

An Algebraic Geometry Approach to Viewing Graph Solvability

Federica Arrigoni, Kathlén Kohn, Andrea Fusiello, Tomas Pajdla

April 7, 2025

Abstract

The concept of viewing graph solvability has gained significant interest in the context of structure-from-motion. A viewing graph is a mathematical structure where nodes are associated to cameras and edges represent the epipolar geometry connecting overlapping views. Solvability studies under which conditions the cameras are uniquely determined by the graph. In this paper we propose a novel framework for analyzing solvability problems based on Algebraic Geometry, demonstrating its potential in understanding structure-from-motion graphs and proving a conjecture that was previously proposed.

1 Introduction

In recent years, there has been a notable increase in interest surrounding the concept of viewing graph solvability in the field of Computer Vision [1–6]. This concept plays a pivotal role in the domain of structure-from-motion (SfM) [7–11], which aims to reconstruct three-dimensional scenes from a multitude of images. A *viewing graph* [1] is a mathematical structure in which the nodes represent the cameras that capture the scene and the edges connect the cameras that have overlapping views. More precisely, an edge is present between two nodes if and only if it is possible to estimate the geometric relationship between the two cameras, encoded in the fundamental matrix (assuming an uncalibrated scenario). This defines a constraint system that is classically considered *solvable* if the information encoded in the fundamental matrices uniquely determines all cameras in the scene, up to a global projective transformation (see Figure 1). Despite significant advances have been recently made both from the theoretical and practical point of view [5, 6], viewing graph solvability still presents open issues, as discussed in the next subsection.

1.1 Related Work

It is well known that a *single* fundamental matrix uniquely determines the two perspective cameras up to a projective transformation [12]. However, when considering *multiple* fundamental matrices attached to the edges of a viewing

graph, there may be cases with many solutions or no solution at all.

A viewing graph is called *solvable* if, for almost all choices of cameras, there are no other sets of cameras yielding the same fundamental matrices (up to global projective transformation). In other terms, it is assumed that a solution exists (i.e., a set of cameras compliant with the given fundamental matrices), and the question is whether such solution is the only one or there are more. The concept of *solving* viewing graph (later called *solvable* by [4]) was first introduced in [1] where small incomplete graphs (up to six cameras) were manually analyzed, by reasoning in terms of how to uniquely recover the missing fundamental matrices from the available ones.

The authors of [1] also derived a *necessary condition* for solvability, namely the property that all the nodes have degree at least two and no two adjacent nodes have degree two. Later, additional necessary conditions were developed: a solvable graph must be biconnected [4]; it must have at least $(11n - 15)/7$ edges, with n being the number of nodes [4]; it must be bearing rigid [13]. The latter means that, as expected, a graph that is solvable with unknown intrinsic parameters is also solvable when they are known [14–16].

Sufficient conditions are also available: in [3] it is proved that those graphs which are constructed from a 3-cycle by adding nodes of degree 2 one at a time are solvable; [4] introduces specific “moves” which can be applied to a graph, possibly transforming it into a complete one, in which case the graph is solvable.

In fact, necessary or sufficient conditions alone are not enough to classify all possible cases so a *characterization* is required. In this respect, the authors of [4, 5] study solvability using principles from Algebraic Geometry. Specifically, a polynomial system of equations was derived in [4], so that solvability can be tested by counting the number of solutions of the system with algebraic geometry tools (e.g., Gröbner basis computation). Building on [4], the authors of [5] improve efficiency by deriving a simplified polynomial system with fewer unknowns. Still, the largest example tested in [5] is a graph with 90 nodes, which is far from the size of structure-from-motion datasets appearing in practice.

The main drawback of [4, 5] is that solving polynomial equations is computationally highly demanding, therefore

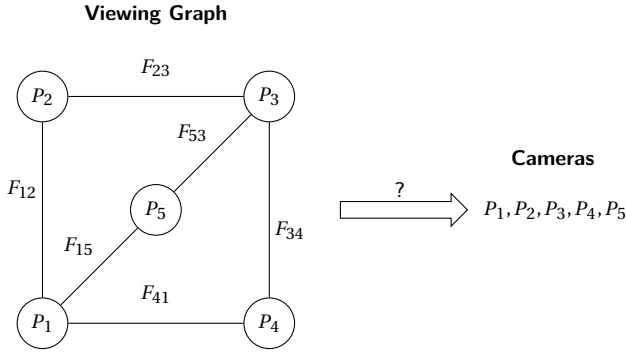


Figure 1: The solvability problem considers the following theoretical question: given a set of fundamental matrices encoded in a graph, how many camera configurations are compliant with such fundamental matrices?

limiting the practical usage of this characterization of solvability. For this reason, the related notion of *finite solvability* has been explored [4]. Specifically, a graph is called *finitely solvable* if, for almost all choices of cameras, there is a *finite* set of cameras that gives the same fundamental matrices (up to global projective transformation). This concept represents a proxy for (unique) solvability since it does not exclude the presence of more than one solution (e.g., two distinct solutions); however, it has been shown to be more practical since it can be deduced from the rank of a suitable matrix. Later, the authors of [6] improved the efficiency of this formulation and developed a method to partition an unsolvable graph into maximal components that are finitely solvable.

Problems related to solvability, which are not addressed in this paper, include the *compatibility* of fundamental matrices, namely, whether a camera configuration exists that produce the given fundamental matrices [17, 18], and the practical task of retrieving cameras from fundamental matrices [19–22].

1.2 Contribution

In the wake of the emerging field of Algebraic Vision [23], in this work we advance the understanding of viewing graphs by focusing on the notion of *finite solvability*. The main contributions can be summarized as follows:

- We derive a new formulation of the problem that is more direct (hence more *intuitive*) than previous work, as our equations explicitly involve cameras and fundamental matrices. Previous sets of equations [4, 6], instead, are harder to interpret, as they involve unknown projective transformations representing the problem ambiguities.
- We show that, by evaluating the rank of the *Jacobian*

matrix of our polynomial equations in a fabricated solution, we can test finite solvability. It is not immediate that this Jacobian check can assess the presence of a *finite* number of solutions overall, being designed as a local analysis. Our proof, based on the Fiber Dimension Theorem [24], confirms a conjecture made in our preliminary work [25].

- Our method for testing finite solvability naturally extends to an algorithm for *graph partitioning* into the maximal components that are finite solvable, to be applied to unsolvable cases with infinitely many solutions. The number of unknowns depends on the number of nodes in the graph, that are typically much inferior to the number of edges used by previous work [4, 6]. This permits us to set the state of the art in terms of *efficiency* on large graphs coming from SfM datasets.

This paper is an extended version of our preliminary study [25]. The manuscript is organized as follows. Section 2 reviews relevant background on solvability and finite solvability. Section 3 presents our theoretical contributions and introduces the set of polynomial equations employed in our formulation. Section 4 details our approach for testing finite solvability and extracting maximal components. Section 5 reports formulas for derivatives, useful both for theory and practice. Experiments on synthetic and real viewing graphs are reported in Section 6, while the conclusion is drawn in Section 7.

2 Background

Let P_1, \dots, P_n denote n uncalibrated cameras, represented by 3×4 full-rank matrices up to scaling, identified with elements of \mathbb{P}^{11} . Let $G = (V, E)$ be an undirected graph with node set $V = \{1, \dots, n\}$ and edge set $E \subseteq \{1, \dots, n\} \times \{1, \dots, n\}$ representing a *viewing graph* of an uncalibrated structure-from-motion problem. We denote the cardinality of the vertex set with $n = |V|$, the number of edges with $m = |E|$ and the fundamental matrix of $(i, j) \in E$ with F_{ij} . We use the following terminology: given a graph $G = (V, E)$, a *configuration* is a map $\mathcal{P} : V \rightarrow \mathbb{P}^{11}$ (full rank) that assigns nodes to cameras. A *framework* is a pair (G, \mathcal{P}) , where $G = (V, E)$ is a graph and \mathcal{P} is a configuration.

Fundamental matrices are equivalence classes of rank two 3×3 matrices up to non-zero scaling, identified with elements of \mathbb{P}^8 . They are assigned to edges via the map $\mathcal{F}_G(\mathcal{P}) = [\dots \mathcal{F}(P_i, P_j) \dots]$ where $\mathcal{F}(P_i, P_j)$ evaluates the fundamental matrix F_{ij} on edge $(i, j) \in E$. Hence:

$$\mathcal{F} : \mathbb{P}^{11} \times \mathbb{P}^{11} \rightarrow \mathbb{P}^8, \quad (P_i, P_j) \mapsto F_{ij}, \quad (1)$$

where F_{ij} is the fundamental matrix defined by cameras P_j and P_i . One way of specifying this map entry-wise is:

$$[F_{ij}]_{h,k} = (-1)^{h+k} \det \begin{bmatrix} P_i^k \\ P_j^h \end{bmatrix}, \quad (2)$$

where P_i^k denotes the 2×4 sub-matrix of camera P_i obtained by removing row k (and similarly for P_j^h with row h). Note that Eq. (2) gives the zero matrix if P_i and P_j have coincident centres, meaning that the map \mathcal{F} is undefined in the projective sense: in this scenario, it is known that the fundamental matrix is not uniquely defined [12]. Therefore, we assume henceforth that cameras have distinct centres.

In particular, since (2) is a polynomial, we see that this is an *algebraic map*, i.e., a function between algebraic varieties given locally by rational functions. For us, an *algebraic variety* is the solution set of a system of polynomial equations.

The **key question** is the following:

given a framework (G, \mathcal{P}_0) , how many configurations \mathcal{P} exist yielding the same fundamental matrices?

In algebraic terms, we want to study the cardinality of the fibers (i.e., pre-images of points) of \mathcal{F}_G ; hence the question can be rephrased as:

what is the cardinality of $\mathcal{F}_G^{-1}(\mathcal{F}_G(\mathcal{P}))$?

In formulating this question, we identify all configurations that are projectively equivalent. For instance, if we state that a configuration is unique, this is always intended up to a global projective transformation, which is an element of PGL_4 , the Projective General Linear Group on \mathbb{P}^3 .

Definition 1 (Solvable framework [4]). Let $\mathcal{P} = \{P_1, \dots, P_n\}$ be a configuration of cameras, and let G be a graph. The framework (G, \mathcal{P}) is called *solvable* if all camera configurations yielding the same fundamental matrices as \mathcal{P} are obtained from \mathcal{P} via a global projective transformation. In other words, $\mathcal{F}_G^{-1}(\mathcal{F}_G(\mathcal{P}))$ is a single point, modulo PGL_4 .

Studying the solvability of frameworks requires considering the actual camera configuration and accounting for special cases, such as collinear centers. To avoid this, a generic configuration is typically considered, leading to another concept of solvability, which is a property of the graph itself.

Definition 2 (Solvable graph[4]). A graph G is called *solvable* if it is solvable for a *generic* configuration of cameras. In other words, G is solvable if and only if, generically, the non-empty fibers of \mathcal{F}_G are points, modulo PGL_4 .

Solvability, in this context, does not concern finding a specific solution (i.e., a camera configuration producing given fundamental matrices). Rather, it focuses on counting the solutions, assuming at least one solution exists. This interpretation is consistent with prior research [4, 13].

Determining the solvability of a graph requires solving a polynomial system of equations [4, 13]. This process is computationally demanding, rendering it prohibitive for large or dense graphs often encountered in practice. A relaxed notion is *finite solvability*, requiring a finite number of solutions, as opposed to one solution.

Definition 3 (Finite solvable graph[4]). A graph G is called *finite solvable* if and only if, generically, the non-empty fibers of \mathcal{F}_G are finite, modulo PGL_4 .

Remark 1. Checking the finite solvability of graphs is computationally more feasible because it can be checked locally. In the formulation of [4], polynomial equations are derived by reasoning on the problem ambiguities, whose solution set forms a smooth algebraic variety endowed with a group structure. This property implies that the dimension of this variety coincides with the dimension of its tangent space at the identity, which can be computed efficiently. Specifically, since the tangent space is a linear space, its dimension reduces to the rank of a linear system of equations. This dimension reveals whether the original polynomial system admits a finite number of solutions or, equivalently, whether the graph is finitely solvable.

Our approach targets the same notion of finite solvability but with a different polynomial system. In contrast to [4] (later improved by [6]), our equations directly involve cameras and fundamental matrices, thereby gaining efficiency and interpretability by design. However, this comes at the price of losing the group property (cameras are not invertible matrices), thus requiring a different mathematical approach.

Remark 2. Although finite solvability is only a necessary condition for solvability, it remains a valuable property. It can be interpreted as a *local solvability*, meaning that the solution is unique within a neighborhood of the given configuration.

3 Theoretical Results

This section is devoted to our theoretical results, which set the basis for the proposed method for checking finite solvability. We first prove a new characterization of the problem and then detail our choice of polynomial equations.

3.1 Characterization of Finite Solvability

Our goal is to study the *generic* finite solvability of a graph G , i.e., we ask whether $\mathcal{F}_G^{-1}(\mathcal{F}_G(\mathcal{P}))$ for *generic* \mathcal{P} is a finite set or an infinite one (modulo PGL_4), which is equivalent to studying the dimension of $\mathcal{F}_G^{-1}(\mathcal{F}_G(\mathcal{P}))$. The dimension of an algebraic variety intuitively quantifies the number of independent parameters required to describe points on the variety, much like the dimension of a vector space or a manifold. While we omit a formal definition here, we note that a variety has dimension 0 if and only if it consists of a finite number of points [24, Chap. 1].

To be more concrete, one can assign random cameras \mathcal{P}_0 to nodes of G and compute the fundamental matrices using $\mathcal{F}_G(\mathcal{P}_0)$ with Equations (1) and (2). The task is then to determine how many camera configurations produce the same fundamental matrices as \mathcal{P}_0 , modulo PGL_4 . This is a *global* question, but it can be addressed through a *local* analysis,

made possible by Proposition 1 that we are going to prove at the end of the section, after recalling some results from Algebraic Geometry.

Lemma 1 (Fiber Dimension Theorem). *If $f : X \rightarrow Y$ is an algebraic map between irreducible varieties (over \mathbb{C}), then*

$$\dim X = \dim f^{-1}(f(x)) + \dim \text{im}(f) \quad (3)$$

for almost all $x \in X$, where im denotes the image of the map.

An *irreducible variety* is a variety that cannot be written as the union of two non-empty proper sub-varieties. Lemma 1 is known as the Fiber Dimension Theorem [24, Chap. 1.6.3]. It establishes that the fiber dimension is constant on generic points, and that this dimension is dual to the dimension of the image parameterized by the map. This relation extends the rank-nullity theorem from Linear Algebra to polynomial maps. In particular, it says that an algebraic map f either has (generically) finite fibers or it has generically infinite fibers. In other words, all generic fibers have the same dimension, hence the behavior of a single fiber is enough to get global information.

Another standard fact in algebraic geometry is that, at a generic point in the domain of an algebraic map, the rank of the Jacobian matrix equals the dimension of the image [24].

Lemma 2 (Lemma 2.4 in Chap. 2.6 of [24]). *If $f : X \rightarrow Y$ is an algebraic map between irreducible varieties, then, for almost all $x \in X$,*

$$\dim \text{im}(f) = \text{rank } Df(x). \quad (4)$$

From these two lemmas it follows immediately that:

Corollary 1. *If $f : X \rightarrow Y$ is an algebraic map between irreducible varieties (over \mathbb{C}), then*

$$\dim f^{-1}(f(x)) = \dim X - \text{rank } Df(x) \quad (5)$$

for almost all $x \in X$.

We are now able to characterize the finite solvability of a graph G in terms of the rank of the Jacobian matrix associated with \mathcal{F}_G , the function that computes the fundamental matrices along the edges of G .

In the following we are going to represent matrices in an affine chart, and consequently work with the affine version¹ of the map \mathcal{F} , denoted by $\mathcal{F}^{\text{aff}} : \mathbb{R}^{11} \times \mathbb{R}^{11} \rightarrow \mathbb{R}^8$. With a little abuse of notation we are not going to distinguish between a projective element and its affine representation, as the map where they appear will be enough to disambiguate.

¹One way of fixing an affine chart in the domain $\mathbb{P}^{11} \times \mathbb{P}^{11}$ is by setting one of the 12 entries in each camera matrix to 1. Similarly, in the codomain \mathbb{P}^8 , we can choose an affine chart by fixing one of the entries of the fundamental matrix to be 1, e.g., the very last entry. That way, the map \mathcal{F}^{aff} becomes a rational map: Its coordinate functions are fractions $\frac{[F_{ij}]_{h,k}}{[F_{33}]_{h,k}}$ of the polynomials in (2).

Proposition 1. *Let $\mathcal{F}_G^{\text{aff}} : (\mathbb{R}^{11})^{|V|} \rightarrow (\mathbb{R}^8)^{|E|}$ be the affine version of the algebraic map that computes the fundamental matrices along a viewing graph G with nodes V and edges E . It has a Jacobian $D\mathcal{F}_G^{\text{aff}}$ made of blocks of size 8×22 , defined as:*

$$[D\mathcal{F}_G^{\text{aff}}]_{i,j} := \frac{\partial \mathcal{F}_G^{\text{aff}}}{\partial P_i, P_j}. \quad (6)$$

Then, for a generic configuration \mathcal{P}_0 , we have:

$$\text{rank}(D\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)) = 11|V| - 15 \iff G \text{ is finite solvable.}$$

Proof. The domain of our map is the set of camera matrices $(\mathbb{R}^{11})^{|V|}$ (interpreted as 3×4 matrices). The map $\mathcal{F}_G^{\text{aff}}$ is well-defined on generic cameras in this domain. Since both domain $X = (\mathbb{R}^{11})^{|V|}$ and codomain $Y = (\mathbb{R}^8)^{|E|}$ of the map $\mathcal{F}_G^{\text{aff}}$ are linear spaces, they are irreducible. So $\mathcal{F}_G^{\text{aff}}$ is an algebraic map between irreducible varieties and Corollary 1 implies that:

$$\dim \mathcal{F}_G^{\text{aff}^{-1}}(\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)) = 11|V| - \text{rank } D\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0).$$

Now, finite solvability means that the generic non-empty fiber is finite modulo PGL_4 , i.e., the fiber is a union of finitely many copies of PGL_4 . Since the latter group has dimension 15, we obtain:

$$\text{finite solvable} \iff \dim \mathcal{F}_G^{\text{aff}^{-1}}(\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)) = 15,$$

hence we get the thesis. \square

3.2 Our Formulation

Our formulation employs a polynomial system where the only unknowns are the camera matrices, by using an implicit homogeneous constraint that links fundamental matrices to cameras (from [12, Chap. 9]). With respect to using the explicit map \mathcal{F} as defined in (2), which is indeed theoretically feasible, this approach yields lower-degree polynomials and eliminates the need to account for the projective scales.

Lemma 3 (Result 9.12 in [12]). *A non-zero matrix F_{ij} is the fundamental matrix corresponding to a pair of cameras P_i and P_j if and only if the matrix $S := P_j^T F_{ij} P_i$ is skew-symmetric.*

Note that any scaling of each of the three terms of the product would clearly leave the result skew-symmetric. The above condition can be rewritten as:

$$S + S^T = 0 \iff P_j^T F_{ij} P_i + P_i^T F_{ij}^T P_j = 0. \quad (7)$$

Since (7) is symmetric, it translates into 10 quadratic equations when considered entry-wise. Observe that these equations have not been used in previous works on solvability [4–6].

We write Sym_4 for the vector space of real symmetric 4×4 matrices, and $\mathbb{P}\text{Sym}_4$ for its projectivization. Note that the

latter is isomorphic to \mathbb{P}^9 since $\dim \text{Sym}_4 = 10$. Let us define:

$$\Phi : \mathbb{P}^{11} \times \mathbb{P}^{11} \times \mathbb{P}^8 \rightarrow \mathbb{P} \text{Sym}_4 \cong \mathbb{P}^9, \quad (8)$$

$$(P_i, P_j, F) \mapsto P_j^\top F P_i + P_i^\top F^\top P_j.$$

Note that Φ is homogeneous in each of its inputs. Lemma 3 states that there is a unique F (in the projective space) such that $\Phi(P_i, P_j, F) = 0$, and this is the fundamental matrix corresponding to the camera pair (P_i, P_j) . In formulae:

$$\Phi(P_i, P_j, F) = 0 \iff F = \mathcal{F}(P_i, P_j) = F_{ij}. \quad (9)$$

Note that Eq. (7) holds for a single edge $(i, j) \in E$. By collecting equations coming from all the edges in the graph G , it results in a *polynomial system*

$$\Phi_G((P_i)_{i \in V}, (F_e)_{e \in E}) = 0. \quad (10)$$

with $\Phi_G : (\mathbb{P}^{11})^{|V|} \times (\mathbb{P}^8)^{|E|} \rightarrow (\mathbb{P}^9)^{|E|}$. Specifically, since we start from fundamental matrices given by a generic configuration \mathcal{P}_0 , our polynomial system is

$$\Phi_G(\mathcal{P}, \mathcal{F}_G(\mathcal{P}_0)) = 0 \quad (11)$$

with unknowns \mathcal{P} . It is clear from the definitions (and Lemma 3) that this system has a unique solution (equal to \mathcal{P}_0) if and only if $\mathcal{F}_G^{-1}(\mathcal{F}_G(\mathcal{P}_0)) = \{\mathcal{P}_0\}$ (modulo PGL_4), which is tantamount to saying that G is solvable.

Similarly to before, we restrict the maps Φ and Φ_G to affine charts. But since Φ vanishes on corresponding camera pairs and their fundamental matrices, this time we do not restrict the codomain to an affine chart (otherwise, we would work with fractions with vanishing denominator; cf. footnote¹). We denote the affine versions of our maps by $\Phi^{\text{aff}} : (\mathbb{R}^{11})^2 \times \mathbb{R}^8 \rightarrow \mathbb{R}^{10}$ and $\Phi_G^{\text{aff}} : (\mathbb{R}^{11})^{|V|} \times (\mathbb{R}^8)^{|E|} \rightarrow (\mathbb{R}^{10})^{|E|}$. The following proposition links the rank of $D\mathcal{F}_G^{\text{aff}}$ to that of the Jacobian of Φ_G^{aff} with respect to cameras, thereby establishing an alternative characterization of finite solvability, which we formalize later in Theorem 1.

Proposition 2. *The map $\mathcal{F}_G^{\text{aff}}$ is implicitly defined by Φ_G^{aff} (in a neighborhood of a solution). Moreover, for a generic configuration \mathcal{P}_0 with fundamental matrices $\mathcal{F}_0 := \mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)$, we have:*

$$\text{rank} \left(\frac{\partial \Phi_G^{\text{aff}}}{\partial (P_i)_{i \in V}}(\mathcal{P}_0, \mathcal{F}_0) \right) = \text{rank}(D\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)). \quad (12)$$

Proof. Consider the function Φ defined in (8). The Jacobian² of Φ , denoted by $D\Phi$, can be reorganized as:

$$D\Phi = \begin{array}{|c|c|} \hline 24 & 9 \\ \hline \frac{\partial \Phi}{\partial P_i, P_j} & \frac{\partial \Phi}{\partial F} \\ \hline \end{array} \quad 10$$

²In fact, this is the Jacobian of $\tilde{\Phi} : (\mathbb{R}^{12})^2 \times \mathbb{R}^9 \rightarrow \mathbb{R}^{10}$, the homogeneous map that induces Φ by identifying collinear points within their projective equivalence classes. However, we omit this distinction to maintain notational simplicity.

Since Φ is linear in F , this means that, for fixed cameras P_i and P_j with distinct centers, the 10×9 matrix $\frac{\partial \Phi}{\partial F}$ has rank 8 with kernel given by $\text{span}\{F\}$.

Since Φ is homogeneous in each of its inputs, we can think of the matrices P_i, P_j, F and $\Phi(P_i, P_j, F)$ in their respective projective spaces instead. Recall that the tangent space of \mathbb{P}^8 at F is the quotient vector space $\mathbb{R}^{3 \times 3} / \text{span}\{F\}$ [24, Chap. 2]. That means, when restricting the domain of Φ to affine charts, which we denote by $\Phi^{\text{aff}} : \mathbb{R}^{11} \times \mathbb{R}^{11} \times \mathbb{R}^8 \rightarrow \mathbb{R}^{10}$, then the 10×8 Jacobian matrix $\frac{\partial \Phi^{\text{aff}}}{\partial F}$ is of full rank 8.

To turn this into an invertible matrix, we fix a generic 8×10 matrix A and consider the composition $A \circ \Phi^{\text{aff}} : \mathbb{R}^{11} \times \mathbb{R}^{11} \times \mathbb{R}^8 \rightarrow \mathbb{R}^8$. Since the Jacobian of Φ^{aff} is divided into two blocks of size 10×22 and 10×8 , then the Jacobian of $A \circ \Phi^{\text{aff}}$ has the following structure:

$$D A \circ \Phi^{\text{aff}} = \begin{array}{|c|c|} \hline 22 & 8 \\ \hline \frac{\partial A \circ \Phi^{\text{aff}}}{\partial P_i, P_j} & \frac{\partial A \circ \Phi^{\text{aff}}}{\partial F} \\ \hline \end{array} \quad 8$$

Since now $\frac{\partial A \circ \Phi^{\text{aff}}}{\partial F}$ is invertible, we can apply the Implicit Function Theorem: there is a function f defined and differentiable in some neighborhood of a solution, such that $A \circ \Phi^{\text{aff}}(P_i, P_j, f(P_i, P_j)) = 0$. This is not a surprise as we already know that \mathcal{F}^{aff} is this function f . However, the Implicit Function Theorem also tells us that the Jacobian of the function $f = \mathcal{F}^{\text{aff}}$ is given by

$$\frac{\partial \mathcal{F}^{\text{aff}}}{\partial P_i, P_j} = - \left(\frac{\partial A \circ \Phi^{\text{aff}}}{\partial F} \right)^{-1} \frac{\partial A \circ \Phi^{\text{aff}}}{\partial P_i, P_j}. \quad (13)$$

(Note that the Jacobian matrices in this equality should be evaluated at a generic camera pair and their corresponding fundamental matrices, just as in (12), but we skip this here for simpler notation.) Since $\frac{\partial A \circ \Phi^{\text{aff}}}{\partial F}$ is invertible, for a generic camera pair, the ranks of $\frac{\partial A \circ \Phi^{\text{aff}}}{\partial P_i, P_j}$ and $\frac{\partial \mathcal{F}^{\text{aff}}}{\partial P_i, P_j}$ are the same. Due to the genericity of the matrix A , this rank is the same as $\text{rank} \frac{\partial \Phi^{\text{aff}}}{\partial P_i, P_j}$.

Observe that, by (13), $\frac{\partial A \circ \Phi^{\text{aff}}}{\partial F}$ serves as a local coordinate change between the explicit coordinates of each fundamental matrix and its implicit coordinates in terms of the skew-symmetric matrix condition, as in the following diagram:

$$\begin{array}{ccc} \mathbb{R}^{11} \times \mathbb{R}^{11} & \xrightarrow{\frac{\partial \mathcal{F}^{\text{aff}}}{\partial P_i, P_j}} & \mathbb{R}^8 \\ & \searrow \frac{\partial A \circ \Phi^{\text{aff}}}{\partial P_i, P_j} & \swarrow \frac{\partial A \circ \Phi^{\text{aff}}}{\partial F} \\ & \mathbb{R}^8 & \end{array} \quad (14)$$

Finally, we consider the map $\mathcal{F}_G^{\text{aff}} : (\mathbb{R}^{11})^{|V|} \rightarrow (\mathbb{R}^8)^{|E|}$ that computes the fundamental matrices along a viewing graph G with nodes V and edges E . Similarly, we extend the function $A \circ \Phi^{\text{aff}}$ to $(A \times \dots \times A) \circ \Phi_G^{\text{aff}} : (\mathbb{R}^{11})^{|V|} \times (\mathbb{R}^8)^{|E|} \rightarrow (\mathbb{R}^8)^{|E|}$.

This gives local coordinate changes between the explicit and implicit coordinates of the fundamental matrices along each of the edges, and so as above we obtain:

$$\frac{\partial \mathcal{F}_G^{\text{aff}}}{\partial (P_i)_{i \in V}} = - \left(\frac{\partial A^{(|E|)} \circ \Phi_G^{\text{aff}}}{\partial (F_e)_{e \in E}} \right)^{-1} \frac{\partial A^{(|E|)} \circ \Phi_G^{\text{aff}}}{\partial (P_i)_{i \in V}}.$$

Hence, $\text{rank} \frac{\partial \mathcal{F}_G^{\text{aff}}}{\partial (P_i)_{i \in V}} = \text{rank} \frac{\partial \Phi_G^{\text{aff}}}{\partial (P_i)_{i \in V}}$ for generic P_i . \square

Finally, we establish our **main result**, which was demonstrated in only one direction in our conference paper [25].

Theorem 1. *A graph G is finite solvable if and only if $\text{rank} \left(\frac{\partial \Phi_G^{\text{aff}}}{\partial (P_i)_{i \in V}}(\mathcal{P}_0, \mathcal{F}_0) \right) = 11|V| - 15$ for a generic configuration \mathcal{P}_0 with fundamental matrices $\mathcal{F}_0 := \mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)$.*

Proof. It follows from Propositions 1 and 2. \square

The condition on the rank of the Jacobian, often referred to as the ‘‘Jacobian check’’ in Algebraic Vision [26], ensures that solutions are finite in the neighborhood of isolated solutions but does not provide guarantees for other fibers. The question of whether a statement can be made about almost all fibers was left open in [25]. Theorem 1 provides the answer, which ultimately relies on the Fiber Dimension Theorem. Figure 2 provides a summary of our results as well as connections between various solvability notions.

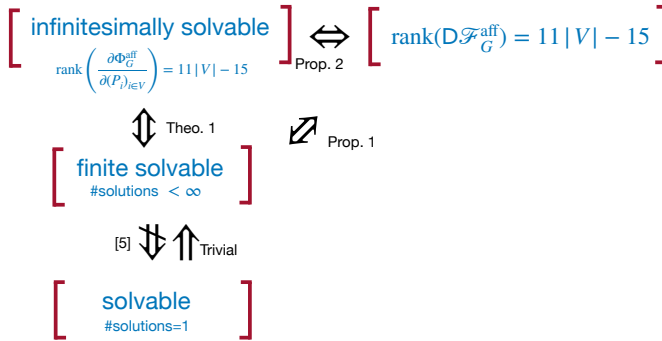


Figure 2: The connections among different concepts of solvability. The rank condition in Theorem 1 was called ‘‘infinitesimally solvable’’ in [25].

4 Proposed Method

In this section we show how Theorem 1 can be used in practice to test finite solvability. We also show how to partition an unsolvable graph into maximal subgraphs (called components) that are finite solvable.

4.1 Testing Finite Solvability

The conclusion of the previous section is that, in order to establish finite solvability of a viewing graph $G = (V, E)$, one can test if $r := \text{rank} \left(\frac{\partial \Phi_G^{\text{aff}}}{\partial (P_i)_{i \in V}} \right)$ is equal to $11|V| - 15$, where this $10|E| \times 11|V|$ -Jacobian is evaluated at a configuration \mathcal{P}_0 and fundamental matrices $\mathcal{F}_G^{\text{aff}}(\mathcal{P}_0)$ (please note that these fundamental matrices are compatible by construction). For computational reasons (that will be clarified in the end of this section), it is preferable to test whether a matrix is full rank rather than determining its exact rank. Therefore, we include 15 additional independent equations in order to fix a basis for PGL_4 , which raises the rank by 15 (making it full rank if and only if G is finite solvable).

In practice, we consider the Jacobian $J_P := \frac{\partial \Phi_G}{\partial (P_i)_{i \in V}}$, which has dimension $10|E| \times 12|V|$ and rank r . The rank is the same as above because it is the codimension of the tangent space of the variety defined by $\Phi_G(\mathcal{P}, \mathcal{F}_G(\mathcal{P})) = 0$ at \mathcal{P}_0 , and that codimension is the same no matter whether one looks at an affine chart or the affine cone over the projective variety. The affine chart is fixed by introducing one additional equation per camera, which raises the rank by $|V|$. Overall, the Jacobian J of the augmented polynomial system has rank $r + 15 + |V|$ and it achieves full-rank $12|V|$ if and only if $r = 11|V| - 15$.

Specifically, the global *projective ambiguity* is fixed, without loss of generality, by arbitrarily choosing the first camera and the first row in the second camera:

$$P_1 = [I_{3 \times 3} \quad 0_{3 \times 1}] \text{ and } [100]P_2 = [0001], \quad (15)$$

resulting in 16 additional equations. Note that $[1 \ 0 \ 0]P_2$ is equivalent to selecting the first row in P_2 . In fact, any pair of cameras can be chosen to fix the projective ambiguity. In practice, we will use two nodes that are endpoint of an edge in the graph (see Section 4.2).

Concerning the selection of the *affine chart*, the scale of each camera can be arbitrarily set, e.g., by fixing the sum of its entries to 1:

$$\mathbf{1}_{12}^T \text{vec}(P_i) = 1, \quad (16)$$

where $\mathbf{1}_{12}$ denotes a vector of ones of length 12. This results in a linear equation for each node, except the first camera used to fix the global ambiguity. In total we add $16 + |V| - 1 = 15 + |V|$ equations. In summary, equations of the form (10) (i.e., $\Phi_G = 0$), (15) and (16) are all collected in a polynomial system, for a total of $10|E| + (|V| - 1) + 16 = 10|E| + |V| + 15$ equations. The unknowns of our polynomial system are the camera matrices, for a total of $12|V|$ unknowns.

The Jacobian matrix of our polynomial system – denoted by J – contains J_P and the derivatives of (15) and (16). J_P is constructed by 10×12 blocks, whose formulas are given in Section 5 – see Equations (28) and (29). The block structure follows the incidence matrix B of the viewing graph, which has one row for every edge and one column for every node. In the row of B that represents the edge (i, j) , there is a -1 in column i and a $+1$ in column j and other entries are zero.

In J_P , the $+1$ is replaced by the 10×12 block (28) and the -1 is replaced by the 10×12 block (29):

$$\left[0 \cdots 0, \frac{\partial \Phi}{\partial (\text{vec } P_i)^\top}, 0 \cdots 0, \frac{\partial \Phi}{\partial (\text{vec } P_j)^\top}, 0 \cdots 0 \right]. \quad (17)$$

Matrices of the form (17) are then stacked for all the edges in the graph to make J_P . Note that this matrix is sparse because B is sparse. As for the derivatives of the additional equations that fix scales and projective ambiguity, they are constant matrices of zero and ones.

To summarize, in our implementation:

1. we assign random cameras \mathcal{P}_0 to nodes of G and compute the fundamental matrices using $\mathcal{F}_G(\mathcal{P}_0)$;
2. we then build the Jacobian as just explained;
3. we test finite solvability by checking whether J is of full rank.

The last point is accomplished by computing the smallest singular value of J , which in turn is equivalent to computing the smallest eigenvalue of $J^\top J$. Checking a given rank, instead, would entail computing more eigenvalues: a number equal to the kernel dimension.

Remark 3. Theorem 1 applies to generic camera configurations, meaning that we establish if the number of generic solutions of \mathcal{F}_G is finite or not. However, additional non-generic solutions may exist – for example, those corresponding to rank-deficient cameras, that symbolic solvers will find. If one is interested in counting all generic solutions then one should incorporate extra equations into the polynomial system to enforce full-rank camera conditions, as was done in [25].

4.2 Finding Maximal Components

We now show how to extract maximal finite-solvable components in the case where a viewing graph is established to be unsolvable. Proposition 3 from [6] states that such components form a *partition* of the edges. In other terms, each edge belongs to exactly one component. However, it is important to remark that a node can belong to more components. For example, an articulation point (or cut vertex) always belongs to two different components (see e.g. the blue and red components of Figure 3a which share a cut vertex).

For this reason, we can not trivially use the *edge-based* methodology from [6], for our formulation is *node-based* (i.e., our unknowns are associated with the nodes in the graph). Indeed, the null space of the Jacobian matrix J , which is nontrivial for an unsolvable graph, is such that any block of 12 rows corresponds to a camera/node in the viewing graph (whereas in [6] there was a correspondence between rows in the null space and edges in the graph). Luckily, it is still possible to identify components from the null space of J with proper modifications. In this context, an important observation is that we need two nodes to fix the

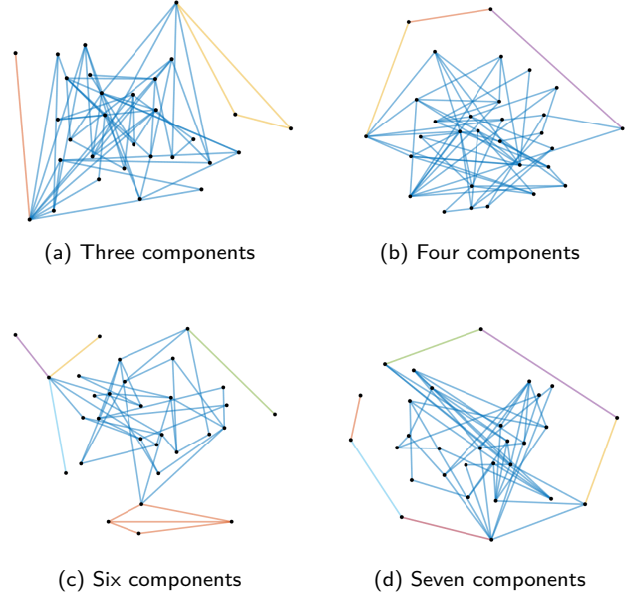


Figure 3: Examples of maximal components on synthetic viewing graphs, where each component is represented with a different color.

global projective ambiguity: in particular, for the purpose of identifying components, it is useful to select two adjacent nodes (i.e., an edge).

Lemma 4. *Let J be the Jacobian matrix constructed as explained above, and let N be the null space of J . A node is in the same component as the edge used to fix the global projective ambiguity \iff the associated rows in N are zero.*

Proof. Following the reasoning from [6], we can prove the thesis based on this observation: if we focus on the component containing the edge used to fix the global projective ambiguity, then the fact that all ambiguities have been fixed in that component, it is equivalent to say that there are no degrees of freedom, i.e., the null space is trivial on that component. \square

According to the above result, we can define an iterative approach to identify components:

- first, we use two adjacent nodes to fix the global projective ambiguity and identify all nodes within such component by selecting the ones corresponding to the zero rows in N ;
- then, we repeatedly apply the same procedure to the remaining part of the graph until there are no more edges to be assigned.

Observe that, although the null space is computed several times (equal to the number of components), the Jacobian matrix has a size that gets smaller and smaller, for the procedure is not re-applied to the whole graph but only to the subgraph containing the remaining edges.

Some visualizations of the components are given in Figure 3, showing established cases of unsolvable graphs, like the “square” topology (3b,3d) or the presence of pendant edges and articulation points (3a,3c,3d).

5 Formulas for Derivatives

In this section we report explicit formulas for the derivatives of our polynomial equations with respect to their unknowns, that are the basis of the Jacobian check implemented within our approach.

5.1 General Formulas

The derivatives of functions involving vectors and matrices ultimately lead back to the partial derivatives of the individual components, and it is all about how to arrange these partial derivatives. There are several conventions, here we follow [27], which allows to apply the chain rule.

Definition 4 ([27]). Let $f : \mathbb{R}^{s \times t} \rightarrow \mathbb{R}^{r \times q}$ be a differentiable function. The derivative of f in X is the matrix $r q \times s t$:

$$Df(X) = \frac{\partial \text{vec } f(X)}{\partial (\text{vec } X)^\top}. \quad (18)$$

where vec denotes the vectorization of the matrix by stacking its columns. In particular, if $f : \mathbb{R}^s \rightarrow \mathbb{R}^q$ then $Df(\mathbf{x})$ coincides with the usual *Jacobian matrix* of f .

We will use the following lemma [27].

Lemma 5. Assuming A, X and B are matrices of sizes $r \times s$, $s \times t$ and $t \times q$, respectively, then the derivatives of the following matrix functions in X are:

$$\begin{aligned} D(AX) &= (I_t \otimes A) \\ D(XB) &= (B^\top \otimes I_s) \\ D(AXB) &= (B^\top \otimes A) \end{aligned} \quad (19)$$

where \otimes denoted the *Kronecker product*.

The above result exploits the “vectorization trick” [28], stating that we can write $\text{vec}(AXB) = (B^\top \otimes A) \text{vec } X$. Furthermore, it can be also shown that [27]:

$$D(X^\top) = K_{s,t} \quad (20)$$

where $K_{s,t}$ is the *commutation matrix*, namely the $st \times st$ matrix such that $\text{vec}(C) = K_{s,t} \text{vec}(C^\top)$ for any C of size $s \times t$. Moreover, for A and B of sizes $r \times s$ and $t \times q$ respectively, we have [27]:

$$B \otimes A = K_{r,t} (A \otimes B) K_{q,s}. \quad (21)$$

The commutation matrix is a permutation, hence it is orthogonal: $K_{s,t} K_{s,t}^\top = I_{st}$. Note also that $K_{s,t} = K_{t,s}^\top$.

When vectorizing symmetric matrices, the vech operator is used to extract only the lower triangular part of the matrix. There exist unique matrices transforming the half-vectorization of a matrix to its vectorization and vice versa

called, respectively, the *duplication matrix* and the *elimination matrix* [27]. In particular for the latter we have: $\text{vech}(X) = L_q \text{vec}(X) \quad \forall X \in \text{Sym}_q(\mathbb{R})$.

The reader is referred to [28] for a review of results involving the Kronecker product and [27] for derivatives of matrix functions.

5.2 Derivatives of Φ with Respect of Cameras

Consider the map Φ defined in Eq. (8). Since the codomain is given by symmetric matrices, we consider – in practice – only the lower triangular part:

$$\Phi(P_i, P_j, F) := \text{vech} \left(P_j^\top F P_i + P_i^\top F_i^\top P_j \right) \quad (22)$$

where the vech operator vectorizes while extracting the non-duplicated entries. Let us define $S := P_j^\top F_i P_i$, then the map Φ rewrites:

$$\Phi(P_i, P_j, F) = \text{vech}(S + S^\top). \quad (23)$$

From the formulae (19) and (20) we get:

$$\frac{\partial \text{vec } S}{\partial (\text{vec } P_i)^\top} = I_4 \otimes (P_j^\top F_{ij}) \quad (24)$$

$$\frac{\partial \text{vec } S^\top}{\partial (\text{vec } P_i)^\top} = K_{4,4} (I_4 \otimes (P_j^\top F_{ij})) \quad (25)$$

$$\frac{\partial \text{vec } S}{\partial (\text{vec } P_j)^\top} = K_{4,4} (I_4 \otimes (F_{ij} P_i)^\top) \quad (26)$$

$$\frac{\partial \text{vec } S^\top}{\partial (\text{vec } P_j)^\top} = I_4 \otimes (F_{ij} P_i)^\top. \quad (27)$$

Hence:

$$\frac{\partial \Phi}{\partial (\text{vec } P_j)^\top} = L_4 (K_{4,4} + I_{16}) (I_4 \otimes (F_{ij} P_i)^\top) \quad (28)$$

$$\frac{\partial \Phi}{\partial (\text{vec } P_i)^\top} = L_4 (K_{4,4} + I_{16}) (I_4 \otimes (P_j^\top F_{ij})) \quad (29)$$

where L_4 is the elimination matrix. Hence the Jacobian of Φ with respect to vectorized cameras $\text{vec}(P_j)$ and $\text{vec}(P_i)$ is the following 10×24 matrix:

$$\begin{aligned} & \frac{\partial \Phi}{\partial [(\text{vec } P_i)^\top | (\text{vec } P_j)^\top]} = \\ & = L_4 (K_{4,4} + I_{16}) \left[(I_4 \otimes (F_{ij} P_i)^\top) | (I_4 \otimes (P_j^\top F_{ij})) \right]. \end{aligned} \quad (30)$$

In the rest of the manuscript, the same Jacobian matrix has been referred to, with a slight abuse of notation, as $\frac{\partial \Phi}{\partial P_i, P_j}$.

5.3 Derivatives of Φ with Respect to Fundamental Matrices

By the definition of elimination matrix and commutation matrix one can rewrite (23) as

$$\Phi(P_i, P_j, F) = L_4 (I_{16} + K_{44}) \text{vec}(S). \quad (31)$$

Since

$$\text{vec}(S) = \text{vec}(P_j^\top F_{ij} P_i) = (P_i^\top \otimes P_j^\top) \text{vec}(F) \quad (32)$$

we get

$$\Phi(P_i, P_j, F) = \underbrace{L_4(I_{16} + K_{44})}_{10 \times 16} (P_i^\top \otimes P_j^\top) \text{vec}(F). \quad (33)$$

Hence, the the Jacobian of Φ with respect to vectorized fundamental matrices can be obtained as the following 10×9 matrix:

$$\frac{\partial \Phi}{\partial (\text{vec } F)^\top} = L_4(K_{4,4} + I_{16})(P_i^\top \otimes P_j^\top). \quad (34)$$

In the rest of the manuscript, the same Jacobian matrix has been referred to, with a slight abuse of notation, as $\frac{\partial \Phi}{\partial F}$.

6 Experiments

In this section we report results on synthetic and real data. We implemented our method in MATLAB R2023b – the code is publicly available³ – and used a MacMini M1 (2020) with 16Gb RAM for our experiments. We compared our approach to the one by Arrigoni et al. [6], that addresses finite solvability as well. We also discuss the efficiency of the analyzed approaches. We did not consider the method by Trager et al. [4] in our comparisons since it was subsumed by [6]. We refer the reader to [6] and Section 6.4 for additional insights on the performance of [4].

6.1 Mining minimally solvable graphs

A graph is called *minimally solvable* if the removal of any edge results in a non-solvable graph. Recall that a necessary condition is that the graph has at least $(11n - 15)/7$ edges, with n being the number of nodes [4]. So, any solvable graph with this number of edges is minimally solvable. In the need of a combinatorial characterization of minimally solvable graphs, we can build a catalog by “mining” them. In this respect, we exhaustively generated all the biconnected graphs (a necessary condition for solvability) with a given number of nodes (up to ten) having $\lceil (11n - 15)/7 \rceil$ edges. Among these candidates, we tested for finite solvability using our method and [6], which always returned the same result, as expected. The results are reported in Table 1. Please note that among the 27 minimal graphs with 9 nodes that passed the test, there are the 10 counterexamples found by [5] that are finite solvable and admit two realizations in \mathbb{R} , showing that finite solvability $\not\Rightarrow$ solvability.

6.2 Synthetic Data

We then analyzed synthetic graphs with $n = 20$ nodes⁴ generated by randomly selecting a fixed percentage of edges from the complete graph (named density), discarding disconnected graphs. We considered density values ranging

³<https://github.com/federica-arrigoni/finite-solvability>

⁴We also tested other values of n , obtaining comparable results.

Table 1: Minimally solvable graphs.

Column “#candidates” reports the number of biconnected graphs with up to 10 nodes and $\lceil (11n - 15)/7 \rceil$ edges; out of these graphs, “#fin_solv” have been found finite solvable.

#nodes	#candidates	#fin_solv
3	1	1
4	1	1
5	2	1
6	9	4
7	20	3
8	161	36
9	433	27
10	5898	756

from 5% to 70%: for each value, 1000 graphs were sampled, for a total of 8000 samples. Both our method and [6] were applied to each graph and they always gave the same output. Results are collected in Table 2: as expected, when the percentage of edges decreases, the graphs are more likely to be unsolvable and the number of components increases. The examples shown in Figure 3 are taken from this experiment.

Table 2: Analysis on 1000 random graphs with 20 nodes and varying density. Column “#fin_solv” reports the number of graphs that passed the finite solvability test. The last column reports the [min max] number of components.

%density	#fin_solv	#comp.
5	0	[10, 25]
10	2	[1, 27]
20	253	[1, 24]
30	826	[1, 5]
40	977	[1, 3]
50	999	[1, 2]
60	1000	[1, 1]
70	1000	[1, 1]

6.3 Real Data

As done in [6], we consider real viewing graphs taken from popular structure-from-motion datasets: the Cornell Arts Quad dataset [7], the 1DSfM dataset [29], and image sequences from [30]. Some statistics about these graphs are reported in Table 3, namely: the number of nodes; the number of edges; the density (i.e., the percentage of available edges with respect to the complete graph). Table 3 also reports the outcome of this experiment: the number of components and the execution times of the competing methods. Note that #components=1 is equivalent to say that

Table 3: Results of our experiments on real SfM datasets [7, 29, 30].

“Time” is the total time (in seconds) spent building the solvability matrix (Arrigoni et al. [6]) or the Jacobian matrix (our approach), testing for finite solvability, and computing the components (only on non-solvable cases).

Name	Dataset			Arrigoni et al. [6]		Our Method	
	#nodes	%density	#edges	#comp.	Time	#comp.	Time
Gustav Vasa	18	72	110	1	0.10	1	0.37
Dino 319	36	37	230	1	0.07	1	0.15
Dino 4983	36	37	231	1	0.06	1	0.06
Folke Filbyter	40	32	250	1	0.06	1	0.05
Jonas Ahls	40	41	321	1	0.09	1	0.20
Park Gate	34	94	529	1	0.13	1	0.12
Toronto University	77	33	974	1	0.30	1	0.18
Sphinx	70	55	1330	1	0.53	1	0.23
Cherub	65	64	1332	1	0.53	1	0.24
Tsar Nikolai I	98	52	2486	1	1.32	1	0.46
Skansen Kronan	131	88	7490	1	10.08	1	2.07
Alcatraz Courtyard	133	92	8058	1	11.59	1	2.30
Buddah Tooth	162	73	9546	1	15.47	1	2.93
Pumpkin	195	65	12276	1	26.16	1	4.63
Ellis Island	240	71	20290	1	74.22	1	11.54
NYC Library	358	32	20662	1	74.44	1	12.10
Madrid Metropolis	370	35	23755	1	99.45	1	15.25
Tower of London	489	20	23844	4	103.83	4	19.53
Piazza del Popolo	345	42	24701	4	108.57	4	18.35
Union Square	853	7	25478	4	124.94	4	27.64
Yorkminster	448	28	27719	1	126.89	1	19.31
Gendarmenmarkt	722	18	48124	4	411.67	4	69.81
Montreal N. Dame	467	48	52417	1	462.81	1	65.53
Roman Forum	1102	12	70153	4	913.57	4	158.16
Alamo	606	53	97184	1	2335.94	1	222.35
Vienna Cathedral	898	26	103530	1	2565.31	1	253.84
Notre Dame	553	68	103932	1	2631.00	1	249.12
Arts Quad	5460	1	221929	1	10979.35	1	898.56
Piccadilly	2446	11	319195	1	25889.20	1	2361.20

the graph is finite solvable. Information on the number of rows/columns of the matrix used by Arrigoni et al. [6] and the one from our formulation is given in Figure 4, whereas explicit formulas are given in Section 6.4.

Results show that there are only five unsolvable cases among the analyzed graphs, all exhibiting four components, in agreement with previous work. One example is visualized in Figure 5. Our method and the one by Arrigoni et al. [6] always gave the same output on all the graphs, as expected. Table 3 also shows that our approach is significantly faster than the state of the art, underlying the advantage of a node-based approach with respect to an edge-based one. Indeed, the matrix employed by our formulation is significantly smaller than the one used by the authors of [6] – this can also be seen in Figure 4 and Table 4. In particular, our direct formulation takes less than 10% of the total running time of [6] on the largest examples (from “Alamo” to “Piccadilly”). This figure becomes 20% for medium size

datasets (from “Skansen Kronan” to “Roman Forum”). For the smallest ones the running time is less than a second and the comparison becomes meaningless.

6.4 Number of Rows/Columns for Different Formulations

Here we summarize the comparison with [6] and [4] in terms of the size of the respective matrices. Recall that $n = |V|$ and $m = |E|$ in a graph G with V vertices and E nodes.

Trager et al. [4]. The solvability matrix of [4] is made of blocks, where each block comprises 20 equations. The number of blocks per node is $d_i(d_i - 1)/2$, where d_i denotes the degree of node i (see Table 2 in [6]). By summing over all the nodes in the graph, a formula is obtained for the number of rows of the solvability matrix (or, equivalently, the

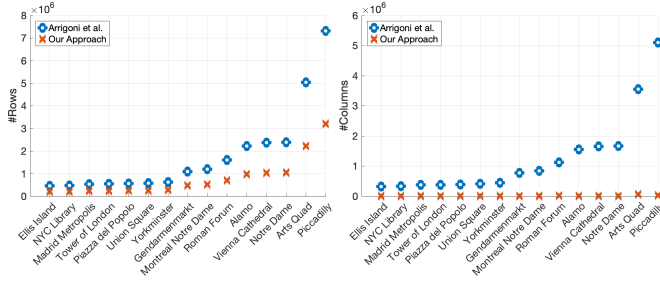


Figure 4: Number of rows and columns of the matrix used by Arrigoni et al. [6] and the one from our formulation on large-scale SfM datasets [7, 29]. Observe that the difference in the number of columns even surpasses one order of magnitude.

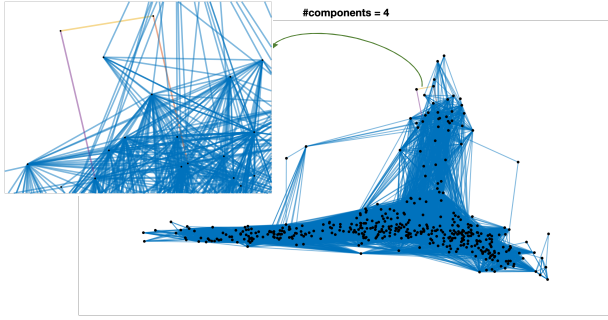


Figure 5: Viewing graph of the Tower of London dataset [29] and maximal components (color-coded). Edges in the largest components are depicted in blue. A zoom is drawn to better visualize the non-solvable part, comprising three edges which resemble the square topology.

number of equations):

$$e_1 = 10 \sum_{i=1}^n (d_i^2 - d_i) + 15 + m = 10 \sum_{i=1}^n d_i^2 - 19m + 15 \quad (35)$$

where $15 + m$ accounts for the additional equations introduced to remove the ambiguities and $\sum_{i=1}^n d_i = 2m$ due to the *degree sum formula* [31]. Exploiting the Cauchy-Schwarz inequality we obtain:

$$\sum_{i=1}^n d_i^2 \geq \frac{1}{n} \left(\sum_{i=1}^n d_i \right)^2 \quad (36)$$

hence, using again the degree sum formula, we get the following lower bound for e_1 :

$$e_1 \geq \frac{10}{n} (2m)^2 - 19m + 15 = 40 \frac{m^2}{n} - 19m + 15. \quad (37)$$

Hence the number of rows grows asymptotically (at least) as $O(n^3)$ for a dense graph (where $m \approx O(n^2)$) and it grows as $O(n)$ for a sparse one (where $m \approx O(n)$).

Arrigoni et al. [6]. The reduced solvability matrix used by [6] is made of blocks of 11 equations. The number of blocks

per node is $d_i - 1$ (therefore it scales linearly in the degree of node i whereas in [4] the growth is quadratic). Hence the total number of rows (i.e., equations) is given by:

$$\begin{aligned} e_2 &= 11 \sum_{i=1}^n (d_i - 1) + 15 + m = 11 \sum_{i=1}^n d_i - 11n + 15 + m \\ &= 23m - 11n + 15. \end{aligned} \quad (38)$$

The above formula implies that the number of rows of the reduced solvability matrix grows asymptotically as $O(n^2)$ for a dense graph and $O(n)$ for a sparse one. Away from the limit case of a perfectly sparse graph with $m = O(n)$, there is an advantage of this formulation with respect to [4]. In concrete terms, it is enough that $m > n$ to ensure that $e_2 \leq e_1$: indeed, after proper simplifications, (37) \geq (38) becomes $40m^2 + 11n^2 \geq 42nm > 42n^2$, which reduces to $40m^2 > 31n^2$, which is always satisfied under the hypothesis $m > n$. The number of columns (i.e., variables) is the same for [4] and [6], and it is given by

$$v_1 = v_2 = 16m. \quad (39)$$

We refer the reader to [6] for additional information on the performance of [4] on real-world datasets.

Our Formulation. As explained in Section 4, our polynomial system employs a total of

$$\begin{aligned} e_3 &= 10m + n + 15 \text{ equations} \\ v_3 &= 12n \text{ unknowns.} \end{aligned} \quad (40)$$

Hence, the number of rows of our Jacobian matrix grows asymptotically as $O(n^2)$ for a dense graph and $O(n)$ for a sparse one. In concrete terms, however, $e_3 \leq e_2$ as soon as $m \geq \frac{12}{13}n \approx n$, which is typically satisfied.

A summary is reported in Table 4. Note that our formulation is the only one where the number of columns scales with the number of nodes (instead of edges) in the graph, as ours is the first node-based method. Observe also that practical datasets are far from the sparse graph approximation, as the number of edges is much larger than the number of nodes.

Table 4: Number of equations/unknowns for the three formulations. The row counts are in decreasing order for typical graphs.

Method	#rows	#columns
Trager et al. [4]	$\geq 40m^2/n - 19m + 15$	$16m$
Arrigoni et al. [6]	$23m - 11n + 15$	$16m$
Ours	$10m + n + 15$	$12n$

7 Conclusion

This paper underscored the viewing graph as a powerful representation of uncalibrated cameras and their geometric relationships. The solvability of the graph corresponds

to the existence of a unique set of cameras, up to a single projective transformation, that conforms to the given fundamental matrices. Our focus was on the relaxed notion of finite solvability, which considers the finiteness of solutions rather than strict uniqueness. This approach is computationally tractable and enables the analysis of large graphs derived from structure-from-motion datasets.

We presented a novel formulation of the problem that provides a more direct approach than previous literature – based on a formula that explicitly establishes links between pairs of cameras through their fundamental matrices, as suggested by the definition of solvability. Building upon this, we developed an algorithm designed to test finite solvability and extract components of unsolvable cases, surpassing the efficiency of previous methods. The core methodology is mathematically sound and extremely simple, as it only requires computing the derivatives of polynomial equations with respect to their unknowns, and checking the rank of the resulting Jacobian matrix. Although the Jacobian check, by definition, applies only to a neighborhood of a particular solution, in our case, this local information extends globally due to the special structure of the problem. We formally established this result—originally conjectured in our preliminary study [25]—using tools from Algebraic Geometry.

The concept of finite solvability, while valuable, represents only a partial step toward a computationally efficient characterization of viewing graph solvability. Its inherent limitation lies in asserting the existence of a finite number of solutions rather than guaranteeing a unique one. The challenge of efficiently verifying uniqueness in large structure-from-motion graphs remains an open question. We hope that our results will inspire further research in this intriguing direction.

Acknowledgements

Federica Arrigoni was supported by PNRR-PE-AI FAIR project funded by the NextGeneration EU program. Tomas Pajdla was supported by the OPJAK CZ.02.01.01/00/22 008/0004590 Roboprox Project. Kathlén Kohn was supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

- [1] Noam Levi and Michael Werman. “The viewing graph”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2003, pp. 518–522.
- [2] Alessandro Rudi, Matia Pizzoli, and Fiora Pirri. “Linear Solvability in the Viewing Graph”. In: *Proceedings of the Asian Conference on Computer Vision*. 2011, pp. 369–381.
- [3] M. Trager, M. Hebert, and J. Ponce. “The Joint Image Handbook”. In: *Proceedings of the International Conference on Computer Vision*. 2015, pp. 909–917.
- [4] Matthew Trager, Brian Osserman, and Jean Ponce. “On the Solvability of Viewing Graphs”. In: *Proceedings of the European Conference on Computer Vision*. 2018, pp. 335–350.
- [5] Federica Arrigoni et al. “Viewing graph solvability via cycle consistency”. In: *Proceedings of the International Conference on Computer Vision*. 2021, pp. 5540–5549.
- [6] Federica Arrigoni, Tomas Pajdla, and Andrea Fusiello. “Viewing Graph Solvability in Practice”. In: *Proceedings of the International Conference on Computer Vision*. 2023, pp. 8147–8155.
- [7] David Crandall et al. “Discrete-Continuous Optimization for Large-Scale Structure from Motion”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2011, pp. 3001–3008.
- [8] Onur Ozyesil et al. “A survey of Structure from Motion”. In: *Acta Numerica* 26 (2017), pp. 305–364.
- [9] Avishek Chatterjee and Venu Madhav Govindu. “Robust Relative Rotation Averaging”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).
- [10] Paul-Edouard Sarlin et al. “Pixel-Perfect Structure-From-Motion With Featuremetric Refinement”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).
- [11] Lalit Manam and Venu Madhav Govindu. “Sensitivity in Translation Averaging”. In: *Neural Information Processing Systems (NeurIPS)*. 2023.
- [12] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press, 2004.
- [13] Federica Arrigoni et al. “Revisiting Viewing Graph Solvability: an Effective Approach Based on Cycle Consistency”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), pp. 1–14.
- [14] Federica Arrigoni and Andrea Fusiello. “Bearing-based Network Localizability: A Unifying View”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.9 (2019), pp. 2049–2069.
- [15] R. Tron et al. “Rigid Components Identification and Rigidity Enforcement in Bearing-Only Localization using the Graph Cycle Basis”. In: *IEEE American Control Conference*. 2015, pp. 3911–3918.
- [16] A. Karimian and Roberto Tron. “Theory and methods for bearing rigidity recovery”. In: *Proceedings of the IEEE Conference on Decision and Control*. 2017, pp. 2228–2235.

- [17] S. Sengupta et al. “A New Rank Constraint on Multi-view Fundamental Matrices, and Its Application to Camera Location Recovery”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 2413–2421.
- [18] Martin Bratelund and Felix Rydell. “Compatibility of Fundamental Matrices for Complete Viewing Graphs”. In: *Proceedings of the International Conference on Computer Vision*. 2023, pp. 3305–3313.
- [19] S.N. Sinha, M. Pollefeys, and L. McMillan. “Camera network calibration from dynamic silhouettes”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2004, pp. I–I.
- [20] Y. Kasten et al. “GPSfM: Global Projective SFM Using Algebraic Constraints on Multi-View Fundamental Matrices”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019, pp. 3259–3267.
- [21] Carlo Colombo and Marco Fanfani. “A closed form solution for viewing graph construction in uncalibrated vision”. In: *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. 2021.
- [22] Rakshith Madhavan, Andrea Fusiello, and Federica Arrigoni. “Synchronization of Projective Transformations”. In: *Proceedings of the European Conference on Computer Vision*. 2024.
- [23] Joe Kileel and Kathlen Kohn. “Snapshot of Algebraic Vision”. In: *arXiv* 2210.11443 (2023).
- [24] Igor. R. Shafarevich. *Basic Algebraic Geometry Vol 1: Varieties in projective space*. 3rd ed. New York: Springer, 2013. DOI: [DOI10 . 1007 / 978 - 3 - 642 - 37956 - 7](https://doi.org/10.1007/978-3-642-37956-7).
- [25] Federica Arrigoni, Andrea Fusiello, and Tomas Pajdla. “A Direct Approach to Viewing Graph Solvability”. In: *Proceedings of the European Conference on Computer Vision*. 2024.
- [26] Timothy Duff et al. “PLMP – Point-Line Minimal Problems in Complete Multi-View Visibility”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.1 (2024), pp. 421–435.
- [27] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. 2nd ed. John Wiley & Sons, 1999.
- [28] Harold V. Henderson and S. R. Searle. “The Vec-Permutation Matrix, the Vec Operator and Kronecker Products: A Review”. In: *Linear and Multilinear Algebra* 9 (1981), pp. 271–288.
- [29] K. Wilson and N. Snavely. “Robust Global Translations with 1DSfM”. In: *Proceedings of the European Conference on Computer Vision*. 2014, pp. 61–75.
- [30] C. Olsson and O. Enqvist. “Stable structure from motion for unordered image collections”. In: *Proceedings of the 17th Scandinavian conference on Image analysis (SCIA'11)*. Springer-Verlag, 2011, pp. 524–535.
- [31] F. Harary. *Graph Theory*. Addison-Wesley, 1972.