\Omega(np^2/(\log^4 n)) \gtrsim n^{1/3}/\log^{3/10} n$. Therefore, the probability in (45) is greater than $1 - O(n^{-6})$ and thus (41) is proved in the current case.

5.2 Proof for the case where $g(c^*) \gtrsim 1/\sqrt{n}$.

We use another version of Chernoff bound for Bernoulli random variables: If $X_1, X_2, \cdots, X_n \sim B(1, \mu)$ i.i.d., then

$$\Pr\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \mu\right| > \delta\mu\right) < 2 \cdot \exp(-\delta^2 \mu n/3) \text{ for all } 0 \leq \delta \leq 1. \tag{46}$$

Applying this inequality to $\{X_{ij}\}_{j\in[n]\setminus\{i\}}$ of (44) yields that with probability $1-\exp(-\Omega(npg(c^*)))$

$$\sum_{j\in N(\boldsymbol{t}_i^*)} I\Big(\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}\Big)=\sum_{j\in[n]\setminus\{i\}} X_{ij}\lesssim npg(c^*). \tag{47}$$

Note that the probability $1-\exp(-\Omega(npg(c^*)))$ exponentially approaches 1 as $n\to\infty$. Indeed, the assumptions $g(c^*)\gtrsim 1/\sqrt{n}$ and $p\gtrsim n^{-1/3}\log^{1/3}n$ imply that $\Omega(npg(c^*))=\Omega(n^{1/6}\log^{1/3}n)$.

Our goal is to upper bound the RHS of (47) by the second term in the RHS of (41). In order to do this we use the following Lemmas, which we prove in Section 5.3.

Lemma 5. *Assuming the setting of Theorem 4, there exists an absolute constant M such that*

$$\frac{1}{c^*}\le M \quad \textit{w.p. } 1-O(n^{-6}). \tag{48}$$

Lemma 6. *Assume the setting of Theorem 4. If $g(c^*)\gtrsim 1/\sqrt{n}$, then*

$$\frac{g(c^*)}{c^*}\lesssim\frac{\epsilon_b\sqrt{\log n}}{p} \quad \textit{w.p. } 1-O(n^{-6}). \tag{49}$$

Given the setting of Theorem 4, we claim that there exists $\boldsymbol{x}_M\in\mathbb{R}^3$ with $\|\boldsymbol{x}_M\|=M$ such that

$$\Phi(\boldsymbol{x}_M)\mathrm{Vol}\Big(\frac{1}{c^*}\Big)\le g(c^*)\le\Phi(\mathbf{0})\mathrm{Vol}\Big(\frac{1}{c^*}\Big) \quad \text{w.p. } 1-O(n^{-6}). \tag{50}$$

The second inequality of (50) is deterministic and follows from the definition of g in (42). The first inequality follows from Lemma 5. Indeed, with the same probability the minimum of Φ in the closed ball $\overline{B(\mathbf{0},1/c^*)}$ is greater than the minimum of Φ in $\overline{B(\mathbf{0},M)}$ and it occurs on the boundary of this ball. Equation (50) implies that $g(c^*)\approx 1/(c^*)^3$ and applying this observation to (49) results in

$$g(c^*)\lesssim\Big(\frac{\epsilon_b\sqrt{\log n}}{p}\Big)^{\frac{3}{4}} \quad \text{w.p. } 1-O(n^{-6}). \tag{51}$$

Combining (51) with (47) yields that with probability $1-O(n^{-6})$,

$$\sum_{j\in N(\boldsymbol{t}_i^*)} I\Big(\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}\Big)\lesssim npg(c^*)\lesssim np\Big(\frac{\epsilon_b\sqrt{\log n}}{p}\Big)^{\frac{3}{4}}=p^{\frac{1}{4}}\epsilon_b^{\frac{3}{4}}n\log^{\frac{3}{8}}n.$$

This concludes Theorem 4, though it remains to prove Lemmas 5 and 6.

5.3 Proofs of Lemmas 5 and 6

We first establish the following inequality, which is necessary for the proofs of both lemmas:

$$\sum_{ij\in E:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}} \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\lesssim\epsilon_b n^2\sqrt{\log n} \quad \text{w.p. } 1-O(n^{-6}). \tag{52}$$

We prove (52) by establishing an inequality involving the left and right derivatives of $f_{ij}(c\boldsymbol{t}_i^*,c\boldsymbol{t}_j^*)$ in c. Since $f_{ij}(\boldsymbol{t}_i,\boldsymbol{t}_j)$ only depends on $\boldsymbol{t}_i-\boldsymbol{t}_j$ and since we assumed that $\boldsymbol{t}_s=\mathbf{0}$, c^* can be defined as follows:

$$c^*=\operatorname*{argmin}_{c\in\mathbb{R}}\sum_{ij\in E}F_{ij}(c), \tag{53}$$

where $F_{ij}(c)=f_{ij}(c\boldsymbol{t}_i^*,c\boldsymbol{t}_j^*)$. This expression implies that

$$\sum_{ij\in E}F'_{ij}(c^{*-})\le 0 \text{ and } \sum_{ij\in E}F'_{ij}(c^{*+})\ge 0. \tag{54}$$

We estimate $F'_{ij}(c)$ for $ij\in E$ in 5 complementary cases.

1. For $ij\in E_g$ and $c>1/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$, $F_{ij}(c)=0$ and thus $F'_{ij}(c)=0$.
2. For $ij\in E_g$ and $c<1/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$, $F_{ij}(c)=1-\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\cdot c$ and thus $F'_{ij}(c)=-\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$.
3. For $ij\in E_b$ and $c>1/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$, $F_{ij}(c)=\sin\alpha\cdot\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\cdot c$, where $0<\alpha\le\pi/2$ and thus $F'_{ij}(c)\le\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$.
4. For $ij\in E_b$ and $c<1/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$, $F_{ij}(c)=\|c\boldsymbol{t}_i^*-c\boldsymbol{t}_j^*-\boldsymbol{\gamma}_{ij}\|$ and thus by the triangle inequality

$$\begin{aligned}\big|F'_{ij}(c)\big|&=\lim_{h\to 0}\left|\frac{\|(c+h)\boldsymbol{t}_i^*-(c+h)\boldsymbol{t}_j^*-\boldsymbol{\gamma}_{ij}\|-\|c\boldsymbol{t}_i^*-c\boldsymbol{t}_j^*-\boldsymbol{\gamma}_{ij}\|}{h}\right|\\&\le\lim_{h\to 0}\left|\frac{\|h\boldsymbol{t}_i^*-h\boldsymbol{t}_j^*\|}{h}\right|=\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|.\end{aligned}$$

5. For $ij\in E$ and $c=1/\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$, the function $F_{ij}(c)$ is not differentiable. However, the left and right derivatives exist and the above equations imply that $\max\left\{\left|F'_{ij}(c^-)\right|,\left|F'_{ij}(c^+)\right|\right\}\leq\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$.

If $F_{ij}(c)$ is differentiable at c^*, then (53) and the first order optimality condition results in $\sum_{ij\in E}F'_{ij}(c^*)=0$. Combining this with the estimates of the first 4 cases above, we obtain that

$$-\sum_{ij\in E_g:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|+\sum_{ij\in E_b}F'_{ij}(c^*)=0. \tag{55}$$

If $F_{ij}(c)$ is not differentiable at c^*, then $c^*=1/\|\boldsymbol{t}_k^*-\boldsymbol{t}_l^*\|$ for some $kl\in E$. Thus, $F_{kl}(c)$ is the only non-differentiable term in $\sum_{ij\in E}F_{ij}(c)$. The combination of the 5 cases above and the second inequality of (54) yield

$$-\sum_{ij\in E_g\backslash kl:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|+\sum_{ij\in E_b\backslash kl}F'_{ij}(c^*)+F'_{kl}(c^{*+})\geq 0. \tag{56}$$

The above estimates for the 5 cases also imply that $|F'_{kl}(c^+)|\leq\|\boldsymbol{t}_k^*-\boldsymbol{t}_l^*\|$ and $|F'_{ij}(c)|\leq\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|$ for $ij\in E_b\backslash kl$. Combining this observation with (56) results in the estimate

$$\sum_{ij\in E_g\backslash kl:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\leq\sum_{ij\in E_b\backslash kl}F'_{ij}(c^*)+F'_{kl}(c^{*+})\leq\sum_{ij\in E_b\cup kl}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|. \tag{57}$$

Since (55) is stronger than (57), we use the weaker result (57) to obtain the following inequality:

$$\begin{aligned}\sum_{ij\in E:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|&=\sum_{ij\in E_g\backslash kl:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|+\sum_{ij\in E_b\cup kl:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\\&\leq\sum_{ij\in E_b\cup kl}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|+\sum_{ij\in E_b\cup kl}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\leq 2\sum_{ij\in E_b\cup kl}(\|\boldsymbol{t}_i^*\|+\|\boldsymbol{t}_j^*\|)\lesssim\epsilon_b n^2\cdot\max_{i\in[n]}\|\boldsymbol{t}_i^*\|.\end{aligned} \tag{58}$$

By the second property of Lemma 18 of [11] and its proof,

$$\max_{i\in[n]}\|\boldsymbol{t}_i^*\|\lesssim\sqrt{\log n}\ \text{ w.p. } 1-O(n^{-6}). \tag{59}$$

This observation and (58) results in (52).

Using (52), we prove Lemma 5 and 6 in Sections 5.3.1 and 5.3.2 respectively.

5.3.1 Proof of Lemma 5

We assume on the contrary that $1/c^*>M$ and use this assumption to derive an inequality for the random variables

$$Y_{ij}=I(ij\in E \text{ and } \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<1/c^*)\cdot\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\ \text{ for fixed } i\in[n] \text{ and } j\in[n]\backslash\{i\}. \tag{60}$$

This inequality uses the constant $\mu_0=\inf_{\|\boldsymbol{x}\|<5}\mathbb{E}[I(\|\boldsymbol{x}-\boldsymbol{y}\|<1/c^*)\cdot\|\boldsymbol{x}-\boldsymbol{y}\|$, where $\boldsymbol{y}\sim N(\mathbf{0},\mathbf{I})$, and is formulated as follows:

$$\frac{1}{2}n^2p\mu_0\lesssim\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}}\sum_{j\in[n]\backslash\{i\}}Y_{ij}\lesssim\frac{n^2p^{7/3}}{\log^4 n}\ \text{ w.p. } 1-O(n^{-6}). \tag{61}$$

We note that (61) results in contradiction w.p. $1-O(n^{-6})$ and thus concludes the proof. Indeed, it implies that with this probability $\mu_0\lesssim p^{4/3}/\log^4 n\to 0$ as $n\to\infty$. Since μ_0 is monotonically increasing as a function of $1/c^*$, $1/c^*\to 0$ as $n\to\infty$, which contradicts our assumption.

The rest of this section proves (61) under the assumption that $1/c^*>M$. We first establish the second inequality of (61) as follows. We first note that

$$\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}}\sum_{j\in[n]\backslash\{i\}}Y_{ij}\leq\sum_{i\in[n]}\sum_{j\in[n]\backslash\{i\}}Y_{ij}=2\sum_{ij\in E(K_n)}Y_{ij}. \tag{62}$$

Subsequently applying (62), the definition of Y_{ij}, (52) and the assumption of Theorem 4 that $\epsilon_b=O(p^{7/3}/\log^{9/2}n)$, we obtain that

$$\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}}\sum_{j\in[n]\backslash\{i\}}Y_{ij}\leq 2\sum_{ij\in E:\ \|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|<\frac{1}{c^*}}\|\boldsymbol{t}_i^*-\boldsymbol{t}_j^*\|\lesssim\epsilon_b n^2\sqrt{\log n}\lesssim\frac{n^2p^{7/3}}{\log^4 n}. \tag{63}$$

To prove the first inequality of (61), we introduce the following notation: Fix $i\in[n]$ and assume that $\|\boldsymbol{t}_i^*\|<5$. Assume further that $\boldsymbol{t}_1^*,\ldots,\boldsymbol{t}_n^*$ are i.i.d. $N(\mathbf{0},\mathbf{I})$ and let Y_{ij} be defined in (60),

$\bar{Y}_i = \sum_{j\in[n]\setminus\{i\}} Y_{ij}/(n-1)$ and $\mu_i = \mathbb{E}(\bar{Y}_i) = p\cdot\mathbb{E}[I(\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < 1/c^*)\cdot\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\|)]$. Applying Hoeffding's inequality to $\{Y_{ij}\}_{j\in[n]\setminus\{i\}}$

$$\bar{Y}_i \geq \frac{1}{2}\mu_i \quad \text{w.p. } 1-2\cdot\exp\left(-\frac{\mu_i^2 n}{2\cdot\max\{Y_{ij}^2\}}\right). \tag{64}$$

Since μ_i is monotonically increasing with respect to $1/c^*$, the assumption that $1/c^* > M$ implies that $\mu_i = \Omega(1)$. Combining this observation with (64) and the definitions of μ_i and μ_0 results in

$$\bar{Y}_i \geq \frac{1}{2}\mu_i \geq \frac{1}{2}\mu_0 p \quad \text{w.p. } 1-2\cdot\exp\left(-\frac{n}{2\cdot\max\{Y_{ij}^2\}}\right). \tag{65}$$

Using the definition of $\bar{Y}_i$, we rewrite (65) as follows: For fixed $i\in[n]$ with $\|\boldsymbol{t}_i^*\| < 5$

$$\sum_{j\in[n]\setminus\{i\}} Y_{ij} \gtrsim np\mu_0 \quad \text{w.p. } 1-2\cdot\exp\left(-\frac{n}{2\cdot\max\{Y_{ij}^2\}}\right). \tag{66}$$

A union bound of (66) over all $i\in[n]$ with $\|\boldsymbol{t}_i^*\| < 5$ has the following form:

$$\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} \sum_{j\in[n]\setminus\{i\}} Y_{ij} \gtrsim \sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} np\mu_0 = \sum_{i\in[n]} I(\|\boldsymbol{t}_i^*\|<5)\cdot np\mu_0$$

$$\text{w.p. } 1-2\sum_{i\in[n]} I(\|\boldsymbol{t}_i^*\|<5)\cdot\exp\left(-\frac{n}{2\cdot\max\{Y_{ij}^2\}}\right). \tag{67}$$

In order to conclude the first inequality of (61) from (66), we first note that the application of (46) yields

$$\sum_{i=1}^{n} I(\|\boldsymbol{t}_i^*\|<5) > n/2 \quad \text{w.p. } 1-2\cdot\exp(-\Omega(n)), \tag{68}$$

and the application of basic inequalities and (59) implies that for all $ij\in E$

$$0 \leq Y_{ij} \leq \max_{ij\in E}\{\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\|\} \leq 2\cdot\max_{i\in[n]}\{\|\boldsymbol{t}_i^*\|\} \lesssim \sqrt{\log n} \quad \text{w.p. } 1-O(n^{-6}). \tag{69}$$

Lemma 5 is concluded by applying (68) and (69) in order to simplify (67) as follows:

$$\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} \sum_{j\in[n]\setminus\{i\}} Y_{ij} \gtrsim \frac{1}{2}n^2 p\mu_0 \quad \text{w.p. } 1-n\cdot\exp\left(-\Omega\left(\frac{n}{\log n}\right)\right) - 2\cdot\exp(-\Omega(n)) - O(n^{-6}).$$

5.3.2 Proof of Lemma 6

To prove the lemma, it suffices to verify w.p. $1-O(n^{-6})$ that

$$\sum_{ij\in E:\ \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}} \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| \gtrsim \frac{1}{2c^*}\cdot npg(c^*)\cdot\frac{n}{2}. \tag{70}$$

Indeed, Lemma 6 clearly follows by combining (52) and (70).

We first bound from below the LHS of (70) by a sum of random variables, which we define as follows. We arbitrarily fix $i\in[n]$ such that $\|\boldsymbol{t}_i^*\| < 5$ and for all $j\in[n]\setminus\{i\}$ let $Z_{ij} = I(ij\in E \text{ and } 1/(2c^*) < \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < 1/c^*)$. We note that

$$\sum_{ij\in E:\ \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}} \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| \geq \sum_{\substack{ij\in E:\ \|\boldsymbol{t}_i^*\|<5\\ \frac{1}{2c^*} < \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}}} \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| = \frac{1}{2}\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} \sum_{\substack{j\in N(\boldsymbol{t}_i^*):\\ \frac{1}{2c^*} < \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| < \frac{1}{c^*}}} \|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\|$$

$$= \frac{1}{2}\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} \sum_{j\in[n]\setminus\{i\}} Z_{ij}\|\boldsymbol{t}_i^* - \boldsymbol{t}_j^*\| \geq \frac{1}{2}\cdot\frac{1}{2c^*}\sum_{\substack{i\in[n]:\\ \|\boldsymbol{t}_i^*\|<5}} \sum_{j\in[n]\setminus\{i\}} Z_{ij}. \tag{71}$$

It remains to bound the RHS of (71) by the RHS of (70) with high probability and conclude the proof. For this purpose, we introduce the following auxiliary function, which uses the random variable $\boldsymbol{y}\sim N(\mathbf{0},\mathbf{I})$,

$$h(c^*) = \inf_{\|\boldsymbol{x}\|<5} \Pr\left(\left\{\frac{1}{2c^*} < \|\boldsymbol{x}-\boldsymbol{y}\| < \frac{1}{c^*}\right\}\right) = \inf_{\|\boldsymbol{x}\|<5} \int_{B(\boldsymbol{x},\frac{1}{c^*})\setminus B(\boldsymbol{x},\frac{1}{2c^*})} \Phi(\boldsymbol{t})d\boldsymbol{t}. \tag{72}$$

In a somewhat similar way to establishing (50), we note that there exists $\boldsymbol{x}_0 \in \mathbb{R}^3$ with $\|\boldsymbol{x}_0\| = 5$, such that

$$C_1 \mathrm{Vol}\Big(\frac{1}{2c^*}\Big) \le h(c^*) \le C_2 \mathrm{Vol}\Big(\frac{1}{c^*}\Big) \quad \text{w.p. } 1 - O(n^{-6}), \tag{73}$$

where $C_1 = \inf_{\|\boldsymbol{x}-\boldsymbol{x}_0\|<M} \Phi(\boldsymbol{x})$, $C_2 = \sup_{\|\boldsymbol{x}-\boldsymbol{x}_0\|<M} \Phi(\boldsymbol{x})$. Thus, equation (50) and (73) imply that

$$g(c^*) \approx h(c^*) \approx \frac{1}{c^{*3}} \quad \text{w.p. } 1 - O(n^{-6}). \tag{74}$$

We further note that $Z_{ij} \sim B(1,\mu_i)$, where $\mu_i \ge p h(c^*)$. Combining this observation with (46) yields that

$$\sum_{j \in [n]\setminus\{i\}} Z_{ij} \gtrsim nph(c^*) \quad \text{w.p. } 1 - 2 \cdot \exp(-\Omega(nph(c^*))). \tag{75}$$

We conclude the proof of (70) as follows. By first applying a union bound for (75) over all i such that $\|\boldsymbol{t}_i^*\| < 5$, and then applying both (74) and (68), we obtain that

$$\sum_{\substack{i \in [n]: \\ \|\boldsymbol{t}_i^*\|<5}} \sum_{j \in [n]\setminus\{i\}} Z_{ij} \gtrsim nph(c^*) \cdot |\{i : \|\boldsymbol{t}_i^*\| < 5\}| \gtrsim npg(c^*) \cdot \frac{n}{2} \tag{76}$$

with probability $P_1 = 1 - n \cdot \exp(-\Omega(npg(c^*))) - 2 \cdot \exp(-\Omega(n)) - O(n^{-6})$. The assumptions $p = \Omega(\sqrt[3]{\log n/n})$ and $g(c^*) \gtrsim 1/\sqrt{n}$ imply that $\Omega(npg(c^*)) \gtrsim \Omega(n^{1/6} \log^{1/3} n)$. Thus, $P_1 = 1 - O(n^{-6})$. Equation (70) and thus the lemma clearly follows from combining (71) and (76).

A Appendix: On the Uniqueness of LUD

In this section we show that under the HLV model with $|E_b| > 0$, the solution of LUD is unique with overwhelming probability. To rigorously formulate the problem, we introduce the notion of self-consistency of a set of vectors in S^2 and show that uniqueness of LUD is equivalent with non-self-consistency of the pairwise directions. We then easily note that the pairwise directions in our setting are non-self-consistent with overwhelming probability.

The definition of self-consistency and an equivalent property, whose proof is straightforward, are formulated as follows.

Definition A.1 (Self-consistency)**.** *Let E be a set of pairs of indices and let n be the minimal integer such that $E \subseteq [n] \times [n]$. A set of vectors $\{\gamma_{ij}\}_{ij \in E} \in S^2$ is self-consistent if there exist $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n \subset \mathbb{R}^3$ that are not all identical such that $(\boldsymbol{t}_i - \boldsymbol{t}_j) = \|\boldsymbol{t}_i - \boldsymbol{t}_j\| \gamma_{ij}$ for each $ij \in E$. Otherwise $\{\gamma_{ij}\}_{ij \in E}$ is non-self-consistent.*

Proposition A.2. *Assume a set E of pairs of indices and let n be the minimal integer such that $E \subseteq [n] \times [n]$. A set of pairwise directions $\{\gamma_{ij}\}_{ij \in E} \in S^2$ is self-consistent if and only if there exist $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n \subset \mathbb{R}^3$ that are not all identical such that $\sum_{ij \in E} f_{ij}(\boldsymbol{t}_i, \boldsymbol{t}_j) = 0$.*

The following theorem reveals the equivalence of non-self-consistency and the uniqueness of LUD.

Theorem A.3. *If $G(V,E)$ is a parallel rigid graph, then the solution of LUD with respect to $G(V,E)$ is unique if and only if the pairwise directions $\{\gamma_{ij}\}_{ij \in E}$ are non-self-consistent.*

Proof of Theorem A.3. Assume that $\{\boldsymbol{\gamma}_{ij}\}_{ij \in E}$ is self-consistent, so there exists a set of vertices $V = \{\boldsymbol{t}_i\}_{i=1}^n$ such that $\boldsymbol{\gamma}_{ij} = (\boldsymbol{t}_i - \boldsymbol{t}_j)/\|\boldsymbol{t}_i - \boldsymbol{t}_j\|$ for each $ij \in E$. Note that for $\hat{\boldsymbol{t}}_i = \boldsymbol{t}_i$ and $\hat{\alpha_{ij}} = \|\boldsymbol{t}_i - \boldsymbol{t}_j\| \ge 1$, $\hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j - \hat{\alpha_{ij}} \gamma_{ij} = 0$. Therefore, $\{c\boldsymbol{t}_i\}_{i=1}^n$ with any $c \ge 1/\min_E \|\boldsymbol{t}_i - \boldsymbol{t}_j\|$ solves (3) and thus the solution of (3) is not unique.

Assume on the contrary that $\{\boldsymbol{\gamma}_{ij}\}_{ij \in E}$ is non-self-consistent. We will show that any two solutions $(\{\hat{\boldsymbol{t}}_i\}_{i=1}^n, \{\hat{\alpha}_{ij}\}_{ij \in E})$ and $(\{\boldsymbol{t}'_i\}_{i=1}^n, \{\alpha'_{ij}\}_{ij \in E})$ of (2) are the same. For $0 \le \lambda \le 1$, define $\boldsymbol{t}_i^\lambda = (1-\lambda)\hat{\boldsymbol{t}}_i + \lambda \boldsymbol{t}'_i$ and $\alpha_{ij}^\lambda = (1-\lambda)\hat{\alpha}_{ij} + \lambda \alpha'_{ij}$. We note that since (2) is a convex optimization problem, for any $0 \le \lambda \le 1$, $(\{\boldsymbol{t}_i^\lambda\}_{i=1}^n, \{\alpha_{ij}^\lambda\}_{ij \in E})$ is also a solution of (2). Therefore, the objective function evaluated at the solution $(\{\boldsymbol{t}_i^\lambda\}_{i=1}^n, \{\alpha_{ij}^\lambda\}_{ij \in E})$, namely $F(\lambda) = \sum_{ij \in E} \|\boldsymbol{t}_i^\lambda - \boldsymbol{t}_j^\lambda - \alpha_{ij}^\lambda \gamma_{ij}\|$, is constant on $[0,1]$. We denote $\hat{\boldsymbol{e}}_{ij} = \hat{\boldsymbol{t}}_i - \hat{\boldsymbol{t}}_j - \hat{\alpha}_{ij} \gamma_{ij}$ and $\boldsymbol{e}'_{ij} = \boldsymbol{t}'_i - \boldsymbol{t}'_j - \alpha'_{ij} \gamma_{ij}$ and rewrite $F(\lambda)$ as

$$F(\lambda) = \sum_{ij \in E} \|\hat{\boldsymbol{e}}_{ij} + \lambda(\boldsymbol{e}'_{ij} - \hat{\boldsymbol{e}}_{ij})\| = \sum_{ij \in E} \sqrt{\|\boldsymbol{e}'_{ij} - \hat{\boldsymbol{e}}_{ij}\|^2 \lambda^2 + 2\lambda \hat{\boldsymbol{e}}_{ij}^T (\boldsymbol{e}'_{ij} - \hat{\boldsymbol{e}}_{ij}) + \|\hat{\boldsymbol{e}}_{ij}\|^2}.$$

Since F is constant, this equation implies that $\hat{\boldsymbol{e}}_{ij}=\boldsymbol{e}'_{ij}$ for all $ij\in E$. That is,

$$\hat{\boldsymbol{t}}_i-\hat{\boldsymbol{t}}_j-\hat{\alpha}_{ij}\boldsymbol{\gamma}_{ij}=\boldsymbol{t}'_i-\boldsymbol{t}'_j-\alpha'_{ij}\boldsymbol{\gamma}_{ij} \quad \text{for } ij\in E. \tag{77}$$

Let $\Delta\boldsymbol{t}_i=\hat{\boldsymbol{t}}_i-\boldsymbol{t}'_i$ for $i\in[n]$ and $\Delta\alpha_{ij}=\hat{\alpha}_{ij}-\alpha'_{ij}$ for $ij\in E$. We rewrite (77) as

$$\Delta\boldsymbol{t}_i-\Delta\boldsymbol{t}_j=\Delta\alpha_{ij}\boldsymbol{\gamma}_{ij} \quad \text{for } ij\in E. \tag{78}$$

Since $\|\boldsymbol{\gamma}_{ij}\|=1$, (78) implies that

$$\Delta\boldsymbol{t}_i-\Delta\boldsymbol{t}_j=\|\Delta\boldsymbol{t}_i-\Delta\boldsymbol{t}_j\|\boldsymbol{\gamma}_{ij} \quad \text{for } ij\in E. \tag{79}$$

The non-self-consistency of $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ implies that the elements of the solution $\{\Delta\boldsymbol{t}_i\}_{i=1}^n$ of (79) are all identical. Consequently, for all $i\in[n]$, $\hat{\boldsymbol{t}}_i-\boldsymbol{t}'_i$ is a constant vector in $\mathbb{R}^3$. The constraint $\sum_i\boldsymbol{t}_i=\boldsymbol{0}$ of (3) implies that the constant vector is zero and thus the solution is unique. □

We remark that under the HLV model with $|E_b|\neq 0$, the non-self-consistency is a necessary condition for exact recovery. Indeed, assume that $V=\{\boldsymbol{t}_i\}_{i=1}^n$, $G(V,E)$ and $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ were generated by the HLV model with $|E_b|\neq 0$ and that $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is self-consistent. Since $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is self-consistent, it is a set of true pairwise directions of a graph $G'(V',E)$. Furthermore, since $|E_b|\neq 0$, $V'\neq V$ and V cannot be obtained from V' by scaling and shifting. That is, LUD outputs V' and cannot recover V.

Proposition A.5 below guarantees with high probability the non-self-consistency of $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$, while assuming the setting of Theorem 1. One can note that its probability is significantly larger than the one of Theorem 1. This proposition thus implies with overwhelming probability the uniqueness of LUD and consequently the well-posedness of exact recovery by LUD. The proof of this result depends on Lemma A.4 below, which demonstrates a necessary condition for self-consistency. Before proceeding with the proof, we introduce the following notation.

Assume that $V=\{\boldsymbol{t}_i\}_{i=1}^n$, $G(V,E)$ and $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ were generated by the HLV model with $|E_b|\neq 0$, where $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is self-consistent. As clarified above, $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is the set of true pairwise directions of a graph $G'(V',E)$, where $V'=\{\boldsymbol{t}'_i\}_{i=1}^n\neq V$ and V cannot be obtained from V' by scaling and shifting. One may view V' as perturbed vertices of V, even though the actual perturbation is of $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$. For $S\subset[n]$, denote $V(S)=\{\boldsymbol{t}_i\}_{i\in S}$, $V'(S)=\{\boldsymbol{t}'_i\}_{i\in S}$ and $E(S)=\{ij\in E: i,j\in S\}$. For $\tilde{E}\subseteq E$, let $\deg(i,\tilde{E})$ denote the degree of node i in the subgraph $G(V,\tilde{E})$. We say that i, $j\in[n]$ are undeformed and denote it by $i\sim j$, if $i\neq j$ and $\exists\ \kappa>0$ such that $\boldsymbol{t}_i-\boldsymbol{t}_j=\kappa(\boldsymbol{t}'_i-\boldsymbol{t}'_j)$. Otherwise, we say that i and j are deformed and denote $i\nsim j$. Note that by definition $i\nsim i$. For each $i\in[n]$, we define the undeformed set $S_i=\{j\in[n]: j\sim i\}$.

Lemma A.4. *Assume that $V=\{\boldsymbol{t}_i\}_{i=1}^n$, $G(V,E)$ and $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ were generated by the HLV model. If $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is self-consistent and $|E_b|\neq 0$, then there exists $j\in[n]$ such that $|S_j|<n/2$.*

Proof. Assume on the contrary that for all $j\in[n]$, $|S_j|\geq n/2$. Since $|E_b|\neq 0$ and $G(V,E)$ is parallel rigid (as assumed by the HLV model), there exists $k,l\in[n]$ such that $k\nsim l$, which implies that $\{k,l\}\cap(S_k\cup S_l)=\emptyset$ and $|S_k\cup S_l|\leq n-2$. Consequently, $|S_k\cap S_l|=|S_k|+|S_l|-|S_k\cup S_l|\geq n/2+n/2-(n-2)=2$. Denote by a and b two of the elements of $S_k\cap S_l$ and note that by definition of the undeformed sets S_k and S_l, $a\sim k$, $b\sim k$, $a\sim l$ and $b\sim l$. Due to the HLV model, the probability that $\{ak,bk,al,bl\}$ lies on a plane in $\mathbb{R}^3$ is zero and thus the graph $G(V(\{a,b,k,l\}),\{ak,bk,al,bl\})$ is parallel rigid in $\mathbb{R}^3$ [18, Figure 4(d)]. Therefore, $V(\{a,b,k,l\})=V'(\{a,b,k,l\})$ up to scale and shift and $k\sim l$, which results in contradiction. □

Proposition A.5. *In the setting of Theorem 1, if $|E_b|\neq 0$, then $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is non-self-consistent with probability $1-\exp(-\Omega(n^{2/3}\log^{1/3}n))$.*

Proof. Assume on the contrary that $\{\boldsymbol{\gamma}_{ij}\}_{ij\in E}$ is self-consistent. By Lemma A.4, there exists $j\in[n]$ such that $|S_j|<n/2$. Note that $\deg(j,E_b)=\deg(j,E(S_j^c))$. Therefore, $n\epsilon_b=\max_{i\in[n]}\deg(i,E_b)\geq\deg(j,E(S_j^c))$. For each $i\in S_j^c\backslash\{j\}$, $I(ij\in E(S_j^c))$ is a Bernoulli random variable $B(1,p)$. Thus, by applying (46) with $\delta=1/2$, $\mu=p$ and the number of terms $|S_j^c|-1=n-|S_j|-1>n/2-1$, we obtain that

$$\deg(j,E(S_j^c))=\sum_{i\in S_j^c\backslash\{j\}} I\big(ij\in E(S_j^c)\big)>\frac{1}{2}\cdot(\frac{n}{2}-1)p \quad \text{w.p. } 1-2e^{-\frac{1}{12}(\frac{n}{2}-1)p}. \tag{80}$$

Combining the assumption $p=\Omega(n^{-1/3}\log^{1/3}n)$ with (80) implies that $n\epsilon_b\geq\deg(j,E(S_j^c))=\Omega(np)$ with probability $1-2\cdot\exp(-\Omega(n^{2/3}\log^{1/3}n)$. This contradicts the assumption of Theorem 1 that $n\epsilon_b=O(np^{7/3}/\log^{9/2}n)$. □

References

[1] M. Arie-Nachimson, S. Z. Kovalsky, I. Kemelmacher-Shlizerman, A. Singer, and R. Basri. Global motion estimation from point matches. In *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, Zurich, Switzerland, October 13-15, 2012*, pages 81–88, 2012.

[2] M. Brand, M. E. Antone, and S. J. Teller. Spectral solution of large-scale extrinsic camera calibration as a graph embedding problem. In *Computer Vision - ECCV 2004, 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part II*, pages 262–273, 2004.

[3] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *J. ACM*, 58(3):11:1–11:37, 2011.

[4] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Information Theory*, 51(12):4203–4215, 2005.

[5] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011.

[6] A. Chatterjee and V. M. Govindu. Efficient and robust large-scale rotation averaging. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 521–528, 2013.

[7] M. Coudron and G. Lerman. On the sample complexity of robust PCA. In *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States.*, pages 3230–3238, 2012.

[8] T. Goldstein, P. Hand, C. Lee, V. Voroninski, and S. Soatto. Shapefit and shapekick for robust, scalable structure from motion. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VII*, pages 289–304, 2016.

[9] V. M. Govindu. Combining two-view constraints for motion estimation. In *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), 8-14 December 2001, Kauai, HI, USA*, pages 218–225, 2001.

[10] V. M. Govindu. Lie-algebraic averaging for globally consistent motion estimation. In *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), 27 June - 2 July 2004, Washington, DC, USA*, pages 684–691, 2004.

[11] P. Hand, C. Lee, and V. Voroninski. Shapefit: Exact location recovery from corrupted pairwise directions. *CoRR*, abs/1506.01437, 2015.

[12] A. Harltey and A. Zisserman. *Multiple view geometry in computer vision (2. ed.)*. Cambridge University Press, 2006.

[13] R. I. Hartley, K. Aftab, and J. Trumpf. L1 rotation averaging using the weiszfeld algorithm. In *The 24th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, Colorado Springs, CO, USA, 20-25 June 2011*, pages 3041–3048, 2011.

[14] G. Lerman, M. B. McCoy, J. A. Tropp, and T. Zhang. Robust computation of linear models by convex relaxation. *Foundations of Computational Mathematics*, 15(2):363–410, 2015.

[15] D. Martinec and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), 18-23 June 2007, Minneapolis, Minnesota, USA*, 2007.

[16] P. Moulon, P. Monasse, and R. Marlet. Global fusion of relative motions for robust, accurate and scalable structure from motion. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1-8, 2013*, pages 3248–3255, 2013.

[17] O. Özyesil and A. Singer. Robust camera location estimation by convex programming. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 2674–2683, 2015.

[18] O. Özyesil, A. Singer, and R. Basri. Stable camera motion estimation using convex programming. *SIAM Journal on Imaging Sciences*, 8(2):1220–1262, 2015.

[19] O. Özyesil, V. Voroninski, R. Basri, and A. Singer. A survey of structure from motion. *Acta Numerica*, 26:305364, 2017.

[20] P. Ravikumar, M. J. Wainwright, G. Raskutti, and B. Yu. High-dimensional covariance estimation by minimizing ℓ_1-penalized log-determinant divergence. *Electron. J. Statist.*, 5:935–980, 2011.

[21] S. Sengupta, T. Amir, M. Galun, T. Goldstein, D. W. Jacobs, A. Singer, and R. Basri. A new rank constraint on multi-view fundamental matrices, and its application to camera location recovery. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, Hawaii, USA, June 22-25, 2017*, pages 4798–4806, 2017.

[22] R. Tron and R. Vidal. Distributed image-based 3-d localization of camera sensor networks. In *Proceedings of the 48th IEEE Conference on Decision and Control, CDC 2009, December 16-18, 2009, Shanghai, China*, pages 901–908, 2009.

[23] R. Tron and R. Vidal. Distributed 3-d localization of camera sensor networks from 2-d image measurements. *IEEE Trans. Automat. Contr.*, 59(12):3325–3340, 2014.

[24] H. Xu, C. Caramanis, and S. Sanghavi. Robust PCA via outlier pursuit. *IEEE Trans. Information Theory*, 58(5):3047–3064, 2012.

[25] T. Zhang and G. Lerman. A novel M-estimator for robust PCA. *Journal of Machine Learning Research*, 15(1):749–808, 2014.