Mathematical Programming 79 (1997) 255–283

Cuts, matrix completions and graph rigidity

Monique Laurent [a,b]

[a] LIENS - Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05, France
[b] CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

Received 6 March 1997; accepted 4 April 1997

Abstract

This paper brings together several topics arising in distinct areas: polyhedral combinatorics, in particular, cut and metric polyhedra; matrix theory and semidefinite programming, in particular, completion problems for positive semidefinite matrices and Euclidean distance matrices; distance geometry and structural topology, in particular, graph realization and rigidity problems.

Cuts and metrics provide the unifying theme. Indeed, cuts can be encoded as positive semidefinite matrices (this fact underlies the approximative algorithm for max-cut of Goemans and Williamson) and both positive semidefinite and Euclidean distance matrices yield points of the cut polytope or cone, after applying the functions $1/\pi \arccos(.)$ or $\sqrt{}$. When fixing the dimension in the Euclidean distance matrix completion problem, we find the graph realization problem and the related question of unicity of realization, which leads to the question of graph rigidity.

Our main objective here is to present in a unified setting a number of results and questions concerning matrix completion, graph realization and rigidity problems. These problems contain indeed very interesting questions relevant to mathematical programming and we believe that research in this area could yield to cross-fertilization between the various fields involved. © 1997 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

Keywords: Cut; Positive semidefinite matrix; Euclidean distance matrix; Matrix completion; Metric polyhedron; Graph rigidity

1. Introduction

Cuts in graphs are very elementary objects in graph theory and combinatorics. Yet they yield a number of interesting and difficult combinatorial optimization problems. For instance, the max-cut problem belongs to the most basic combinatorial problems, whose NP-hardness was already recognized in the fundamental paper of Karp [30]. This optimization problem has received a lot of attention throughout the years. It has been

Published by Elsevier Science B.V.
PII S0025-5610(97)00043-9

attacked from various angles ranging from enumeration methods, integer linear programming, and, more recently, continuous optimization and more specifically semidefinite programming. The latter approach is based essentially on the observation that every cut can be represented by a matrix which is positive semidefinite with an all ones diagonal. By this, the max-cut problem can be naturally relaxed by optimizing over the set of positive semidefinite matrices with an all-ones diagonal. A crucial result established by Goemans and Williamson [18] is that this relaxation yields an efficiently computable and very good approximation for the max-cut problem.

This fact, that cuts can be viewed as matrices with special properties, sets a natural bridge with linear algebra and, more specifically, combinatorial matrix theory. A problem which has occupied researchers in this area for quite some time is the so-called *completion problem* for various matrix properties. This problem asks whether the unspecified entries of a partially defined matrix can be completed so as to obtain a fully defined conventional matrix satisfying a desired property. In our discussion, the matrix property in question is positive semidefiniteness.

Results concerning the positive semidefinite completion problem have been obtained in the literature of linear algebra, that involve notions that are well known to researchers in the area of mathematical programming; in particular, the notions of cuts and metrics. Moreover, finding a positive semidefinite completion for a partial matrix is a typical instance of a semidefinite programming problem.

One of our objectives in this paper is to present some of the results about the positive semidefinite completion problem that are most relevant to mathematical programming techniques and notions. We believe indeed that this combinatorial matrix theory problem contains interesting combinatorial and optimization problems yielding to cross-fertilization between the two fields.

Cuts, metrics and positive semidefinite matrices are also closely related to the so-called Euclidean distance matrices; that is, the matrices whose entries can be realized as the pairwise squared Euclidean distances among a set of points (in arbitrary dimension). Euclidean distance matrices are a central notion in the area of distance geometry; their study was initiated by Cayley last century and continued in the 1930s, in particular, by Menger and Schoenberg (cf. the classic book by Blumenthal [8], and the monograph by Crippen and Havel [11] treating also applications). We mention here several results concerning the completion problem for Euclidean distance matrices.

If one fixes the dimension of the space in which the points realizing the (partial) matrix are to be found, one finds the graph realization problem, a well-studied problem in distance geometry, having important applications, in particular, to the area of molecular chemistry. This is the following problem: Given a graph $G = (V, E)$ with weights $d \in \mathbb{R}_+^E$ on its edges and a prescribed dimension $k \geqslant 1$, is it possible to find points $p_i \in \mathbb{R}^k$ $(i \in V)$ such that the square of the Euclidean distance between p_i and p_j is equal to the prescribed weight d_{ij} for every edge $ij \in E$? This problem turns out to be NP-complete in any fixed dimension $k \geqslant 1$. In contrast, the complexity of the same problem with unprescribed dimension is not known! It is not even known if the problem belongs to NP. Yet, the latter problem becomes polynomial-time solvable if one allows approximations (since it can be formulated as a semidefinite programming problem). Moreover, its exact version is polynomial-time solvable for some classes of graphs, such as chordal graphs.

Beside the problem mentioned above of the existence of a realization in the k-space for a weighted graph, the problem of unicity of such a realization arises naturally (unicity up to congruence). This is also an NP-hard problem as shown by Saxe [47]. However, if one restrict oneself to searching for generic realizations (i.e., realizations in which the coordinates of the points are algebraically independent over the rational field $\mathbb{Q}$), then the problem becomes tractable at least in small dimension $k \leqslant 2$. In this context we find the problem of characterizing generic rigid graphs, a well-studied problem in the area of structural topology.

The paper is organized as follows. Section 2 contains definitions and preliminaries on graphs and matrices and, in Section 3, we recall some basic links between cuts and positive semidefinite matrices and Goemans-Williamson's result with some applications. We treat in Sections 4 and 5 the completion problem for positive semidefinite matrices and for Euclidean distance matrices. These sections are organized as follows: in Sections 4.1 and 5.1 we expose necessary conditions for the existence of a completion and a characterization of the graphs for which these conditions are also sufficient, and the Sections 4.2 and 5.2 contain complexity results for these completion problems. When fixing the dimension in the Euclidean distance matrix completion problem, one finds the graph realization problem and the related question of unicity of realization; this leads to the question of characterizing rigid graphs, considered in Section 6.

2. Preliminaries

2.1. Graphs

All graphs are assumed here to be simple (i.e., without loops and parallel edges). We set $V_n := \{1, \ldots, n\}$ and $E_n := \{ij \mid 1 \leqslant i < j \leqslant n\}$. Hence, $K_n = (V_n, E_n)$ is the complete graph on n nodes. Given a graph $G = (V_n, E)$, where $E \subseteq E_n$, its *suspension graph* ∇G is defined as the graph with node set $V_{n+1} := V_n \cup \{n+1\}$ and with edge set $E(\nabla G) := E \cup \{(i, n+1) \mid i \in V_n\}$.

Let C be a circuit on $n-1$ nodes; then, $W_n := \nabla C$ denotes the *wheel* on n nodes, obtained by adding a new node (the *center* of the wheel) adjacent to all nodes on the circuit C. (Cf. Fig. 1(a) for a picture of the wheel W_7.)

Let $G = (V_n, E)$ be a graph. Given a subset $U \subseteq V_n$, $G[U]$ denotes the subgraph of G *induced* by U, with node set U and with edge set $\{uv \in E \mid u, v \in U\}$. One says that U is a *clique* in G when $G[U]$ is a complete graph. For a subset $S \subseteq V_n$ the *cut* $\delta(S)$ consists of the edges of G having one end node in S and the other one in $V_n \setminus S$.

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs such that the set $K := V_1 \cap V_2$ induces a clique (possibly empty) in both G_1 and G_2 and there is no edge between a node of $V_1 \setminus K$ and a node of $V_2 \setminus K$. Then, the graph $G := (V_1 \cup V_2, E_1 \cup E_2)$ is called

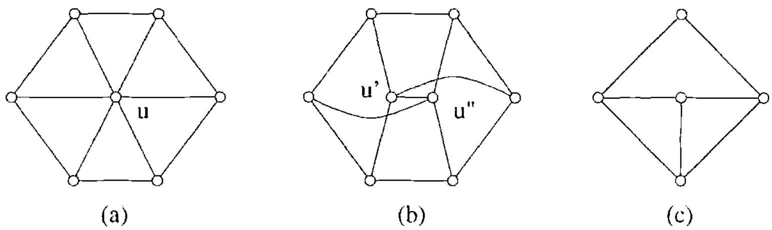

Fig. 1. (a) The wheel W_7. (b) Splitting node u in W_7. (c) The graph $\widehat{W_4}$.

the *clique sum* of G_1 and G_2. One also says that G is their *k-clique sum* if $k = |K|$. Call a graph *prime* if it cannot be decomposed as a clique sum of two (smaller) graphs. Tarjan [50] proposes an algorithm for decomposing a graph into prime pieces by means of clique sums, that runs in time $O(nm)$ for a graph with n nodes and m edges.

A graph H is said to be a *minor* of a graph G if H can be obtained from G by repeatedly deleting and/or contracting edges and deleting isolated nodes. *Deleting* an edge e in G simply means discarding it from the edge set of G. *Contracting* edge $e = uv$ means identifying both end nodes of e and discarding multiple edges and loops if some are created during the identification of u and v.

Call *splitting* the operation converse to that of contracting an edge. Hence, splitting a node u (of degree $\geqslant 2$) in a graph means replacing u by two adjacent nodes u' and u'' and replacing every edge uv in an arbitrary manner, either by $u'v$, or by $u''v$ (but in such a way that each of u' and u'' is adjacent to at least one node). Figs. 1(a) and (b) show the wheel on 7 nodes and a splitting of it, while (c) shows the graph $\widehat{W_4}$ obtained by splitting one node in $W_4 = K_4$.

Subdividing an edge $e = uv$ means inserting a new node w and replacing edge e by the two edges uw and wv. Hence, this is a special case of splitting. A graph that can be constructed from a given graph G by subdividing its edges is called a *homeomorph* of G. Note that splitting a node of degree 2 or 3 amounts to subdividing one of the edges incident to that node and, thus, a graph has no K_4-minor if and only if it contains no homeomorph of K_4 as a subgraph. (Such graph is also called a (simple) *series-parallel* graph.)

Finally, we introduce some classes of graphs that will play an important role in this paper. A graph G is said to be *chordal* if every circuit of G with length $\geqslant 4$ has a chord; a *chord* of a circuit C is an edge joining two nonconsecutive nodes of C. We also consider the class $\mathcal{G}_{\mathrm{wh}}$ which consists of the graphs that do not contain a wheel W_n ($n \geqslant 5$) or a splitting of a wheel W_n ($n \geqslant 4$) as an induced subgraph.

Chordal graphs, graphs with no K_4-minor, and graphs in the class $\mathcal{G}_{\mathrm{wh}}$ have a relatively simple structure as they can be decomposed by means of clique sums into "easy" pieces. Indeed, a graph is chordal if and only if it can be decomposed by means of clique sums into cliques [13]; and a graph G has no K_4-minor if and only if $G = K_3$, or G is a subgraph of a clique k-sum ($k = 0, 1, 2$) of two smaller graphs (i.e., with less nodes than G), each having no K_4-minor [14]. Johnson and McKee [29] show that $\mathcal{G}_{\mathrm{wh}}$ consists precisely of the graphs that can be obtained by means of clique sums from chordal graphs and graphs with no K_4-minor; interestingly, they derive this decomposition result from a result of Barrett et al. [5] concerning the PSD completion problem (cf. Theorem 7).

2.2. Matrices

An $n \times n$ symmetric matrix A is said to be *positive semidefinite* (then we write $A \succeq 0$) if $x^{\mathrm{T}}Ax \geqslant 0$ for all $x \in \mathbb{R}^n$ and *positive definite* if $x^{\mathrm{T}}Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. An $n \times n$ matrix $D = (d_{ij})$ is called a *Euclidean distance matrix* if there exist vectors $p_1, \ldots, p_n \in \mathbb{R}^k$ (for some $k \geqslant 1$) such that $d_{ij} = \|p_i - p_j\|^2$ for all $i, j = 1, \ldots, n$; one says then that the vectors $p_1, \ldots, p_n$ form a *realization* (or *embedding*) of D. (Here, $\|x\| := \sqrt{\sum_{i=1}^k (x_i)^2}$ denotes the Euclidean norm of $x \in \mathbb{R}^k$.) We let PSD_n and EDM_n denote, respectively, the sets of positive semidefinite matrices and Euclidean distance matrices of order n. Moreover, $\mathcal{E}_n$ denotes the subset of PSD_n consisting of the positive semidefinite matrices whose diagonal entries are all equal to 1. The set $\mathcal{E}_n$ is called an *elliptope* in [36] (elliptope standing for *ellip*soid and poly*tope*) and matrices in $\mathcal{E}_n$ are called *correlation matrices* (e.g., in [10,39,38]), a terminology borrowed from statistics. In what follows we sometimes abbreviate "positive semidefinite" to PSD and "Euclidean distance" to EDM. Finally, for a graph $G = (V_n, E)$, we let $\mathcal{E}(G)$ (resp. $\mathrm{EDM}(G)$) denote the projection of $\mathcal{E}_n$ (resp. of EDM_n) on the subspace $\mathbb{R}^E$ indexed by the edge set of G.

Given vectors $v_1, \ldots, v_n \in \mathbb{R}^n$, their *Gram matrix* is the $n \times n$ matrix with entries $v_i^{\mathrm{T}} v_j$ for $i, j = 1, \ldots, n$. Every Gram matrix is obviously positive semidefinite and, as is well known, every positive semidefinite matrix can be represented as a Gram matrix.

3. Cuts and positive semidefinite matrices

Given a graph $G = (V, E)$ and a weight function $w \in \mathbb{R}_+^E$ on its edges, the *max-cut problem* consists of finding a cut $\delta(S)$ whose weight $w(\delta(S)) := \sum_{e \in \delta(S)} w_e$ is maximum. NP-hardness of this basic combinatorial problem can be derived by a simple reduction from the partition problem [30]. The max-cut problem [1] has been extensively studied in the past decade. Much effort has been made, in particular, for developing algorithms permitting to solve efficiently some special instances of max-cut and to find quickly good approximate solutions for general graphs. These algorithms use essentially tools from linear programming and polyhedral combinatorics and, for the most recent ones, from spectral theory and semidefinite programming. The polyhedral approach has led to the study of the cut polytope $\mathrm{CUT}^{\square}(G)$ (defined as the convex hull of the

[1] Detailed information on max-cut can be found in the survey [43]; cf. also the book [12] for a global overview on cuts and related topics, and the annotated bibliography [35].

incidence vectors of the cuts in G) and its linear relaxations, in particular, the metric polytope[2] $\mathrm{MET}^{\square}(G)$ defined by the following inequalities:

$$x(F) - x(C \setminus F) \leqslant |F| - 1 \quad \text{for } F \subseteq C,\ C \text{ cycle in } G,\ |F| \text{ odd}. \tag{1}$$

The semidefinite programming approach is based on a representation of cuts by positive semidefinite matrices as we now see. For a subset $S \subseteq V_n$ consider its ± 1-incidence vector $x_S \in \mathbb{R}^n$ with coordinates $x_S(i) := 1$ if $i \in S$ and $x_S(i) := -1$ if $i \in V_n \setminus S$. Then the quantity

$$\frac{1}{2} \sum_{1 \leqslant i < j \leqslant n} w_{ij}(1 - x_S(i)x_S(j))$$

is equal to the weight of the cut $\delta(S)$ (after setting $w_{ij} := 0$ if ij is not an edge of G). Hence the max-cut problem can be formulated as

$$\max_{x \in \{\pm 1\}^n} \frac{1}{2} \sum_{1 \leqslant i < j \leqslant n} w_{ij}(1 - x_i x_j). \tag{2}$$

Here a cut is encoded by a matrix xx^{T} (where $x \in \{\pm 1\}^n$) which is positive semidefinite with an all-ones diagonal. Therefore, the set $\mathcal{E}_n$ consisting of the positive semidefinite matrices with an all-ones diagonal forms (up to an affine transformation) a relaxation of the cut polytope $\mathrm{CUT}^{\square}(K_n)$ of the complete graph, and hence the program:

$$\max_{X=(x_{ij}) \in \mathcal{E}_n} \frac{1}{2} \sum_{1 \leqslant i < j \leqslant n} w_{ij}(1 - x_{ij}) \tag{3}$$

yields an upper bound for the value of the max-cut problem (2).

Goemans and Williamson [18] show that problem (3) yields in fact a very good approximation of max-cut. We recall this fact in some detail as it will form the basis of some results for matrix completion problems exposed later. In what follows $\mathrm{mc}(w)$ denotes the optimum value of the max-cut problem (2) and $\mathrm{sd}(w)$ that of the relaxed problem (3).

Theorem 1 (Goemans and Williamson [18]). *Given nonnegative edge weights w, we have*

$$\frac{\mathrm{mc}(w)}{\mathrm{sd}(w)} \geqslant \alpha, \quad \textit{where } \alpha := \min_{0 \leqslant \theta \leqslant \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta};$$

the quantity α can be estimated as $0.87856 < \alpha < 0.87857$.

The proof uses the fact that every positive semidefinite matrix can be represented as a Gram matrix. Thus the quantity $\mathrm{sd}(w)$ can be reformulated as

[2] Barahona and Mahjoub [4] show that $\mathrm{CUT}^{\square}(G) \subseteq \mathrm{MET}^{\square}(G)$, with equality if and only if G has no K_5-minor. Moreover, they show that an inequality (1) defines a facet of $\mathrm{MET}^{\square}(G)$ (or $\mathrm{CUT}^{\square}(G)$) if and only if C is a chordless circuit in G. In particular, $\mathrm{MET}^{\square}(K_n)$ is defined by the following inequalities: $x_{ij} - x_{ik} - x_{jk} \leqslant 0$ and $x_{ij} + x_{ik} + x_{jk} \leqslant 2$ for $i, j, k \in V_n$, known as the *triangle inequalities*.

$$\mathrm{sd}(w) = \max \quad \frac{1}{2} \sum_{1 \leqslant i<j \leqslant n} w_{ij}(1 - v_i^{\mathrm{T}} v_j) \tag{4}$$
$$\text{s.t.} \quad v_1, \ldots, v_n \text{ unit vectors in } \mathbb{R}^n.$$

Proof. Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be unit vectors realizing the maximum in (4). The key step consists of constructing a random cut whose weight is close to the value of max-cut. It goes as follows: Select a random unit vector $r \in \mathbb{R}^n$ and set $S := \{i \in V_n \mid v_i^{\mathrm{T}} r \geqslant 0\}$. Let E_S denote the expected weight of the cut $\delta(S)$. Then,

$$E_S \leqslant \mathrm{mc}(w). \tag{5}$$

On the other hand, $E_S = \sum w_{ij} p_{ij}$ where p_{ij} is the probability that edge ij belongs to cut $\delta(S)$. The quantity p_{ij} is equal to the probability that a random hyperplane separates the two vectors v_i and v_j and thus to $\pi^{-1} \arccos(v_i^{\mathrm{T}} v_j)$. Therefore,

$$E_S = \frac{1}{\pi} \sum_{1 \leqslant i<j \leqslant n} w_{ij} \arccos(v_i^{\mathrm{T}} v_j). \tag{6}$$

Hence, $E_S \geqslant \alpha \sum_{1 \leqslant i<j \leqslant n} w_{ij}(1 - v_i^{\mathrm{T}} v_j)/2$ by definition of α. This implies that $E_S \geqslant \alpha \cdot \mathrm{sd}(w)$ which, combined with (5), gives the result. □

Problem (3) is a typical instance of a semidefinite programming problem. It can be solved efficiently (with arbitrary precision), e.g., using the ellipsoid method or interior-point methods (cf., e.g., [1,18]). Thus the upper bound $\mathrm{sd}(w)$ for max-cut can be efficiently computed. Moreover Goemans and Williamson indicate how to "derandomize" their procedure in order to obtain a (deterministic) cut whose weight is at least the expected weight.

As an application (of the proof) of Theorem 1, we have a "recipe" for constructing linear inequalities that are satisfied by the pairwise angles of a set of n unit vectors. Namely, every linear inequality valid for the cut polytope $\mathrm{CUT}^{\square}(K_n)$ of the complete graph yields such an inequality.

Theorem 2. *Let $v_1, \ldots, v_n$ be n unit vectors in $\mathbb{R}^m$ ($n \geqslant 3$, $m \geqslant 1$). Let $w \in \mathbb{R}^{E_n}$ and $w_0 \in \mathbb{R}$ such that the inequality $w^{\mathrm{T}} x \leqslant w_0$ is valid for the cut polytope $\mathrm{CUT}^{\square}(K_n)$. Then,*

$$\sum_{1 \leqslant i<j \leqslant n} w_{ij} \arccos(v_i^{\mathrm{T}} v_j) \leqslant \pi w_0.$$

Proof. The result follows from relations (5) and (6) and the fact that $\mathrm{mc}(w) \leqslant w_0$. □

For instance, we find the well-known geometrical fact that the pairwise angles between three vectors satisfy the triangle inequalities. Fejes Tóth [16] asked more generally what is the maximum value for the sum of angles between a set of n vectors; he conjectured that this maximum sum is $\lceil n/2 \rceil \lfloor n/2 \rfloor \pi$ and proved this conjecture for $n \leqslant 6$. This conjectured value is indeed correct in view of Theorem 2 since the maximum cardinality of a cut in K_n is $\lceil n/2\rceil\lfloor n/2\rfloor$. In the case of three vectors the following stronger result holds (which can be checked by trigonometric manipulation; cf. [6]).

Theorem 3. *The following assertions are equivalent for $\alpha, \beta, \gamma \in [0, \pi]$.*
(i) *The matrix*

$$A := \begin{pmatrix} 1 & \cos\alpha & \cos\beta \\ \cos\alpha & 1 & \cos\gamma \\ \cos\beta & \cos\gamma & 1 \end{pmatrix}$$

is positive semidefinite.
(ii) $\alpha \leqslant \beta+\gamma$, $\beta \leqslant \alpha+\gamma$, $\gamma \leqslant \alpha+\beta$ *and* $\alpha+\beta+\gamma \leqslant 2\pi$.

The next result is a reformulation of Theorems 2 and 3.

Corollary 4. *Let $X = (x_{ij})$ be an $n \times n$ symmetric matrix with an all-ones diagonal and let $x = (x_{ij})_{1\leqslant i<j\leqslant n}$ denote the vector in $\mathbb{R}^{\binom{n}{2}}$ consisting of the upper triangular entries of X. Then,*

$$X \in \mathcal{E}_n \ (\text{i.e., } X \succeq 0) \quad \Longrightarrow \quad \frac{1}{\pi}\arccos x \in \mathrm{CUT}^{\square}(K_n).$$

Moreover, this implication holds as an equivalence in the case $n = 3$.

To conclude, we have just observed some links between cuts and positive semidefinite matrices in the set $\mathcal{E}_n$. On the one hand, each cut (and, thus, each point in the cut polytope) can be encoded by a matrix in $\mathcal{E}_n$; on the other hand, the function $\pi^{-1}\arccos$ applied to the entries of a matrix in $\mathcal{E}_n$ yields a point belonging to the cut polytope. In the next section we further explore the latter property, in the setting of the completion problem for positive semidefinite matrices, a problem arising in combinatorial matrix theory.

4. The positive semidefinite completion problem

A *partial matrix* is a matrix $X = (x_{ij})$ whose entries are specified only on a subset of the positions but in such a way that x_{ji} is specified and equal to x_{ij} whenever x_{ij} is specified. We consider here the following problem:

The PSD completion problem (PSD). Given a partial matrix X, is it possible to choose the unspecified entries of X in such a way that the resulting matrix is positive semidefinite?

We may restrict ourselves to the case when all diagonal entries are specified and (up to rescaling) equal to 1. Note that rescaling may introduce square roots, which may cause problems from an algorithmic point of view. The problem can then be reformulated as follows: *Given a graph $G = (V_n, E)$ (whose edge set E corresponds to the set of specified entries) and given a vector $x \in \mathbb{R}^E$, does x belong to the projected elliptope $\mathcal{E}(G)$?* Thus we find the question of testing membership in the convex body $\mathcal{E}(G)$.

This problem has received considerable attention in the literature of linear algebra. This is due to its many applications, in particular, to probability, statistics, system engineering, geophysics, etc. Results have been obtained along the following lines: finding necessary conditions for membership in $\mathcal{E}(G)$ and identifying the graphs for which these necessary conditions are also sufficient; finding a positive semidefinite completion satisfying certain requirement like having maximum determinant; etc.

Concerning the latter question, the following is proved in [19]: Given a partial matrix X with specified diagonal entries, if X can be completed to a positive definite matrix, then there exists a unique positive definite completion whose determinant is maximal and this matrix is characterized by the fact that its inverse has zeros precisely in the positions corresponding to unspecified entries in X. Moreover, in the case when the graph of specified entries is chordal, this maximum determinant as well as the unspecified entries can be expressed explicitly in terms of the specified entries (see [26] and references therein).

We focus here on the first question concerning finding necessary conditions for membership in the elliptope $\mathcal{E}(G)$. Results are summarized in Section 4.1. We then consider in Section 4.2 complexity issues for the PSD completion problem. The exact complexity of this problem is not known. However, the problem can be solved in polynomial time if one allows an arbitrary small precision. Moreover, for chordal graphs, the problem can be solved exactly by a polynomial-time combinatorial algorithm.

A detailed treatment of the material exposed in this section and the next one concerning completion problems for PSD and EDM matrices can be found in the survey paper [34].

4.1. Necessary conditions

Let $G = (V_n, E)$ be a graph and let $x \in \mathbb{R}^E$ be a vector for which we wish to test membership in $\mathcal{E}(G)$. For a clique K in G let x_K denote the projection of x on the edge set of K. Then,

$$\text{(PSDK)} \qquad x_K \in \mathcal{E}(K) \quad \text{for each clique } K \text{ in } G$$

is a necessary condition for $x \in \mathcal{E}(G)$ (because every principal submatrix of a positive semidefinite matrix is positive semidefinite). Clearly, $x \in [-1, 1]^E$ if $x \in \mathcal{E}(G)$ and thus we can parametrize x as $x = \cos(\pi a)$ where $a \in [0, 1]^E$ (that is, $a_e := \pi^{-1} \arccos x_e$ for all $e \in E$). From Corollary 4 we have that

$$\text{(PSDC)} \qquad a \in \mathrm{CUT}^{\square}(G)$$

is also a necessary condition for $x \in \mathcal{E}(G)$. This condition was formulated in [32]. A weaker condition[3] has been found earlier in [6] involving the linear relaxation of $\mathrm{CUT}^{\square}(G)$ by the metric polytope $\mathrm{MET}^{\square}(G)$; namely, the condition:

$$\text{(PSDM)} \qquad a \in \mathrm{MET}^{\square}(G).$$

None of the conditions (PSDK), (PSDM), or (PSDC) suffices for characterizing $\mathcal{E}(G)$ in general. For instance, let $G = (V, E)$ be a nonchordal graph and let C be a chordless circuit in G of length $\geqslant 4$. Define $x \in \mathbb{R}^E$ by setting $x_e := 1$ for all edges e in C except $x_{e_0} := -1$ for one edge e_0 in C, and $x_e := 0$ for all remaining edges in G. Then, as noted in [19], x satisfies (PSDK) but $x \notin \mathcal{E}(G)$ (for instance, because (PSDM) is violated). As another example, consider the 4×4 matrix X with diagonal entries 1 and with off-diagonal entries $-1/2$. Then, X is not positive semidefinite; hence, the vector $x := (-1/2, \dots, -1/2) \in \mathbb{R}^{E(K_4)}$ does not belong to $\mathcal{E}(K_4)$, while $\pi^{-1} \arccos x = (2/3, \dots, 2/3)$ belongs to $\mathrm{MET}^{\square}(K_4) = \mathrm{CUT}^{\square}(K_4)$.

Hence arises the question of characterizing the graphs G for which the conditions (PSDK), (PSDM), (PSDC) (taken together or separately) suffice for the description of $\mathcal{E}(G)$. Let $\mathcal{P}_{\mathrm{K}}$ (resp. $\mathcal{P}_{\mathrm{M}}$, $\mathcal{P}_{\mathrm{C}}$) denote the class of graphs G for which the condition (PSDK) (resp. (PSDM), (PSDC)) is sufficient for the description of $\mathcal{E}(G)$. Moreover, let $\mathcal{P}_{\mathrm{KM}}$ (resp. $\mathcal{P}_{\mathrm{KC}}$) denote the class of graphs G for which the two conditions (PSDK) and (PSDM) (resp. (PSDK) and (PSDC)) taken together suffice for the description of $\mathcal{E}(G)$. Obviously,

$$\mathcal{P}_{\mathrm{M}} \subseteq \mathcal{P}_{\mathrm{C}} \quad \text{and} \quad \mathcal{P}_{\mathrm{KM}} \subseteq \mathcal{P}_{\mathrm{KC}}.$$

The classes $\mathcal{P}_{\mathrm{K}}$, $\mathcal{P}_{\mathrm{M}}$ and $\mathcal{P}_{\mathrm{C}}$ are described below.

As observed above, every graph $G \in \mathcal{P}_{\mathrm{K}}$ must be chordal. Grone et al. [19] show that $\mathcal{P}_{\mathrm{K}}$ consists precisely of the chordal graphs. One quick way to derive this result is by showing that $\mathcal{P}_{\mathrm{K}}$ is closed under taking clique sums, since cliques (trivially) belong to $\mathcal{P}_{\mathrm{K}}$ and every chordal graph can be build from cliques by taking clique sums. Another proof will be given in Section 4.2.

Theorem 5 (Grone et al. [19]). *For a graph $G = (V, E)$, we have*

$$\mathcal{E}(G) = \{x \in \mathbb{R}^E \mid x_K \in \mathcal{E}(K) \ \forall K \text{ clique in } G\}$$

if and only if G is chordal.

We turn to the description of the classes $\mathcal{P}_{\mathrm{M}}$ and $\mathcal{P}_{\mathrm{C}}$. By Theorem 3, the graph K_3 belongs to $\mathcal{P}_{\mathrm{M}}$ and, as was observed earlier, the graph K_4 does not belong to $\mathcal{P}_{\mathrm{C}}$. More generally, circuits belong to $\mathcal{P}_{\mathrm{M}}$ [6]. It is shown in [32] that the classes $\mathcal{P}_{\mathrm{M}}$ and $\mathcal{P}_{\mathrm{C}}$

---

[3] Barrett et al. [6] proved necessity of (PSDM) by first showing the result from Theorem 3 about K_3 and then deriving the general result by induction on the length of cycle C in (1). In other words, they rediscovered the fact (proved by Barahona [3]) that projecting the triangle inequalities on the edge set of a graph G yields the inequalities (1).

are identical and consist precisely of the graphs with no K_4-minor. The proof is based on the decomposition result for graphs with no K_4-minor mentioned in Section 2. It consists of verifying that both classes $\mathcal{P}_{\mathrm{M}}$ and $\mathcal{P}_{\mathrm{C}}$ are closed under taking minors and that $\mathcal{P}_{\mathrm{M}}$ is closed under taking clique sums.

Theorem 6 (Laurent [32]). *The following assertions are equivalent for a graph G:*
(i) $\mathcal{E}(G) = \{x = \cos(\pi a) \mid a \in \mathrm{MET}^{\square}(G)\}$.
(ii) $\mathcal{E}(G) = \{x = \cos(\pi a) \mid a \in \mathrm{CUT}^{\square}(G)\}$.
(iii) *G has no K_4-minor.*

Let us now consider the classes $\mathcal{P}_{\mathrm{KM}}$ and $\mathcal{P}_{\mathrm{KC}}$. Clearly, it suffices here to assume that (PSDK) holds for all cliques of size $\geqslant 4$ (as the cliques of size $\leqslant 3$ are taken care of by (PSDM) or (PSDC)). Several equivalent characterizations for the graphs in $\mathcal{P}_{\mathrm{KM}}$ have been discovered by Barrett et al. [5]; more precisely, they show the equivalence of assertions (i), (iii), (iv), (v) in Theorem 7. Building upon their result, Johnson and McKee [29] show the equivalence of (i) and (vi); in other words, the graphs in $\mathcal{P}_{\mathrm{KM}}$ arise from the graphs in $\mathcal{P}_{\mathrm{K}}$ and $\mathcal{P}_{\mathrm{M}}$ by taking clique sums. Laurent [33] observes moreover the equivalence of (i) and (ii); hence, the two classes $\mathcal{P}_{\mathrm{KM}}$ and $\mathcal{P}_{\mathrm{KC}}$ coincide even though the cut condition (PSDC) is stronger than the metric condition (PSDM). The survey [34] contains a full proof of Theorem 7 which is at several places simpler and shorter than the original one from [5].

The results from Theorem 7 are interesting from a purely graph theoretical point of view; indeed, among other characterizations, they provide a decomposition result for the class of graphs $\mathcal{G}_{\mathrm{wh}}$, which is defined by excluded induced configurations. Hence, they make the link between interesting graph theoretic properties and matrix properties and, therefore, they are a good illustration of a fruitful interaction between combinatorial and algebraic aspects.

Theorem 7. *The following assertions are equivalent for a graph G:*
(i) *$G \in \mathcal{P}_{\mathrm{KM}}$, i.e., $\mathcal{E}(G)$ consists of the vectors $x = \cos(\pi a)$ such that $a \in \mathrm{MET}^{\square}(G)$ and $x_K \in \mathcal{E}(K)$ for every clique K in G.*
(ii) *$G \in \mathcal{P}_{\mathrm{KC}}$, i.e., $\mathcal{E}(G)$ consists of the vectors $x = \cos(\pi a)$ such that $a \in \mathrm{CUT}^{\square}(G)$ and $x_K \in \mathcal{E}(K)$ for every clique K in G.*
(iii) *$G \in \mathcal{G}_{\mathrm{wh}}$, i.e., no induced subgraph of G is W_n ($n \geqslant 5$) or a splitting of W_n ($n \geqslant 4$).*
(iv) *Every induced subgraph of G that contains a homeomorph of K_4 contains a clique of size 4.*
(v) *There exists a chordal graph G' containing G as a subgraph and having no new clique of size 4.*
(vi) *G can be obtained by means of clique sums from chordal graphs (or cliques) and graphs with no K_4-minor.*

Remark that the result from Theorem 5 can be stated for partial matrices with arbitrary diagonal entries. Namely, if X is a partial matrix with specified diagonal entries and whose specified off-diagonal entries form a chordal graph and if every fully specified principal submatrix of X is positive semidefinite, then X can be completed to a positive semidefinite matrix. This is not the case for Theorems 6 and 7 as the conditions (PSDM) and (PSDC) can only be formulated for matrices with an all-ones diagonal.

4.2. *Computing positive semidefinite completions*

We group here several observations concerning the complexity of problem (PSD), the completion problem for positive semidefinite matrices. This problem contains as a subproblem the problem of testing membership in the elliptope $\mathcal{E}(G)$ of a graph.

Although polynomial-time solvable for some classes of graphs (e.g., for chordal graphs as we see below), the exact complexity status of problem (PSD) is not known. This problem is in fact a typical instance of the following feasibility problem for semidefinite programming. (For two $n \times n$ matrices A and B, one sets $\langle A, B\rangle := \sum_{i,j=1}^{n} a_{ij}b_{ij}$.)

The semidefinite programming feasibility problem (F). Given rational $n \times n$ matrices $A_1, \ldots, A_m$ and vectors $b_1, \ldots, b_m \in \mathbb{Q}^n$, decide if there exists a matrix $X \in \mathrm{PSD}_n$ such that $\langle A_i, X\rangle = b_i$ for all $i = 1, \ldots, m$.

It is one of the major open questions in the field of semidefinite programming to determine the complexity status of problem (F). It is not known whether (F) is in NP. Some complexity results are given by Ramana [44]. In particular, he develops an exact duality theory which enables him to show that (F) belongs to NP if and only if (F) belongs to co-NP. Therefore, if NP $\neq$ co-NP then (F) is neither NP-complete nor co-NP-complete.

However the problem becomes easy if one allows approximations; more precisely, the weak membership problem can be solved in polynomial time. In the case of the elliptope $\mathcal{E}(G)$, this is the problem: Given $x \in \mathbb{Q}^E$ and $\varepsilon > 0$, decide whether, (i) $x \in S(K, \varepsilon)$ ("x is almost in K"), or (ii) $x \notin S(K, -\varepsilon)$ ("x is almost in the complement of K"), where K stands for $\mathcal{E}(G)$. (We remind that $S(K,\varepsilon) = \{y \mid \exists x \in K \text{ with } \|x - y\| < \varepsilon\}$ and $S(K, -\varepsilon) = \mathbb{R}^n \setminus S(\mathbb{R}^n \setminus K, \varepsilon)$.)

Moreover, the problem of finding a positive semidefinite completion of a partial matrix X (if one exists) can be answered by solving, for instance, the following semidefinite programming problem:

$$\begin{aligned} \max \quad & \sum_{i,j \mid ij \notin E} y_{ij} \\ \text{s.t.} \quad & Y = (y_{ij}) \in \mathrm{PSD}_n \\ & y_{ii} = x_{ii} \quad \forall i = 1, \ldots, n \\ & y_{ij} = x_{ij} \quad \forall ij \in E. \end{aligned}$$

The weak version[4] of this optimization problem can be solved in polynomial time. This can be done using the ellipsoid method or interior-point algorithms (cf., e.g., [22,41,1]); specific algorithms are discussed in [28]. However, such algorithms can only give approximate solutions and, thus, are not guaranteed to find exact completions.

On the other hand, for the class of chordal graphs, a combinatorial algorithm has been found in [19] that permits to solve the PSD completion problem in an exact manner and in polynomial time. This algorithm exploits the following properties of chordal graphs.

Let $G = (V_n, E)$ be a graph. An ordering $v_1, \ldots, v_n$ of the nodes of G is called a *perfect elimination ordering* if, for every $i = 1, \ldots, n-1$, the set of nodes v_j (for $j > i$) that are adjacent to v_i induces a clique. It is well known that G is chordal if and only if it has a perfect elimination ordering. Moreover, such an ordering can be found, if one exists, in time $O(n+m)$ [46] (n is the number of nodes and m the number of edges). From this follows that, if G is chordal, then one can construct a sequence of graphs $G_0 := G$, $G_1, \ldots, G_p = K_n$, where each G_i is chordal and G_{i+1} is obtained from G_i by adding one edge. (Indeed, if G is not complete and if $v_1, \ldots, v_n$ is a perfect elimination ordering of its nodes, let i be the largest index in $[1,n]$ for which there exists $j > i$ such that v_i and v_j are not adjacent; then, adding edge v_iv_j to G yields a new graph G' which is again chordal as $v_1, \ldots, v_n$ remains a perfect elimination ordering for G'.)

As we now see this property permits to show the result from Theorem 5 (namely, that chordal graphs belong to the class $\mathcal{P}_{\mathrm{K}}$) and to construct an explicit PSD completion for $x \in \mathcal{E}(G)$ when G is chordal. We also deal here with positive definite (PD) completions. An obvious necessary condition for $x \in \mathbb{R}^E$ to have a PD completion is the following restrictive form of (PSDK):

$$\text{(PSDK}*) \qquad x_K \in \operatorname{int}\mathcal{E}(K) \text{ for each clique } K \text{ in } G$$

(where, for a set $A \subseteq \mathbb{R}^m$, $\operatorname{int} A$ denotes its relative interior).

Let G be a chordal graph and let $G_0 := G, G_1, \ldots, G_p = K_n$ be a sequence of chordal graphs bringing G to the complete graph, adding one edge at a time. The complete graph K_n obviously belongs to class $\mathcal{P}_{\mathrm{K}}$. Hence, the fact that $G \in \mathcal{P}_{\mathrm{K}}$ will follow by induction if we can show the following: *Let H be a chordal graph and u, v be two nonadjacent nodes in H; then the graph H belongs to $\mathcal{P}_{\mathrm{K}}$ whenever the graph $H + uv$ (obtained by adding edge uv to H) belongs to $\mathcal{P}_{\mathrm{K}}$.* Observe that, under these assumptions, there exists a unique maximal clique K in $H + uv$ containing both nodes u and v. Hence, the above statement is an immediate consequence of the result in Lemma 8. This shows, therefore, that $G \in \mathcal{P}_{\mathrm{K}}$.

The proof of Lemma 8 is constructive; it gives an efficient algorithm for computing a rational PD or PSD completion (if one exists) for a partial rational matrix with only one unspecified entry. Therefore, by the above discussion, we have an inductive procedure[5]

[4] The weak version of the optimization problem: $\max c^{\mathrm{T}}x$ for $x \in K$ (over a convex set $K \subset \mathbb{R}^n$) can be formulated as follows: Given a rational $\varepsilon > 0$, either (i) find a vector $y \in \mathbb{Q}^n$ such that $y \in S(K, \varepsilon)$ and $c^{\mathrm{T}}x \leqslant c^{\mathrm{T}}y + \varepsilon$, or (ii) assert that $S(K, -\varepsilon) = \emptyset$; cf. [22].

[5] We expose here the case of partial matrices with an all-ones diagonal, but this procedure applies more generally to partial matrices with arbitrary specified diagonal entries.

for constructing a PD (or PSD) completion of a vector x indexed by the edge set of a chordal graph G. This procedure can be carried out in polynomial time for rational data.

Lemma 8. *Let $H := K_n \backslash e$ be the graph obtained by deleting one edge e in the complete graph K_n. Let x be a vector indexed by the edge set of H. If x satisfies* (PSDK$*$) *(resp.* (PSDK)*), then $x \in \operatorname{int}\mathcal{E}(H)$ (resp. $x \in \mathcal{E}(H)$). Moreover, if x is rational valued, then we can find a rational PD (resp. PSD) completion of x.*

Proof. Suppose that e is the edge $(1,n)$; then H has two maximal cliques on $K_1 := \{1,\dots,n-1\}$ and $K_2 := \{2,\dots,n\}$. Assume that x satisfies (PSDK). Then we can find vectors $u_1,\dots,u_n$ such that $x_{ij} = u_i^{\mathrm{T}} u_j$ for all $i,j \in K_1$ and all $i,j \in K_2$, which shows that $x \in \mathcal{E}(H)$. It remains to show that $x \in \operatorname{int}\mathcal{E}(H)$ if x satisfies (PSDK$*$) and that x has a rational completion if x is rational valued. For this, let X denote the partial symmetric matrix corresponding to x, of the form

$$X = \begin{pmatrix} 1 & a^{\mathrm{T}} & z \\ a & A & b \\ z & b^{\mathrm{T}} & 1 \end{pmatrix}$$

and set

$$X_1 := \begin{pmatrix} 1 & a^{\mathrm{T}} \\ a & A \end{pmatrix}, \qquad X_2 := \begin{pmatrix} A & b \\ b^{\mathrm{T}} & 1 \end{pmatrix},$$

where A is a symmetric $(n-2) \times (n-2)$ matrix, $a, b \in \mathbb{R}^{n-2}$ and $z \in \mathbb{R}$ is the free entry to be determined; matrix X_i corresponds to x_{K_i} for $i = 1,2$. Suppose first that x satisfies (PSDK$*$), i.e., that both X_1 and X_2 are positive definite. In order to show that $x \in \operatorname{int}\mathcal{E}(H)$, it suffices to construct z for which $\det X > 0$. By assumption, we have that $\det A > 0$, $\det(X_1) = (1 - a^{\mathrm{T}}A^{-1}a)\cdot \det A > 0$, and $\det(X_2) = (1 - b^{\mathrm{T}}A^{-1}b)\cdot\det A > 0$. Moreover,

$$\begin{aligned} \det X &= \det A \cdot \det\left(\begin{pmatrix} 1 & z \\ z & 1 \end{pmatrix} - (a \quad b)^{\mathrm{T}} A^{-1} (a \quad b)\right) \\ &= \det A \cdot \det \begin{pmatrix} 1 - a^{\mathrm{T}}A^{-1}a & z - a^{\mathrm{T}}A^{-1}b \\ z - b^{\mathrm{T}}A^{-1}a & 1 - b^{\mathrm{T}}A^{-1}b \end{pmatrix}. \end{aligned}$$

Hence, $\det X > 0$ if we choose $z := a^{\mathrm{T}}A^{-1}b$. Observe moreover that $z \in \mathbb{Q}$ if x is rational valued.

We now turn to the general case when x satisfies (PSDK) but not necessarily (PSDK$*$). The only thing which remains to be shown is that, if x is rational valued, then it admits at least one rational valued PSD completion. If the value $u_1^{\mathrm{T}}u_n$ happens to be rational, then we are done as the Gram matrix of $u_1,\dots,u_n$ is a rational PSD completion of x. Observe that $u_1^{\mathrm{T}}u_n$ is indeed rational if u_1 (or u_n) belongs to the linear span of $\{u_2,\dots,u_{n-1}\}$ (since then there exists a rational vector λ such that $X_1\lambda = 0$ and $\lambda_1 \neq 0$). So, we can now assume that neither u_1 nor u_n belongs to the linear span of $\{u_2, \ldots, u_{n-1}\}$. Say, $\{u_2, \ldots, u_p\}$ is a maximal linearly independent subset of $\{u_2, \ldots, u_{n-1}\}$. By the reasoning above, we can assign a rational value z to the $(1,n)$-entry in the submatrix X' of X with row/column indices $1, 2, \ldots, p, n$ which makes X' positive definite. We now verify that this value z makes X itself positive semidefinite. By construction, we can find vectors $w_1, w_2, \ldots, w_p, w_n$ whose Gram matrix is X'. Recall that the Gram matrix of $u_2, \ldots, u_n$ is X_2. Let T be an orthogonal transformation mapping w_i to u_i for $i = 2, \ldots, p, n$. Then, X coincides with the Gram matrix of the vectors $Tw_1, u_2, \ldots, u_n$, since one can easily verify that $(Tw_1)^T u_i = u_1^T u_i$ for $i = p+1, \ldots, n-1$. $\square$

As we just saw, the PD and PSD completion problems can be solved efficiently when the graph of specified entries is chordal. What about the case when this graph has no K_4-minor or, more generally, belongs to the class $\mathcal{G}_{\mathrm{wh}}$?

A nice feature is that the graphs in $\mathcal{G}_{\mathrm{wh}}$ can be recognized efficiently. Indeed, one can test if a graph has no K_4-minor in time $O(n)$ [51]. Then, in order to test whether a graph G belongs to $\mathcal{G}_{\mathrm{wh}}$ it suffices, in view of Theorem 7, to decompose G by means of clique sums into indecomposable pieces (which can be done in time $O(nm)$ using the algorithm of Tarjan [50]) and to check whether all pieces are cliques or without a K_4-minor.

For graphs with no K_4-minor, the existence of a PSD completion is characterized by the metric condition (PSDM). Checking membership in the metric polytope $\mathrm{MET}^{\square}(G)$ can be done in polynomial time (for any graph) [4]. Yet there is some difficulty in checking condition (PSDM). Indeed, even if x is rational valued, $\pi^{-1} \arccos x$ is quite unlikely to remain rational! Thus we can only work with a rational approximation of $\pi^{-1} \arccos x$. This means that we may encounter problems of numerical stability for deciding whether (PSDM) holds when $\pi^{-1} \arccos x$ happens to be very close to the boundary of $\mathrm{MET}^{\square}(G)$.

Yet, assuming that we can compute with infinite precision, we would like to mention a combinatorial method for computing completions in the case of graphs with no K_4-minor or, more generally, graphs in $\mathcal{G}_{\mathrm{wh}}$.

Let G be a graph with no K_4-minor and let x be a vector indexed by the edge set of G for which we wish to compute a PSD completion (if one exists). We use the following property: One can find a set F of additional edges such that, when adding F to G, one obtains a new graph G' which is chordal and has no K_4-minor. Moreover, such set F can be found in time $O(n)$ [51]. We now proceed as follows. Check whether $a := \pi^{-1} \arccos x \in \mathrm{MET}^{\square}(G)$ (using, e.g., the algorithm of [4]). If not, then we know that x has no PSD completion. Else, compute a vector $b \in \mathrm{MET}^{\square}(G')$ whose projection on G is the starting vector a. Then, $y := \cos(\pi a)$ belongs to $\mathcal{E}(G')$ (as G' has no K_4-minor). Now, as G' is chordal, we can compute a PSD completion of y by the techniques exposed earlier. Thus, we find in this way a completion of x.

Finally, if G is a graph in $\mathcal{G}_{\mathrm{wh}}$, we decompose it using clique sums into cliques and pieces with no K_4-minor. For each piece H, we can compute a PSD completion y_H for the projection x_H of x on the edge set of H (if one exists). Now, we have a partial matrix y whose entries are specified on a clique sum of complete graphs. As this graph is chordal, we can compute a PSD completion of y and, thus, of x.

5. The Euclidean distance matrix completion problem

We now consider the following problem:

The Euclidean distance matrix completion problem (EDM). Given a graph $G = (V_n, E)$ and a vector $d \in \mathbb{R}^E_+$, decide if there exist vectors $p_1, \dots, p_n \in \mathbb{R}^k$ (for some $k \geqslant 1$) such that

$$\sqrt{d_{ij}} = \|p_i - p_j\| \quad \text{for all } ij \in E. \tag{7}$$

In other words, decide if the partial matrix with an all-zeros diagonal and with off-diagonal entries $d_{ij} = d_{ji}$ for $ij \in E$ can be completed to a Euclidean distance matrix; that is, if d belongs to the projected cone EDM(G).

We remind that $p = (p_1, \dots, p_n) \in \mathbb{R}^{kn}$ is called a *realization* of the weighted graph (G, d) if (7) holds. Problem (EDM) is therefore the problem of finding a realization of a weighted graph in the Euclidean space (of arbitrary dimension). If such a realization exists, then one can be found in the space of dimension $k \leqslant \lfloor(\sqrt{8|E|+1} - 1)/2\rfloor$; this bound (better than the trivial value n) was found by Barvinok [7] (cf. Chapter 31 in [12] for a simple proof).

One may also consider the following problem of finding a realization of a weighted graph in the Euclidean space of a fixed dimension:

The graph realization problem in the Euclidean k-space (EDMk). Given a graph $G = (V, E)$, a vector $d \in \mathbb{Q}^E_+$ and an integer $k \geqslant 1$, decide if there exist vectors $p_1, \dots, p_n \in \mathbb{R}^k$ such that $d_{ij} = \|p_i - p_j\|^2$ for $ij \in E$.

Euclidean distance matrices are a classic notion in distance geometry, whose study has a long history. Interest in them and the associated completion problems (EDM) and (EDMk) has been renewed recently in view of their many applications. They are, for instance, central in the theory of multidimensional scaling (cf. the survey [37]) and problem (EDMk) in dimension $k = 2, 3$ arises in chemistry for the determination of molecular conformations. Indeed, new techniques from nuclear magnetic resonance spectroscopy permit to partially determine interatomic distances in a molecule; the question being then to reconstruct the three dimensional shape of the molecule from these partial data (cf. [11,53,23]). Problem (EDMk) is, in fact, known in the literature under several other names; e.g., as the position-location problem in [54], the molecule problem in [25], or the "fundamental problem of distance geometry" in [11].

We group in Section 5.1 results dealing with the EDM completion problem in arbitrary dimension and we consider in Section 5.2 complexity aspects of problems (EDM) and (EDMk).

5.1. Necessary conditions

Research has been done on the EDM completion problem along the same lines as for the PSD completion problem. In particular, three necessary conditions have been found that are analogue to the conditions (PSDK), (PSDM) and (PSDC). Namely, for a graph $G = (V_n, E)$ and a vector $d \in \mathbb{R}^E_+$, each of the three conditions:

(EDMK) $\quad d_K \in \mathrm{EDM}(K)$ for every clique K in G,

(EDMM) $\quad \sqrt{d} \in \mathrm{MET}(G)$,

(EDMC) $\quad \sqrt{d} \in \mathrm{CUT}(G)$

is necessary [6] for $d \in \mathrm{EDM}(G)$. Here, $\mathrm{MET}(G)$ denotes the *metric cone*, which is defined by the inequalities (1) with zero right hand side (i.e., the inequalities: $x_e - x(C \setminus \{e\}) \leqslant 0$ for $e \in C$, C cycle in G) and $\mathrm{CUT}(G)$ denotes the *cut cone* of G, defined as the conic hull of the incidence vectors of the cuts in G. The classes of graphs for which the conditions (EDMK), (EDMM), (EDMC) are sufficient turn out to be the same as in the PSD case; compare Theorems 9, 10 and 11 with Theorems 5, 6 and 7.

Theorem 9 (Bakonyi and Johnson [2]). *For a graph $G = (V, E)$, we have*

$$\mathrm{EDM}(G) = \{d \in \mathbb{R}^E \mid d_K \in \mathrm{EDM}(K)\ \forall K \text{ clique in } G\}$$

if and only if G is chordal.

Theorem 10 (Laurent [33]). *The following assertions are equivalent for a graph G:*
(i) $\mathrm{EDM}(G) = \{d \in \mathbb{R}^E_+ \mid \sqrt{d} \in \mathrm{CUT}(G)\}$.
(ii) $\mathrm{EDM}(G) = \{d \in \mathbb{R}^E_+ \mid \sqrt{d} \in \mathrm{MET}(G)\}$.
(iii) *G has no K_4-minor.*

Theorem 11 (Johnson et al. [27] and Laurent [33]). *The following assertions are equivalent for a graph G:*
(i) $\mathrm{EDM}(G) = \{d \in \mathbb{R}^E_+ \mid \sqrt{d} \in \mathrm{MET}(G)$ *and* $d_K \in \mathrm{EDM}(K)\ \forall K$ *clique in* $G\}$.
(ii) $\mathrm{EDM}(G) = \{d \in \mathbb{R}^E_+ \mid \sqrt{d} \in \mathrm{CUT}(G)$ *and* $d_K \in \mathrm{EDM}(K)\ \forall K$ *clique in* $G\}$.
(iii) *No induced subgraph of G is a wheel W_n ($n \geqslant 5$) or a splitting of a wheel W_n ($n \geqslant 4$).*

In fact, these results can be derived from the corresponding results for the PSD completion problem, since positive semidefinite and Euclidean distance matrices are very closely related notions. We recall below two well-known operations, due to Schoenberg [48,49], that permit to link PSD and EDM matrices. Based on these operations, it is observed in [33] how the results from Theorems 9, 10, 11 can be derived, respectively, from those from Theorems 5, 6, 7 (one of the tools is Lemma 14).

[6] Necessity of (EDMK) and (EDMM) is obvious; necessity of (EDMC) follows from the fact that every ℓ_2-embeddable metric is ℓ_1-embeddable (cf. [33]).

Given a graph $G = (V_n, E)$ and its suspension graph ∇G, let $d \in \mathbb{R}^{E(\nabla G)}$ and let $x \in \mathbb{R}^{V_n \cup E}$ be defined by

$$x_{ii} := d_{i,n+1} \text{ for } i \in V_n, \quad x_{ij} := \tfrac{1}{2}(d_{i,n+1} + d_{j,n+1} - d_{ij}) \text{ for } ij \in E. \tag{8}$$

Then, it can be easily verified that

$$d \in \mathrm{EDM}(\nabla G) \iff x \text{ can be completed to a PSD matrix.} \tag{9}$$

(Indeed, a PSD completion of x can be represented as the Gram matrix of some vectors $p_1, \ldots, p_n$; then vectors $p_1, \ldots, p_n, p_{n+1} := 0$ satisfy relation (7) and thus provide an EDM completion of d.) This relation permits, in particular, to establish a one-to-one linear correspondence between the cone $\mathrm{EDM}(\nabla G)$ and the elliptope $\mathcal{E}(G)$. The next result was proved by Schoenberg [49] in the case of the complete graph and extended to arbitrary graphs in [33].

Proposition 12. *Let $G = (V_n, E)$ be a graph and let $d \in \mathbb{R}^E$. Then, $d \in \mathrm{EDM}(G)$ if and only if $\exp(-\lambda d) := (\exp(-\lambda d_e))_{e \in E} \in \mathcal{E}(G)$ for all $\lambda > 0$.*

As examples of applications, we show below that the condition (EDMM) suffices for the description of $\mathrm{EDM}(K_3)$ and we link the two conditions (PSDM) and (EDMM).

Lemma 13. $\mathrm{EDM}(K_3) = \{d \in \mathbb{R}^3_+ \mid \sqrt{d} \in \mathrm{MET}(K_3)\}$.

Proof. Consider the matrices

$$D := \begin{pmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{pmatrix} \quad \text{and} \quad X := \begin{pmatrix} b & \frac{b+c-a}{2} \\ \frac{b+c-a}{2} & c \end{pmatrix}$$

defined from D by (8). By relation (9), $D \in \mathrm{EDM}(K_3)$ if and only if X is positive semidefinite. Now, the latter holds if and only if $\det X \geqslant 0$, i.e., if $4bc - (b+c-a)^2 \geqslant 0$. The latter condition can be rewritten as: $a^2 - 2a(b+c) + (b-c)^2 \leqslant 0$, which is equivalent to $b + c - 2\sqrt{bc} = (\sqrt{b} - \sqrt{c})^2 \leqslant a \leqslant b + c + 2\sqrt{bc} = (\sqrt{b} + \sqrt{c})^2$. Hence, we find the condition that $\sqrt{d} \in \mathrm{MET}(K_3)$. $\square$

Lemma 14. *Let $G = (V_n, E)$ be a graph and $d \in \mathbb{R}^E_+$. Then,*

$$\sqrt{d} \in \mathrm{MET}(G) \implies \frac{1}{\pi} \arccos(\mathrm{e}^{-\lambda d}) \in \mathrm{MET}^{\square}(G) \textit{ for all } \lambda > 0.$$

Proof. Note that it suffices to show the result in the case when $G = K_n$ (as the general result will then follow by taking projections) and for $n = 3$ (as $\mathrm{MET}(K_n)$ and $\mathrm{MET}^{\square}(K_n)$ are defined by inequalities that involve only three points). Now, we have: $\sqrt{d} \in \mathrm{MET}(K_3) \iff d \in \mathrm{EDM}(K_3)$ (by Lemma 13); $d \in \mathrm{EDM}(K_3) \iff \exp(-\lambda d) \in \mathcal{E}(K_3)$ for all $\lambda > 0$ (by Proposition 12); finally, $\exp(-\lambda d) \in \mathcal{E}(K_3) \iff \pi^{-1} \arccos(\mathrm{e}^{-\lambda d}) \in \mathrm{MET}^{\square}(K_3)$ (by Theorem 3). $\square$

5.2. The graph realization problem

We group here several observations concerning the complexity of the completion problems (EDM) and (EDMk).

A first observation is that the EDM completion problem contains the PSD completion problem as a special instance. Indeed, in view of relation (9), the PSD completion problem for graph G is equivalent to the EDM completion problem for the suspension graph ∇G. In particular, problem (EDM) can also be formulated as a semidefinite programming problem and, thus, solved in polynomial time with an arbitrary precision, while its exact complexity is not known.

On the other hand, problem (EDMk) has been shown to be NP-complete for every integer $k \geqslant 1$ by Saxe [47]. Moreover, (EDMk) remains NP-complete if the data d are assumed to take their values in the set $\{1, 2\}$.

In dimension $k = 1$, the proof is particularly simple and consists of reducing problem (EDM1) from the partition problem, which is well known to be NP-complete. For this, let $a_1, \dots, a_n$ be positive integers (to be partitioned) and consider the circuit $C = (1, \dots, n)$ with edge weights $d_{i,i+1} := (a_i)^2$ for $i = 1, \dots, n$ (setting $n + 1 = 1$). Then there exist scalars $p_1, \dots, p_n \in \mathbb{R}$ such that $|p_{i+1} - p_i| = a_i$ for all $i = 1, \dots, n$ if and only if the sequence $a_1, \dots, a_n$ can be partitioned (namely, $\sum_{i \in S} a_i = \sum_{i \in [1,n] \setminus S} a_i$, where $S := \{i \mid a_i = p_{i+1} - p_i\}$).

However, both problems can be solved in polynomial time for chordal graphs. In the case of the complete graph, they can be answered by checking positive semidefiniteness and computing the rank of an associated matrix. Namely, if D is a symmetric matrix of order $n + 1$ with zero diagonal and if X denotes the symmetric matrix of order n whose entries are defined by relation (8), then D can be realized by vectors in the k-dimensional space if and only if X is positive semidefinite and has rank less than or equal to k. More generally, Bakonyi and Johnson [2] show that when the entries are specified on a chordal graph, the same step-by-step technique as the one exposed in Section 4.2 in the PSD case also applies for constructing EDM completions. They prove that if D is a partial matrix with zero diagonal whose specified entries form a chordal graph and such that every fully specified principal submatrix can be realized in the k-space, then D can be completed to an EDM matrix having a realization in the k-space.

Several algorithms have been proposed in the literature for the solution of the graph realization problem (EDMk), in particular in dimension $k \leqslant 3$, which is the case most relevant to practical applications. The problem can be naturally formulated as a nonlinear global optimization problem: min $f(p)$ s.t. $p = (p_1, \dots, p_n) \in \mathbb{R}^{kn}$, where the cost function $f(.)$ can, for instance, be chosen as: $f(p) = \sum_{ij \in E} (\|p_i - p_j\|^2 - d_{ij})^2$. So $f(.)$ is nonnegative and zero if and only if p provides a realization of the weighted graph (G, d). This optimization problem is hard to solve (the function $f(.)$ may have many local minimizers). Hendrickson [25] describes an approach for solving this problem based on a divide-and-conquer strategy; the basic steps consist of finding subgraphs having a unique realization, treating each of them separately (after possibly breaking

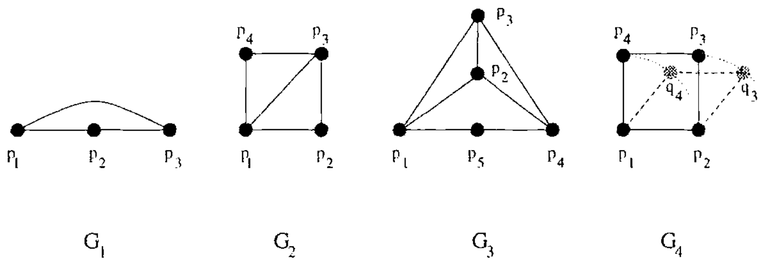

Fig. 2. Four frameworks in the plane.

them into smaller pieces) and trying to combine the solutions. Pardalos and Liu [42] propose an approach based on tabu search. Further work has been done, in particular, by chemists for the molecular conformation problem; a good overview can be found in [11] and [23].

Another question of interest which comes up in connection with the graph realization problem is that of unicity. For instance, as mentioned above, finding subgraphs having a unique realization is an important step in Hendrickson's algorithm [25]. If $p = (p_1, \ldots, p_n) \in \mathbb{R}^{kn}$ is a realization of the weighted graph (G, d) and if T is an orthogonal transformation of $\mathbb{R}^k$, then $T(p_1), \ldots, T(p_n)$ is obviously another realization of (G, d). Two realizations p and q are said to be *congruent* if $q_i = T(p_i)\ \forall i = 1, \ldots, n$, for some orthogonal transformation T. The following additional terminology is commonly used: The pair (G, p) which consists of a graph G and locations of its vertices at points $p_1, \ldots, p_n \in \mathbb{R}^k$, is called a *framework* (in the k-space) (edge weights are then defined in the obvious way).

The graph realization problem – Unicity (Uk). Given a graph $G = (V_n, E)$ with edge weights $d \in \mathbb{Q}^E$, an integer $k \geqslant 1$ and a realization $p \in \mathbb{R}^{kn}$ of (G, d), does there exist another realization of (G, d) which is not congruent to p?

Saxe [47] shows that problem (Uk) is NP-complete even if the edge weights d are supposed to be $\{1, 2\}$-valued.

Let us note at this point that non-unicity may occur in various ways. Consider for instance the four graphs from Fig. 2 viewed as frameworks in the plane $\mathbb{R}^2$.

When searching for another (noncongruent) realization of the framework (G_i, p) we may assume that p_1 and p_2 are fixed (in order to avoid translations and rotations) and moreover, in the case of G_2 and G_3, that p_3 is fixed (in order to avoid reflections). Then one can easily see that G_1 has no other realization in $\mathbb{R}^2$; G_2 has exactly one other realization; while G_3 and G_4 have an infinity of other realizations (cf. Fig. 3). In the case of G_4 there is a continuous deformation bringing p to q. This is not true for G_3 as one cannot move continuously from p to q while preserving edge lengths. Hence nonunicity has a "discrete" nature in the case of frameworks G_i for $i = 1, 2, 3$. The notion of rigidity, which will be discussed in the next section, permits to capture

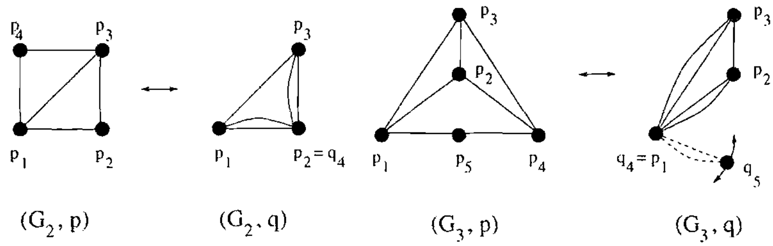

Fig. 3. Other realizations for G_2 and G_3.

these different behaviours. In fact, the frameworks (G_i, p) $(i = 1, 2, 3)$ are rigid while (G_4, p) is flexible. Precise definitions for rigidity and flexibility are given in Section 6; briefly said, these notions depend only on properties of the graph itself not on the choice of a specific generic realization (i.e., whose coordinates are algebraically independent over the rationals). Rigidity turns out to be a somewhat simpler notion than that of unicity of realization, at least in dimension $k \leqslant 2$ where it can be fully characterized.

6. Rigidity of graphs

Suppose we have a graph G whose vertices are positioned at points in the plane $\mathbb{R}^2$. Its edges are viewed as rigid rods that can rotate freely at their end nodes but are incompressible and inextendible. For instance, a triangle is a rigid structure since the three rods determine the positions of the vertices (up to a Euclidean motion). On the other hand, the square is flexible since it can be deformed continuously while preserving edge lengths; but adding one edge to the square makes it rigid (cf. Fig. 2). Determining whether a given framework is rigid or flexible is a central question in structural topology, with obvious applications to engineering. It can be asked in any dimension k; however, a complete answer is known only in dimension $k \leqslant 2$. After giving precise definitions for rigidity and flexibility, we recall here some of the main known results about rigid graphs.

6.1. Rigid and flexible frameworks

We define here the notions of rigidity and flexibility for graphs; a detailed treatment can be found, e.g., in [45] and [52]. Let (G, p) be a framework in $\mathbb{R}^k$ consisting of a graph $G = (V_n, E)$ and a vector $p = (p_1, \ldots, p_n) \in \mathbb{R}^{kn}$. A natural question which was already posed in Section 5.2 is whether there exists another realization of the framework (G, p); in other words, whether there exists $q \in \mathbb{R}^{kn}$ not congruent to p and satisfying the edge conditions:

$$\|q_i - q_j\| = \|p_i - p_j\| \quad \text{for every edge } ij \in E. \tag{10}$$

The mapping $f : \mathbb{R}^{kn} \longrightarrow \mathbb{R}^m$ defined by $f(p) := (\ldots, \|p_i - p_j\|^2, \ldots)_{ij \in E}$ is called the *edge function* of G. Thus we are interested in the set $f^{-1}(f(p))$ (which is a smooth manifold). The set M_p which consists of the realizations $q \in \mathbb{R}^{kn}$ that are congruent to p, is a smooth submanifold of $f^{-1}(f(p))$ whose dimension is equal to $k + \binom{k}{2} - \binom{k-k'}{2}$ if $p_1, \ldots, p_n$ span an affine subspace of dimension k' (as such a subspace admits k translations and $\binom{k}{2} - \binom{k-k'}{2}$ distinct orthogonal transformations). Hence, (G,p) has a unique realization (up to congruence) if $M_p = f^{-1}(f(p))$.

The framework (G,p) is said to be *flexible* if there exists a differentiable function $x : [0,1] \longrightarrow \mathbb{R}^{kn}$ such that (i) $x(0) = p$; (ii) $\|x_i(t) - x_j(t)\| = \|p_i - p_j\|$ for all edges $ij \in E$; and (iii) $x(t)$ is not congruent to p for $0 < t \leqslant 1$. Such a path x is called a *flexing* of (G,p) and (G,p) is said to be *rigid* if it is not flexible.

A desirable feature of a rigid framework (G,p) would be that every realization $q \in f^{-1}(f(p))$ sufficiently close to p is, in fact, congruent to p. For instance, the rigid framework G_2 from Fig. 2 presents this feature. This property is not clear from the definition of rigidity given here. However, it follows from results in algebraic geometry that this property does hold. Hence, it is indeed the case that a framework (G,p) is rigid in $\mathbb{R}^k$ if and only if the two manifolds M_p and $f^{-1}(f(p))$ coincide near p. Hence, the rigidity or flexibility of (G,p) is governed by the two manifolds M_p and $f^{-1}(f(p))$ (near p) and can be determined by comparing their dimensions.

It turns out that one can compute the dimension of the manifold $f^{-1}(f(p))$ near p when p is generic or, more generally, regular for the edge function $f(.)$. We now define the notions of generic and regular points. Let $R(G,p)$ denote the matrix of order $m \times kn$ whose rows are indexed by the edges of G and having a group of k columns for each node of G, the ijth row has entry $p_i - p_j$ at node i, entry $p_j - p_i$ at node j and entry 0 elsewhere; $R(G,p)$ is called the *rigidity matrix* of (G,p). Note that $R(G,p)$ coincides (up to a factor 2) with the matrix $df(p)$ of partial derivatives of the edge function f at p. Let r denote the maximum rank of the rigidity matrix $R(G,p)$ for $p \in \mathbb{R}^{kn}$; points p where this maximum rank is attained are said to be *regular* for the edge function f. Then, the implicit function theorem implies that, for any regular point p, $f^{-1}(f(p))$ is a $(kn - r)$-dimensional manifold near p.

Call $p \in \mathbb{R}^{kn}$ *generic* if the coordinates of $p_1, \ldots, p_n$ are algebraically independent over the field $\mathbb{Q}$ of rationals. Every generic $p \in \mathbb{R}^{kn}$ is obviously regular for the edge function f. Moreover, if $p \in \mathbb{R}^{kn}$ is generic, then the subspace of $\mathbb{R}^k$ spanned affinely by $p_1, \ldots, p_n$ has dimension $k' := \min(k, n-1)$. Since (G,p) is rigid if and only if both manifolds M_p and $f^{-1}(f(p))$ have the same dimension near p, we obtain that, for p generic, (G,p) is rigid if $\operatorname{rank} R(G,p) = kn - \binom{k+1}{2} + \binom{k-k'}{2}$ and flexible if $\operatorname{rank} R(G,p) < kn - \binom{k+1}{2} + \binom{k-k'}{2}$. Therefore, the generic realizations of a given graph G in $\mathbb{R}^k$ are, either all rigid, or all flexible. Note that generic points form a dense subset of $\mathbb{R}^{kn}$. Hence every graph G has a typical (or generic) behaviour; one says that G is *rigid* in $\mathbb{R}^k$ if (G,p) is rigid for any generic $p \in \mathbb{R}^{kn}$ and *flexible* otherwise. This gives the following result.

Proposition 15. *A graph G with n nodes is rigid in $\mathbb{R}^k$ if and only if the generic rank of the rigidity matrix $R(G,p)$ is equal to the quantity $S(n,k)$, where $S(n,k) := kn - \binom{k+1}{2}$ if $n \geqslant k$ and $S(n,k) := \binom{n}{2}$ if $n < k$.*

Therefore, there is an efficient randomized algorithm for deciding rigidity of a graph G: pick randomly p in $\mathbb{R}^{kn}$, then with probability one the value of the rank of $R(G,p)$ permits to decide correctly about the rigidity of G. It is however interesting to characterize rigidity by some purely graph theoretical properties. So far a combinatorial characterization of rigid graphs is known only in dimension $k \leqslant 2$, where it uses graph connectivity and matroidal features. In dimension $k \geqslant 3$, only partial results are known. We review below some of the main known results.

6.2. *Rigidity in the plane*

Let us begin with the easy task of characterizing rigidity in the 1-dimensional space $\mathbb{R}$. Obviously, any framework (G,p) in $\mathbb{R}$ where G is not connected is not rigid, since each connected component can be moved separately. Moreover, the rigidity matrix $R(G,p)$ has rank $\leqslant n-1$, with equality if and only if G is connected (note that $R(G,p)$ coincides up to rescaling with the node-edge incidence matrix of an orientation of G). Therefore,

Proposition 16. *A graph is rigid in the line $\mathbb{R}$ if and only if it is connected.*

We now consider rigidity in the plane $\mathbb{R}^2$; it has been characterized by Laman [31], whose results were later extended by Lovász and Yemini [40]. For a graph G with n nodes, the quantity

$$\varphi(G) := 2n - 3 - \operatorname{rank} R(G,p),$$

where $p \in \mathbb{R}^{2n}$ is any generic realization in $\mathbb{R}^2$, is called the *degree of freedom* of G. Thus, G is rigid when $\varphi(G) = 0$. For a subset $Y \subseteq E$ of edges, V_Y denotes the set of nodes that are incident to some edge in Y. The following result is shown in [40].

Theorem 17. *The degree of freedom of a graph $G = (V_n, E)$ with $n \geqslant 2$ nodes is given by*

$$\varphi(G) = 2n - 3 - \min \sum_{i=1}^{k} (2|V_{E_i}| - 3),$$

where the minimum is taken over all partitions $(E_1, \ldots, E_k)$ of E into nonempty subsets.

Corollary 18. *A graph $G = (V_n, E)$ on $n \geqslant 2$ nodes is rigid if and only if*

$$2n - 3 \leqslant \sum_{i=1}^{k} (2|V_{E_i}| - 3)$$

for every partition of E into nonempty subsets $E_1, \ldots, E_k$.

Therefore, one can compute the degree of freedom of a graph (and thus decide rigidity) in polynomial time. Indeed, the function: $g(Y) := 2|V_Y| - 3$ $(Y \subseteq E)$ is submodular and nonnegative on nonempty sets. Hence,

$$\min\Big(\sum_{i=1}^{k} g(E_i) \;\Big|\; E_1, \ldots, E_k \text{ is a partition of } E \text{ into nonempty subsets}\Big)$$

can be computed using the ellipsoid method, as mentioned in [21]. Gabow and Westermann [17] propose a simpler combinatorial algorithm for computing the degree of freedom of a graph on n nodes in time $O(n^2)$.

We now mention further combinatorial features of rigid graphs and, in particular, Laman's characterization for minimally rigid graphs. Consider the rigidity matrix $R(G,p)$ where $p \in \mathbb{R}^{2n}$ is generic. For a subset $Y \subseteq E$, one says that Y is *generic independent* if the corresponding set of rows of $R(G,p)$ is linearly independent (this definition makes sense as it does not depend on the specific choice of p generic). By definition,

$$\varphi(G) = 2n - 3 - \max(|Y| \mid Y \subseteq E \text{ is generic independent}).$$

The next result can be derived from Theorem 17 but it can also be checked directly.

Proposition 19. *The following assertions are equivalent for a graph $G = (V, E)$ with at least two nodes.*

(i) *The set E is generic independent.*

(ii) $|Y| \leqslant 2|V_Y| - 3$ *for every nonempty subset $Y \subseteq E$.*

(iii) *Doubling an arbitrary edge in G results in a graph that can be decomposed as the union of two forests.*

Proof. The implication (i) $\Longrightarrow$ (ii) follows from the definitions.

(ii) $\Longrightarrow$ (iii) Given an edge $e \in E$, let G' denote the graph obtained from G by adding e' in parallel with e. One checks easily that $|Z| \leqslant 2|V_Z| - 2$ for every subset Z of edges of G'. Hence, the set $E \cup \{e'\}$ is independent in the matroid defined as the union of two copies of the graphic matroid of G'; that is, $E \cup \{e'\}$ is union of two forests.

(iii) $\Longrightarrow$ (i) It suffices to show the existence of $p \in \mathbb{R}^{2n}$ for which the rigidity matrix $R(G,p)$ has rank $|E|$. We start with a preliminary result. Given a graph $H = (V, E)$ and vectors $d_e \in \mathbb{R}^2$ $(e \in E)$, let $S(H,d)$ denote the $|E| \times 2|V|$ matrix whose entries in the row indexed by edge $e = ij$ are d_e at column i, $-d_e$ at column j and 0 elsewhere. Then, one can check that $S(H,d)$ has full rank $|E|$ if H is the union of two forests and d is generic. Let $d \in \mathbb{R}^{2|E|}$ be generic. By applying the above fact to the graph H obtained from G by adding an edge in parallel with an edge e of G, we obtain that there exist vectors $u_1^e, \ldots, u_n^e \in \mathbb{R}^2$ such that $u_i^e \neq u_j^e$ and $d_f^{\mathrm{T}}(u_h^e - u_k^e) = 0$ for all edges $f = hk \in E$. By taking a suitable linear combination of the vectors u^e $(e \in E)$, we can find $u = (u_1, \ldots, u_n) \in \mathbb{R}^{2n}$ such that $u_i \neq u_j$ and $d_e^{\mathrm{T}}(u_i - u_j) = 0$ for every edge $e = ij \in E$. If $u_i = (x_i, y_i)$, set $p_i := (-y_i, x_i)$ for $i = 1, \ldots, n$. Then, $(p_i - p_j)^{\mathrm{T}}(u_i - u_j) = 0$ and $d_e^{\mathrm{T}}(u_i - u_j) = 0$ for every edge $e = ij \in E$. Hence, $p_i - p_j$ and d_e are both orthogonal to the nonzero vector $u_i - u_j$. Therefore, $p_i - p_j = \beta_e d_e$ for some scalar $\beta_e \neq 0$. Since the matrix $S(G, d)$ has rank $|E|$, we deduce that the matrix $R(G, p)$ too has rank $|E|$, which shows that the set E is generic independent. □

One can check in polynomial time whether a graph is union of two forests, using the matroid partition algorithm of Edmonds [15] (or the faster algorithm by Gabow and Westermann [17]); thus one can test generic independence of E by running $|E|$ times this algorithm. The next result follows as an immediate consequence of Proposition 19; it was first obtained by Laman [31].

Proposition 20. *The following assertions are equivalent for a graph $G = (V, E)$ with at least two nodes.*

(i) *G is minimally rigid (that is, G is rigid and $G \setminus e$ is not rigid for every edge e of G).*

(ii) *G is rigid and E is generic independent.*

(ii) *$|E| = 2|V| - 3$ and $|Y| \leqslant 2|V_Y| - 3$ for every nonempty subset $Y \subseteq E$.*

Even though rigidity in the plane can be characterized by some purely graph theoretical properties, it is yet of interest to try to relate it to some other graph features such as connectivity. Clearly, every rigid graph in $\mathbb{R}^2$ must be 2-connected. Indeed, if G has a cut node, then one can continuously deform the graph by rotating around this node while preserving edge lengths, i.e., one can find a flexing. On the other hand, the graph $G_2 = K_4 \setminus e$ from Fig. 2 is minimally rigid while not 3-connected. Lovász and Yemini [40] show that 6-connectivity suffices, in fact, for ensuring rigidity.

Theorem 21. *Every 6-connected graph is rigid in the plane.*

They also give an example of a 5-connected graph which is not rigid; cf. Fig. 4. (Nonrigidity can be demonstrated with the help of Corollary 18, taking as classes of the partition the eight complete graphs K_5 and the remaining 20 edges.)

6.3. Rigidity in the space

In contrast with the case of the plane, no characterization is known for rigid graphs in the space $\mathbb{R}^k$ for $k \geqslant 3$. Some necessary conditions can be easily derived from the treatment above. First, k-connectivity is an obvious necessary condition for rigidity in $\mathbb{R}^k$. By the definitions, a graph $G = (V_n, E)$ is rigid in $\mathbb{R}^k$ if and only if there exists a subset $F \subseteq E$ such that F is generic independent with $|F| = S(n, k)$. ($S(n, k)$ is defined in Proposition 15.) Therefore, if G is rigid in $\mathbb{R}^k$, then there exists $F \subseteq E$ such that $|F| = S(n, k)$ and $|Y| \leqslant S(|V_Y|, k)$ for all $\emptyset \neq Y \subseteq F$. As we saw above, this condition is sufficient for ensuring rigidity in the case $k = 1$ (then it is equivalent to G being

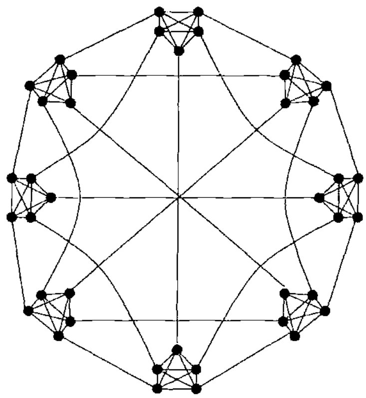

Fig. 4. A 5-connected graph nonrigid in $\mathbb{R}^2$.

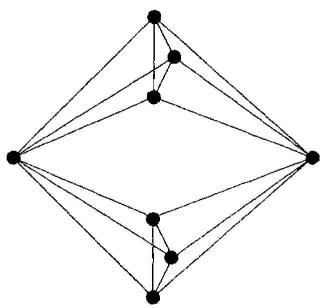

Fig. 5. A nonrigid graph in $\mathbb{R}^3$.

connected) and $k = 2$ (by Laman's result). However, sufficiency is lost if $k \geqslant 3$. For instance, the graph from Fig. 5 satisfies the necessary condition for $k = 3$ but is not rigid in $\mathbb{R}^3$ (as it has a node cutset of size 2).

Characterizing rigidity in $\mathbb{R}^3$ seems a hard problem. However some results are known for some classes of graphs. For instance, Roth [45] could characterize the planar graphs that are rigid in $\mathbb{R}^3$. Such graphs are necessarily 3-connected and, thus, arise from 3-dimensional polytopes.

Theorem 22. *Let P be a convex polytope in $\mathbb{R}^3$ and let (G,p) be the associated framework, where G is the 1-skeleton graph of P and p consists of the vertices of P. Then,* $\operatorname{rank} R(G,p) = m$, *the number of edges of P. In particular, (G,p) is rigid if and only if every face of P is a triangle.*

Corollary 23. *A planar graph is rigid in $\mathbb{R}^3$ iff it is 3-connected and triangulated.*

Bolker and Roth [9] have investigated the class of complete bipartite graphs $K_{m,n}$ $(1 \leqslant m \leqslant n)$. Clearly, if $mn < (m+n)k - \binom{k+1}{2}$, then $K_{m,n}$ is flexible in $\mathbb{R}^k$ (since rank $R(G,p) \leqslant mn$). For $k=3$, this yields that $K_{m,n}$ is flexible in $\mathbb{R}^3$ except if $m=n=1$, or $m=4$ and $n \geqslant 6$, or $m,n \geqslant 5$ in which cases $K_{m,n}$ can be shown to be rigid in $\mathbb{R}^3$ [9].

To conclude, let us briefly go back to the question of unicity of graph realizations considered in Section 5.2. Rigidity and $(k+1)$-connectivity of the underlying graph are obvious necessary conditions for a framework (G,p) to have a unique realization. Hendrickson [24] shows that a stronger necessary condition is that G must be *redundantly rigid*, which means that G remains rigid after deletion of any single edge. He shows, moreover, that redundant rigidity is a generic property. More precisely, if G is not redundantly rigid with at least $k+1$ vertices then, for every generic $p \in \mathbb{R}^{kn}$, the framework (G,p) has another realization in $\mathbb{R}^k$ not congruent to p.

Hence, there are many open questions concerning rigid graphs. An important open problem is the characterization of the rigid graphs in the space $\mathbb{R}^3$ (more generally, $\mathbb{R}^k$, $k \geqslant 3$) and determining the complexity status of this problem. Another open question is to characterize the frameworks for which Hendrickson's conditions ($(k+1)$-connectivity and redundant rigidity) suffice for ensuring the unicity of the realization.

References

[1] F. Alizadeh, Interior point methods in semidefinite programming with applications in combinatorial optimization, *SIAM Journal on Optimization* 5 (1995) 13–51.

[2] M. Bakonyi and C.R. Johnson, The Euclidian distance matrix completion problem, *SIAM Journal on Matrix Analysis and Applications* 16 (1995) 646–654.

[3] F. Barahona, On cuts and matchings in planar graphs, *Mathematical Programming* 60 (1993) 53–68.

[4] F. Barahona and A.R. Mahjoub, On the cut polytope, *Mathematical Programming* 36 (1986) 157–173.

[5] W.W. Barrett, C.R. Johnson and R. Loewy, The real positive definite completion problem: Cycle completability, *Memoirs of the American Mathematical Society* 584 (1996) 69 pp.

[6] W. Barrett, C.R. Johnson and P. Tarazaga, The real positive definite completion problem for a simple cycle, *Linear Algebra and its Applications* 192 (1993) 3–31.

[7] A.I. Barvinok, Problems of distance geometry and convex properties of quadratic maps, *Discrete and Computational Geometry* 13 (1995) 189–202.

[8] L.M. Blumenthal, *Theory and Applications of Distance Geometry* (Oxford University Press, Oxford, 1953; Chelsea, New York, 2nd ed., 1970).

[9] E.D. Bolker and B. Roth, When is a bipartite graph a rigid framework? *Pacific Journal of Mathematics* 90 (1980) 27–44.

[10] J.P.R. Christensen and J. Vesterstrøm, A note on extreme positive definite matrices, *Mathematische Annalen* 244 (1979) 65–68.

[11] G.M. Crippen and T.F. Havel, *Distance Geometry and Molecular Conformation* (Research Studies Press, Taunton, Somerset, UK, 1988).

[12] M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Algorithms and Combinatorics, Vol. 15 (Springer, Berlin, 1997).

[13] G.A. Dirac, On rigid circuit graphs, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 25 (1961) 71–76.

[14] R. Duffin, Topology of series-parallel networks, *Journal of Mathematical Analysis and Applications* 10 (1965) 303–318.

[15] J. Edmonds, Minimum partition of a matroid into independent subsets, *Journal of Research of the National Bureau of Standards B* 69 (1965) 67–72.

[16] L. Fejes Tóth, Über eine Punktverteilung auf der Kugel, *Acta Mathematica Academiae Scientiarum Hungaricae* 10 (1959) 13–19.

[17] H.N. Gabow and H.H. Westermann, Forests, frames and games: Algorithms for matroid sums and applications, in: *Proc. 20th Ann. ACM Symp. on the Theory of Computing* (1988) 407–421.
[18] M.X. Goemans and D.P. Williamson, 0.878-approximation algorithms for MAX CUT and MAX 2SAT, in: *Proc. 26th Ann. ACM Symp. on the Theory of Computing* (1994) 422–431.
[19] R. Grone, C.R. Johnson, E.M. Sá and H. Wolkowicz, Positive definite completions of partial hermitian matrices, *Linear Algebra and its Applications* 58 (1984) 109–124.
[20] R. Grone, S. Pierce and W. Watkins, Extremal correlation matrices, *Linear Algebra and its Applications* 134 (1990) 63–70.
[21] M. Grötschel, L. Lovász and A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica* 1 (1981) 169–197.
[22] M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Algorithms and Combinatorics, Vol. 2 (Springer, Berlin, 1988).
[23] B. Hendrickson, The molecule problem: Determining conformation from pairwise distances, Ph.D. Thesis, Tech. Rept. 90-1159, Dept. of Computer Science, Cornell University, Ithaca, NY, 1990.
[24] B. Hendrickson, Conditions for unique graph realizations, *SIAM Journal on Computing* 21 (1992) 65–84.
[25] B. Hendrickson, The molecule problem: Exploiting structure in global optimization, *SIAM Journal on Optimization* 5 (1995) 835–857.
[26] C.R. Johnson, Matrix completion problems: a survey, in: C.R. Johnson, ed., *Matrix Theory and Applications*, Proceedings of Symposia in Applied Mathematics, Vol. 40 (American Mathematical Society, Providence, RI, 1990) 171–198.
[27] C.R. Johnson, C. Jones and B. Kroschel, The distance matrix completion problem: Cycle completability, *Linear and Multilinear Algebra* 39 (1995) 195–207.
[28] C.R. Johnson, B. Kroschel and H. Wolkowicz, An interior-point method for approximate positive semidefinite completions, CORR Rept. 95-11, University of Waterloo, 1995, to appear in *Computational Optimization and Applications*.
[29] C.R. Johnson and T.A. McKee, Structural conditions for cycle completable graphs, *Discrete Mathematics* 159 (1996) 155–160.
[30] R.M. Karp, Reducibility among combinatorial problems, in: R.E. Miller and J.W. Thatcher, eds., *Complexity of Computer Computations* (Plenum Press, New York, 1972) 85–103.
[31] G. Laman, On graphs and rigidity of plane skeletal structures, *Journal of Engineering Mathematics* 4 (1970) 331–340.
[32] M. Laurent, The real positive semidefinite completion problem for series-parallel graphs, *Linear Algebra and its Applications* 252 (1997) 347–366.
[33] M. Laurent, A connection between positive semidefinite and Euclidean distance matrix completion problems, *Linear Algebra and its Applications*(1997), to appear.
[34] M. Laurent, A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems, in: P. Pardalos and H. Wolkowicz, eds., *Topics in Semidefinite and Interior-Point Methods*, The Fields Institute for Research in Mathematical Science, Communications Series, Providence, RI (1997), to appear.
[35] M. Laurent, Max-cut problem, in: M. Dell'Amico, F. Maffioli and S. Martello, eds., *Annotated Bibliographies in Combinatorial Optimization* (John Wiley & Sons, New York, 1997), to appear.
[36] M. Laurent and S. Poljak, On a positive semidefinite relaxation of the cut polytope, *Linear Algebra and its Applications* 223–224 (1995) 439–461.
[37] J. de Leeuw and W. Heiser, Theory of multidimensional scaling, in: P.R. Krishnaiah and L.N. Kanal, eds, *Handbook of Statistics*, Vol. 2 (North-Holland, Amsterdam, 1982) 285–316.
[38] Chi-Kwong Li and Bit-Shun Tam, A note on extreme correlation matrices, *SIAM Journal on Matrix Analysis and its Appplications* 15 (1994) 903–908.
[39] R. Loewy, Extreme points of a convex subset of the cone of positive semidefinite matrices, *Mathematische Annalen* 253 (1980) 227–232.
[40] L. Lovász and Y. Yemini, On generic rigidity in the plane, *SIAM Journal on Algebraic and Discrete Methods* 3 (1982) 91–98.
[41] Y.E. Nesterov and A.S. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming: Theory and Algorithms* (SIAM Publications, Philadelphia, PA. 1994).
[42] P.M. Pardalos and X. Lin, A tabu based pattern search method for the distance geometry problem, in: F. Giannessis et al., eds., *Mathematical Programming* (Kluwer Academic Publishers, Dordrecht, 1997).

[43] S. Poljak and Z. Tuza, Maximum cuts and largest bipartite subgraphs, in: W. Cook, L. Lovász and P.D. Seymour, eds., *Combinatorial Optimization*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 20 (AMS, Providence, RI, 1995) 181–244.

[44] M.V. Ramana, An exact duality theory for semidefinite programming and its complexity implications, *Mathematical Programming* 77 (1997) 129–162.

[45] B. Roth, Rigid and flexible frameworks, *The American Mathematical Monthly* 88 (1981) 6–21.

[46] D.J. Rose, R.E. Tarjan and G.S. Lueker, Algorithmic aspects of vertex elimination on graphs, *SIAM Journal on Computing* 5 (1976) 266–283.

[47] J.B. Saxe, Embeddability of weighted graphs in k-space is strongly NP-hard, in: *Proc. 17th Allerton Conf. in Communications, Control and Computing* (1979) 480–489.

[48] I.J. Schoenberg, Remarks to M. Fréchet's article "Sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert", *Annals of Mathematics* 36 (1935) 724–732.

[49] I.J. Schoenberg, Metric spaces and positive definite functions, *Transactions of the American Mathematical Society* 44 (1938) 522–536.

[50] R.E. Tarjan, Decomposition by clique separators, *Discrete mathematics* 55 (1985) 221–232.

[51] J.A. Wald and C.J. Colbourn, Steiner trees, partial 2-trees and minimum IFI networks, *Networks* 13 (1983) 159–167.

[52] W. Whiteley, Matroids and rigid structures, in: N. White, ed., *Matroid Applications*, Encyclopedia of Mathematics and its Applications, Vol. 40 (Cambridge University Press, Cambridge, MA, 1992) 1–53.

[53] K. Wüthrich, Protein structure determination in solution by nuclear magnetic resonance spectroscopy, *Science* 243 (1989) 45–50.

[54] Y. Yemini, Some theoretical aspects of position-location problems, in: *Proc. 20th Ann. Symp. on Foundations of Computer Science* (1979) 1–8.