Embedded in the Shadow of the Separator

Frank Göring, Christoph Helmberg, and Markus Wappler*

September 2005, revised March 2007

Abstract

Eigenvectors to the second smallest eigenvalue of the Laplace matrix of a graph, also known as Fiedler vectors, are the basic ingredient in spectral graph partitioning heuristics. Maximizing this second smallest eigenvalue over all nonnegative edge weightings with bounded total weight yields the *absolute algebraic connectivity* introduced by Fiedler, who proved tight connections of this value to the connectivity of the graph. Our objective is to gain a better understanding of the connections between separators and the eigenspace of this eigenvalue by studying the dual semidefinite optimization problem to the absolute algebraic connectivity. Exploiting optimality conditions we show that this problem is equivalent to finding an embedding of the n nodes of the graph in $n-$space so that their barycenter is the origin, the distance between adjacent nodes is bounded by one and the nodes are spread as much as possible (the sum of the squared norms is maximized). For connected graphs we prove that for any separator in the graph, at least one of the two separated node sets is embedded in the shadow (with the origin being the light source) of the convex hull of the separator. Furthermore, we show that there always exists an optimal embedding whose dimension is bounded by the tree width of the graph plus one.
Keywords: spectral graph theory, semidefinite programming, eigenvalue optimization, embedding, graph partitioning, tree-width
MSC 2000: 05C50; 90C22, 90C35, 05C10, 05C78

1 Introduction

Let $G = (N, E)$ be an undirected graph with node set $N = \{1, \dots, n\}$ and edge set $E \subseteq \{\{i, j\} : i, j \in N, i \neq j\}$. Edge $\{i, j\}$ will be abbreviated by ij if there is no danger of confusion. The adjacency matrix $A \in \mathbb{R}^{n \times n}$ of the graph is defined as the (symmetric) matrix having $a_{ij} = 1$ if $ij \in E$ and 0 otherwise. The *Laplace matrix* or *Laplacian* of the graph is the matrix $L = \operatorname{diag}(Ae) - A$, where e denotes the vector of all ones of appropriate dimension and $\operatorname{diag}(v)$ denotes the diagonal matrix having v on its main diagonal. For

*Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany. {frank.goering, helmberg, markus.wappler}@mathematik.tu-chemnitz.de

symmetric matrices $H \in \mathbb{R}^{n\times n}$ we order the eigenvalues by $\lambda_1(H) \le \lambda_2(H) \le \cdots \le \lambda_n(H)$. Because the Laplacian L is positive semidefinite and $Le = 0$, we have $\lambda_1(L) = 0$ with eigenvector e. Fiedler [6, 7] showed that the second smallest eigenvalue $\lambda_2(L)$ is tightly related to edge and vertex connectivity of the graph. In particular, $\lambda_2(G)$ is positive if and only if G is connected. Therefore, Fiedler called $\lambda_2(L)$ the *algebraic connectivity* of the graph. Eigenvectors to $\lambda_2(L)$, often referred to as Fiedler vectors, have been used quite successfully in heuristics for graph partitioning in parallel computing [20, 21, 14], in clustering of geometric objects [1] or hyperlinks in the world wide web [12] or even computer vision [15]. The second smallest eigenvalue allows to derive various bounds in graph partitioning or bandwidth optimization [13, 11]; further properties of the Laplacian spectrum are presented in [8, 9, 10, 2, 29] and [17, 18] give a survey on the Laplacian spectrum of graphs. See also [3, 19] for related applications of spectral graph theory in combinatorial optimization.

By the Courant-Fischer Theorem, the eigenvalue $\lambda_2(L)$ and its eigenvectors may be characterized as optimal solutions to the optimization problem

$$\lambda_2(L) = \min_{v\in\mathbb{R}^n, v^T e=0, \|v\|=1} v^T L v.$$

The usefulness of λ_2 and its eigenvectors in graph partitioning should relate to this characterization in some way. In order to get a better understanding of these connections, it seems natural to study the eigenspace of λ_2 for weighted matrices on the same support (i.e., arc weighted graphs on the same edge set) that are extremal in the sense, that for their distribution of the weight, λ_2 is maximal. The optimal λ_2 with respect to the support of the graph was introduced by Fiedler [7] under the name "absolute algebraic connectivity". We study the semidefinite dual of this optimization problem. Due to complementarity, the optimal solutions of the dual are restricted to the eigenspace of the optimal λ_2 and so all properties of dual optimal solutions directly provide information on the structure of the eigenspace associated with the absolute algebraic connectivity. It turns out that the dual may be interpreted as an embedding problem of the nodes of G in $\mathbb{R}^n$, see (4). The same optimization problem appears in [24] in connection with finding the fastest mixing Markov process on a graph; this work also mentions interest in low-dimensional solutions of this problem within the field of maximum variance unfolding in machine learning [26, 27].

We show that optimal embeddings of (4) have structural properties tightly connected to the separator structure of the graph (Th. 3). In particular, if a subset $S \subset N$ of nodes separates the graph into two disconnected node sets C_1, C_2 forming a partition of $N \setminus S$, then for one of the two sets, say C_1, all nodes are in the "shadow" of the convex hull of the nodes in S as seen from the origin, i.e., the straight line segment between any node of C_1 and the origin intersects the convex hull of S. Since any nonzero projection of the embedding to a one-dimensional subspace yields an eigenvector to the optimal λ_2 (Rem. 2), this offers good geometric insight into the usefulness of Fiedler vectors for graph partitioning.

The embedding may also be interpreted as a variant of vector labelings of graphs as introduced in [16]. On first sight, strong similarities exist with respect to the Colin de

Verdière number $\mu(G)$, see the excellent survey [25]. But the strong Arnold-property is not required in our context, so no direct connection to $\mu(G)$ should be expected. Yet, similar to maximizing the corank in the Colin de Verdière number, one may ask for an optimal embedding of minimal dimension. Besides general interest in the existence of low dimensional optimal solutions of semidefinite programs [22] such solutions are also sought in the machine learning applications [26, 27] mentioned above. Even though we are still far from answering this question to our full satisfaction, we are able to exhibit an intriguing bound based on the tree width of the graph. Indeed, we show in the proof of Th. 5, that there is always an optimal embedding whose dimension is bounded by the cardinality of a "central" node of an arbitrary tree decomposition of G. This bound is tight for some particular graph classes (see Ex. 8). None the less, the bound seems to be far too pessimistic, e.g., for planar graphs. Therefore it is conceivable that significantly better bounds can be obtained by minor related approaches.

The paper is organized as follows. In Section 2 we derive the embedding problem as the dual problem to the eigenvalue optimization problem of determining the absolute algebraic connectivity and present an overview of our main results on this embedding together with some examples. The proof of the first result (the Separator-Shadow Theorem) is given in Section 3. Section 4 is devoted to optimality preserving manipulations of optimal embeddings for reducing the dimension of embeddings. These are rotations and foldings around separators that contain the origin in their convex hull and allow, in Section 5, to design an algorithm that gives rise to the proof of the tree width bound on the minimal dimension of optimal solutions.

We use basic notions and notation from graph theory and semidefinite programming ([28]). In particular, for symmetric $H \in \mathbb{R}^{n\times n}$, $H \succeq 0$ is used to denote positive semidefiniteness; for matrices $A, B \in \mathbb{R}^{m\times n}$, $\langle A, B\rangle = \sum_{ij} A_{ij}B_{ij}$ is the canonical inner product; in the case of vectors $a, b \in \mathbb{R}^n$ we will simply use a^Tb; $\|\cdot\|$ refers to the usual Euclidean norm; e denotes the vector of all ones of appropriate size; for a set $\mathcal{S} \subset \mathbb{R}^n$, $\operatorname{conv}\mathcal{S}$ refers to the convex hull of $\mathcal{S}$ and $\operatorname{cone}\mathcal{S} = \{\lambda x : x \in \operatorname{conv}\mathcal{S}, \lambda \geq 0\}$. The projection on a closed convex set C is denoted by $p_C(\cdot)$.

2 Optimal Embeddings and Main Results

In the remainder of the paper we assume, that the graph $G = (N, E)$ is connected and $n \geq 2$. Let

$$\widehat{\mathcal{W}} = \{\hat{w} \in \mathbb{R}^E_+ : \sum_{ij\in E} \hat{w}_{ij} = \hat{w}^Te = 1\}$$

denote the set of all possible nonnegative edge weightings that sum up to 1. For a particular $\hat{w} \in \widehat{\mathcal{W}}$, let $A_{\hat{w}}$ denote the weighted adjacency matrix, i.e., $A_{ij} = \hat{w}_{ij}$ for $ij \in E$ and 0 otherwise, and $L_{\hat{w}} = \operatorname{diag}(A_{\hat{w}}e) - A_{\hat{w}}$ the corresponding weighted Laplacian. For $i, j \in V$, $i \neq j$ define $E_{ij} \in \mathbb{R}^{n\times n}$ as the matrix having the two diagonal elements $(E_{ij})_{ii} = (E_{ij})_{jj} = 1$, the two offdiagonal elements $(E_{ij})_{ij} = (E_{ij})_{ji} = -1$ and all other elements equal to zero.

Then we may rewrite the Laplacian as

$$L_{\hat{w}} = \sum_{ij \in E} \hat{w}_{ij} E_{ij}.$$

The matrix $L_{\hat{w}}$ is positive semidefinite (because E_{ij} is positive semidefinite and $\hat{w}_{ij} \geq 0$ for all $ij \in E$) and has an eigenvalue zero with eigenvector e (because $E_{ij}e = 0$). Our basic optimization problem is to determine the absolute algebraic connectivity

$$\hat{a}(G) = |E| \max_{\hat{w} \in \widehat{\mathcal{W}}} \lambda_2(L_{\hat{w}}), \tag{1}$$

where $\hat{a}(G)$ denotes the absolute algebraic connectivity introduced in [7]. The maximum is attained, because a continuous function is maximized over a compact set. Since G is assumed to be connected, the result of Fiedler [6] for $\hat{w} = \frac{1}{|E|}e$ asserts $\lambda_2(L_{\hat{w}}) = \frac{1}{|E|}\lambda_2(L) > 0$, so the optimum value is strictly positive. In order to reformulate the optimization problem as a semidefinite program it will be convenient to shift the smallest eigenvalue 0 to a sufficiently large value. Thus, (1) may be rewritten as the following semidefinite program,

$$\begin{array}{rl} \frac{\hat{a}(G)}{|E|} = \max & \lambda \\ \text{s.t.} & \sum_{ij \in E} \hat{w}_{ij} E_{ij} + \hat{\mu} ee^T - \lambda I \succeq 0, \\ & \sum_{ij \in E} \hat{w}_{ij} = 1, \\ & \hat{w} \geq 0, \ \lambda, \hat{\mu} \text{ free.} \end{array}$$

Because the optimal value is strictly greater than zero by the connectedness of G, we may rescale the problem by $1/\lambda$ and equivalently minimize the sum of the scaled weights $w_{ij} = \hat{w}_{ij}/\lambda$ instead,

$$\begin{array}{rl} \frac{|E|}{\hat{a}(G)} = \min & \sum_{ij \in E} w_{ij} \\ \text{s.t.} & \sum_{ij \in E} w_{ij} E_{ij} + \mu ee^T \succeq I, \\ & w \geq 0, \ \mu \text{ free.} \end{array} \tag{2}$$

Note that by the considerations above, $w = \frac{1}{\lambda_2(L) - \varepsilon}e$, $\mu = 1 + \varepsilon$ is a strictly feasible solution for $\lambda_2(L) > \varepsilon > 0$. Therefore the program attains its optimal solution and semidefinite duality theory together with strict feasibility asserts that the optimal value of its dual semidefinite program is also attained. The dual reads

$$\begin{array}{rll} \frac{|E|}{\hat{a}(G)} = \max & \langle I, X \rangle & \\ \text{s.t.} & \left\langle ee^T, X \right\rangle = 0, & \\ & \left\langle E_{ij}, X \right\rangle \leq 1 & \text{for } ij \in E, \\ & X \succeq 0. & \end{array} \tag{3}$$

Now consider a Gram representation of X via a matrix $V \in \mathbb{R}^{n \times n}$ with $X = V^T V$ and denote column i of V by v_i, $V = [v_1, \ldots, v_n]$. Then,

$$X_{ij} = v_i^T v_j \quad \text{and} \quad \langle E_{ij}, X \rangle = \|v_i\|^2 - 2v_i^T v_j + \|v_j\|^2 = \|v_i - v_j\|^2.$$

Since $0 = \langle ee^T, X \rangle = e^T X e = e^T V^T V e$ and $Ve = \sum v_i$, the dual semidefinite program (3) translates directly to

$$\begin{array}{rl} \frac{|E|}{\hat{a}(G)} = \max & \sum_{i \in N} \|v_i\|^2 \\ \text{s.t.} & (\sum_{i \in N} v_i)^2 = 0, \\ & \|v_i - v_j\|^2 \le 1 \qquad \text{for } ij \in E, \\ & v_i \in \mathbb{R}^n \text{ for } i \in N. \end{array} \tag{4}$$

Thus, the dual problem to (1) is equivalent to finding an embedding of the nodes of the graph in $n-$space so that their barycenter is at the origin (we will call this the *equilibrium constraint*), the distances of adjacent nodes are bounded by one (the *distance constraints*), and the sum of their squared norms is maximized.

Remark 1 *Together with the KKT conditions (we do not list feasibility constraints again)*

$$\begin{array}{rcll} v_j + \sum_{ij \in E} w_{ij}(v_i - v_j) + \mu \sum_{i \in N} v_i & = & 0 & \forall j \in N \\ w_{ij}(1 - \|v_i - v_j\|^2) & = & 0 & \forall ij \in E \\ \mu(\sum_{i \in N} v_i)^2 & = & 0 & \end{array}$$

the embedding problem suggests the following physical interpretation of optimal primal and dual solutions. Consider each node as having a point mass of unit size and imagine each edge being a mass free rope of length one that connects the points. Now the optimum solution of (4) *corresponds to an equilibrium solution of this net spread within a force field that acts with force* v *on a point of mass one at position* v*. The* w_{ij} *are the forces acting along rope* ij*. Indeed, all* $w_{ij} > 0$ *are on the same scale as the force field, because* $w_{ij} > 0$ *only if* $\|v_i - v_j\|^2 = 1$*. So the first line of the KKT conditions asserts that these forces are in equilibrium in each point (*$\mu \sum v_i = 0$ *by feasibility, so this term does not enter). If an optimal two dimensional embedding exists, such a physical situation is encountered when spreading the net on a disk rotating around its center (the centripetal force is* $m\omega^2 r$*, where* m *is the mass,* ω *the angular frequency and* r *the radius). In [24] the same problem (and interpretation) was derived starting from the problem of determining the fastest mixing Markov chain.*

We illustrate this for an example graph on 30 vertices, see Fig. 1, that was generated by picking the vertices randomly in the unit square and by connecting two points by an edge if their Euclidean distance is at most 0.3*. The edge weights corresponding to an optimal solution of problem* (2) *are given in grey shades in Fig. 2 (white is weight 0, black is maximum weight). The optimal embedding of* (4) *is displayed in Fig. 2. It was computed using SeDuMi [23] and is in fact two dimensional in this case. The origin is indicated by the small circle in the center.*

Remark 2 *The projections of optimal embeddings onto one-dimensional subspaces yield eigenvectors to the algebraic connectivity. To see this, suppose* $V = [v_1, \ldots, v_n]$ *is an*

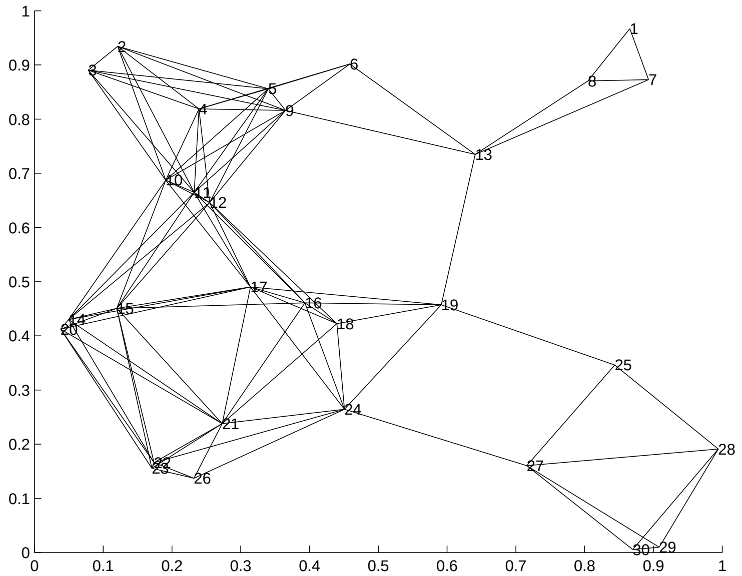

Figure 1: Original graph: The 30 vertices, picked randomly in $[0,1]^2$, are connected by an edge if the Euclidean distance is at most 0.3.

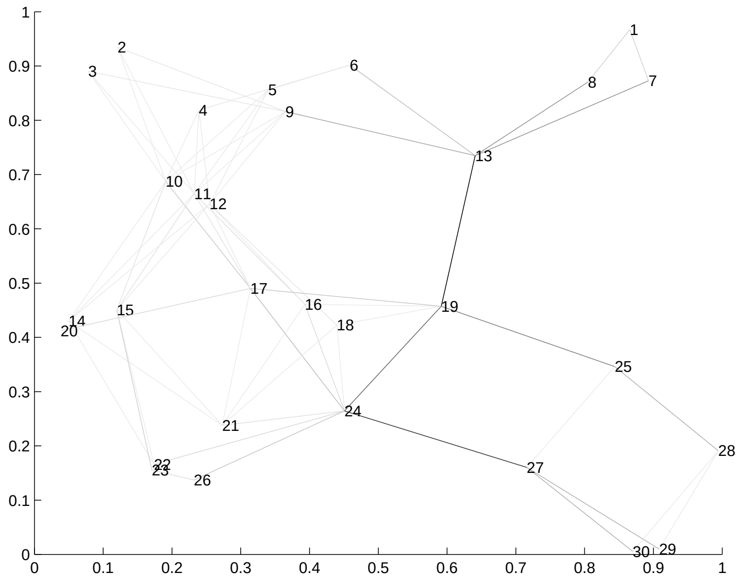

Figure 2: Graph with optimal edge weights.

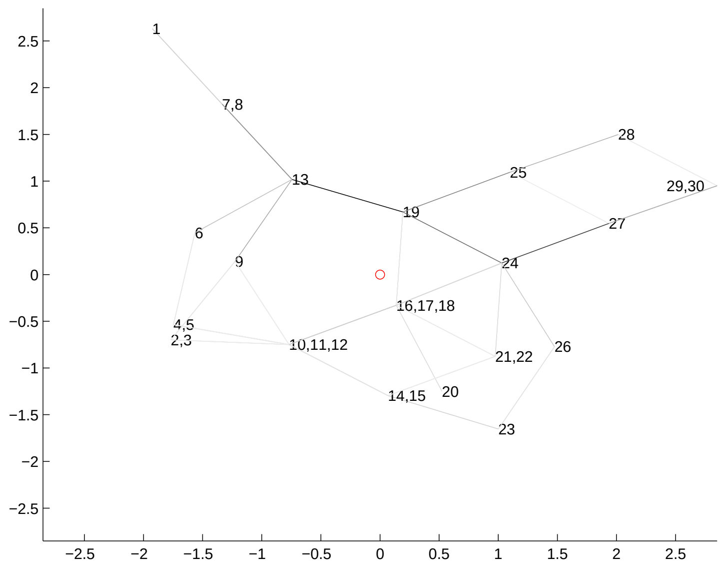

Figure 3: Optimal embedding (the central circle indicates the origin).

optimal embedding of (4) *and* $\hat{w}$ *are the corresponding optimal weights giving rise to the algebraic connectivity* $|E|\lambda_2(L_{\hat{w}})$ *in* (1). *For any* $p \in \operatorname{span}\{v_1, \ldots, v_n\}$ *with* $\|p\| = 1$ *the vector* $u = V^T p$ *is the projection of the optimal embedding onto the one-dimension subspace* $\{\alpha p : \alpha \in \mathbb{R}\}$ *and an eigenvector to* $\lambda_2(L_{\hat{w}})$. *Indeed,* $X = V^T V$ *is then an optimal solution of* (3), $w = \hat{w}/\lambda_2(L_{\hat{w}})$ *together with* $\mu = \lambda_2(L_{\hat{w}}) + 1$ *an optimal solution of* (2) *and by complementarity and* $\langle ee^T, X\rangle = 0$ *we obtain*

$$
\begin{aligned}
0 &= \left\langle X, \sum_{ij \in E} w_{ij} E_{ij} + \mu ee^T - I \right\rangle \lambda_2(L_{\hat{w}}) \\
&= \left\langle X, \sum_{ij \in E} \lambda_2(L_{\hat{w}}) w_{ij} E_{ij} - \lambda_2(L_{\hat{w}}) I \right\rangle \\
&= \left\langle V^T V, L_{\hat{w}} - \lambda_2(L_{\hat{w}}) I \right\rangle \\
&= \left\langle I, V(L_{\hat{w}} - \lambda_2(L_{\hat{w}}) I) V^T \right\rangle.
\end{aligned}
$$

So each column of V^T *and hence* $u = V^T p$ *is in the eigenspace of* $L_{\hat{w}}$ *to eigenvalue* $\lambda_2(L_{\hat{w}})$.

Our main results show that structural properties of optimal embeddings v_i, $i \in N$, of (4) are tightly linked to the separator structure of the underlying graph, where a (node-)*separator* of G is a subset $S \subset N$ of nodes, whose removal disconnects the graph into at least two connected components. Often we will not discern every single component arising this way but simply speak of two or more disconnected sets of nodes. The first result states that for each separator S all but at most one of its components must be embedded so that any ray emanating from the origin first has to hit $\operatorname{conv}\{v_s : s \in S\}$ before it can reach a node of these components, i.e., considering the origin as a light source and $\text{conv}\{v_s : s \in S\}$ as a solid object, all but one of the components must be embedded in the shadow of the separator.

Theorem 3 (Separator-Shadow) *Let $v_i \in \mathbb{R}^n$ $(i \in N)$ be an optimal solution of (4) for a connected graph $G = (N, E)$ and let S be a separator in G giving rise to disconnected sets C_1 and C_2. For at least one C_j*

$$\text{conv}\{0, v_i\} \cap \text{conv}\{v_s : s \in S\} \neq \emptyset \qquad \forall i \in C_j.$$

In words, the straight line segments $\text{conv}\{0, v_i\}$ of all nodes $i \in C_j$ intersect the convex hull of the points in S.

We encourage the reader to check the separator-shadow property on some of the separators in Figure 3, e.g., for $S = \{13\}$ or $S = \{19, 24\}$.

Considering a separator S with the property $0 \notin \text{conv}\{v_s : s \in S\}$, the Separator-Shadow Theorem ensures that all but one of the components are embedded in the subspace spanned by the separator. Thus, if all minimal separators are small in size, there seems to be hope that there also exists an optimal embedding of small dimension. Our second main result confirms this expectation. In order to state it, we first recall the definitions of tree-decomposition and tree-width as given in [5].

Definition 4 *For a graph $G = (N, E)$ a* tree-decomposition *of G is a tree $T = (\mathcal{N}, \mathcal{E})$ with $\mathcal{N} \subseteq 2^N$ and $\mathcal{E} \subseteq \binom{\mathcal{N}}{2}$ satisfying the following requirements:*
(i) $N = \bigcup_{U \in \mathcal{N}} U$.
(ii) For every $e \in E$ there is a $U \in \mathcal{N}$ with $e \subseteq U$.
(iii) If $U_1, U_2, U_3 \in \mathcal{N}$ with U_2 on the $T-$path from U_1 to U_3, then $U_1 \cap U_3 \subseteq U_2$.
The width of T is the number $\max\{|U| - 1 : U \in \mathcal{N}\}$. The tree-width *$tw(G)$ is the least width of any tree-decomposition of G.*

For example, trees have tree-width one (each edge forms one set U, so $\mathcal{N} = E$ and $\mathcal{E} = \{\{e_1, e_2\} : e_1, e_2 \in E, e_1 \cap e_2 \neq \emptyset\}$). In general, it is NP-complete to determine the tree-width, but any valid tree-decomposition provides an upper bound.

Theorem 5 *For each connected graph G there exists an optimal embedding of (4) of dimension at most $tw(G) + 1$.*

It is not difficult to devise examples where optimal embeddings of much higher dimensions exist, as well. A simple one, that will also be helpful in the remainder of the paper, is the star $K_{1,n}$.

Example 6 *For a star $K_{1,n} = (\{0, 1, \ldots, n\}, \{\{0, i\} : i = 1, \ldots, n\})$ with $n \geq 2$ one optimal solution embeds the center node 0 in the origin and all other nodes in the vertices of a regular n-1 dimensional simplex with $\|v_i\| = 1$, $i = 1, \ldots, n$ for an objective value of n (optimality follows from setting $w_{ij} = 1$ and $\mu = 1$ in (2)). For even $n \geq 2$ a one* *dimensional optimal embedding is given by assigning the center node 0 to the origin, half the outer nodes to +1 and the other half to -1. For odd* $n \geq 3$ *one possibility to find a two dimensional optimal embedding is to put node 0 into the origin, node* 1 *to* $(1,0)$*, even nodes* $i \geq 2$ *to* $(-\frac{1}{n-1}, \sqrt{1-(\frac{1}{n-1})^2})$ *and the odd nodes* $i \geq 3$ *to* $(-\frac{1}{n-1}, -\sqrt{1-(\frac{1}{n-1})^2})$.

Solving (3) by interior point methods will, in fact, always generate optimal embeddings of (4) of maximum dimension, because interior point methods generate maximally complementary solutions [4]. So the next question is whether it is difficult to find optimal embeddings satisfying the bound of Theorem 5. For general graphs there is little hope to find a tree-decomposition giving the tree-width of the graph, but for a given optimal embedding and some tree-decomposition of width t our proof of Theorem 5 allows to transform the embedding algorithmically by a sequence of optimality-preserving rotations and foldings into an optimal embedding of dimension at most $t+1$.

The bound of Theorem 5 on the minimum dimension of optimal embeddings is not tight for all graphs. Already in the example above, any star $K_{1,n}$ with even $n \geq 2$ has an optimal embedding in one dimension. For certain classes of graphs the bound of Theorem 5 on the minimum dimension of optimal embeddings is, in fact, far off (e.g. planar grid graphs have one-dimensional optimal embeddings), but in general the bound cannot be improved as is shown by the second of the following two examples.

Example 7 (Complete Graphs) *For* $K_n = (\{1,\dots,n\}, \{\{i,j\} : 1 \leq i < j \leq n\})$ *we show that the unique optimal embedding is the regular* $n-1$ *dimensional simplex with all points lying on the ball of radius* $r_n = \sqrt{\frac{n-1}{2n}}$*. The optimal* X *is given by* $X_{ii} = r_n^2 = \frac{n-1}{2n}$ *for* $1 \leq i \leq n$*,* $X_{ij} = X_{ji} = -\frac{r_n^2}{n-1} = -\frac{1}{2n}$ *for* $1 \leq i < j \leq n$*, and the optimal weights are* $w_{ij} = \frac{1}{n}$ *for* $1 \leq i < j \leq n$*. Choosing* $\mu = \frac{1}{n}$ *we compute* $L_w + \mu ee^T - I = 0$*, so* (w,μ) *is feasible for* (2) *with objective* $\frac{n-1}{2}$*. Likewise,* X *is feasible for* (3) *and* $\langle I, X\rangle = \frac{n-1}{2}$*, so optimality is shown. Furthermore, since* $w_{ij} > 0$ *for all* ij*, the constraints* $\langle E_{ij}, X\rangle = 1$ *hold for all optimal* X*, i.e., the embedding must have all points pairwise at distance one. So the regular* $n-1$ *dimensional simplex is the only optimal embedding. Note that the tree-width of* K_n *is* $n-1$*, thus the complete graphs are not tight with respect to the bound of Th. 5.*

Example 8 (Graphs with tight dimension bound) *We append to* K_n *three independent vertices that are completely linked to* K_n *resulting in a graph* $G(n) = (\{1,\dots,n+3\}, E(n) = \{\{i,j\} : 1 \leq i \leq n, i < j \leq n+3\})$*. The tree-width of* $G(n)$ *is* n *and for* $n \geq 4$ *the minimal dimension of an optimal embedding of* $G(n)$ *is* $n+1$*. In fact, we show that, for* $n \geq 4$*, the vertices of* K_n *are again arranged as a centrally symmetric* $n-1$ *dimensional simplex with all points lying on a ball of radius* $r_n = \sqrt{\frac{n-1}{2n}}$ *and the three new points are arranged centrally symmetric on a circle orthogonal to this simplex with radius* $\bar{r} = \sqrt{\frac{n+1}{2n}}$*. The optimum of* (3) *is obtained by extending the optimum of Ex. 7 with* $X_{ii} = \bar{r}^2 = \frac{n+1}{2n}$ *for* $n < i \leq n+3$*,* $X_{ij} = X_{ji} = -\frac{\bar{r}^2}{2} = -\frac{n+1}{4n}$ *for* $n < i < j \leq n+3$*, and* $X_{ij} = X_{ji} = 0$ *for*

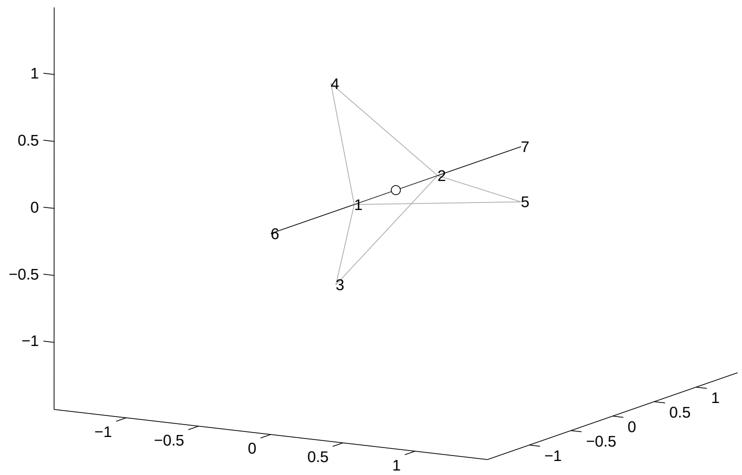

Figure 4: A graph with tree-width 2 and optimal embedding of dimension at least three, see Ex. 8 (the central circle indicates the origin).

$1 \le i \le n$, $n < j \le n+3$. *The optimal weights are* $w_{ij} = \frac{1}{n}$ *for* $1 \le i \le n$, $n < j \le n+3$ *and* $w_{ij} = \frac{1}{n} - \frac{3}{n^2}$ *for* $1 \le i < j \le n$ *(use Rem. 1 and symmetry). Setting* $\mu = \frac{1}{n}$ *the slack matrix of* (2) *computes to*

$$Z = L_w + \frac{1}{n}ee^T - I = \begin{bmatrix} \frac{3}{n^2}J_n & 0 \\ 0 & \frac{1}{n}J_3 \end{bmatrix} \succeq 0, \tag{5}$$

where J_k *denotes the square matrix of all ones of order* k*. Therefore* (w, μ) *is feasible for* (2)*, the objective value is* $3n\frac{1}{n} + \frac{n(n-1)}{2}(\frac{1}{n} - \frac{3}{n^2}) = 1 + \frac{n}{2} + \frac{3}{2n}$*. Likewise,* X *is positive semidefinite because it is a Gram matrix. Furthermore,* X *satisfies all distance constraints and has the same objective value. Hence the primal and the dual solution are optimal.*

Now take any optimal embedding v_i, $i = 1, \ldots, n$ *and set* $V = [v_1, \ldots, v_n]$*. Since* $w > 0$*, all optimal embeddings must have all edge lengths equal to one,* $\|v_i - v_j\| = 1$ *for all* $ij \in E(n)$*. By* (5) *and semidefinite complementarity it holds that* $\langle V^T V, Z \rangle = 0$*, thus* $\sum_{i=1}^{n} v_i = 0$ *and* $\sum_{i=n+1}^{n+3} v_i = 0$*. So the embedding of* K_n *must be centrally symmetric like in Ex. 7, and by the distance constraints each of the three additional vertices must be embedded orthogonal to the embedding of* K_n *with distance* $\bar{r}$ *to the origin. As the three vectors have to sum up to zero, this can only be done in two additional dimensions. This completes the proof.*

For $n = 1$ *the construction yields a star with one central and three exterior nodes and the bound is also tight. For* $n = 2$ *the embedding described above is not optimal (it would collapse to the image of the star), for* $n = 3$ *the embedding is optimal but not of minimal dimension. Without going into details, the cases* $n = 2, 3$ *can be extended to tight examples by appending to each node of* K_n *yet another node by a single edge, see Fig. 4 for an* *illustration of the resulting embedding for* $n = 2$.

3 The proof of the Separator-Shadow Theorem 3

Our proof of the Separator-Shadow Theorem will be indirect. Given a feasible embedding that does not satisfy the statement of the theorem, we improve it by folding appropriate components out of the current space in opposite directions (see Figures 6 and 7 below). This requires some preparations. First note that a feasible embedding cannot be full dimensional and so there is always space for folding.

Observation 9 *Given* $v_i \in \mathbb{R}^n$ *(*$i \in N$*) feasible for* (4)*, there is a vector* $h \in \mathbb{R}^n$*,* $\|h\| = 1$*, with* $v_i \in \mathcal{H} = \{x \in \mathbb{R}^n : h^T x = 0\}$ *for* $i \in N$*.*

Proof. The n vectors v_i satisfy $\sum_{i \in N} v_i = 0$, so they are linearly dependent and therefore $\dim(\text{span}\,\{v_1, \ldots, v_n\}) \leq n - 1$. ∎

Given the linear subspace $\mathcal{H} = \{x \in \mathbb{R}^n : h^T x = 0\}$, a normalized $b \in \mathcal{H}$, and some $\beta \in \mathbb{R}$, we next describe the operation of folding the flat halfspace $\{x \in \mathcal{H} : b^T x < \beta\}$ along the affine subspace $\mathcal{B} = \{x \in \mathbb{R}^n : h^T x = 0, b^T x = \beta\}$ by rotating it around $\mathcal{B}$ into direction h by an angle γ and show that distances between folded points are not longer than before. The latter fact will help to ensure feasibility with respect to the distance constraints of (4). In stating this operation we make use of the fact that due to $\|h\| = \|b\| = 1$ and $h^T b = 0$ the projection of a point $x \in \mathbb{R}^n$ onto $\mathcal{B}$ is computed by

$$p_{\mathcal{B}}(x) = x + (\beta - b^T x)b - h^T x h. \tag{6}$$

Therefore the rotation of $x \in \mathcal{H}$ around $\mathcal{B}$ uses the radius $\|x - p_{\mathcal{B}}(x)\| = |\beta - b^T x|$.

Observation 10 (Folding a Flat Halfspace) *Given* $h \in \mathbb{R}^n$ *with* $\|h\| = 1$*,* $\mathcal{H} = \{x \in \mathbb{R}^n : h^T x = 0\}$*, given* $b \in \mathcal{H}$ *with* $\|b\| = 1$*,* $\beta \in \mathbb{R}$*, and* $\mathcal{B} = \{x \in \mathcal{H} : b^T x = \beta\}$*. Define the continuous map* $\varphi : \mathcal{H} \times [-\pi, \pi] \to \mathbb{R}^n$ *by*

$$\varphi(x, \gamma) = \begin{cases} p_{\mathcal{B}}(x) - (\beta - b^T x)[b \cos\gamma + h \sin\gamma] & \text{if } b^T x < \beta, \\ x & \text{if } b^T x \geq \beta. \end{cases}$$

For all $\gamma \in [-\pi, \pi]$ *and all* $x \in \mathcal{H}$*,*

(i) $p_{\mathcal{B}}(x) = p_{\mathcal{B}}(\varphi(x, \gamma))$ *and* $\|x - p_{\mathcal{B}}(x)\| = \|\varphi(x, \gamma) - p_{\mathcal{B}}(x)\|$*,*
(ii) $\|\varphi(x, \gamma) - \varphi(y, \gamma)\| = \|x - y\|$ *for all* $y \in \mathcal{B}$*,*
(iii) $\|\varphi(x, \gamma) - \varphi(y, \gamma)\| \leq \|x - y\|$ *for all* $y \in \mathcal{H}$*.*

Proof. (i) follows by direct calculation from (6).

If $b^T x \geq \beta$ and $b^T y \geq \beta$ the points are not transformed. If $b^T x < \beta$ and $b^T y \leq \beta$ both points are subject to the same orthogonal transformation which preserves distances. This implies (ii). If, w.l.o.g., $b^T x < \beta \leq b^T y$, the intersection of the line segment between x and

y and $\mathcal{B}$ determines a unique point $z \in \mathcal{B} \cap \operatorname{conv}\{x, y\}$. The triangle inequality and (ii) yield $\|\varphi(x,\gamma) - y\| \leq \|\varphi(x,\gamma) - z\| + \|z - y\| = \|x - z\| + \|z - y\| = \|x - y\|$, so (iii) holds. ■

Next, we need to trace the objective value as we fold a subset of nodes. Any such operation can be viewed as a combination of a rotation around the barycenter of the nodes and a uniform translation without rotation. The following two observations show that rotations around the barycenter do not affect the cost function while the change induced by a translation is easily tracked via the barycenter alone.

Observation 11 (Rotation around the Barycenter) *Given $v_i \in \mathbb{R}^n$ ($i \in C \subseteq N$), $\bar{v} = \frac{1}{|C|}\sum_{i\in C} v_i$, set $v_i' = Q(v_i - \bar{v}) + \bar{v}$ where Q is an orthogonal matrix. Then,*

$$\sum_{i\in C} \|v_i'\|^2 = \sum_{i\in C} \|v_i\|^2 .$$

Proof.

$$\begin{aligned}
\sum_{i\in C} \|v_i'\|^2 &= \sum_{i\in C} \|Q(v_i - \bar{v})\|^2 + 2\bar{v}^T Q \underbrace{\sum_{i\in C}(v_i - \bar{v})}_{=0} + |C|\,\|\bar{v}\|^2 \\
&= \sum_{i\in C} \|v_i - \bar{v}\|^2 + 2\bar{v}^T \sum_{i\in C}(v_i - \bar{v}) + |C|\,\|\bar{v}\|^2 \\
&= \sum_{i\in C} \|v_i - \bar{v} + \bar{v}\|^2 = \sum_{i\in C} \|v_i\|^2 .
\end{aligned}$$

■

Observation 12 (Translation) *Given $d \in \mathbb{R}^n$, $v_i \in \mathbb{R}^n$ and $v_i' = v_i + d$ ($i \in C \subseteq N$), $\bar{v} = \frac{1}{|C|}\sum_{i\in C} v_i$. Then,*

$$\sum_{i\in C} \|v_i'\|^2 = \sum_{i\in C} \|v_i\|^2 + |C|(2\bar{v} + d)^T d.$$

Proof. $\sum_{i\in C} \|v_i + d\|^2 = \sum_{i\in C} \|v_i\|^2 + 2|C|\bar{v}^T d + |C| d^T d.$ ■

Putting these together, we now describe the cost change arising in folding a subset of the nodes.

Observation 13 (The Cost of Folding) *Given h, b, $\beta > 0$, $\mathcal{H}$, $\mathcal{B}$, and φ as in Obs. 10, given $v_i \in \{x \in \mathcal{H} : b^T x < \beta\}$ ($i \in C \subseteq N$), $\bar{v} = \frac{1}{|C|}\sum_{i\in C} v_i$, $\gamma \in [-\pi, \pi]$, set for $i \in C$*

$$v_i' = \varphi(v_i, \gamma).$$

Then,

$$\sum_{i\in C} \|v_i'\|^2 = \sum_{i\in C} \|v_i\|^2 + 2|C| r (1 - \cos\gamma)\beta \qquad \text{with } r = \beta - b^T \bar{v} > 0.$$

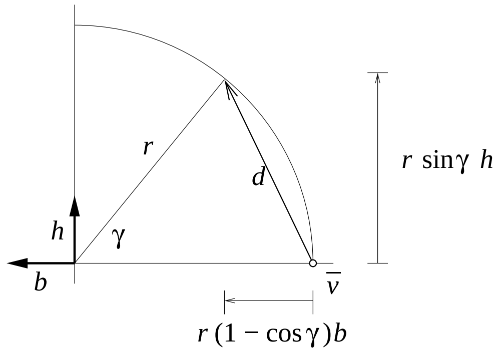

Figure 5: Rotation around an Affine Subspace.

Proof. The rotation around $\mathcal{B}$ may be split into a rotation of the points in C around their barycenter $\bar{v}$ as in Obs. 11 and a translation as analyzed in Obs. 12. The corresponding displacement d for rotating $\bar{v}$ around $\mathcal{B}$ by angle γ is $d = r(\sin\gamma)h + r(1-\cos\gamma)b$ where $r = \beta - b^T\bar{v} > 0$ is the radius (see Fig. 5). By $\bar{v}^T h = 0$, $b^T h = 0$, and Obs. 12 the cost function changes by

$$\begin{aligned}|C|(2\bar{v}^T d + d^T d) &= |C|\left(2r(1-\cos\gamma)\bar{v}^T b + r^2[\sin^2\gamma + (1-\cos\gamma)^2]\right)\\ &= |C|\left(2r(1-\cos\gamma)\bar{v}^T b + r^2[2 - 2\cos\gamma]\right)\\ &= 2|C|r(1-\cos\gamma)(\bar{v}^T b + r)\\ &= 2|C|r(1-\cos\gamma)\beta.\end{aligned}$$

■

Proof of Theorem 3. Let $h \in \mathbb{R}^n$ with $\|h\| = 1$ satisfy $h^T v_i = 0$ for all $i \in N$ as in Obs. 9 and let $\mathcal{S} = \text{conv}\{v_s : s \in S\}$. Assume, for contradiction, that the theorem is not true. Then there is a node in C_1, call it node 1, and a node in C_2, call it node 2, embedded in v_1 and v_2 respectively, that satisfy $\text{conv}\{0, v_1\} \cap \mathcal{S} = \text{conv}\{0, v_2\} \cap \mathcal{S} = \emptyset$. By convex separation each set $\text{conv}\{0, v_j\}$ can be separated from $\mathcal{S}$ by a separating hyperplane within the subspace $\text{span}\{v_i : i \in N\}$. So for $j \in \{1, 2\}$ there are vectors $b_j \in \text{span}\{v_i : i \in N\}$ (these satisfy $b_j^T h = 0$) and scalars $\beta_j > 0$ so that $b_j^T x \geq \beta_j$ for all $x \in \mathcal{S}$ and $b_j^T x < \beta_j$ for all $x \in \text{conv}\{0, v_j\}$.

Next we show that we can find a convex combination of these two inequalities by choosing an appropriate $\alpha \in [0, 1]$ so that for $b(\alpha) = (1-\alpha)b_1 + \alpha b_2$, $\beta(\alpha) = (1-\alpha)\beta_1 + \alpha\beta_2$ the open halfspace $\{x : b(\alpha)^T x < \beta(\alpha)\}$ contains points of both C_1 and C_2 (illustrated in Fig. 6). Indeed, for $\alpha = 0$ the halfspace contains v_1 and so a point of C_1, for $\alpha = 1$ it contains v_2 which belongs to C_2, and it contains the origin for all $\alpha \in [0, 1]$. Suppose, for contradiction, that in sweeping α through $[0, 1]$ the halfspace looses the last point of C_1 before it encounters the first point of C_2 at some particular $\bar{\alpha}$. Then the corresponding hyperplane defined by $b(\alpha)^T x = \beta(\alpha) > 0$ would separate 0 strictly from $\text{conv}\{v_i : i \in N\}$; but this contradicts the feasibility of the v_i as the origin is a convex combination of the v_i by the equilibrium constraint $\frac{1}{n}\sum_{i\in N} v_i = 0$.

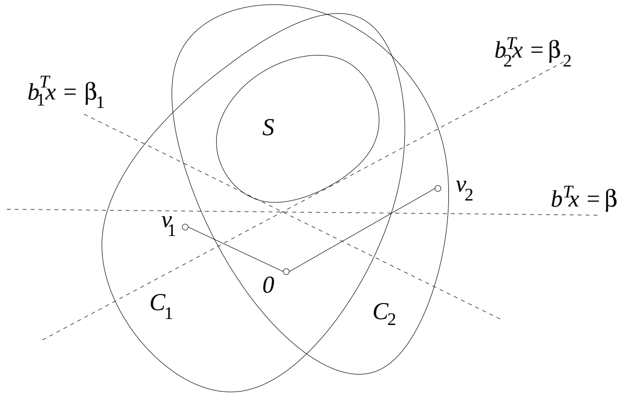

Figure 6: Initial setting in case 1 of the separator-shadow proof.

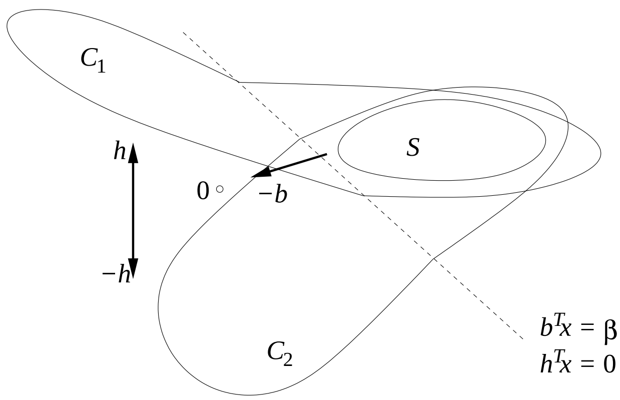

Figure 7: Improving movement in Case 1 of the separator-shadow proof.

Having found appropriate b and β note that $b^T h = 0$ holds and by scaling b and β we may assume w.l.o.g. $\|b\| = 1$. Let, for $j \in \{1, 2\}$, $M_j = \{i \in C_j : b^T v_i < \beta\}$, $m_j = |M_j| > 0$, and $\bar{v}_j = \frac{1}{m_j} \sum_{i \in M_j} v_i$. Next, consider rotating independently for each j the points in M_j around the affine subspace $\mathcal{B} = \{x \in \mathbb{R}^n : h^T x = 0, b^T x = \beta\}$ as specified in Obs. 13. Because the points in M_1 and M_2 are not adjacent and distances to the remaining points are not increased by Obs. 10(iii), the edge constraints in (4) remain satisfied. We show, that rotating the points in M_1 in direction h and the points in M_2 against direction h by sufficiently small angels γ_1 and γ_2 improves the solution (see Fig. 7). As in the proof of Obs. 13 denote, for $j \in \{1, 2\}$, radius and displacement of $\bar{v}_j$ by

$$r_j = \beta - b^T \bar{v}_j > 0 \qquad \text{and} \qquad d_j = r_j[(\sin \gamma_j)h + (1 - \cos \gamma_j)b]$$

yielding the improvement $2m_j r_j (1 - \cos \gamma_j)\beta$. Rotation j adds $m_j d_j$ to the barycenter of all points and has to be compensated in order to maintain feasibility with respect to the equilibrium constraint. Shifts of the global barycenter in the direction of h can be avoided by requiring $m_1 d_1^T h = -m_2 d_2^T h$, i.e., given γ_1 choose γ_2 in dependence of γ_1 so that $m_1 r_1 \sin \gamma_1 = -m_2 r_2 \sin \gamma_2$. After carrying out these rotations it therefore remains to shift all points by

$$d = -(m_1 d_1^T b + m_2 d_2^T b)b/n = -[m_1 r_1 (1 - \cos \gamma_1) + m_2 r_2 (1 - \cos \gamma_2)]b/n$$

for feasibility in (4). Using Obs. 12, the total objective improvement is

$$\begin{aligned} &\sum_{j \in \{1,2\}} 2m_j r_j (1 - \cos \gamma_j)\beta - n d^T d = \\ &= \sum_{j \in \{1,2\}} 2m_j r_j (1 - \cos \gamma_j)\beta - \frac{1}{n}[m_1 r_1 (1 - \cos \gamma_1) + m_2 r_2 (1 - \cos \gamma_2)]^2. \end{aligned}$$

This is positive for γ_1 and $\gamma_2(\gamma_1)$ close enough to zero, yielding a contradiction to the optimality of the embedding. ■

4 Separators containing the Origin

The freedom for squeezing optimal embeddings into lower dimensions that will be needed for the proof of Theorem 5 in Section 5, is offered by separators that contain the origin in the convex hull of their embedded nodes. Example 6 of the star $K_{1,n}$ may help to illustrate the main idea: Alluding to the physical interpretation, we will rearrange, in a first step, the cumulated force vectors of the separated node sets so that they are balanced in just one or two additional dimensions with respect to this central separator. In a second step, we will show how to combine this with reducing the dimension of each component. The result will be that we either find a particularly large component that governs the dimension of the entire embedding or no such component exists and we succeed in flattening the embedding to a space exceeding the dimension of the separator by at most two.

We start with an optimal embedding v_i $(i \in V)$ of G and fix some h and $\mathcal{H}$ as specified in Obs. 9. Let $S \subset V$ be a separator in G satisfying

$$0 \in \mathcal{S} = \operatorname{conv}\{v_i : i \in S\}$$

separating G into m disconnected sets $C_j \subset N$, $j \in M = \{1, \ldots, m\}$. For each $j \in M$ the *cumulated vector* is denoted by $\bar{v}_j = \sum_{i \in C_j} v_i$. Typically, we will not modify the embedding on the linear subspace spanned by the vectors of the separator,

$$\mathcal{L} = \operatorname{span} \mathcal{S}.$$

Modifications will be restricted to its orthogonal complement $\mathcal{L}^\perp$, so mostly our illustrations are given with respect to the embedding obtained by projecting the v_i onto $\mathcal{L}^\perp$. In the projected embedding $p_{\mathcal{L}^\perp}(v_i)$ $(i \in V)$, all nodes $i \in S$ are embedded in the origin and, like in the case of the star, the projected cumulated vectors $p_{\mathcal{L}^\perp}(\bar{v}_j)$, $j \in M$, pointing out of the origin in various directions, are in equilibrium, i.e., $\sum_{j \in M} p_{\mathcal{L}^\perp}(\bar{v}_j) = 0$ by feasibility. We note for later use that in any such configuration, none of the vectors may be longer than the sum of the others. Indeed, set $\bar{\delta}_j = \|p_{\mathcal{L}^\perp}(\bar{v}_j)\|$ $(j \in M)$, then $\sum_{j \in M} p_{\mathcal{L}^\perp}(\bar{v}_j) = 0$ implies

$$\sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j \geq \bar{\delta}_{\hat{\jmath}} \text{ for } \hat{\jmath} \in M. \tag{7}$$

The following fundamental fact will be used repeatedly (the equilibrium constraint may get violated initially but this will be taken care of later). For each $j \in M$ the vector $p_{\mathcal{L}^\perp}(\bar{v}_j)$ can be rotated around the origin freely within $\mathcal{L}^\perp$ while preserving all distances between nodes in $C_j \cup S$ by applying to all $p_{\mathcal{L}^\perp}(v_i)$, $i \in C_j$, an orthogonal transformation Q_j with $\mathcal{L}$ contained in its invariant subspace (i.e., Q_j restricted to $\mathcal{L}$ is the identity). Furthermore, such transformations do not influence the objective value, as distances to $0 \in \mathcal{L}$ are preserved. We complete step one by showing that the vectors $p_{\mathcal{L}^\perp}(\bar{v}_j)$ with their lengths $\bar{\delta}_j = \|p_{\mathcal{L}^\perp}(\bar{v}_j)\|$ can always be rotated into at most three normalized directions d_1, d_2, d_3 so that the equilibrium constraint holds again in $\mathcal{L}^\perp$ (by definition, the equilibrium constraint stays valid within $\mathcal{L}$).

Observation 14 *Given scalars $\bar{\delta}_j \geq 0$ for $j \in M = \{1, \ldots, m\}$, $m \geq 2$, so that, for each $\hat{\jmath} \in M$, $\sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j \geq \bar{\delta}_{\hat{\jmath}}$. There exist vectors $d_1, d_2, d_3 \in \mathbb{R}^2$ with $\|d_1\| = \|d_2\| = \|d_3\| = 1$ and an assignment $\kappa : M \to \{1, 2, 3\}$ so that $\sum_{j \in M} \bar{\delta}_j d_{\kappa(j)} = 0$. This also holds if in addition $|\{j \in M : \kappa(j) = 1\}| = 1$ is required.*

Proof. If $|M| = 2$ then $\bar{\delta}_1 = \bar{\delta}_2$ and the claim holds for $d_1 = -d_2$ and κ correspondingly. Otherwise let $\hat{\jmath} \in M$ be the smallest number so that $\sum_{j=1}^{\hat{\jmath}-1} \bar{\delta}_j < \frac{1}{2} \sum_{j \in M} \bar{\delta}_j \leq \sum_{j=1}^{\hat{\jmath}} \bar{\delta}_j$, set $\kappa(\hat{\jmath}) = 1$, $\kappa(j) = 2$ for $\hat{\jmath} > j \in M$ and $\kappa(j) = 3$ for $\hat{\jmath} < j \in M$. Set $\hat{\delta}_h = \sum_{j \in M, \kappa(j) = h} \bar{\delta}_j$, $h \in \{1, 2, 3\}$. Note that $\hat{\delta}_1 \leq \hat{\delta}_2 + \hat{\delta}_3$, $\hat{\delta}_2 \leq \hat{\delta}_1 + \hat{\delta}_3$, $\hat{\delta}_3 \leq \hat{\delta}_1 + \hat{\delta}_2$. Assume, w.l.o.g., that $\hat{\delta}_1 \leq \hat{\delta}_2 \leq \hat{\delta}_3$. Set $d_1(\alpha) = (\cos \alpha, -\sin \alpha)^T$ for $0 \leq \alpha \leq \pi$, $d_2(\alpha) = (\cos \gamma(\alpha), \sin \gamma(\alpha))^T$ where $\gamma(\alpha)$ is defined implicitly by $\hat{\delta}_2 \sin \gamma(\alpha) = \hat{\delta}_1 \sin \alpha$, and $d_3 = (-1, 0)^T$. Then $b(\alpha) =$

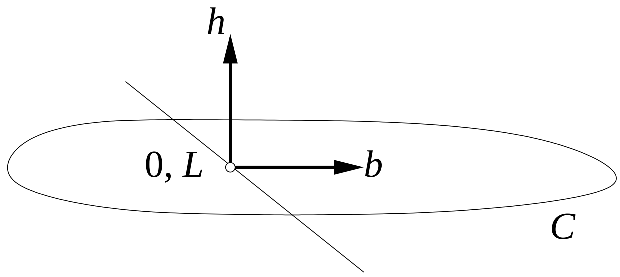

Figure 8: Initial setting before the transformation of C in Obs. 15-19.

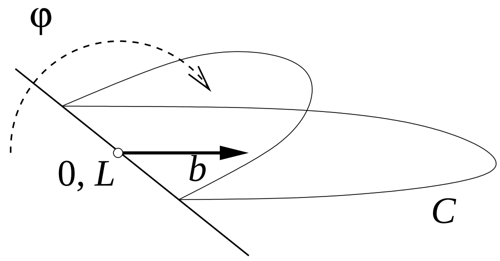

Figure 9: φ folds C into the halfspace specified by b (Obs. 15).

$\hat{\delta}_1 d_1(\alpha) + \hat{\delta}_2 d_2(\alpha) + \hat{\delta}_3 d_3$ satisfies $[b(\alpha)]_2 = 0$ for all $0 \le \alpha \le \pi$, $[b(0)]_1 \ge 0$ and $[b(\pi)]_1 \le 0$, so by continuity of $b(\alpha)$ there is an $\hat{\alpha} \in [0, \pi]$ with $b(\hat{\alpha}) = 0$. ■

Let us now turn towards reducing the dimension of the node sets. If span $\{v_i : i \in C_j\} \subseteq \mathcal{L}$ for $j \in M$ then the embedding is good enough for our purposes. Assume therefore that there is some $j \in M$ with span $\{v_i : i \in C_j\} \not\subseteq \mathcal{L}$. In manipulating the embedding of C_j we will again only apply orthogonal transformations (sometimes we will simultaneously use separate ones for each point in C_j) that contain $\mathcal{L}$ in their invariant subspace. Therefore all distances of points in C_j will preserve their distance to the origin and to the embedding of S. In consequence, optimality is guaranteed if feasibility can be maintained. In particular, feasibility of the distance constraints is ensured whenever distances within C_j are not increased. Our manipulations may, however, increase the length of $p_{\mathcal{L}^\perp}(\bar{v}_j)$ and thus $\bar{\delta}_j$. But by Obs. 14 it suffices that condition (7) is satisfied at the end in order to restore the equilibrium constraint, as well.

The goal is to squeeze the entire embedding of component C_j into the flat halfspace spanned by $\mathcal{L}$ and one additional direction $b_j \in \mathcal{H} \cap \mathcal{L}^\perp$ with $\|b_j\| = 1$. This works as follows. We first fold all nodes into the flat halfspace $\{x \in \mathcal{H} : b^T x \ge 0\}$ via Obs. 10 (put $b = b_j$ and $\beta = 0$), see figures 8 and 9; this leaves $\mathcal{L} \subseteq \mathcal{B}$ untouched as required. Then we collapse this flat halfspace into the even flatter halfspace cone$(\mathcal{L} \cup \{b_j\})$ as if collapsing an umbrella by rotating the ribs towards its handle. These two operations are concatenated to a continuous transformation $u_i(t)$ of the embedding for $t \in [0, 1]$ and we will see that the norm $\bar{\delta}_j(t) = \|p_{\mathcal{L}^\perp}(\bar{u}_j(t))\|$ of the cumulated vector $\bar{u}_j(t) = \sum_{i \in C_j} u_i(t)$ is nondecreasing throughout, so that we can easily stop the transformation at an appropriate t to ensure

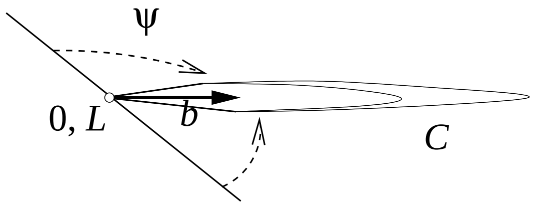

Figure 10: ψ collapses C into the flat halfspace spanned by $\mathcal{L}$ and direction b (Obs. 17).

condition (7). We start with the folding operation.

Observation 15 (Transformation Part 1, Folding) *Given $j \in M$, $b_j \in \mathcal{H} \cap \mathcal{L}^\perp$ with $\|b_j\| = 1$, define $\varphi_i : [0,1] \to \mathbb{R}^n$ for $i \in C_j$ by*

$$\varphi_i(t) = \begin{cases} v_i - (v_i^T b_j)b_j + (v_i^T b_j)[b_j \cos t\pi + h \sin t\pi] & \text{if } v_i^T b_j < 0 \\ v_i & \text{if } v_i^T b_j \geq 0. \end{cases}$$

Then, for $i \in C_j$,
(i) $\varphi_i(0) = v_i$,
(ii) $\varphi_i(1) \in \{x \in \mathcal{H} : b^T x \geq 0\}$,
and for all $t \in [0,1]$ it holds that
(iii) $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(\varphi_i(t))$ *and* $\|p_{\mathcal{L}^\perp}(v_i)\| = \|p_{\mathcal{L}^\perp}(\varphi_i(t))\|$,
(iv) $\|\varphi_i(t) - v\| = \|v_i - v\|$ *for* $v \in \mathcal{L} \supseteq \{v_s : s \in S\}$,
(v) $\|\varphi_i(t) - \varphi_k(t)\| \leq \|v_i - v_k\|$ *for* $k \in C_j$.

Proof. (i,ii) follow from direct calculation and $v_i \in \mathcal{H}$, (iii,iv,v) from Obs. 10(i,ii,iii) using $b = b_j$, $\beta = 0$, $\varphi_i(t) = \varphi(v_i, t\pi)$, and the fact that $\mathcal{S} \subseteq \mathcal{L} \subseteq \mathcal{B}$. ■

Next, we show that the length of the projected cumulated vector increases throughout this first transformation.

Observation 16 *For φ_i ($i \in C_j$) as defined in Obs. 15, define $\bar{\varphi}_j : [0,1] \to \mathbb{R}^n$ by*

$$\bar{\varphi}_j(t) = \sum_{i \in C_j} \varphi_i(t).$$

The length $\|p_{\mathcal{L}^\perp}(\bar{\varphi}_j(t))\|$ is nondecreasing in $t \in [0,1]$.

Proof. The choice of b_j ensures $\mathcal{B} = \{x \in \mathcal{H} : b_j^T x = 0\} \supseteq \mathcal{L}$ and $\mathcal{B}^\perp = \operatorname{span}\{h, b_j\}$. By definition of the φ_i in Obs. 15 we obtain

$$\begin{aligned} \|p_{\mathcal{L}^\perp}(\bar{\varphi}_j(t))\|^2 &= \|p_{\mathcal{L}^\perp}(p_{\mathcal{B}}(\bar{\varphi}_j(0)))\|^2 + \|p_{\mathcal{B}^\perp}(\bar{\varphi}_j(t))\|^2 \\ &= \|p_{\mathcal{L}^\perp}(p_{\mathcal{B}}(\bar{\varphi}_j(0)))\|^2 + \\ &\quad \left\| \sum_{i \in C_j, v_i^T b_j < 0} (v_i^T b_j)[b_j \cos t\pi + h \sin t\pi] + \sum_{i \in C_j, v_i^T b_j \geq 0} (v_i^T b_j) b_j \right\|^2. \end{aligned}$$

As b_j and h are orthogonal it remains to study the monotonicity of

$$\Big[\sum_{i\in C_j, v_i^T b_j<0} v_i^T b_j \cos t\pi + \sum_{i\in C_j, v_i^T b_j\geq 0} v_i^T b_j\Big]^2 + \Big[\sum_{i\in C_j, v_i^T b_j<0} v_i^T b_j \sin t\pi\Big]^2 =$$

$$= \Big[\sum_{i\in C_j, v_i^T b_j<0} v_i^T b_j\Big]^2 (\cos^2 t\pi + \sin^2 t\pi) + \Big[\sum_{i\in C_j, v_i^T b_j\geq 0} v_i^T b_j\Big]^2 +$$

$$\underbrace{2\Big[\sum_{i\in C_j, v_i^T b_j<0} v_i^T b_j\Big]\Big[\sum_{i\in C_j, v_i^T b_j\geq 0} v_i^T b_j\Big]}_{\leq 0} \cos t\pi$$

The last term is clearly nondecreasing. ■

The collapsing transformation starts from the points $\varphi_i(1)$ and runs as follows.

Observation 17 (Transformation Part 2, Collapsing)
Given the setting of Obs. 15. For $i \in C_j$*, set* $\delta_i = \|p_{\mathcal{L}^\perp}(\varphi_i(1))\|$*, determine* $0 \leq \gamma_i \leq \frac{\pi}{2}$ *and* $q_i \in \mathcal{L}^\perp$*,* $q_i^T b_j = 0$*,* $\|q_i\| = 1$ *so that* $p_{\mathcal{L}^\perp}(\varphi_i(1)) = \delta_i(q_i \cos\gamma_i + b_j \sin\gamma_i)$ *and define* $\psi_i : [0,1] \to \mathbb{R}^n$ *by*

$$\psi_i(t) = p_{\mathcal{L}}(\varphi_i(1)) + \delta_i \left[q_i \cos(\gamma_i + t[\frac{\pi}{2} - \gamma_i]) + b_j \sin(\gamma_i + t[\frac{\pi}{2} - \gamma_i])\right].$$

Then, for $i \in C_j$*,*
(i) $\psi_i(0) = \varphi_i(1)$*,*
(ii) $\psi_i(1) = p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\| b_j \in \mathcal{L} + \{\beta b_j : \beta \geq 0\}$
and for all $t \in [0,1]$ *it holds that*
(iii) $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(\psi_i(t))$ *and* $\|p_{\mathcal{L}^\perp}(v_i)\| = \|p_{\mathcal{L}^\perp}(\psi_i(t))\|$*,*
(iv) $\|\psi_i(t) - v\| = \|v_i - v\|$ *for* $v \in \mathcal{L} \supseteq \{v_s : s \in S\}$*,*
(v) $\|\psi_i(t) - \psi_k(t)\| \leq \|v_i - v_k\|$ *for* $k \in C_j$*.*

Proof. First note that Obs.15(iii) implies $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(\varphi_i(1))$ and $\delta_i = \|p_{\mathcal{L}^\perp}(v_i)\|$. Now (i) and (ii) follow from direct calculation and (iii) and (iv) are proved in the same way as (i) and (ii) of Obs. 10. It remains to prove (v).

Because of Obs. 15(v) it suffices to prove $\|\psi_i(t) - \psi_k(t)\|^2 \leq \|\varphi_i(1) - \varphi_k(1)\|^2$ for $i, k \in C_j$. For this we need to show $\psi_i(t)^T \psi_k(t) \geq \varphi_i(1)^T \varphi_k(1)$ which leads to the condition

$$\begin{aligned} f_{ik}(t) &= (q_i^T q_k)[\cos(\gamma_i + t[\frac{\pi}{2} - \gamma_i])\cos(\gamma_k + t[\frac{\pi}{2} - \gamma_k])] + \\ &\quad \sin(\gamma_i + t[\frac{\pi}{2} - \gamma_i])\sin(\gamma_k + t[\frac{\pi}{2} - \gamma_k]) \\ &\geq (q_i^T q_k)[\cos\gamma_i \cos\gamma_k] + \sin\gamma_i \sin\gamma_k = f_{ik}(0). \end{aligned} \tag{8}$$

We prove that $f_{ik}(t)$ is nondecreasing in $t \in [0,1]$. In the case $q_i^T q_k < 0$ both cosine terms in $f_{ik}(t)$ are non increasing and the sine terms are non decreasing. In the remaining case we use the angle addition formulas to find

$$f_{ik}(t) = q_i^T q_k \cos((1-t)[\gamma_i - \gamma_k]) + (1 - q_i^T q_k)\sin(\gamma_i + t[\frac{\pi}{2} - \gamma_i])\sin(\gamma_k + t[\frac{\pi}{2} - \gamma_k]).$$

But $0 \leq q_i^T q_k \leq 1$ and so the cosine and sine terms are non decreasing. ■

Again, we continue with showing that during this transformation the length of the projected cumulated vector is nondecreasing.

Observation 18 *For ψ_i $(i \in C_j)$ as defined in Obs. 17, define $\bar{\psi}_j : [0,1] \to \mathbb{R}^n$ by*

$$\bar{\psi}_j(t) = \sum_{i \in C_j} \psi_i(t).$$

The length $\|p_{\mathcal{L}^\perp}(\bar{\psi}_j(t))\|$ is nondecreasing in $t \in [0,1]$.

Proof. Using the functions f_{ik} introduced in (8) we may write

$$\begin{aligned}
\Big\| \sum_{i \in C_j} p_{\mathcal{L}^\perp}(\psi_i(t)) \Big\|^2 &= \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(\psi_i(t))\|^2 + \sum_{i,k \in C_j, i<k} 2(p_{\mathcal{L}^\perp}(\psi_i(t)))^T (p_{\mathcal{L}^\perp}(\psi_k(t))) \\
&= \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(\psi_i(t))\|^2 + \sum_{i,k \in C_j, i<k} \delta_i \delta_k f_{ik}(t)
\end{aligned}$$

and we have shown in the proof of Obs. 17 that each $f_{ik}(t)$ is nondecreasing in $t \in [0,1]$. ■

We concatenate both transformations into one and summarize our findings on this collapsing transformation.

Observation 19 (Collapsing Transformation)
Given $j \in M$, $b_j \in \mathcal{H} \cap \mathcal{L}^\perp$ with $\|b_j\| = 1$ define $u_i : [0,1] \to \mathbb{R}^n$ for $i \in C_j$ by

$$u_i(t) = \begin{cases} \varphi_i(2t) & \text{for } t \in [0, \frac{1}{2}], \\ \psi_i(2[t - \frac{1}{2}]) & \text{for } t \in (\frac{1}{2}, 1], \end{cases} \tag{9}$$

with φ_i and ψ_i as given in Obs. 15 and Obs. 17. Then, for $i \in C_j$,
(i) $u_i(0) = v_i$,
(ii) $u_i(1) = p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\| b_j \in \mathcal{L} + \{\beta b_j : \beta \geq 0\}$,
and for all $t \in [0,1]$ it holds that
(iii) $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(u_i(t))$ *and* $\|p_{\mathcal{L}^\perp}(v_i)\| = \|p_{\mathcal{L}^\perp}(u_i(t))\|$,
(iv) $\|u_i(t) - v\| = \|v_i - v\|$ *for* $v \in \mathcal{L} \supseteq \{v_s : s \in S\}$,
(v) $\|u_i(t) - u_k(t)\| \leq \|v_i - v_k\|$ *for* $k \in C_j$.
Furthermore, for

$$\bar{u}_j(t) = \sum_{i \in C_j} u_i(t)$$

the length $\|p_{\mathcal{L}^\perp}(\bar{u}_j(t))\|$ is nondecreasing in $t \in [0,1]$ and $\|p_{\mathcal{L}^\perp}(\bar{u}_j(1))\| = \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(v_i)\|$.

Proof. The result follows from Obs. 15, 17 and Obs. 16, 18. ■

Suppose that the lengths $\bar{\delta}_j = \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(v_i)\|$ $(j \in M)$ of the collapsed sets satisfy (7). Then, in order to obtain an embedding that is also in equilibrium with respect to the subspace $\mathcal{L}^\perp$, we only have to choose the collapsing direction b_j of each component C_j according to the vectors d_k (embedded in $\mathcal{L}^\perp$) with the assignment κ of Obs. 14, $b_j = d_{\kappa(j)}$. This will yield an optimal embedding of dimension at most $\dim \mathcal{L} + 2$ as described in following lemma.

Lemma 20 *Let $v_i \in \mathbb{R}^n$ for $i \in N$ be an optimal solution of* (4) *for a connected graph $G = (N, E)$ and let $S \subset N$ with $0 \in \mathcal{S} = \operatorname{conv}\{v_s : s \in S\}$ be a separator in G giving rise to disconnected sets $C_j \subset N$, $j \in M = \{1, \ldots, m\}$. Put $\mathcal{L} = \operatorname{span} \mathcal{S}$ and, for $j \in M$, $\bar{\delta}_j = \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(v_i)\|$.*

If $\bar{\delta}_{\hat{\jmath}} \le \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$ for all $\hat{\jmath} \in M$ then there exist vectors $d_1, d_2, d_3 \in \mathcal{L}^\perp$, $\|d_1\| = \|d_2\| = \|d_3\| = 1$ with $\dim \operatorname{span}\{d_1, d_2, d_3\} \le 2$, $b_j \in \{d_1, d_2, d_3\}$, $j \in M$, so that the embedding v_i', $i \in N$, with

$$v_i' = \begin{cases} v_i & \text{for } i \in S, \\ p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\| b_j & \text{for } i \in C_j. \end{cases}$$

is also an optimal embedding of (4). *Furthermore, such an embedding exists with $b_j = d_1$ for at most one $j \in M$.*

Proof. Choose h and $\mathcal{H}$ as specified in Obs. 9. If $\bar{\delta}_j = 0$ for all $j \in M$ then the statement holds for $d_1 = d_2 = d_3 = h$ because $v_i' = v_i \in \mathcal{L}$ for $i \in N$. So we may assume $\bar{\delta}_j > 0$ for at least two $j \in M$. In the case $\dim(\mathcal{H} \cap \mathcal{L}^\perp) = 1$ we must have $|S| = n - 2$, $m = 2$, and $|C_1| = |C_2| = 1$, so $b_1 = d_1 = -b_2 = -d_2 = -d_3$ with $d_1 = p_{\mathcal{L}^\perp}(v_i)/\|p_{\mathcal{L}^\perp}(v_i)\|$ satisfies all requirements. It remains to consider the case $\dim(\mathcal{H} \cap \mathcal{L}^\perp) \ge 2$.

By Obs. 14 we find three vectors $d_1, d_2, d_3 \in \mathcal{H} \cap \mathcal{L}^\perp$ of norm one and an assignment $\kappa : M \to \{1, 2, 3\}$ satisfying $\sum_{j \in M} \bar{\delta}_j d_{\kappa(j)} = 0$ and $\{j \in M : \kappa(j) = 1\} = 1$. For $j \in M$ set $b_j = d_{\kappa(j)}$ and let $u_i(t)$, $i \in C_j$, be the transformations of Obs. 19 for the respective b_j. Then $v_i' = u_i(1)$ for $i \in C_j$, $j \in M$ by Obs. 19(ii). The distance constraints are satisfied for the new embedding because for $\{i, k\} \in E$ either

$i, k \in S$:	$\|v_i' - v_k'\| = \|v_i - v_k\|$	by definition,
$i \in C_j$ for some $j \in M, k \in S$:	$\|v_i' - v_k'\| = \|v_i - v_k\|$	by Obs. 19(iv),
$i, k \in C_j$ for some $j \in M$:	$\|v_i' - v_k'\| \le \|v_i - v_k\|$	by Obs. 19(v).

The equilibrium constraint is satisfied on $\mathcal{L}$, because $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(v_i')$ for all $i \in N$ (by definition for $i \in S$ and by Obs. 19(iii) otherwise). It is also satisfied on $\mathcal{L}^\perp$, because

$$\sum_{i \in N} p_{\mathcal{L}^\perp}(v_i') = \sum_{i \in S} \underbrace{p_{\mathcal{L}^\perp}(v_i')}_{=0} + \sum_{j \in M} \sum_{i \in C_j} p_{\mathcal{L}^\perp}(v_i') = \sum_{j \in M} \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(v_i)\| b_j = \sum_{j \in M} \bar{\delta}_j d_{\kappa(j)} = 0$$

by construction of the d_j. Finally, the objective value has not changed because $\|v_i\| = \|v_i'\|$ for all $i \in N$ (by definition for $i \in S$ and by Obs. 19(iii) otherwise). ∎

If one set $\hat{\jmath} \in M$ is 'heavier' than the other sets, $\bar{\delta}_{\hat{\jmath}} > \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$, the need to recover feasibility in the equilibrium constraint will not allow to collapse $\hat{\jmath}$ in full. We can, however, collapse all other sets and compensate this by carrying through the transformation in $\hat{\jmath}$ up to the $t_{\hat{\jmath}} \in [0, 1]$ when $\|p_{\mathcal{L}^\perp}(\bar{u}_{\hat{\jmath}}(t_{\hat{\jmath}}))\| = \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$. Even though this may lead to a slight increase in the overall dimension if $t_{\hat{\jmath}} < \frac{1}{2}$, it will help later to reduce the number of components we have to worry about.

Lemma 21 *Given the setting of Lemma 20 assume that there is a $\hat{\jmath} \in M$ with $\bar{\delta}_{\hat{\jmath}} > \sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j$. There exists an $h \in \operatorname{span}\{v_i : i \in N\}^\perp$ and an optimal embedding v'_i $(i \in N)$ of (4) with*

$$\begin{array}{ll} v'_i \in \operatorname{span}\{h, v_i : i \in C_{\hat{\jmath}}\} & \textit{for } i \in C_{\hat{\jmath}}, \\ v'_i = v_i & \textit{for } i \in S, \\ v'_i = p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\|\bar{b} & \textit{for } i \in C_j \textit{ with } j \in M \setminus \{\hat{\jmath}\}, \end{array}$$

where $\bar{b} = -\frac{p_{\mathcal{L}^\perp}(\bar{v}'_{\hat{\jmath}})}{\|p_{\mathcal{L}^\perp}(\bar{v}'_{\hat{\jmath}})\|}$ if $\bar{v}'_{\hat{\jmath}} = \sum_{i\in C_{\hat{\jmath}}} v'_i \notin \mathcal{L}$ and $\bar{b} = 0$ otherwise.

Furthermore, if there is some direction $\hat{b} \in \operatorname{span}\{v_i : i \in C_{\hat{\jmath}}\} \cap \mathcal{L}^\perp \setminus \{0\}$ with $\hat{b}^T v_i \geq 0$ for $i \in C_{\hat{\jmath}}$, then such an embedding exists with $v'_i \in \operatorname{span}\{v_i : i \in C_{\hat{\jmath}}\}$ for $i \in C_{\hat{\jmath}}$.

Proof. If $\sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j = 0$ then we may choose $h = \bar{b} = 0$ and not transform the embedding at all to obtain the result. Therefore assume $\sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j \neq 0$. Choose h and $\mathcal{H}$ as specified in Obs. 9. Since $\bar{\delta}_{\hat{\jmath}} > 0$ we can find a $b_{\hat{\jmath}} \in \mathcal{L}^\perp \cap \operatorname{span}\{v_i : i \in C_{\hat{\jmath}}\}$ with $\|b_{\hat{\jmath}}\| = 1$. Let $u_i(t)$ $(i \in C_{\hat{\jmath}})$ denote the transformations of Obs. 19 for this $b_{\hat{\jmath}}$, set $\bar{u}_{\hat{\jmath}}(t) = \sum_{i\in C_{\hat{\jmath}}} u_i(t)$. By Obs. 19, the function $\|p_{\mathcal{L}^\perp}(\bar{u}_{\hat{\jmath}}(t))\|$ is continuous and nondecreasing. As the equilibrium constraint is satisfied for the v_i $(i \in N)$, we have

$$\begin{aligned} \|p_{\mathcal{L}^\perp}(\bar{u}_{\hat{\jmath}}(0))\| &\overset{\text{Obs.19(i)}}{=} \Big\|\sum_{i\in C_{\hat{\jmath}}} p_{\mathcal{L}^\perp}(v_i)\Big\| \overset{\text{equilib.}}{=} \Big\|\sum_{i\in N\setminus C_{\hat{\jmath}}} p_{\mathcal{L}^\perp}(v_i)\Big\| \\ &\overset{p_{\mathcal{L}^\perp}(v_i)=0\ (i\in S)}{=} \Big\|\sum_{j\in M\setminus\{\hat{\jmath}\}} \sum_{i\in C_j} p_{\mathcal{L}^\perp}(v_i)\Big\| \\ &\leq \sum_{j\in M\setminus\{\hat{\jmath}\}} \Big\|\sum_{i\in C_j} p_{\mathcal{L}^\perp}(v_i)\Big\| \overset{\text{(by def.)}}{=} \sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j \end{aligned}$$

and, by assumption, $\|p_{\mathcal{L}^\perp}(\bar{u}_{\hat{\jmath}}(1))\| = \bar{\delta}_{\hat{\jmath}} > \sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j$. So there is a $t_{\hat{\jmath}} \in [0,1]$ with

$$\|p_{\mathcal{L}^\perp}(\bar{u}_{\hat{\jmath}}(t_{\hat{\jmath}}))\| = \sum_{j\in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j. \tag{10}$$

Set $v'_i = u_i(t_{\hat{\jmath}})$ for $i \in C_{\hat{\jmath}}$ and put

$$\bar{v}'_{\hat{\jmath}} = \sum_{i\in C_{\hat{\jmath}}} v'_i = \bar{u}_{\hat{\jmath}}(t_{\hat{\jmath}}) \quad \text{and} \quad \bar{b} = -p_{\mathcal{L}^\perp}(\bar{v}'_{\hat{\jmath}})/\|p_{\mathcal{L}^\perp}(\bar{v}'_{\hat{\jmath}})\|. \tag{11}$$

For $j \in M \setminus \{\hat{\jmath}\}$ set $b_j = \bar{b}$ and let $u_i(t)$, $i \in C_j$, be the transformations of Obs. 19 for the respective b_j. Then, by Obs. 19 (ii), $v'_i = u_i(1)$ for $i \in C_j$ with $j \in M \setminus \{\hat{\jmath}\}$. The equilibrium constraint is satisfied for the embedding v'_i, $i \in N$, because it holds on $\mathcal{L}$ due

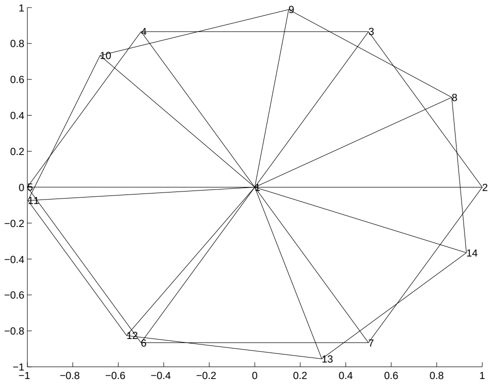

Figure 11: Optimal two dimensional embedding of two wheels with identical hub, see Rem. 22. The construction of the proof of Lemma 21 would yield a three dimensional embedding.

to $p_{\mathcal{L}}(v_i) = p_{\mathcal{L}}(v_i')$ for $i \in N$ by Obs. 19 (iii) and it holds on $\mathcal{L}^\perp$, because

$$\begin{aligned}
\sum_{i \in N} p_{\mathcal{L}^\perp}(v_i') \quad &= \quad \sum_{i \in S} \underbrace{p_{\mathcal{L}^\perp}(v_i')}_{=0} + \sum_{i \in C_{\hat{\jmath}}} p_{\mathcal{L}^\perp}(v_i') + \sum_{j \in M \setminus \{\hat{\jmath}\}} \sum_{i \in C_j} p_{\mathcal{L}^\perp}(v_i') \\
&\overset{\text{(by def.)}}{=} \quad p_{\mathcal{L}^\perp}(\bar{v}_{\hat{\jmath}}') + \sum_{j \in M \setminus \{\hat{\jmath}\}} \sum_{i \in C_j} \|p_{\mathcal{L}^\perp}(v_i)\| \bar{b} \\
&\overset{(11)}{=} \quad \left(\|p_{\mathcal{L}^\perp}(\bar{v}_{\hat{\jmath}}')\| - \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j \right) \frac{p_{\mathcal{L}^\perp}(\bar{v}_{\hat{\jmath}}')}{\|p_{\mathcal{L}^\perp}(\bar{v}_{\hat{\jmath}}')\|} \quad \overset{(10)}{=} \quad 0.
\end{aligned}$$

Feasibility of v_i', $i \in N$, with respect to the distance constraints and optimality follow from Obs. 19 (iv,v) as in the proof of Lemma 20.

Finally, suppose $\hat{b}$ exists as described in the statement of the Lemma. Then we may choose $b_{\hat{\jmath}} = \frac{\hat{b}}{\|\hat{b}\|}$ and by construction (9) of the u_i $(i \in C_{\hat{\jmath}})$, $u_i(t) = v_i$ for $t \in [0, \frac{1}{2}]$ (see Obs. 15 for φ_i) and $u_i(t) \in \mathcal{L} + \operatorname{span}\{\hat{b}, v_i\}$ for $t \in [\frac{1}{2}, 1]$ (see Obs. 17 for ψ_i). This completes the proof. ■

Remark 22 *A solution corresponding to the modified solution of this lemma is not necessarily an optimal embedding of minimal dimension. Consider, e.g., the graph consisting of two wheels with identical hub and rims of k and $k+1$ nodes with $k \geq 6$, see Fig. 11.*

Lemma 21 does not provide a bound on the dimension of the embedding but will tell us which component we have to think about next in order to get to such a bound. The case discussed in Lemma 20, on the other hand, yields a bound of $|S| + 1$, because $\dim(\mathcal{L}) \leq |S| - 1$ due to $0 \in \mathcal{S} = \operatorname{conv}\{v_i : i \in S\}$. In order to arrive at the result of Theorem 5, however, we will need a further refinement of Lemma 20 in the case $\dim(\mathcal{L}) = |S| - 1$. In order to set up the scene, assume that the v_i $(i \in N)$ are already embedded in dimension

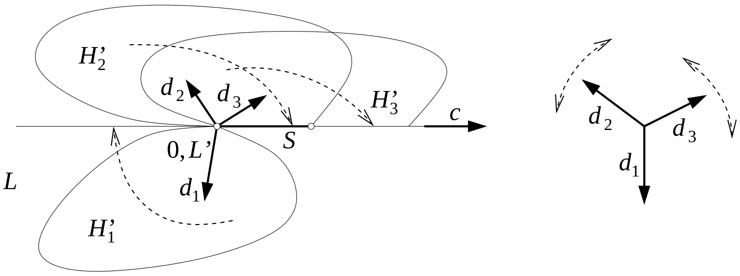

Figure 12: Transformation in the proof of Lemma 25.

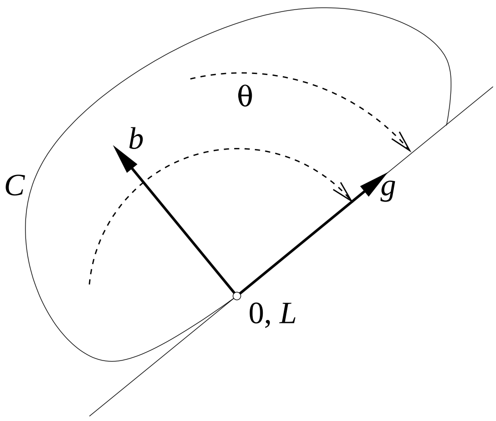

Figure 13: θ squeezes C spanned by $\mathcal{L}' + \operatorname{span}\{g\}$ and nonnegative d into the boundary halfspace spanned by $\mathcal{L}$ and nonnegative g. (Obs. 23).

$|S|+1$ as described in Lemma 20 with each node set C_j collapsed to some flat halfspace $\mathcal{L} + \{\delta d_{\kappa(j)} : \delta \geq 0\}$ and denote by H_1 the set C_j that is the only one assigned to direction d_1. We are interested in the case that H_1 is not connected to some $\hat{s} \in S$, so $S' = S \setminus \{\hat{s}\}$ is a separator for H_1 in G. By Theorem 3 we must have $0 \in \mathcal{S}' = \operatorname{conv}\{v_i : i \in S'\}$ with $\mathcal{L}' = \operatorname{span} \mathcal{S}'$ a linear subspace of dimension $\dim(\mathcal{L}') = \dim(\mathcal{L}) - 1$ and $v_{\hat{s}} \neq 0$ (otherwise the dimension of the embedding would be $|S|$ already). Figure 12 depicts the situation when projected onto $\mathcal{L}'^{\perp}$ with $H_i = \bigcup_{j \in M : \kappa(j) = i} C_j$ the set of nodes which are embedded in direction d_i, $i = 1, 2, 3$. It will turn out, that the transformation indicated in this illustration will yield an optimal embedding of dimension at most $|S|$, so we can get rid of one more dimension. For this purpose we introduce yet another transformation comparable to closing a fan, see Fig. 13. In the following observation think of vector $g = \pm p_{\mathcal{L}'^{\perp}}(v_{\hat{s}}))/\|p_{\mathcal{L}'^{\perp}}(v_{\hat{s}})\|$ as spanning the missing direction in $\mathcal{L}$ and d as the additional direction d_j for spanning the embedding of some node set $C = H_j$.

Observation 23
Given a linear subspace $\mathcal{L}' \subset \mathbb{R}^n$, $d, g \in \mathcal{L}'^{\perp}$ with $\|d\| = \|g\| = 1$, $d^T g = 0$ and $v_i \in \{x \in \mathcal{L}' + \operatorname{span}\{d, g\} : d^T x \geq 0\}$ $(i \in C \subseteq N)$. For $i \in C$, set $\delta_i = \|p_{\mathcal{L}'^{\perp}}(v_i)\|$, determine

$\gamma_i \in [0,\pi]$ *so that* $p_{\mathcal{L}'^\perp}(v_i) = \delta_i(g\cos\gamma_i + d\sin\gamma_i)$ *and define continuous maps* $\theta_i : [0,1] \to \mathbb{R}^n$

$$\theta_i(t) = p_{\mathcal{L}'}(v_i) + \delta_i\left[g\cos(\gamma_i - t\gamma_i) + d\sin(\gamma_i - t\gamma_i)\right].$$

Then, for $i \in C$,
(i) $\theta_i(0) = v_i$,
(ii) $\theta_i(1) = p_{\mathcal{L}'}(v_i) + \|p_{\mathcal{L}'^\perp}(v_i)\|g$,
and for all $t \in [0,1]$ *it holds that*
(iii) $p_{\mathcal{L}'}(v_i) = p_{\mathcal{L}'}(\theta_i(t))$ *and* $\|p_{\mathcal{L}'^\perp}(v_i)\| = \|p_{\mathcal{L}'^\perp}(\theta_i(t))\|$,
(iv) $\|\theta_i(t) - v\| \le \|v_i - v\|$ *for* $v \in \mathcal{L}' + \{\beta g : \beta > 0\}$,
(v) $\|\theta_i(t) - \theta_k(t)\| \le \|v_i - v_k\|$ *for* $k \in C$.
Furthermore, for $\bar\theta_C(t) = \sum_{i\in C}\theta_i(t)$,
(vi) $p_{\mathcal{L}'^\perp}(\bar\theta_C(t)) \in \operatorname{span}\{g\} + \{\beta d : \beta \ge 0\}$ *for* $t \in [0,1]$,
(vii) $g^T\bar\theta_C(t)$ *is strictly increasing in* $t \in [0,1]$ *if* $\gamma_i \in (0,\pi]$ *and* $\delta_i > 0$ *for some* $i \in C$,
(viii) $\bar\theta_C(1) = \sum_{i\in C} p_{\mathcal{L}'}(v_i) + g\sum_{i\in C}\delta_i$.

Proof. (i,ii,iii) follow from direct calculation and by exploiting that $g, d \in \mathcal{L}'^\perp$ are orthonormal vectors. In order to prove (iv), apply the same arguments used in the proofs of Obs. 10(ii,iii) to $v \in \mathcal{L}'$ and to $v \in \mathcal{L}' + \{\beta g : \beta \ge 0\}$, respectively.

For proving (v), i.e., $\|\theta_i(t) - \theta_k(t)\|^2 \le \|v_i - v_k\|^2$ for $i, k \in C$, it suffices to show that

$$f_{ik}(t) = \theta_i(t)^T\theta_k(t) \ge v_i^T v_k \overset{\text{(i)}}{=} \theta_i(0)^T\theta_k(0) = f_{ik}(0),$$

or, as $g, d \in \mathcal{L}'^\perp$ are orthonormal vectors, that the function

$$f_{ik}(t) = p_{\mathcal{L}'}(v_i)^T p_{\mathcal{L}'}(v_k) + \delta_i\delta_k[\cos(\gamma_i - t\gamma_i)\cos(\gamma_k - t\gamma_k) + \sin(\gamma_i - t\gamma_i)\sin(\gamma_k - t\gamma_k)]$$

is nondecreasing in $t \in [0,1]$. By the angle addition formulas and since the cosine is an even function,

$$\cos(\gamma_i - t\gamma_i)\cos(\gamma_k - t\gamma_k) + \sin(\gamma_i - t\gamma_i)\sin(\gamma_k - t\gamma_k) = \cos((1-t)|\gamma_i - \gamma_k|).$$

The right hand side is non decreasing and, thus, f_{ik} is nondecreasing.

(vi) and (viii) follow from direct computation and for (vii) it suffices to observe that

$$g^T\bar\theta_C(1) = \sum_{i\in C}\delta_i\cos(\gamma_i - t\gamma_i)$$

is strictly increasing because $\delta_i \ge 0$ for all $i \in C$ and $\cos(\gamma_i - t\gamma_i)$ is strictly increasing in $t \in [0,1]$ whenever $\gamma_i \in (0,\pi]$. ■

The next observation will serve to find the correct balancing of the parameters for each H_j in order to guarantee the equilibrium constraint on the subspace spanned by g and appropriately chosen d_j.

Observation 24 *Given continuous functions* $\lambda_j : [0,1] \to \mathbb{R}^2$ *(*$j \in \{1,2,3\}$*) and* $\sigma \in \mathbb{R}$ *with*

(i) $[\lambda_1(t)]_1$ *is strictly decreasing,* $[\lambda_2(t)]_1$ *and* $[\lambda_3(t)]_1$ *are strictly increasing,*
(ii) $[\lambda_i(t)]_2 \geq 0$ *for* $t \in [0,1]$ *and* $i = 1,2,3$,
(iii) $[\lambda_1(0)]_1 + [\lambda_2(0)]_1 + [\lambda_3(0)]_1 + \sigma = 0$,
(iv) $[\lambda_i(0)]_2 < [\lambda_j(0)]_2 + [\lambda_k(0)]_2$ *for pairwise distinct* $i,j,k \in \{1,2,3\}$,
(v) $[\lambda_1(1)]_2 = [\lambda_2(1)]_2 = [\lambda_3(1)]_2 = 0$.

There exist $t_1, t_2, t_3 \in [0,1]$ *and pairwise distinct* $\hat{\imath}, \hat{\jmath}, \hat{k} \in \{1,2,3\}$ *satisfying*

(vi) $[\lambda_1(t_1)]_1 + [\lambda_2(t_2)]_1 + [\lambda_3(t_3)]_1 + \sigma = 0$,
(vii) $[\lambda_{\hat{\imath}}(t_{\hat{\imath}})]_2 = [\lambda_{\hat{\jmath}}(t_{\hat{\jmath}})]_2 + [\lambda_{\hat{k}}(t_{\hat{k}})]_2$.

Proof. Due to continuity, the monotonicity property (i), and the initial condition (iii) there exists a continuous nondecreasing function $\tau : [0,\bar{\tau}] \to [0,1]$ defined implicitly via

$$[\lambda_1(t)]_1 + [\lambda_2(\tau(t))]_1 + [\lambda_3(\tau(t))]_1 + \sigma = 0,$$

where

$$\bar{\tau} = \max\{t \in [0,1] : [\lambda_1(t)]_1 + [\lambda_2(t')]_1 + [\lambda_3(t')]_1 + \sigma = 0 \text{ for some } t' \in [0,1]\}.$$

By definition, $(t'_1, t'_2, t'_3) = (\bar{\tau}, \tau(\bar{\tau}), \tau(\bar{\tau}))$ satisfies (vi) and by monotonicity at least one of t'_1, t'_2, t'_3 is equal to one. Then (v) and (ii) imply that there are pairwise distinct $i,j,k \in \{1,2,3\}$ with $[\lambda_i(t_i)]_2 \geq [\lambda_j(t_j)]_2 + [\lambda_k(t_k)]_2$. By the initial condition (iv) and the continuity of the λ_j and τ, there must be a smallest $t_1 \in (0,1]$ so that $t_2 = t_3 = \tau(t_1)$ satisfy (vi) and (vii). ■

Lemma 25 *Given the setting of Lemma 20 assume that* $\bar{\delta}_{\hat{\jmath}} \leq \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$ *holds for all* $\hat{\jmath} \in M$ *and let* $\bar{\jmath} \in M$ *be the only index with* $b_{\bar{\jmath}} = d_1$ *within the new embedding of Lemma 20. If at most* $|S| - 1$ *nodes of* S *are adjacent to nodes in* $C_{\bar{\jmath}}$*, then there is an optimal embedding of dimension at most* $|S|$.

Proof. Let v_i, $i \in N$ be the optimal embedding resulting from Lemma 20 with normalized vectors $d_1, d_2, d_3 \in \mathcal{L}^\perp$ satisfying $\dim \operatorname{span}\{d_1, d_2, d_3\} \leq 2$ and an assignment $\kappa : M \to \{1,2,3\}$ with $b_j = d_\kappa(j)$ for $j \in M$. Set $H_k = \bigcup_{j \in M : \kappa(j) = k} C_j$ for $k \in \{1,2,3\}$. Then,

$$v_i \in \mathcal{L} + \{\beta d_j : \beta \geq 0\} \quad \text{for } i \in H_j,\ j \in \{1,2,3\}. \tag{12}$$

Together with $\mathcal{L} = \operatorname{span} \mathcal{S}$ and $0 \in \mathcal{S} = \operatorname{conv}\{v_s : s \in S\}$ the dimension of this embedding is bounded by $\dim \mathcal{L} + \dim \operatorname{span}\{d_1, d_2, d_3\}$ and $\dim \mathcal{L} \leq |S| - 1$. If $\dim \mathcal{L} < |S| - 1$ or $\dim \operatorname{span}\{d_1, d_2, d_3\} < 2$ then the statement holds, so we may assume $\dim \mathcal{L} = |S| - 1$ and $\dim \operatorname{span}\{d_1, d_2, d_3\} = 2$. Next suppose there is a $j \in \{1,2,3\}$ with $v_i^T d_j = 0$ for all $i \in H_j$, w.l.o.g. assume this to hold for $j = 1$. Then the equilibrium constraint on $\mathcal{L}^\perp$ simplifies to $\sum_{i \in H_2} \|p_{\mathcal{L}^\perp}(v_i)\| d_2 = \sum_{i \in H_3} \|p_{\mathcal{L}^\perp}(v_i)\| d_3$. Thus, the embedding on $\mathcal{L}^\perp$ is restricted to a one dimensional subspace and the dimension of the embedding is again bounded by $|S|$. So it remains to consider the case

$$\text{for each } j \in \{1, 2, 3\}, \quad v_i^T d_j > 0 \text{ for some } i \in H_j. \tag{13}$$

By assumption there is a node $\hat{s} \in S$ not adjacent to any node in $H_1 = C_{\bar{j}}$. Put $S' = S \setminus \{\hat{s}\}$. This set S' separates H_1 from G. We have $0 \in \mathcal{S}' = \operatorname{conv}\{v_s : s \in S'\}$, because otherwise the Separator-Shadow Theorem 3 would imply $v_i \in \mathcal{L}' = \operatorname{span} \mathcal{S}'$ for $i \in H_1$, in contradiction to (13). Now $0 \in \mathcal{S}'$ yields $\dim \mathcal{L}' = |S'| - 1$ and as $\dim \mathcal{L} = |S| - 1$ we find a vector $\hat{g}$ with

$$0 \neq \hat{g} = \frac{p_{\mathcal{L}'^\perp}(v_{\hat{s}})}{\|p_{\mathcal{L}'^\perp}(v_{\hat{s}})\|} \in \mathcal{L} \cap \mathcal{L}'^\perp \quad \text{and} \quad \hat{g}^T v_s = 0 \text{ for } s \in S'. \tag{14}$$

Set $g_1 = -\hat{g}$ and $g_2 = g_3 = \hat{g}$, then by (12)

$$\text{for each } j \in \{1, 2, 3\}, \quad v_i \in \{x \in \mathcal{L}' + \operatorname{span}\{d_j, g_j\} : d_j^T x \geq 0\} \text{ for all } i \in H_j.$$

Therefore we may use Obs. 23 for $j \in \{1, 2, 3\}$ with $C = H_j$, $d = d_j$, $g = g_j$ to define transformations $\theta_i(t)$ for $i \in H_j$ and $\bar{\theta}_j(t) = \bar{\theta}_{H_j}(t)$. Observe that $\mathcal{S}' \subset \mathcal{L}'$ and $\mathcal{S} \subset \mathcal{L}' + \{\beta g_j : \beta \geq 0\}$ for $j \in \{2, 3\}$, so Obs. 23 (iv,v) establish that for $j \in \{1, 2, 3\}$ and $t_j \in [0, 1]$ the distance constraints of edges incident to nodes $i \in H_j$ remain satisfied for embedding $\theta_i(t_j)$ and the objective value remains unchanged due to Obs. 23 (iii) by $0 \in \mathcal{L}'$. Also note, that replacing d_j by some other normalized $d'_j \in \mathcal{L}^\perp$ will not affect distance constraints but only the equilibrium constraint. So it remains to find appropriate $t_j \in [0, 1]$ and normalized $d'_j \in \mathcal{L}^\perp$ so that the equilibrium constraint holds while the dimension of the embedding is reduced by at least one. For this purpose, define for $j \in \{1, 2, 3\}$ the function $\lambda_j : [0, 1] \to \mathbb{R}^2$ by

$$\lambda_j(t) = \begin{pmatrix} \hat{g}^T \bar{\theta}_j(t) \\ d_j^T \bar{\theta}_j(t) \end{pmatrix} \qquad \text{for } t \in [0, 1].$$

We show that the λ_j and $\sigma = \hat{g}^T v_{\hat{s}}$ satisfy the requirements of Obs. 24. Obs. 24(i) holds because of Obs. 23(vii) and (13). Obs. 24(ii) follows from Obs. 23(vi). Obs. 24(iii) is implied by the feasibility of equilibrium constraint on the linear subspace spanned by $\hat{g}$ for the embedding v_i, $i \in N$; for this, use Obs. 23(i), (14) and the definition of σ. Suppose Obs. 24(iv) does not hold and assume, w.l.o.g., that $\lambda_1(0) \geq \lambda_2(0) + \lambda_3(0)$, then by (12) and Obs. 23(i) this is equivalent to

$$\sum_{i \in H_1} \|p_{\mathcal{L}^\perp}(v_i)\| \geq \sum_{i \in H_2 \cup H_3} \|p_{\mathcal{L}^\perp}(v_i)\|$$

and together with the equilibrium constraint

$$\sum_{i \in H_1} \|p_{\mathcal{L}^\perp}(v_i)\| d_1 + \sum_{i \in H_2} \|p_{\mathcal{L}^\perp}(v_i)\| d_2 + \sum_{i \in H_2} \|p_{\mathcal{L}^\perp}(v_i)\| d_3 = 0$$

this implies $d_1 = -d_2 = -d_3$ in contradiction to $\dim \operatorname{span}\{d_1, d_2, d_3\} = 2$. Thus, Obs. 24(iv) holds. Finally, Obs. 24(v) follows from Obs. 23(viii). Hence, there exist $t_1, t_2, t_3 \in [0, 1]$ and pairwise distinct $\hat{\imath}, \hat{\jmath}, \hat{k} \in \{1, 2, 3\}$ so that Obs. 24(vi,vii) hold. Now,

$$\text{choose } \hat{d} \in \mathcal{L}^\perp, \|\hat{d}\| = 1, \quad \text{set } d'_{\hat{\imath}} = -d'_{\hat{\jmath}} = -d'_{\hat{k}} = \hat{d} \tag{15}$$

and

$$v'_i = \begin{cases} v_i & i \in S, \\ p_{\mathcal{L}'}(v_i) + \delta_i[g_j \cos(\gamma_i - t_j\gamma_i) + d'_j \sin(\gamma_i - t_j\gamma_i)] & i \in H_j, j \in \{1, 2, 3\}. \end{cases}$$

Since only the d_j have been replaced by d'_j, $j \in \{1, 2, 3\}$, the distance constraints are still valid for the new embedding v'_i, $i \in N$, and the objective value is unchanged. Furthermore, setting

$$\bar{\theta}'_j(t) = \sum_{i \in H_j} p_{\mathcal{L}'}(v_i) + \delta_i[g_j \cos(\gamma_i - t\gamma_i) + d'_j \sin(\gamma_i - t\gamma_i)] \quad \text{for } j \in \{1, 2, 3\},$$

we see that the functions λ_j, $j \in \{1, 2, 3\}$, also satisfy

$$\lambda_j(t) = \begin{pmatrix} \hat{g}^T \bar{\theta}'_j(t) \\ d'^T_j \bar{\theta}'_j(t) \end{pmatrix} \qquad \text{for } t \in [0, 1].$$

Therefore Obs. 24(vi,vii) still hold for t_1, t_2, t_3 and $\hat{\imath}, \hat{\jmath}, \hat{k}$ yielding

$$\begin{aligned} 0 &= \sigma + \hat{g}^T(\bar{\theta}'_1(t_1) + \bar{\theta}'_2(t_2) + \bar{\theta}'_2(t_2)) = \hat{g}^T(v_{\hat{s}} + \sum_{j \in \{1,2,3\}} \sum_{i \in H_j} v'_i) \overset{(14)}{=} \hat{g}^T(\sum_{i \in N} v'_i) \\ 0 &= d'^T_{\hat{\imath}} \bar{\theta}'_{\hat{\imath}}(t_{\hat{\imath}}) - d'^T_{\hat{\jmath}} \bar{\theta}'_{\hat{\jmath}}(t_{\hat{\jmath}}) - d'^T_{\hat{k}} \bar{\theta}'_{\hat{k}}(t_{\hat{k}}) \overset{(15)}{=} \hat{d}^T \sum_{j \in \{1,2,3\}} \sum_{i \in H_j} v'_i \overset{\hat{d} \in \mathcal{L}^\perp}{=} \hat{d}^T(\sum_{i \in N} v'_i) \end{aligned}$$

So the equilibrium constraint holds on the linear subspaces spanned by $\hat{g}$ and $\hat{d}$. It also holds on $\mathcal{L}'$ because $p_{\mathcal{L}'}(v_i) = p_{\mathcal{L}'}(v'_i)$ for $i \in N$ and the embedding v_i was feasible. Since $v'_i \in \mathcal{L}' + \operatorname{span}\{\hat{g}, \hat{d}\} = \mathcal{L} + \operatorname{span}\{\hat{d}\}$ for $i \in N$, the new embedding satisfies the equilibrium constraint on the entire space. Therefore it is an optimal embedding of dimension at most $\dim \mathcal{L} + 1 = |S|$. ■

5 The proof of the Tree-Width Theorem 5

We will show that for any tree-decomposition $T = (\mathcal{N}, \mathcal{E})$ of G (see Def. 4) there is always an optimal embedding of dimension at most $\max\{|U| : U \in \mathcal{N}\}$. As this also holds for a tree-decomposition giving the tree-width of G, this will prove the theorem.

Note that in a tree decomposition any $U \in \mathcal{N}$ and any $U \cap U'$ with $\{U, U'\} \in \mathcal{E}$ is a separator of G (see e.g. Lemma 12.3.1 in [5]). In the proof we will show that for any optimal embedding v_i, $i \in N$, we can find a separator S of the form $U \in \mathcal{N}$ or $\{U, U'\} \in \mathcal{E}$ with $0 \in \text{conv}\{v_s : s \in S\}$ so that either Lemma 20 or Lemma 25 yield an optimal embedding of dimension at most $|S|$.

The first step asserts, that for any optimal embedding any tree-decomposition has "zero-nodes" containing the origin in their convex hull.

Lemma 26 *Given a tree decomposition $T = (\mathcal{N}, \mathcal{E})$ of a connected graph $G = (N, E)$ and an optimal embedding $v_i \in \mathbb{R}^n$ $(i \in N)$ of* (4)*. There is an $S \in \mathcal{N}$ with $0 \in \text{conv}\{v_s : s \in S\}$.*

Proof. Consider a subtree $T' = (\mathcal{N}', \mathcal{E}')$ of T with $|\mathcal{N}'|$ minimal so that $0 \in \text{conv}\{v_i : i \in \bigcup_{U \in \mathcal{N}'} U\}$. Such a tree exists since the condition holds for $T' = T$ by the equilibrium constraint. Let the convex combination giving the origin be described by $C = \bigcup_{U \in \mathcal{N}'} U$ and $\alpha \in \mathbb{R}^C_+$ with $\alpha^T e = 1$ so that $\sum_{i \in C} \alpha_i v_i = 0$.

Assume, for contradiction, that $|\mathcal{N}'| > 1$. Then there is an edge $\{U, U'\} \in \mathcal{E}'$ and $S' = U \cap U'$ is a separator of G. Deleting edge $\{U, U'\}$ from T' splits T' into two nonempty subtrees $T'_j = (\mathcal{N}'_j, \mathcal{E}'_j)$ for $j \in \{1, 2\}$ with $0 \notin \text{conv}\{v_i : i \in \bigcup_{U \in \mathcal{N}'_j} U\}$ by assumption. Set $N'_j = \bigcup_{U \in \mathcal{N}'_j} U$. Because $S' \subseteq N'_j$ for $j \in \{1, 2\}$ we obtain $0 \notin \mathcal{S}' = \text{conv}\{v_i : i \in S'\}$. Thus, the Separator-Shadow Th. 3 applied to S' implies that, w.l.o.g., for all $i \in N'_1$, $\text{conv}\{v_i, 0\} \cap \mathcal{S}' \neq \emptyset$. But then the origin must be contained in the convex hull of subtree T'_2 as we show next. Put $C_1 = N'_1 - S'$, $C_2 = N'_2$ and set, for $j \in \{1, 2\}$, $\bar{\alpha}_j = \sum_{i \in C_j} \alpha_i$ and $\bar{v}_j = \frac{1}{\bar{\alpha}_j} \sum_{i \in C_j} \alpha_i v_i \in \text{conv}\{v_i : i \in N'_j\}$. Then $0 = \bar{\alpha}_1 \bar{v}_1 + \bar{\alpha}_2 \bar{v}_2 \in \text{conv}\{\bar{v}_1, \bar{v}_2\}$ (by definition of the α_i) and $\emptyset \neq \mathcal{S}' \cap \text{conv}\{\bar{v}_1, 0\} \subset \text{conv}\{\bar{v}_1, \bar{v}_2\}$ (by the separator-shadow property), so there is a $p \in \mathcal{S}' \subset \text{conv}\{v_i : i \in N'_2\}$ with $0 \in \text{conv}\{p, \bar{v}_2\} \subset \text{conv}\{v_i : i \in N'_2\}$, a contradiction to the minimality of $|\mathcal{N}'|$. Hence, T' consists of only one node. ■

We will call a node $U \in \mathcal{N}$ a *zero-node* (with respect to the embedding v_i, $i \in N$) if $0 \in \text{conv}\{v_i : i \in U\}$ and an edge $\{U, U'\} \in \mathcal{E}$ a *zero-edge* (with respect to the embedding v_i, $i \in N$) if $0 \in \text{conv}\{v_i : i \in U \cap U'\}$. Note, for a zero-edge both endpoints are zero-nodes.

Observation 27 *The subgraph $T' = (\mathcal{N}', \mathcal{E}')$ of $T = (\mathcal{N}, \mathcal{E})$ induced by the zero-nodes of an optimal embedding v_i $(i \in N)$ of* (4) *is a tree and $\mathcal{E}'$ is the set of zero-edges.*

Proof. Suppose that there are two zero-nodes U and U' that are not connected in T' or that they are connected in T' by an edge that is not a zero-edge. In both cases there is an edge $\{S, S'\} \in \mathcal{E}$ with $0 \notin \mathcal{S} = \text{conv}\{v_i : i \in S \cap S'\}$ on the path connecting U and U' in T. But then the Separator-Shadow Th. 3 implies, w.l.o.g., that $\text{conv}\{v_i, 0\} \cap \mathcal{S} \neq \emptyset$ for all $i \in U$. This can be worked out to contradict the assumption that U is a zero-node. ■

Hence, for a given tree-decomposition any optimal embedding induces a *zero-tree* (with respect to the embedding v_i, $i \in N$) consisting of the zero-nodes and zero-edges.

The algorithmic idea is to pick a zero-node U, transform the embedding for $S = U$ as suggested in lemmas 20, 21, 25 and to check whether the resulting dimension is at most $|U|$. If it is not, it will turn out that in the zero-tree of the new optimal embedding, U has a unique incident zero-edge $\{U, U'\}$ leading to the that part of the graph whose embedding cannot yet be flattened out sufficiently with respect to U. We then go on transforming the new optimal embedding with respect to the separator $U \cap U'$ which may again lead to a sufficiently flat optimal embedding or, in failing to find one, lead on to U' via the the part that is not flat enough. Now, at some point this algorithm might turn back in U' and try to cross this last edge a second time. Happily, this will immediately allow to produce an optimal embedding that is sufficiently flat. As going on in one direction will only be possible for a finite number of times, this will complete the proof.

We start with the convenient case, where all parts can be flattened out sufficiently.

Lemma 28 *Given a tree decomposition $T = (\mathcal{N}, \mathcal{E})$ of a connected graph $G = (N, E)$, an optimal embedding $v_i \in \mathbb{R}^n$ $(i \in N)$ of (4), and a zero-node $S \in \mathcal{N}$ whose deletion splits T into m subtrees $T_j = (\mathcal{N}_j, \mathcal{E}_j)$ $(j \in M = \{1, \dots, m\})$. Put*

$$\begin{aligned} \mathcal{L} &= \operatorname{span}\{v_s : s \in S\}, \\ C_j &= \bigcup_{U \in \mathcal{N}_j} U \setminus S, \\ \bar{\delta}_j &= \sum_{i \in C_j} \|p_{\mathcal{L}}(v_i)\| \quad (j \in M). \end{aligned}$$

If $\bar{\delta}_{\hat{\jmath}} \le \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$ for $\hat{\jmath} \in M$ then there is an optimal embedding v_i' $(i \in N)$ of dimension at most $|U'|$ for some $U' \in \{S, U : \{S, U\} \in \mathcal{E}\}$.

Proof. We distinguish two cases. In the first case assume that S has a neighbor U' in T with $|U'| > |S|$ and apply Lemma 20 with respect to S and the C_j $(j \in M)$. The resulting optimal embedding v_i' has dimension at most $\dim \mathcal{L} + 2$ and since $\dim \mathcal{L} \le |S| - 1$ $(0 \in \operatorname{conv}\{v_s : s \in S\})$ the dimension is at most $|U'|$.

In the second case all neighbors U of S in T satisfy $|U| \le |S|$. By definition, no two nodes in $\mathcal{N}$ are identical, so each set C_j is separated from S by a subset $S_j = S \cap U_j$ induced by an edge $\{S, U_j\} \in \mathcal{E}$ with $|S_j| < |S|$. Therefore we may apply Lemma 25 with respect to S and the C_j $(j \in M)$ and obtain an optimal embedding of dimension at most $|S|$. ■

If, however, one of the sets is too big to be flattened out, we can find a unique edge that leads us towards a more balanced center in the big set.

Lemma 29 *Given the setting of Lemma 28, assume that $\bar{\delta}_{\hat{\jmath}} > \sum_{j \in M \setminus \{\hat{\jmath}\}} \bar{\delta}_j$ for a $\hat{\jmath} \in M$. Let v_i' $(i \in N)$ be an optimal embedding arising from Lemma 21 for this S and the C_j $(j \in M)$. The (unique) edge $\{S, \widehat{U}\} \in \mathcal{E}$ with $\widehat{U} \in \mathcal{N}_{\hat{\jmath}}$ is a zero-edge with respect to this new optimal embedding.*

Proof. Since $\bar{\delta}_{\hat{\jmath}} > 0$ neither the subtree $T_{\hat{\jmath}}$ nor $C_{\hat{\jmath}}$ are empty, so there is an edge $\{S, \widehat{U}\} \in \mathcal{E}$ with $\widehat{U} \in \mathcal{N}_{\hat{\jmath}}$. Suppose, for contradiction, that it is not a zero edge with respect to the embedding v_i' $(i \in N)$. Then $S' = S \cap \widehat{U}$ separates G into $C_{\hat{\jmath}}$ and $N \setminus (S' \cup C_{\hat{\jmath}})$. By assumption, $0 \notin \operatorname{conv}\{v_s : s \in S'\}$ and $0 \in \operatorname{conv}\{v_s : s \in S\}$, so the Separator-Shadow

Th. 3 applied with respect to the separator S' implies that $v_i \in \text{cone}\{v_s : s \in S'\} \subset \mathcal{L}$ for $i \in C_{\hat{\jmath}}$. But then $\bar{\delta}_{\hat{\jmath}} = 0$. ■

Note that $\widehat{U}$ is a zero-node of the embedding v_i' and we could continue with transforming v_i' with respect to $\widehat{U}$ ending up in Lemma 28 or Lemma 29 again. However, in order to ensure that no edge is crossed twice, we need to look at the zero-edge itself first.

Lemma 30 *Given the setting of Lemma 29 with $\{S, \widehat{U}\} \in \mathcal{E}$ being the zero-edge with respect to embedding v_i' $(i \in N)$ satisfying $\widehat{U} \in \mathcal{N}_{\hat{\jmath}}$. Deleting this edge in T splits T into two subtrees $T_j' = (\mathcal{N}_j', \mathcal{E}_j')$ with $j \in M' = \{S, \widehat{U}\}$ so that $S \in \mathcal{N}_S'$ and $\widehat{U} \in \mathcal{N}_{\widehat{U}}'$. Put*

$$\begin{aligned} S' &= S \cap \widehat{U}, \\ \mathcal{L}' &= \text{span}\{v_s : s \in S'\}, \\ C_j' &= \bigcup_{U \in \mathcal{N}_j'} U \setminus S', \\ \bar{\delta}_j' &= \sum_{i \in C_j} \|p_{\mathcal{L}'}(v_i')\| \quad (j \in M'). \end{aligned}$$

If $\bar{\delta}_S' \geq \bar{\delta}_{\widehat{U}}'$ then there is an optimal embedding v_i'' $(i \in N)$ of dimension at most $|S|$.

Proof. If $\bar{\delta}_S' = \bar{\delta}_{\widehat{U}}'$ then Lemma 20 applied to embedding v_i' with respect to S' and C_j' for $j \in M'$ yields an optimal embedding

$$v_i'' = \begin{cases} v_i' & \text{for } i \in S', \\ p_{\mathcal{L}'}(v_i) + \|p_{\mathcal{L}'^\perp}(v_i)\| b & \text{for } i \in C_S' \\ p_{\mathcal{L}'}(v_i) - \|p_{\mathcal{L}'^\perp}(v_i)\| b & \text{for } i \in C_{\widehat{U}}'. \end{cases}$$

for some normalized $b \in \mathcal{L}'^\perp$ and the dimension is bounded by $\dim \mathcal{L}' + 1 \leq |S'| \leq |S|$.

For $\bar{\delta}_S > \bar{\delta}_{\widehat{U}}$ remember that the v_i' were constructed via Lemma 21. So with the definitions of $\bar{b}$ and the v_i' given there, we have

$$v_i' = p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\| \bar{b} \quad \text{for } i \in C_S' = \bigcup_{j \in M \setminus \{\hat{\jmath}\}} C_j \cup S \setminus S'.$$

If $\bar{b} = 0$ then all these v_i' lie in $\mathcal{L}$ and by applying Lemma 21 to v_i' with respect to S' and C_j' for $j \in M'$, the space of the new optimal embedding v_i'', $i \in N$, will be $\mathcal{L}$ enlarged by some direction h at most, so its dimension is bounded by $\dim \mathcal{L} + 1 \leq |S|$.

If $\bar{b} \neq 0$, then $p_{\mathcal{L}^\perp}(v_i) \neq 0$ for some $i \in C_S'$ and using $\bar{b} \in \mathcal{L}^\perp$, $\|\bar{b}\| = 1$, we get

$$\bar{b}^T v_i' = \bar{b}^T p_{\mathcal{L}}(v_i) + \|p_{\mathcal{L}^\perp}(v_i)\| \bar{b}^T \bar{b} = \|p_{\mathcal{L}^\perp}(v_i)\| \geq 0 \quad \text{for } i \in C_S'.$$

Since $\mathcal{L}' \subseteq \mathcal{L}$ we obtain $\bar{b} \in \text{span}\{v_i' : i \in C_S'\} \cap \mathcal{L}'^\perp \setminus \{0\}$ and $\bar{b}^T v_i' \geq 0$ for $i \in C_S'$. So we are in the special case of Lemma 21. Thus, applying Lemma 21 to v_i' with respect to S' and C_j' for $j \in M'$ yields a new optimal embedding v_i'', $i \in N$, with $v_i'' \in \text{span}\{v_i' :$

$i \in C'_S\} \subseteq \mathcal{L} + \operatorname{span}\{\bar{b}\}$ for $i \in C'_S$ and therefore $v''_i \in \mathcal{L} + \operatorname{span}\{\bar{b}\}$ for all $i \in N$. The dimension of this new embedding is again bounded by $|S|$. ■

In proving finiteness of the algorithm below we will see that the values $\bar{\delta}'_j$ of Lemma 30 do not change if the algorithm turns back in $\hat{U}$ to cross the same edge again, so that the condition $\bar{\delta}'_S \geq \bar{\delta}'_{\hat{U}}$ will be met the second time at the latest.

Algorithm 31
Input: *a connected graph* $G = (V, E)$*, a tree decomposition* $T = (\mathcal{N}, \mathcal{E})$ *of* G*, an optimal embedding* v_i*,* $i \in N$*, of* (4)*.*
Step 0: *Set* S *to a zero-vertex of* T *with respect to the embedding.*
Step 1: *Using the notation of Lemma 28 with respect to* S*, determine* $\bar{\delta}_j$ *for* $j \in M$*.*
Step 2: *If* $\bar{\delta}_{\hat{\jmath}} \leq \sum_{j \in M\setminus\{\hat{\jmath}\}} \bar{\delta}_j$ *for* $\hat{\jmath} \in M$ *apply the proof of Lemma 28 to find an optimal embedding of dimension at most the width of* T *plus one and stop.*
Step 3: *Transform, as described in Lemma 29, the optimal embedding to* v'_i $(i \in N)$ *and compute the corresponding zero-edge* $\{S, \widehat{U}\}$*. Determine* $\bar{\delta}'_S$ *and* $\bar{\delta}'_{\widehat{U}}$ *in the notation of Lemma 30.*
Step 4: *If* $\bar{\delta}_S \geq \bar{\delta}_{\widehat{U}}$ *apply the proof of Lemma 30 to find an optimal embedding of dimension at most the width of* T *plus one and stop.*
Step 5: *Set* $S \leftarrow \widehat{U}$*,* $v_i \leftarrow v'_i$ *for* $i \in N$ *and goto Step 1.*

Theorem 32 *Let* $G = (N, E)$ *be a connected graph and* $T = (\mathcal{N}, \mathcal{E})$ *a tree decomposition of* G*. Algorithm 31 is correct and stops with an optimal embedding for* (4) *of dimension at most width of* T *plus one in at most* $|\mathcal{N}|$ *iterations.*

Proof. Step 0 can be carried through by Lemma 26.

If in Step 1 the set M is empty ($\mathcal{N} = \{S\}$), then the condition in Step 2 is vacuously satisfied and the transformation of Lemma 28 is the identity. But in this case $S = |N|$ and any optimal embedding has dimension at most $|N| - 1$ by Obs. 9, so the algorithm stops correctly.

We will prove that the algorithm steps over any edge at most once without stopping and this will yield the iteration bound. Suppose $\{S, \widehat{U}\}$ is the first edge of T to be considered a second time and that the algorithm just stepped from S to $\widehat{U}$ and now considers stepping back to S. Then $\widehat{U}$ transforms, by means of Lemma 29 the embedding v'_i that was generated by S via Lemma 29. By construction (see Lemma 21), both transformations have $\mathcal{L}' = \operatorname{span}\{v'_i : i \in S \cap \widehat{U}\}$ as an invariant subspace. Therefore the numbers $\bar{\delta}'_S$ and $\bar{\delta}'_{\widehat{U}}$ of Lemma 30 computed in Step 3 have identical values in both cases (but with names interchanged), so the condition of Step 4 is certainly satisfied the second time and the algorithm stops.

The correctness of the statement regarding the dimension of the optimal embedding at termination is a consequence of the respective lemmas 28 and 30. ■

Acknowledgements. We thank two anonymous referees for their careful and considerate reports that helped to restructure and improve the presentation of the paper significantly.

References

[1] K. L. Boyer and K. Sengupta. Modelbase partitioning using property matrix spectra. *Computer Vision and Image Understanding*, 70(2):177–196, May 1998.

[2] T. F. Chan, P. Ciarlet, and W. K. Szeto. On the optimality of the median cut spectral bisection graph partitioning method. *SIAM J. Sci. Comput.*, 18(3):943–948, May 1997.

[3] D. Cvetković, M. Doob, and H. Sachs. *Spectra of Graphs. Theory and application.* J. A. Barth Verlag, Leipzig, 3rd edition, 1995.

[4] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Operations Research Letters*, 20:213–221, 1997.

[5] R. Diestel. *Graph Theory.* Springer, Berlin, 2nd edition, 2000.

[6] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23:298–305, 1973.

[7] M. Fiedler. Laplacian of graphs and algebraic connectivity. *Combinatorics and Graph Theory*, 25:57–70, 1989.

[8] M. Fiedler. Absolute algebraic connectivity of trees. *Linear and Multilinear Algebra*, 26:85–106, 1990.

[9] M. Fiedler. A geometric approach to the Laplacian matrix of a graph. *IMA Vol. Math. Appl.*, 50:73–98, 1993.

[10] M. Fiedler. Some minimax problems for graphs. *Discrete Mathematics*, 121:65–74, 1993.

[11] C. Helmberg, B. Mohar, S. Poljak, and F. Rendl. A spectral approach to bandwidth and separator problems in graphs. *Linear and Multilinear Algebra*, 39:73–90, 1995.

[12] X. Ji and H. Zha. Extracting shared topics of multiple documents. *Lect. Notes Comput. Sci.*, 2637:100–110, 2003.

[13] M. Juvan and B. Mohar. Laplace eigenvalues and bandwidth-type invariants of graphs. *J. Graph Theory*, 17(3):393–407, 1993.

[14] A. Kaveh and H. A. Rahimi Bondarabady. Bisection for parallel computing using Ritz and Fiedler vectors. *Acta Mechanica*, 167:131–144, Feb. 2004.

[15] J. Keuchel, C. Schnörr, C. Schellewald, and D. Cremers. Binary partitioning, perceptual grouping, and restoration with semidefinite programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1364–1379, Nov. 2003.

[16] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, IT-25(1):1–7, Jan. 1979.

[17] B. Mohar. The Laplacian spectrum of graphs. In *Graph Theory, Combinatorics, and Applications*, pages 871–898. John Wiley and Sons, 1991.

[18] B. Mohar. Graph Laplacians. In B. et al., editor, *Topics in algebraic graph theory*, volume 102 of *Encyclopedia of Mathematics and Its Applications*, pages 113–136. Cambridge University Press, 2004.

[19] B. Mohar and S. Poljak. Eigenvalues in combinatorial optimization. In Brualdi, Friedland, and Klee, editors, *Combinatorial and Graph Theoretic Problems in Linear Algebra*. IMA Volume in Mathematics and Applications, Vol. 50, 1993.

[20] A. Pothen, H. D. Simon, and K. Liou. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.*, 11:430–452, 1990.

[21] H. D. Simon, A. Sohn, and R. Biswas. Harp: A dynamic spectral partitioner. *Journal of Parallel and Distributed Computing*, 50:83–103, 1998.

[22] A. M.-C. So, Y. Ye, and J. Zhang. A unified theorem on SDP rank reduction. Technical report, Department of Management Science and Engineering, Stanford University, 2006.

[23] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones (updated for version 1.05). Manual, Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, Oct. 2001.

[24] J. Sun, S. Boyd, L. Xiao, and P. Diaconis. The fastest mixing markov process on a graph and a connection to a maximum variance unfolding problem. *SIAM Review*, 48(4):681–699, 2006.

[25] H. van der Holst, L. Lovász, and A. Schrijver. The Colin de Verdière graph parameter. In Lovász et al., editor, *Graph theory and combinatorial biology*, volume 7 of *Bolyai Soc. Math. Stud.*, pages 29–85. János Bolyai Mathematical Society, Budapest, 1999.

[26] K. Q. Weinberger and L. K. Saul. Unsupervised learning of image manifolds by semidefinite programming. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-04)*, volume 2, pages 988–995, Washington D.C., 2004.

[27] K. Q. Weinberger, F. Sha, and L. K. Saul. Learning a kernel matrix for nonlinear dimensionality reduction. In *Proceedings of the Twenty First International Conference on Machine Learning (ICML-04)*, pages 839–846, Banff, Canada, 2004.

[28] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming*, volume 27 of *International Series in Operations Research and Management Science*. Kluwer Academic Publishers, Boston/Dordrecht/London, 2000.

[29] C. W. Wu. On Rayleigh-Ritz ratios of a generalized Laplacian matrix of directed graphs. *Linear Algebra Appl.*, 402:207–227, 2005.