

# INVARIANT THEORY AND SCALING ALGORITHMS FOR MAXIMUM LIKELIHOOD ESTIMATION

Carlos Améndola, Kathlén Kohn, Philipp Reichenbach, Anna Seigal

**ABSTRACT.** We uncover connections between maximum likelihood estimation in statistics and norm minimization over a group orbit in invariant theory. We focus on Gaussian transformation families, which include matrix normal models and Gaussian graphical models given by transitive directed acyclic graphs. We use stability under group actions to characterize boundedness of the likelihood, and existence and uniqueness of the maximum likelihood estimate. Our approach reveals promising consequences of the interplay between invariant theory and statistics. In particular, existing scaling algorithms from statistics can be used in invariant theory, and vice versa.

## 1. INTRODUCTION

The task of fitting data to a model is fundamental in statistics. A statistical model is a set of probability distributions. We seek a point in a model that best fits some empirical data. A widespread approach is to maximize the likelihood of observing the data as we range over the model. A point that maximizes the likelihood is called a *maximum likelihood estimate* (MLE). There are several ways to compute an MLE for different statistical models, usually via optimization approaches that find a local maximum [MBT14, Myu03]. There is growing interest in understanding when algorithms to find an MLE are guaranteed to work, and under which conditions an MLE exists or is unique. In this paper, we approach such questions using invariant theory.

Invariant theory studies actions of groups on vector spaces or, more generally, on algebraic varieties. An important concept is the orbit of a point under the group action, which is the set of all points that differ from the original point by a transformation in the group. The *capacity* of a point is the infimal norm along its orbit. If the orbit is closed, the capacity is attained; otherwise the capacity is attained only on the orbit closure. Points with zero capacity are called unstable; they form the null cone, a classical object in invariant theory dating back to Hilbert [Hil93], which is of particular interest for moduli spaces of algebraic objects. More recently, algorithmic approaches to stability questions have been taken, with a special focus on testing null cone membership [AZGL<sup>+</sup>18, BGO<sup>+</sup>17, BFG<sup>+</sup>19, DM17, GGOW16, IQS18]. A number of applied problems have been cast within an invariant theoretic framework, including questions in quantum information theory, complexity theory and analytic inequalities, see e.g. [BFG<sup>+</sup>19, §1.2].

There is a close connection between statistical models and group actions, dating back to Fisher [Fis34]. We build a bridge between invariant theory and maximum likelihood estimation. In this paper, we study this connection in the setting of multivariate Gaussian models. We define *Gaussian group models*, multivariate Gaussian models whose

concentration matrices are of the form  $g^\top g$ , where  $g$  lies in a group. Examples of Gaussian group models are matrix normal models and Gaussian graphical models defined by transitive directed acyclic graphs.

The connection between invariant theory and maximum likelihood estimation also holds for discrete statistical models, as we discuss in our companion paper [AKRS20]. There, we show that maximum likelihood estimation in log-linear models is equivalent to computing the capacity under a torus action. Both Gaussian group models and log-linear models fall within the framework of exponential families.

**Main contributions.** We show that finding the MLE can be cast as the problem of computing the capacity, see Propositions 3.4 and 3.13. Viewing maximum likelihood estimation as a norm minimization problem allows us to build a correspondence between notions of stability from invariant theory and MLE properties:

$$\left\{ \begin{array}{c} \text{unstable} \\ \text{semistable} \\ \text{polystable} \\ \text{stable} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{likelihood unbounded from above} \\ \text{likelihood bounded from above} \\ \text{MLE exists} \\ \text{MLE exists uniquely} \end{array} \right\}$$

For some models we prove an exact equivalence between the four notions of stability on the left and the four properties of the MLE on the right, see Theorem 3.15 for complex Gaussian group models. For real statistical models, we prove real analogues of the correspondence that hold at two levels of generality; see Theorems 3.6 and 3.10. The two levels of generality correspond to non-reductive and reductive groups.

While invariant theory often focuses on reductive groups, Gaussian group models are natural to study in both settings. For matrix normal models, which are given by reductive groups, we use descriptions of the null cone to give improved bounds on the number of samples generically required for a bounded likelihood function, see Theorem 4.8 and Corollary 4.9. Gaussian models defined by transitive directed acyclic graphs are in general given by non-reductive groups. For such models, our results translate to exact conditions for MLE existence in terms of linear independence of the rows of the sample matrix, see Theorem 5.3.

Our connection between invariant theory and maximum likelihood estimation leads to the algorithmic consequences that we detail below.

**Algorithmic implications.** Scaling algorithms are iterative algorithms existing both in statistics and in invariant theory. They are characterized by update steps, which are given by a group action in many instances. For matrix normal models, we show the equivalence of two alternating algorithms: operator scaling from invariant theory for null cone membership testing [Gur04, GGOW16], and the flip-flop algorithm from statistics for maximum likelihood estimation [Dut99, LZ05]; see the left of Figure 1 and Section 4.5. This equivalence enables us to obtain a complexity analysis for the flip-flop algorithm (see Theorem 4.15) by directly adapting the result for the corresponding null cone membership problem from [BGO<sup>+</sup>17, Theorem 1.1].

We now describe how this can be extended to more general scaling algorithms, see the right hand side of Figure 1. The flip-flop algorithm can be thought of as an instance of *iterative proportional scaling (IPS)* (or iterative proportional fitting (IPF)), a

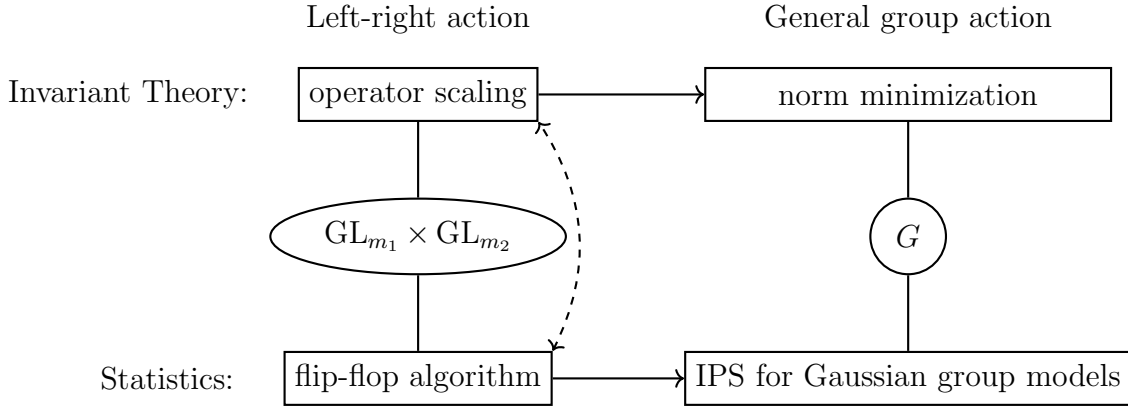


FIGURE 1. Overview of different scaling algorithms. For the invariant theory algorithms, we use matrices of determinant one, e.g.  $SL_{m_1} \times SL_{m_2} \subseteq GL_{m_1} \times GL_{m_2}$ .

family of methods to find the MLE in a statistical model [FM81, Cra98]. For Gaussian group models, we can find an MLE via the geodesically convex optimization approaches from [BFG<sup>+</sup>19] that minimize the norm over an orbit. These algorithms can be thought of as generalizations of operator scaling. We therefore regard them as IPS for Gaussian group models.

The connection between norm minimization in invariant theory and IPS in statistics is discussed for torus actions and discrete models in our companion paper [AKRS20]. There, [AKRS20, Figure 4] gives the analogue of Figure 1 for the setting of a discrete model and a torus action (rather than a Gaussian model and a general group action). The starting point of both Figures is Sinkhorn scaling [Sin64], an alternating method that involves the left-right action of a product of two tori. The alternating idea from Sinkhorn’s scaling generalizes to products of groups, e.g. to operator scaling and the flip-flop algorithm in Figure 1.

We see that algorithms in invariant theory can be used in maximum likelihood estimation, and vice versa. In statistics, many iterative algorithms for finding the MLE are well-known. It is a more recent question to understand when they converge, i.e. when an MLE exists, and when convergence is to a unique solution, i.e. when the MLE is unique. The historical progression is the opposite in invariant theory: the distinction between different types of stability is classical, while more recent approaches use algorithms to test instability. Our results are intended to stimulate further research to deepen the connection between the fields.

**Organization.** To address readers with different backgrounds, we present preliminaries from invariant theory and statistics in Section 2. We consider the general setting of a Gaussian group model in Section 3. We then study matrix normal models in Section 4, followed by transitive directed acyclic graphs, in Section 5.

## 2. PRELIMINARIES

**2.1. Maximum likelihood estimation.** A statistical model is a set of probability distributions. In this paper we consider multivariate Gaussian distributions with mean zero. The density function of an  $m$ -dimensional Gaussian with mean zero and covariance matrix  $\Sigma$  is

$$f_{\Sigma}(y) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}y^{\top}\Sigma^{-1}y\right),$$

where  $y \in \mathbb{R}^m$  and  $\Sigma$  is in the cone of  $m \times m$  positive definite matrices, which we denote by  $\text{PD}_m$ . We often consider the concentration matrix  $\Psi = \Sigma^{-1}$ . A Gaussian model is determined by a set of concentration matrices, i.e. a subset of  $\text{PD}_m$ .

A maximum likelihood estimate (MLE) is a point in the model that maximizes the likelihood of observing some data  $y = (y_1, \dots, y_n)$ , where  $n$  is the sample size. That is, an MLE maximizes the *likelihood function*

$$(1) \quad L_y(\theta) = f_{\theta}(y_1) \cdots f_{\theta}(y_n),$$

where the model is parametrized by  $\theta \in \Theta$ . It is often convenient to work with the *log-likelihood function*  $\ell_y = \log L_y$ , which has the same maximizers.

For Gaussian models  $\mathcal{M} \subseteq \text{PD}_m$ , the data is a tuple  $Y = (Y_1, \dots, Y_n) \in (\mathbb{R}^m)^n$ . The likelihood function (1) is

$$L_Y(\Psi) = \prod_{i=1}^n f_{\Psi^{-1}}(Y_i).$$

The log-likelihood function can be written, up to additive and multiplicative constants, as

$$(2) \quad \ell_Y(\Psi) = \log \det(\Psi) - \text{tr}(\Psi S_Y),$$

where  $S_Y = \frac{1}{n} \sum_{i=1}^n Y_i Y_i^{\top}$  is the sample covariance matrix, an  $m \times m$  positive semi-definite matrix. It is well-known that the unique maximizer of the likelihood over the positive definite cone is  $\hat{\Psi} = S_Y^{-1}$ , if  $S_Y$  is invertible. If  $S_Y$  is not invertible, the likelihood function is unbounded and the MLE does not exist, see [Sul18, Proposition 5.3.7].

The minimum number of samples needed for an MLE to generically exist is the *maximum likelihood threshold* (mlt) of a model. The minimum number of samples needed for the likelihood to be generically bounded is denoted by  $\text{mlt}_b$ . By *generically*, we mean that a property holds away from an algebraic hypersurface. Hence, it will hold almost surely, i.e. outside of a set of Lebesgue measure zero. As an example, the discussion above says that  $\text{mlt} = \text{mlt}_b = m$  when the Gaussian model is the full positive definite cone,  $\mathcal{M} = \text{PD}_m$ .

**2.2. Invariant theory.** This section gives a friendly guide to our invariant theory setting, following [Wal17]. We explain how our seemingly special setting fits into usual terminology of invariant theory in Remark 2.4.

Invariant theory studies actions of a group  $G$  and notions of stability with respect to this action. In this article we work with linear actions on a real or complex vector space. Such a linear action corresponds to a representation  $\varrho: G \rightarrow \text{GL}_m(\mathbb{K})$ , i.e. each group element  $g \in G$  is assigned an invertible matrix in  $\text{GL}_m(\mathbb{K})$  where  $\mathbb{K}$  is  $\mathbb{R}$  or  $\mathbb{C}$ . The

group element  $g \in G$  acts on  $\mathbb{K}^m$  by left multiplication with the matrix  $\varrho(g)$ . For a vector  $v \in \mathbb{K}^m$ , we define the capacity to be  $\text{cap}(v) := \inf_{g \in G} \|g \cdot v\|^2$ . Here and throughout the paper  $\|\cdot\|$  denotes the Euclidean norm for vectors and Frobenius norm for matrices. We now define the four notions of stability for such an action.

**Definition 2.1.** Let  $v \in \mathbb{K}^m$ . We denote the orbit of  $v$  by  $G \cdot v$ , the orbit closure with respect to the Euclidean topology by  $\overline{G \cdot v}$  and the stabilizer of  $v$  by  $G_v$ . We say  $v$  is

- (a) *unstable*, if  $0 \in \overline{G \cdot v}$ , i.e.  $\text{cap}(v) = 0$ .
- (b) *semistable*, if  $0 \notin \overline{G \cdot v}$ , i.e.  $\text{cap}(v) > 0$ .
- (c) *polystable*, if  $v \neq 0$  and  $G \cdot v$  is closed.
- (d) *stable*, if  $v$  is polystable and  $G_v$  is finite.

The set of unstable points is called the *null cone* of the group action.

The orbit and orbit closure of  $v$  only depend on the group  $\varrho(G)$ . Thus, when studying the notions from Definition 2.1(a)–(c) we can assume  $G \subseteq \text{GL}_m$  after restricting to the image of  $\varrho$ . We call  $G \subseteq \text{GL}_m$  *Zariski closed* if  $G$  is the zero locus of a set of polynomials in the matrix entries. The transpose of  $g \in G$  is denoted by  $g^\top$  and the Hermitian transpose by  $g^*$ . We say that a group  $G$  is *self-adjoint* if  $g \in G$  implies  $g^\top \in G$  (for  $\mathbb{K} = \mathbb{R}$ ), or if  $g \in G$  implies  $g^* \in G$  (for  $\mathbb{K} = \mathbb{C}$ ).

Next, we introduce the moment map and state the Kempf-Ness theorem, a crucial ingredient for many of our results. We consider  $G \subseteq \text{GL}_m(\mathbb{K})$ , a Zariski closed and self-adjoint subgroup. For each vector  $v \in \mathbb{K}^m$ , we study the map

$$\gamma_v: G \longrightarrow \mathbb{R}, \quad g \longmapsto \|gv\|^2,$$

and note that the infimum of  $\gamma_v$  is the capacity of  $v$ . Since  $G$  is defined by polynomial equations, we can consider its tangent space  $T_{I_m}G \subseteq \mathbb{K}^{m \times m}$  at the identity matrix  $I_m$ , and we can compute the differential of the map  $\gamma_v$  at the identity:

$$D_{I_m}\gamma_v: T_{I_m}G \longrightarrow \mathbb{R}, \quad \dot{g} \longmapsto 2 \operatorname{Re}[\operatorname{tr}(\dot{g}vv^*)].$$

The *moment map*  $\mu$  assigns this differential to each vector  $v$ , i.e.

$$\mu: \mathbb{K}^m \longrightarrow \operatorname{Hom}_{\mathbb{R}}(T_{I_m}G, \mathbb{R}), \quad v \longmapsto D_{I_m}\gamma_v.$$

The moment map vanishes at a vector  $v$  if and only if the identity matrix  $I_m$  is a critical point of the map  $\gamma_v$ . Now we are ready to formulate the Kempf-Ness theorem, which is due to [KN79] for  $\mathbb{K} = \mathbb{C}$ . The first proof for  $\mathbb{K} = \mathbb{R}$  was given in [RS90].

**Theorem 2.2** (Kempf-Ness). *Let  $G \subseteq \text{GL}_m(\mathbb{K})$  be a Zariski closed self-adjoint subgroup with moment map  $\mu$ , where  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ . If  $\mathbb{K} = \mathbb{R}$ , let  $K$  be the set of orthogonal matrices in  $G$ . If  $\mathbb{K} = \mathbb{C}$ , let  $K$  be the set of unitary matrices in  $G$ . For  $v \in \mathbb{K}^m$ , we have:*

- (a) *The vector  $v$  is of minimal norm in its orbit if and only if  $\mu(v) = 0$ .*
- (b) *If  $\mu(v) = 0$  and  $w \in G \cdot v$  is such that  $\|v\| = \|w\|$ , then  $w \in K \cdot v$ .*
- (c) *If the orbit  $G \cdot v$  is closed, then there exists some  $w \in G \cdot v$  with  $\mu(w) = 0$ .*
- (d) *If  $\mu(v) = 0$ , then the orbit  $G \cdot v$  is closed.*
- (e) *The vector  $v$  is polystable if and only if there exists  $0 \neq w \in G \cdot v$  with  $\mu(w) = 0$ .*

(f) *The vector  $v$  is semistable if and only if there exists  $0 \neq w \in \overline{G \cdot v}$  with  $\mu(w) = 0$ .*

*Proof.* Parts (a)–(d) are [Wal17, Theorems 3.26 and 3.28] while part (e) is a direct consequence of (c) and (d). Part (f) follows from the fact that any orbit closure  $\overline{G \cdot v}$  contains a unique closed orbit, which is not the zero orbit if and only if  $v$  is semistable. For  $\mathbb{K} = \mathbb{C}$  this can be found in [Wal17, Theorem 3.20] and for  $\mathbb{K} = \mathbb{R}$  we refer to [RS90, Section 9.3] or [BL17, Theorem 1.1(iii)]. For the latter, note that [BL17, Condition (1)] is satisfied in our setting by [Wal17, Theorem 2.16].

The assumption that  $G$  is connected, which appears in [Wal17, Theorem 3.26], is not needed here, by the following argument. If  $G^\circ$  is the identity component of  $G$ , then the quotient group  $G/G^\circ$  is finite and its elements can be represented by unitary matrices, by the polar decomposition [Wal17, Theorem 2.16]. Hence (a)–(f) above depend only on  $G^\circ$ .  $\square$

The following result relates the capacity over  $\mathbb{C}$  to the capacity over  $\mathbb{R}$ .

**Proposition 2.3.** *Let  $G_{\mathbb{R}}$  be a Zariski closed self-adjoint subgroup of  $\mathrm{GL}_m(\mathbb{R})$  and denote by  $G_{\mathbb{C}}$  its Zariski closure in  $\mathrm{GL}_m(\mathbb{C})$ . Let  $\mathrm{cap}_{\mathbb{K}}(v)$  be the capacity of  $v \in \mathbb{K}^m$  under  $G_{\mathbb{K}}$  and denote the null cone under left multiplication with  $G_{\mathbb{K}}$  by  $\mathcal{N}_{\mathbb{K}}$ . Then, for  $v \in \mathbb{R}^m$ , we have the equality of capacities  $\mathrm{cap}_{\mathbb{R}}(v) = \mathrm{cap}_{\mathbb{C}}(v)$ . In particular,  $\mathcal{N}_{\mathbb{R}} = \mathcal{N}_{\mathbb{C}} \cap \mathbb{R}^m$ .*

*Proof.* The group  $G_{\mathbb{C}} \subseteq \mathrm{GL}_m(\mathbb{C})$  is self-adjoint by [Wal17, Lemma 3.29]. The capacity  $\mathrm{cap}_{\mathbb{K}}(v)$  is attained at all elements of minimal norm in the closed orbit contained in  $\overline{G_{\mathbb{K}} \cdot v}$ , by Kempf-Ness. Hence we can reduce to studying a closed orbit  $G_{\mathbb{R}} \cdot v$ . If  $w$  is of minimal norm in  $G_{\mathbb{R}} \cdot v$ , then it is of minimal norm in  $G_{\mathbb{C}} \cdot v$  by [Wal17, Lemma 3.31] or [RS90, Lemma 8.1]. Thus,  $G_{\mathbb{C}} \cdot w$  is closed by Kempf-Ness and hence  $\|w\|^2 = \mathrm{cap}_{\mathbb{C}}(v)$ .  $\square$

*Remark 2.4.* We relate our special setting to the usual setting from invariant theory, where one considers a linearly reductive group  $G$  over  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ . For such a group, any finite dimensional rational representation  $\varrho: G \rightarrow \mathrm{GL}(V)$  over  $\mathbb{K}$  on a vector space  $V$  is semisimple (also called fully reducible), i.e. the representation decomposes into irreducible representations. Moreover,  $\varrho(G) \subseteq \mathrm{GL}(V)$  is a closed algebraic subgroup, see e.g. [Mil17, Theorem 5.39]. Hence, there exists an inner product on  $V$  such that  $\varrho(G) \subseteq \mathrm{GL}(V)$  is self-adjoint, see [Mos55, Theorem 7.1] for  $\mathbb{K} = \mathbb{C}$  and [Mos55, Theorem 7.2] for  $\mathbb{K} = \mathbb{R}$ .

### 3. GAUSSIAN GROUP MODELS

We construct Gaussian models from representations  $G \rightarrow \mathrm{GL}(V)$  of a group  $G$  on a real vector space  $V$ . This extends the idea that log-linear models are orbits of the action by a torus which is utilized in [AKRS20]. Our construction only depends on the image of the group  $G$  inside  $\mathrm{GL}(V)$ . We view each group element as an  $m \times m$  invertible matrix by fixing an isomorphism  $V \cong \mathbb{R}^m$ . The *Gaussian group model* given by  $G$  is the multivariate Gaussian model consisting of all distributions of mean zero whose concentration matrices lie in the set

$$\mathcal{M}_G = \{g^\top g \mid g \in G\}.$$

Equivalently, we take  $\mathcal{M}_G$  to be the model consisting of distributions whose covariance matrices are of the form  $gg^\top$ . This is an instance of a *transformation family*, a statistical

model on which a group acts transitively, see [BNBJJ82]. Our construction includes familiar examples of statistical models.

**Example 3.1.** When  $G$  is the general linear group  $\text{GL}(V)$ , every concentration matrix lies in  $\mathcal{M}_G$  and we get a standard multivariate Gaussian of dimension  $\dim(V)$ , see Section 2.1.  $\diamond$

**Example 3.2.** When  $G$  is the torus of diagonal matrices  $\text{GT}(V)$ , the concentration matrices  $g^\top g$  are also diagonal and the Gaussian group model consists of  $\dim(V)$  independent univariate Gaussian variables.  $\diamond$

Given two matrices  $A_k \in \mathbb{R}^{m_k \times m_k}$ , the Kronecker product  $A_1 \otimes A_2$  is a an  $m_1 m_2 \times m_1 m_2$  matrix. Its rows are indexed by  $(i_1, i_2)$ , and its columns by  $(j_1, j_2)$ , where the indices  $i_k$  and  $j_k$  range from 1 to  $m_k$ . The entry of  $A_1 \otimes A_2$  at index  $((i_1, i_2), (j_1, j_2))$  is  $(A_1)_{i_1 j_1} (A_2)_{i_2 j_2}$ .

**Example 3.3.** Consider the subset of  $\text{GL}_{m_1 m_2}$  given by the image of

$$\begin{aligned} \text{GL}_{m_1} \times \text{GL}_{m_2} &\longrightarrow \text{GL}_{m_1 m_2} \\ (g_1, g_2) &\longmapsto g_1 \otimes g_2. \end{aligned}$$

The concentration matrices in the Gaussian group model are those of the form

$$(g_1 \otimes g_2)^\top (g_1 \otimes g_2) = g_1^\top g_1 \otimes g_2^\top g_2,$$

a Kronecker product of an  $m_1 \times m_1$  concentration matrix and an  $m_2 \times m_2$  concentration matrix. These Gaussian group models are known as matrix normal models, which we discuss in detail in Section 4. This setting can be extended to tensor normal models under the map  $(g_1, \dots, g_d) \mapsto g_1 \otimes \dots \otimes g_d$ .  $\diamond$

We discuss further examples in the context of directed graphical models in Section 5. Now, we describe maximum likelihood estimation for the Gaussian group model given by  $G$ . Given  $n$  samples  $Y = (Y_1, \dots, Y_n) \in V^n$ , we write  $S_Y = \frac{1}{n} \sum_{i=1}^n Y_i Y_i^\top$  for the sample covariance matrix. As in (2), the log-likelihood function is

$$(3) \quad \ell_Y(\Psi) = \log \det(\Psi) - \text{tr}(\Psi S_Y).$$

We maximize the log-likelihood function over the set  $\mathcal{M}_G$  of concentration matrices in the model. An MLE is a concentration matrix in  $\mathcal{M}_G$  that maximizes the log-likelihood.

Next, we describe how finding the MLE relates to finding the capacity of the tuple  $Y$ . A consequence of our results is that algorithms to find the capacity can be used to find the MLE in Gaussian group models. For example, we can apply methods described in [BFG<sup>+</sup>19] to the settings of Theorems 3.10 and 3.15.

**3.1. Equivalence of optimization problems.** We compare the maximization of the log-likelihood to the minimization of the norm  $\|g \cdot Y\|^2$  where  $Y$  is a tuple of samples and  $g$  is an element of the group. The action of the group  $G$  on the tuple  $Y = (Y_1, \dots, Y_n) \in V^n$  is given by  $g \cdot Y = (gY_1, \dots, gY_n)$ , i.e. when considering the action on  $V^n$  the group  $G$  is diagonally embedded in  $\text{GL}(V^n)$ . We can rewrite the norm as

$$(4) \quad \|g \cdot Y\|^2 = \sum_{i=1}^n (gY_i)^\top gY_i = n \text{tr}(g^\top g S_Y).$$

We compare this expression for the norm with the log-likelihood in (3). The term appearing with  $S_Y$  in the trace is  $\Psi$  in the log-likelihood and  $g^\top g$  in the norm. This explains our choice to let the Gaussian group model consist of distributions with concentration matrix  $g^\top g \in \mathcal{M}_G$ .

Combining the expressions for the norm and the log-likelihood, we see that maximizing the log-likelihood over concentration matrices in the model  $\mathcal{M}_G$  is equivalent to minimizing

$$-\ell_Y(g^\top g) = \frac{1}{n} \|g \cdot Y\|^2 - \log \det(g^\top g)$$

over  $g \in G$ . We show that this minimization can be done in two steps. First, we minimize the norm over the subgroup  $G_{\text{SL}}^\pm$ , consisting of matrices in  $G$  of determinant  $\pm 1$ . Then, we find the scalar multiple of this matrix that minimizes the overall expression. For this, we require that the group  $G$  is closed under non-zero scalar multiples, i.e. if  $g \in G$  then  $\lambda g \in G$  for all non-zero real scalars  $\lambda$ .

**Proposition 3.4.** *Let  $Y \in V^n$  be a tuple of samples. If the group  $G \subseteq \text{GL}(V)$  is closed under non-zero scalar multiples, the supremum of the log-likelihood (3) over  $\mathcal{M}_G$  is the double infimum*

$$-\inf_{\lambda \in \mathbb{R}_{>0}} \left( \frac{\lambda}{n} \left( \inf_{h \in G_{\text{SL}}^\pm} \|h \cdot Y\|^2 \right) - \dim(V) \log \lambda \right).$$

*The MLEs, if they exist, are the matrices  $\lambda h^\top h$ , where  $h$  minimizes  $\|h \cdot Y\|$  under the action of  $G_{\text{SL}}^\pm$  on  $V^n$ , and  $\lambda \in \mathbb{R}_{>0}$  is the unique value minimizing the outer infimum.*

*Proof.* Maximizing  $\ell_Y(\Psi)$  over  $\mathcal{M}_G$  is equivalent to minimizing

$$\begin{aligned} f: G &\rightarrow \mathbb{R} \\ g &\mapsto \frac{1}{n} \|g \cdot Y\|^2 - \log \det(g^\top g), \end{aligned}$$

since  $f(g)$  only depends on the positive definite matrix  $g^\top g$ . We write  $g \in G$  as  $g = \tau h$ , where  $\tau \in \mathbb{R}_{>0}$  and  $h \in G_{\text{SL}}^\pm$ . Using  $g^\top g = \tau^2 h^\top h$ , and setting  $\lambda := \tau^2$  and  $m := \dim(V)$ , we have

$$f(g) = \frac{\tau^2}{n} \|h \cdot Y\|^2 - \log \det(\tau^2 h^\top h) = \frac{\tau^2}{n} \|h \cdot Y\|^2 - \log(\tau^{2m}) = \frac{\lambda}{n} \|h \cdot Y\|^2 - m \log(\lambda).$$

The minimum value of the function  $\lambda \mapsto \lambda C - \log(\lambda)$  is  $\log(C) + 1$  for  $C > 0$ , which increases as  $C$  increases. Hence, to minimize  $f$ , we can first find the minimal norm in the orbit closure and then minimize the univariate function in  $\lambda$ , i.e.

$$\inf_{g \in G} f(g) = \inf_{\lambda \in \mathbb{R}_{>0}} \left( \frac{\lambda}{n} \left( \inf_{h \in G_{\text{SL}}^\pm} \|h \cdot Y\|^2 \right) - m \log \lambda \right).$$

Furthermore, an MLE is a matrix  $\hat{\Psi} \in \mathcal{M}_G$  that maximizes  $\ell_Y(\Psi)$ . Comparing  $\ell_Y(\Psi)$  with the infimum in the claim, we see that the MLEs are all matrices  $\hat{\Psi} = g^\top g = \lambda h^\top h$ , where  $g = \sqrt{\lambda} h$ , and  $h$  and  $\lambda$  minimize the inner and outer infima respectively.  $\square$



The group  $G_{\text{SL}}^\pm$  may split into two parts:  $G_{\text{SL}}^+$  consisting of matrices in  $G$  of determinant one, and  $G_{\text{SL}}^-$  consisting of matrices of determinant  $-1$ . If we prefer to optimize over one part, say  $G_{\text{SL}}^+$ , we can compute the capacity of  $Y$  under  $G_{\text{SL}}^\pm$  by doing two minimizations. A fixed matrix  $h' \in G_{\text{SL}}^-$  gives a bijection between  $G_{\text{SL}}^+$  and  $G_{\text{SL}}^-$  via  $h \mapsto hh'$ . Hence we can minimize  $\|h \cdot Y\|$  over  $G_{\text{SL}}^\pm$  by minimizing both  $\|h \cdot Y\|$  and  $\|h \cdot (h' \cdot Y)\|$  over  $G_{\text{SL}}^+$ . However, we can ignore neither  $G_{\text{SL}}^+$  nor  $G_{\text{SL}}^-$ . The following is an example of a group, closed under non-zero scalar multiples, such that the norm  $\|h \cdot Y\|$  can be attained at one but not the other.

**Example 3.5.** Let the group  $G$  consist of non-zero scalar multiples of block-diagonal  $6 \times 6$  matrices of the form

$$(5) \quad \begin{bmatrix} M & 0 & 0 \\ 0 & S_1 M S_1^{-1} & 0 \\ 0 & 0 & S_2 M S_2^{-1} \end{bmatrix}, \quad \text{where} \quad S_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

and  $M \in O_2$  is an orthogonal  $2 \times 2$  matrix. The component  $G_{\text{SL}}^+$  consists of matrices in (5) where  $M$  is special orthogonal, while the component  $G_{\text{SL}}^-$  consists of matrices in (5) where  $M$  is orthogonal with determinant  $-1$ . Note that although the group  $G$  contains matrices of determinant  $-1$ , it does not contain any orthogonal matrices of determinant  $-1$ .

The norm of  $\|g \cdot Y\|$ , for a tuple of samples  $Y$ , can be expressed in terms of the sample covariance matrix  $S_Y$ . Consider the tuple of four samples given by

$$Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 2\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 2\sqrt{5} \\ 0 & 0 & \frac{6\sqrt{5}}{5} & \frac{8\sqrt{5}}{5} \end{bmatrix}, \quad \text{with} \quad S_Y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_1^2 \end{bmatrix}.$$

The capacity problem can be rewritten as minimizing the trace  $\text{tr}(g^\top g S_Y)$  over matrices  $g \in G_{\text{SL}}^\pm$ , by (4), to give

$$\inf_{h \in G_{\text{SL}}^\pm} \|h \cdot Y\|^2 = 4 \cdot \inf_{M \in O_2} \left[ \text{tr}((S_1 M S_1^{-1})^\top (S_1 M S_1^{-1}) S_2) + \text{tr}((S_2 M S_2^{-1})^\top (S_2 M S_2^{-1}) S_1^2) \right].$$

We can parametrize the  $2 \times 2$  special orthogonal matrices by  $P$  and the  $2 \times 2$  orthogonal matrices of determinant  $-1$  by  $Q$  where

$$P = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \quad Q = \begin{bmatrix} -a & -b \\ -b & a \end{bmatrix}, \quad \text{with} \quad a, b \in \mathbb{R}, \quad \text{and} \quad a^2 + b^2 = 1.$$

Then the minimization problems over  $G_{\text{SL}}^+$  and  $G_{\text{SL}}^-$  can be rewritten as

$$\begin{aligned} \inf_{h \in G_{\text{SL}}^+} \frac{1}{4} \|h \cdot Y\|^2 &= \min_{a^2 + b^2 = 1} \left( 13a^2 - \frac{44}{3}ab + \frac{419}{12}b^2 \right), \\ \inf_{h \in G_{\text{SL}}^-} \frac{1}{4} \|h \cdot Y\|^2 &= \min_{a^2 + b^2 = 1} \left( \frac{71}{3}a^2 - \frac{28}{3}ab + \frac{97}{4}b^2 \right). \end{aligned}$$

Note that both infima can only be attained for  $a$  and  $b$  having the same sign, because of the negative coefficients of  $ab$ ; we assume  $a, b \geq 0$ . Substituting  $b = \sqrt{1 - a^2}$  in the latter minimum, we see that

$$\frac{71}{3}a^2 + \frac{97}{4}(1 - a^2) - \frac{28}{3}a\sqrt{1 - a^2} \geq \frac{97}{4} + \left(\frac{71}{3} - \frac{97}{4}\right) - \frac{28}{3} \cdot \frac{1}{2} = 19.$$

In contrast, setting  $a = 1$  and  $b = 0$  in the former minimum gives a value of 13. Hence  $\inf_{h \in G_{\text{SL}}^+} \|h \cdot Y\|^2 < \inf_{h \in G_{\text{SL}}^-} \|h \cdot Y\|^2$ . Multiplying  $Y$  by a fixed matrix in  $G_{\text{SL}}^-$  gives a tuple of samples where the strict inequality is reversed, and the minimum is attained only at the negative component  $G_{\text{SL}}^-$ .  $\diamond$

**3.2. Relating stability to the MLE.** We use Proposition 3.4 to prove the following correspondence between stability notions and MLE existence.

**Theorem 3.6.** *Consider a tuple  $Y \in V^n$  of samples, and a group  $G \subseteq \text{GL}(V)$  that is closed under non-zero scalar multiples. The stability under the action of  $G_{\text{SL}}^\pm$  on  $V^n$  is related to ML estimation for the Gaussian group model  $\mathcal{M}_G$  as follows.*

- (a)  $Y$  unstable  $\Leftrightarrow \ell_Y$  not bounded from above
- (b)  $Y$  semistable  $\Leftrightarrow \ell_Y$  bounded from above
- (c)  $Y$  polystable  $\Rightarrow$  MLE exists

*Proof.* If  $Y$  is unstable then  $C := \inf_{h \in G_{\text{SL}}^\pm} \|h \cdot Y\|^2 = 0$ . Hence the outer infimum from Proposition 3.4 equals  $-\infty$ , so the supremum of  $\ell_Y$  is infinite. Conversely, if  $Y$  is semistable, then  $C > 0$  and thus the outer infimum from Proposition 3.4 is some real number and  $\ell_Y$  is bounded from above. This gives parts (a) and (b).

If  $Y$  is polystable, then the infimum  $C > 0$  is attained for some  $h \in G_{\text{SL}}^\pm$  and  $\lambda h^\top h$  is an MLE, where  $\lambda \in \mathbb{R}_{>0}$  minimizes the outer infimum in Proposition 3.4.  $\square$

*Remark 3.7.* Assume that  $G$  contains an orthogonal matrix of determinant  $-1$ , say  $o \in G$ . Then minimizing the norm  $\|h \cdot Y\|$  over  $G_{\text{SL}}^\pm$  is equivalent to minimizing it over  $G_{\text{SL}}^+$ . Hence, in this case, Proposition 3.4 and Theorem 3.6 both hold for  $G_{\text{SL}}^+$  as well as  $G_{\text{SL}}^\pm$ . This is because we can write  $g \in G$  as  $g = \tau o h$ , where  $\tau \in \mathbb{R}_{>0}$  and  $h \in G_{\text{SL}}^+$ , and then follow the computations in the proof of Proposition 3.4.

If we add more assumptions on our group  $G$ , we can strengthen Theorem 3.6 using the Kempf-Ness theorem over  $\mathbb{R}$ . First, we assume that  $G$  is Zariski closed, i.e. the zero locus of a set of polynomials in the matrix entries. Since we want the group  $G$  to be closed under non-zero scalar multiples, it can be expressed as the zero locus of homogeneous polynomials. Secondly, we assume that  $G$  is self-adjoint, i.e.  $g^\top \in G$  for every  $g \in G$ . On the statistics side, this implies that the set of concentration matrices in  $\mathcal{M}_G$  is equal to the set of covariance matrices in the model. These additional assumptions hold for Examples 3.1, 3.2, and 3.3.

**Lemma 3.8.** *Let  $G \subseteq \text{GL}(V)$  be a Zariski closed self-adjoint group, closed under non-zero scalar multiples. If there is an element of  $G$  with negative determinant, then  $G$  contains an orthogonal matrix of determinant  $-1$ . In particular, Proposition 3.4 and Theorem 3.6 still hold after replacing  $G_{\text{SL}}^\pm$  by  $G_{\text{SL}}^+$ .*

*Proof.* Pick  $g \in G$  with  $\det(g) < 0$ . Since  $G$  is Zariski closed and self-adjoint, the polar decomposition can be carried out in  $G$ , by [Wal17, Theorem 2.16]. In particular, there is an orthogonal matrix  $o \in G$  and a positive definite matrix  $p \in G$  such that  $g = op$ . Then  $\det(g) < 0$  implies  $\det(o) < 0$ , i.e.  $\det(o) = -1$ . The second part of the claim follows from Remark 3.7.  $\square$

As a consequence of Lemma 3.8 we work with  $G_{\text{SL}}^+$  (instead of  $G_{\text{SL}}^\pm$ ) in the following.

**Proposition 3.9.** *Let  $Y \in V^n$  be a tuple of samples, and  $G \subseteq \text{GL}(V)$  a Zariski closed self-adjoint group which is closed under non-zero scalar multiples. If  $\lambda h^\top h$  is an MLE given  $Y$ , with  $h \in G_{\text{SL}}^+$  and  $\lambda \in \mathbb{R}_{>0}$ , then all MLEs given  $Y$  are of the form  $g^\top(\lambda h^\top h)g$ , where  $g$  is in the  $G_{\text{SL}}^+$ -stabilizer of  $Y$ .*

*Proof.* By Proposition 3.4 for  $G_{\text{SL}}^+$ , the matrix  $h$  minimizes the norm of  $Y$  under the action of  $G_{\text{SL}}^+$  and hence so does  $hg$  for any  $g$  in the  $G_{\text{SL}}^+$ -stabilizer of  $Y$ . Therefore,  $\lambda(hg)^\top hg = g^\top(\lambda h^\top h)g$  is another MLE. Conversely, by Proposition 3.4 any MLE is of the form  $\lambda(h')^\top h'$  with  $h' \in G_{\text{SL}}^+$  such that

$$\|h' \cdot Y\|^2 = \inf_{\tilde{h} \in G_{\text{SL}}^+} \|\tilde{h} \cdot Y\|^2 = \|h \cdot Y\|^2.$$

Since  $G \subseteq \text{GL}(V)$  is Zariski closed and self-adjoint,  $G_{\text{SL}}^+ \subseteq \text{GL}(V)$  is Zariski closed and self-adjoint and so is its diagonal embedding into  $\text{GL}(V^n)$ . Thus we can apply Kempf-Ness, Theorem 2.2(b). For the  $G_{\text{SL}}^+$  action on  $V^n$ , there is an orthogonal matrix  $o \in G_{\text{SL}}^+$  with  $o \cdot (h \cdot Y) = h' \cdot Y$ . Hence,  $g := h^{-1} o^{-1} h'$  is in the  $G_{\text{SL}}^+$ -stabilizer of  $Y$  and using  $h' = ohg$  we deduce  $\lambda(h')^\top h' = g^\top(\lambda h^\top h)g$ .  $\square$

With these extra assumptions on the group  $G$ , we obtain a stronger version of Theorem 3.6. Moreover, with these assumptions we are in the setting of [BFG<sup>+</sup>19], so we can use their algorithmic methods to compute the capacity in order to find an MLE. We discuss these connections to algorithms for matrix normal models in Section 4.

**Theorem 3.10.** *Let  $Y \in V^n$  be a tuple of samples, and  $G \subseteq \text{GL}(V)$  a Zariski closed self-adjoint group that is closed under non-zero scalar multiples. The stability under the action of  $G_{\text{SL}}^+$  on  $V^n$  is related to ML estimation for the Gaussian group model  $\mathcal{M}_G$  as follows.*

- (a)  $Y$  unstable  $\Leftrightarrow \ell_Y$  not bounded from above
- (b)  $Y$  semistable  $\Leftrightarrow \ell_Y$  bounded from above
- (c)  $Y$  polystable  $\Leftrightarrow$  MLE exists
- (d)  $Y$  stable  $\Rightarrow$  finitely many MLEs exist  $\Leftrightarrow$  unique MLE exists

*Proof.* We recall that the action of  $G_{\text{SL}}^+$  on  $V^n$  is given by the diagonal embedding into  $\text{GL}(V^n)$ , and that this turns  $G_{\text{SL}}^+$  into a Zariski closed self-adjoint subgroup of  $\text{GL}(V^n)$  by the assumptions on  $G \subseteq \text{GL}(V)$ .

By Theorem 3.6, it remains to prove the converse implication in (c) and condition (d). If an MLE given  $Y$  exists, then the log-likelihood function  $\ell_Y$  is bounded from above and attains its maximum. Hence the double infimum from Proposition 3.4 is attained, and there exists  $h \in G_{\text{SL}}^+$  such that  $h \cdot Y$  has minimal norm in the orbit of  $Y$  under  $G_{\text{SL}}^+$ . Hence the orbit is closed by Kempf-Ness, Theorem 2.2(d), and  $Y$  is polystable.

We now prove condition (d). If  $Y$  is stable, its stabilizer  $\text{Stab}_Y$  is finite. Then there are only finitely many MLEs given  $Y$ , by Proposition 3.9. It remains to show that a tuple  $Y$  cannot have finitely many MLEs unless it has a unique MLE. A tuple  $Y$  with finitely many MLEs is polystable, by condition (c). Moreover, we can relate the stabilizers of  $Y$  and  $h \cdot Y$  by  $\text{Stab}_{h \cdot Y} = h \text{Stab}_Y h^{-1}$ . Combining Propositions 3.4 and 3.9, we can relate the MLEs given  $Y$  to the MLEs given  $h \cdot Y$  via

$$\{\text{MLEs given } h \cdot Y\} = (h^{-1})^\top \{\text{MLEs given } Y\} h^{-1}.$$

Hence, to study the stabilizer and MLE of a polystable  $Y$  we can assume that  $Y$  is of minimal norm in its orbit under  $G_{\text{SL}}^+$ . One of the MLEs given  $Y$  is then  $\lambda I$ , where  $\lambda > 0$  minimizes the outer infimum in Proposition 3.4, and  $I$  is the identity matrix of size  $\dim(V)$ .

We show that the set  $\{g^\top g \mid g \in \text{Stab}_Y\}$  is either the identity matrix or infinite. This implies that  $Y$  either has a unique MLE or infinitely many MLEs, because the MLEs given  $Y$  are the matrices  $g^\top (\lambda I^\top I) g = \lambda g^\top g$ , where  $g \in \text{Stab}_Y$ , by Proposition 3.9. The group  $\text{Stab}_Y$  is self-adjoint by [Wal17, Corollary 2.25]. If it is contained in the set of orthogonal matrices, then  $\{g^\top g \mid g \in \text{Stab}_Y\}$  consists only of the identity matrix. Otherwise, let  $h \in \text{Stab}_Y$  be non-orthogonal. Then  $h^\top \in \text{Stab}_Y$  and hence  $h^\top h \in \text{Stab}_Y$ , and this positive definite matrix is not equal to the identity matrix. The matrix  $h^\top h$  has infinite order, since the eigenvalues of  $(h^\top h)^N$  are the  $N$ th powers of the eigenvalues of  $h^\top h$ , and there exist eigenvalues that are not equal to one. Since  $(h^\top h)^N \in \text{Stab}_Y$  and  $((h^\top h)^N)^\top ((h^\top h)^N) = (h^\top h)^{2N}$ , the set  $\{g^\top g \mid g \in \text{Stab}_Y\}$  is infinite.  $\square$

*Remark 3.11.* In the setting of a Zariski closed self-adjoint group  $G$  closed under non-zero scalar multiples, the results in Proposition 3.4, Proposition 3.9, and Theorem 3.10 are unchanged if we replace  $G_{\text{SL}}^+$  by the larger subgroup  $G_{\text{SL}}^\pm$ , by the same argument as in Lemma 3.8. In fact, we can also replace  $G_{\text{SL}}^+$  by the smaller group  $G_{\text{SL}}^\circ$ , the identity component of  $G_{\text{SL}}^+$ . This is because the quotient group  $G_{\text{SL}}^+/G_{\text{SL}}^\circ$  is finite and every equivalence class has an orthogonal matrix representative, by the polar decomposition [Wal17, Theorem 2.16]. The same argument holds for any Zariski-closed self-adjoint subgroup  $H$  of  $G$  with the same identity component as  $G_{\text{SL}}^+$ . We may not have such choices for groups that are not Zariski closed and self-adjoint, see Example 3.5.

We note that the converse of Theorem 3.10(d) does not hold by Example 4.2 from the next section. We also stress the importance of the assumption that the group  $G$  is self-adjoint for condition (d). This assumption is needed to conclude that the MLE is unique from the fact that there are finitely many MLEs. Indeed, the following example exhibits a Zariski closed group  $G$ , closed under non-zero scalar multiples, for which there exist samples  $Y$  with a finite number of MLEs in the Gaussian group model given by  $G$ , but not a unique MLE.

**Example 3.12.** Let  $G$  be generated by  $-I$  and all non-zero scalar multiples of a non-orthogonal matrix  $M$  with  $M^2 = I$ . For example, we can take

$$M = \begin{bmatrix} 1/2 & 3 \\ 1/4 & -1/2 \end{bmatrix}.$$

The group consists of non-zero scalar multiples of the matrices  $M$  and  $I_2$ . The MLEs to the Gaussian group model  $\mathcal{M}_G$  given samples  $Y$  are given by group elements  $h \in G_{\text{SL}}^\pm$  that minimize the norm  $\|h \cdot Y\|$ , by Proposition 3.4. Since scaling the matrix by some  $\lambda$  scales its determinant by  $\lambda^2$ , the subset  $G_{\text{SL}}^+$  consists of  $\pm I_2$ , and the subset  $G_{\text{SL}}^-$  consists of the matrices  $\pm M$ . Consider the single sample

$$Y = \begin{bmatrix} 6 \\ 1 \end{bmatrix}.$$

Then  $\|M \cdot Y\|^2 = \|Y\|^2$ , and the sample  $Y$  has exactly two distinct MLEs.  $\diamond$

**3.3. Complex Gaussian models.** Invariant theory is more classical over the field of complex numbers than over the reals numbers. We see in this section that several of our results can be simplified and strengthened when working over  $\mathbb{C}$ . The statistical consequences concern statistical models over the complex numbers, as in [Woo56, Goo63, AHSE95].

We consider a complex vector space  $V$  and a subgroup  $G \subseteq \text{GL}_{\mathbb{C}}(V)$  of the complex general linear group on  $V$ . To view the group elements in  $G$  as invertible matrices we fix an isomorphism  $V \cong \mathbb{C}^m$ . The complex *Gaussian group model*  $\mathcal{M}_G$  consists of all multivariate distributions of mean zero whose concentration matrix is of the form  $g^*g$  for some  $g \in G$ . The log-likelihood function becomes

$$(6) \quad \ell_Y(\Psi) = \log \det(\Psi) - \text{tr}(\Psi S_Y), \quad \text{where } S_Y := \frac{1}{n} \sum_{i=1}^n Y_i Y_i^*.$$

For the action of the group  $G$  on a tuple  $Y = (Y_1, \dots, Y_n) \in V^n$  given by  $g \cdot Y = (gY_1, \dots, gY_n)$ , the norm becomes

$$\|g \cdot Y\|^2 = \sum_{i=1}^n (gY_i)^* gY_i = n \text{tr}(g^*g S_Y).$$

Hence, as before, maximizing the log-likelihood over concentration matrices in the complex Gaussian group model  $\mathcal{M}_G$  is equivalent to minimizing

$$-\ell_Y(g^*g) = \frac{1}{n} \|g \cdot Y\|^2 - \log \det(g^*g).$$

Analogously to Proposition 3.4, this can be done in two steps. Since we now work over  $\mathbb{C}$ , we only need to compute the capacity under the subgroup  $G_{\text{SL}}^+ \subseteq G$  of matrices with determinant one, instead of using  $G_{\text{SL}}^\pm$ . In particular, the situation described in Example 3.5 cannot happen over  $\mathbb{C}$ , and we do not need to consider the extra assumptions in Remark 3.7.

**Proposition 3.13.** *Let  $Y \in V^n$  be a tuple of complex samples. If the group  $G$  is closed under non-zero complex scalar multiples, the supremum of the log-likelihood (6) over the model  $\mathcal{M}_G$  is the double infimum*

$$- \inf_{\lambda \in \mathbb{R}_{>0}} \left( \frac{\lambda}{n} \left( \inf_{h \in G_{\text{SL}}^+} \|h \cdot Y\|^2 \right) - \dim(V) \log \lambda \right).$$

The MLEs, if they exist, are the matrices  $\lambda h^* h$ , where  $h$  minimizes  $\|h \cdot Y\|$  under the action of  $G_{\text{SL}}^+$  on  $V^n$ , and  $\lambda \in \mathbb{R}_{>0}$  is the unique value minimizing the outer infimum.

*Proof.* The proof is analogous to the proof of Proposition 3.4. The only difference is that we can write  $g \in G$  as  $g = \tau h$ , where  $\tau \in \mathbb{C} \setminus \{0\}$  and  $h \in G_{\text{SL}}^+$ . Then we see that

$$-\ell_Y(g^* g) = \frac{|\tau|^2}{n} \|h \cdot Y\|^2 - \dim(V) \log(|\tau|^2).$$

Setting  $\lambda = |\tau|^2$  and continuing as in the proof of Proposition 3.4, shows the claim.  $\square$

Using the same assumptions as in Proposition 3.13, we see that Theorem 3.6 holds over  $\mathbb{C}$  after replacing  $G_{\text{SL}}^\pm$  by  $G_{\text{SL}}^+$ . The most important difference between the real and the complex setting is that Theorem 3.10(d) is an equivalence over  $\mathbb{C}$ . In Example 4.2, we will see that this is not true over  $\mathbb{R}$ . In the remainder of this section, we prove this equivalence for complex Gaussian group models given by self-adjoint groups  $G$ , i.e. groups that satisfy  $g^* \in G$  for every  $g \in G$ . We first give an analogue of Proposition 3.9 over  $\mathbb{C}$ .

**Proposition 3.14.** *Let  $Y \in V^n$  be a tuple of complex samples, and  $G \subseteq \text{GL}_{\mathbb{C}}(V)$  be a Zariski closed self-adjoint group, which is closed under non-zero complex scalar multiples. If  $\lambda h^* h$  is an MLE given  $Y$ , with  $h \in G_{\text{SL}}^+$  and  $\lambda \in \mathbb{R}_{>0}$ , then all MLEs given  $Y$  are of the form  $g^*(\lambda h^* h)g$ , where  $g$  is in the  $G_{\text{SL}}^+$ -stabilizer of  $Y$ .*

*Proof.* This is proven analogously as Proposition 3.9 using the complex version of Kempf-Ness Theorem 2.2 and Proposition 3.13 instead of Proposition 3.4.  $\square$

**Theorem 3.15.** *Consider a tuple  $Y \in V^n$  of complex samples, and let  $G \subseteq \text{GL}_{\mathbb{C}}(V)$  be a Zariski closed self-adjoint group, which is closed under non-zero complex scalar multiples. The stability under the action of  $G_{\text{SL}}^+$  on  $V^n$  is related to ML estimation for the complex Gaussian group model  $\mathcal{M}_G$  as follows.*

- (a)  $Y$  unstable  $\Leftrightarrow \ell_Y$  not bounded from above
- (b)  $Y$  semistable  $\Leftrightarrow \ell_Y$  bounded from above
- (c)  $Y$  polystable  $\Leftrightarrow$  MLE exists
- (d)  $Y$  stable  $\Leftrightarrow$  finitely many MLEs exist  $\Leftrightarrow$  unique MLE exists

*Proof.* We prove that uniqueness of the MLE given  $Y$  implies that  $Y$  is stable. The proofs of the other parts of the theorem are the same as in the real setting in Theorems 3.6 and 3.10.

Let us assume that the MLE given  $Y$  exists uniquely. We see from (c) that  $Y$  is polystable. Hence, we need to show that the  $G_{\text{SL}}^+$ -stabilizer of  $Y$ , denoted by  $\text{Stab}_Y$ , is finite. For  $h \in G_{\text{SL}}^+$  we have  $\text{Stab}_{h \cdot Y} = h \text{Stab}_Y h^{-1}$  and, from Proposition 3.14, we have

$$\{\text{MLEs given } h \cdot Y\} = (h^{-1})^* \{\text{MLEs given } Y\} h^{-1}.$$

As in the real setting, this allows us to assume that  $Y$  is of minimal norm in its orbit under  $G_{\text{SL}}^+$ . Then  $\lambda I$  is the MLE given  $Y$ , where  $\lambda \in \mathbb{R}_{>0}$  minimizes the outer infimum in Proposition 3.13. Since the matrix  $\lambda I$  is the unique MLE, the stabilizer  $\text{Stab}_Y$  is contained in the group of unitary matrices in  $G$ , by Proposition 3.14. In particular,

$\text{Stab}_Y$  is  $\mathbb{C}$ -compact. As the subgroup  $\text{Stab}_Y$  is also Zariski closed (defined by the equations  $gY = Y$ ) we conclude that  $\text{Stab}_Y$  is finite.  $\square$

#### 4. MATRIX NORMAL MODELS

In this section we study matrix normal models, which we have already seen in Example 3.3. Consider the multivariate Gaussian of dimension  $m = m_1 m_2$ . A matrix normal model is a sub-model consisting of covariance matrices that factor as a Kronecker product  $\Sigma_1 \otimes \Sigma_2$  where  $\Sigma_i \in \text{PD}_{m_i}$ . Setting  $\Psi_1 := \Sigma_1^{-1}$  and  $\Psi_2 := \Sigma_2^{-1}$ , we can write the log-likelihood function (2) for the matrix normal model as

$$(7) \quad \ell_Y(\Psi_1, \Psi_2) = m_2 \log \det(\Psi_1) + m_1 \log \det(\Psi_2) - \frac{1}{n} \text{tr} \left( \Psi_1 \sum_{i=1}^n Y_i \Psi_2 Y_i^\top \right).$$

An MLE is a concentration matrix  $\hat{\Psi}_1 \otimes \hat{\Psi}_2 \in \text{PD}_{m_1} \otimes \text{PD}_{m_2}$  that maximizes the log-likelihood. Unless specified, we refer to matrix normal models over the real numbers and abbreviate  $\text{GL}_m(\mathbb{R})$  and  $\text{SL}_m(\mathbb{R})$  to  $\text{GL}_m$  and  $\text{SL}_m$  respectively.

**4.1. Relating norm minimization to ML estimation.** We describe how to specialize our results for Gaussian group models from Section 3 to matrix normal models. For this, consider the left-right action of  $\text{GL}_{m_1} \times \text{GL}_{m_2}$  on  $(\mathbb{R}^{m_1 \times m_2})^n$  given by

$$(8) \quad g \cdot Y := (g_1 Y_1 g_2^\top, \dots, g_1 Y_n g_2^\top),$$

where  $Y = (Y_1, \dots, Y_n)$  is a sample tuple in  $(\mathbb{R}^{m_1 \times m_2})^n$  and  $g = (g_1, g_2) \in \text{GL}_{m_1} \times \text{GL}_{m_2}$ . The left-right action induces the representation

$$\varrho: \text{GL}_{m_1} \times \text{GL}_{m_2} \rightarrow \text{GL}_{m_1 m_2}, \quad (g_1, g_2) \mapsto g_1 \otimes g_2$$

and the matrix normal model arises as the Gaussian group model of  $G := \varrho(\text{GL}_{m_1} \times \text{GL}_{m_2})$ .

The subgroup  $G \subseteq \text{GL}_{m_1 m_2}$  is Zariski closed, self-adjoint and closed under non-zero scalar multiples. Therefore, our results from the previous section apply to the action of  $G_{\text{SL}}^+$ . However, it is possible and more convenient to directly work with the left-right action of  $\text{SL}_{m_1} \times \text{SL}_{m_2}$ . The following theorem makes this precise.

**Theorem 4.1.** *Let  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  be a matrix tuple. The supremum of the log-likelihood  $\ell_Y$  in (7) over  $\text{PD}_{m_1} \times \text{PD}_{m_2}$  is given by the double infimum*

$$(9) \quad - \inf_{\lambda \in \mathbb{R}_{>0}} \left( \frac{\lambda}{n} \left( \inf_{h \in \text{SL}_{m_1} \times \text{SL}_{m_2}} \|h \cdot Y\|^2 \right) - m_1 m_2 \log \lambda \right).$$

*The MLEs, if they exist, are the matrices of the form  $\lambda h_1^\top h_1 \otimes h_2^\top h_2$ , where  $h = (h_1, h_2)$  minimizes  $\|h \cdot Y\|$  under the left-right action of  $\text{SL}_{m_1} \times \text{SL}_{m_2}$ , and  $\lambda \in \mathbb{R}_{>0}$  is the unique value that minimizes the outer infimum.*

*If there are several MLEs given  $Y$ , they are related via the stabilizer of  $Y$  in  $\text{SL}_{m_1} \times \text{SL}_{m_2}$ . More precisely, every  $(g_1, g_2)$  in the stabilizer of  $Y$  yields an MLE  $\lambda g_1^\top h_1^\top h_1 g_1 \otimes g_2^\top h_2^\top h_2 g_2$  and, conversely, every MLE given  $Y$  is of this form.*

The stability under the left-right action of  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  is related to ML estimation via:

- (a)  $Y$  unstable  $\Leftrightarrow \ell_Y$  not bounded from above
- (b)  $Y$  semistable  $\Leftrightarrow \ell_Y$  bounded from above
- (c)  $Y$  polystable  $\Leftrightarrow$  MLE exists
- (d)  $Y$  stable  $\Rightarrow$  MLE exists uniquely

*Proof.* The subgroup  $H := \varrho(\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}) \subseteq G$  is Zariski closed, self-adjoint and shares the same identity component as  $G_{\mathrm{SL}}^+$ . Thus Propositions 3.4, 3.9 and Theorem 3.10 apply to  $H$  as well, by Remark 3.11. Furthermore, the kernel of  $\varrho$  when restricted to  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  is finite. Hence, the stability notions in Definition 2.1(a)–(d) coincide for  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  and  $H$ , so we can consider  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  instead of its image  $H$  under  $\varrho$ .  $\square$

We have seen in Theorem 3.15 that over the complex numbers, the converse of Theorem 4.1(d) also holds. However, over the reals there exist matrix tuples  $Y$  with a unique MLE but an infinite stabilizer, as the following example shows.

**Example 4.2.** We set  $m_1 = m_2 = n = 2$  and take  $Y \in (\mathbb{R}^{2 \times 2})^2$ , where

$$Y_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We prove that the MLE given  $Y$  is unique although the stabilizer of  $Y$  is infinite.

We first show that  $Y$  is polystable under the left-right action of  $\mathrm{SL}_2 \times \mathrm{SL}_2$ . Note that any matrix in  $\mathrm{SL}_2$  has Frobenius norm at least  $\sqrt{2}$ . Indeed, if  $\sigma_1$  and  $\sigma_2$  are the singular values of  $g$ , then  $\|g\|^2 = \sigma_1^2 + \sigma_2^2$ , where  $\sigma_1\sigma_2 = 1$ . By the arithmetic mean - geometric mean inequality, we have  $\|g\|^2 \geq 2$ . Therefore  $Y_1$  and  $Y_2$  have minimal Frobenius norm in  $\mathrm{SL}_2$  and thus  $Y$  is of minimal norm in its orbit. By Kempf-Ness, Theorem 2.2(d), the matrix tuple  $Y$  is polystable.

The stabilizer of  $Y$  consists of matrices  $(g_1, g_2) \in \mathrm{SL}_2 \times \mathrm{SL}_2$  with  $g_1 Y_i g_2^\top = Y_i$ . For  $Y_1$ , this gives  $g_1 g_2^\top = I_2$ , i.e.  $g_2^\top = g_1^{-1}$ . Then, from  $Y_2$ , we obtain  $g_1 Y_2 = Y_2 g_1$ , and so

$$g_1 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{with } a^2 + b^2 = 1,$$

i.e.  $g_1 \in \mathrm{SO}_2(\mathbb{R})$  and hence  $g_2 = g_1^{-\top} = g_1$ . Thus the stabilizer of  $Y$  is contained in the infinite set  $\{(g, g) \mid g \in \mathrm{SO}_2\}$ . In fact, we have equality, as  $\mathrm{SO}_2$  is commutative and  $Y_1, Y_2 \in \mathrm{SO}_2$ .

Since  $Y$  is of minimal norm in its orbit, we use Theorem 4.1 to conclude that  $\lambda I_2 \otimes I_2$  is an MLE. Any other MLE is given by  $\lambda g_1^\top I_2 g_1 \otimes g_2^\top I_2 g_2$  for some  $(g_1, g_2)$  in the stabilizer of  $Y$ . Since the stabilizer is contained in  $\mathrm{SO}_2 \times \mathrm{SO}_2$ , the MLE is unique.

We remark that for the complex matrix normal model the MLEs involve  $g^*g$  rather than  $g^\top g$ , by Proposition 3.14, hence from the complex stabilizer  $\{(g, g) \mid g \in \mathrm{SO}_2(\mathbb{C})\}$  we obtain infinitely many MLEs.  $\diamond$

The following example shows that all stability conditions in Theorem 4.1(a)–(d) can occur.



**Example 4.3.** We set  $m_1 = m_2 = 2$ , and study stability under  $\mathrm{SL}_2 \times \mathrm{SL}_2$  on  $(\mathbb{R}^{2 \times 2})^n$ . We use the matrices

$$Y_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad Y_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y_4 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

- (a) The matrix  $Y_4$  is unstable and the matrix tuple  $(Y_4, Y_4)$  is unstable as well.
- (b) The orbit of the matrix tuple  $(Y_1, Y_4)$  is contained in  $\{(g, M) \mid g \in \mathrm{SL}_2, M \neq 0\}$ . In particular,  $(Y_1, Y_4)$  is semistable as  $\mathrm{SL}_2$  is closed. Moreover, for any  $g \in \mathrm{SL}_2$  and  $M \in \mathbb{R}^{2 \times 2} \setminus \{0\}$  we have

$$\|(g, M)\|^2 = \|g\|^2 + \|M\|^2 \geq 2 + \|M\|^2 > 2,$$

where we used  $\|g\|^2 \geq 2$ , see Example 4.2. On the other hand, we have

$$\left( \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix}, \begin{pmatrix} \varepsilon^{-1} & 0 \\ 0 & \varepsilon \end{pmatrix} \right) \cdot (Y_1, Y_4) = \left( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \varepsilon^2 \\ 0 & 0 \end{pmatrix} \right),$$

which tends to  $(Y_1, 0)$  as  $\varepsilon \rightarrow 0$ . Since  $\|(Y_1, 0)\|^2 = 2$  the capacity of  $(Y_1, Y_4)$  is not attained by an element in the orbit of  $(Y_1, Y_4)$ , and  $Y$  is not polystable.

- (c) The matrix  $Y_1 = I_2$  is polystable by Kempf-Ness, Theorem 2.2(d), as it is an  $\mathrm{SL}_2$  matrix of minimal norm. An MLE is given by  $\lambda I_2 \otimes I_2$ , where  $\lambda$  is the minimizer of the outer infimum in (9). Furthermore,  $Y_1$  is not stable, because its stabilizer is  $\{(g, g^{-\top}) \mid g \in \mathrm{SL}_2\}$ . There are infinitely many MLEs given  $Y$ , of the form  $\lambda g^\top g \otimes g^{-1} g^{-\top}$  for  $g \in \mathrm{SL}_2$ , see Theorem 4.1.
- (d) We show that  $Y = (Y_1, Y_2, Y_3)$  is stable. First, any tuple  $(M_1, M_2, M_3)$  in the orbit of  $Y$  satisfies  $M_1, M_2 \in \mathrm{SL}_2$  and  $\det(M_3) = -1$ . Any  $2 \times 2$  matrix of determinant  $\pm 1$  has Frobenius norm at least  $\sqrt{2}$ , by the same argument as in Example 4.2. Therefore,  $Y$  is of minimal norm in its orbit, and hence polystable by Theorem 2.2(d). It remains to show that the stabilizer of  $Y$  is finite. The discussion from Example 4.2 ensures that the stabilizer of  $Y$  is contained in  $\{(g, g) \mid g \in \mathrm{SO}_2\}$ . Given  $g \in \mathrm{SO}_2$ , the condition  $gY_3g^\top = Y_3$  implies  $gY_3 = Y_3g$ . This holds exactly for  $g = \pm I_2$ . Therefore, the stabilizer of  $Y$  is the finite set  $\{(I_2, I_2), (-I_2, -I_2)\}$ .  $\diamond$

**4.2. Boundedness of the likelihood via semistability.** We give new conditions that guarantee the boundedness of the likelihood in a matrix normal model. To do this, we use the equivalence of the boundedness of the likelihood with the semistability of a matrix tuple under left-right action, see Theorem 4.1(b). We consider matrix tuples in  $(\mathbb{R}^{m_1 \times m_2})^n$  where we may assume by duality that  $m_1 \geq m_2$ . The null cone of the complex left-right action of  $\mathrm{SL}_{m_1}(\mathbb{C}) \times \mathrm{SL}_{m_2}(\mathbb{C})$  on matrix tuples was described in [BD06, Theorem 2.1]. We prove the real analogue of this result and, with this, give a characterization of the matrix tuples with unbounded log-likelihood in Theorem 4.4. This has been derived in [DKH20, Theorems 3.1(i) and 3.3(i)] using a different method.

The dimension of the complex null cone is given in [BD06]. By translating this result to the real numbers, we derive a new upper bound on the maximum likelihood threshold  $\mathrm{mlt}_b$ , the minimum number of samples needed for the likelihood function to be generically bounded from above; see Corollary 4.9. This translates in invariant theory to finding the minimum sample size  $n$  such that the null cone does not fill its ambient space. In

addition, we recover lower and upper bounds from the literature in Corollaries 4.5, 4.10 and 4.11.

**Theorem 4.4.** *Consider  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$ , a tuple of  $n$  samples from a matrix normal model. The log-likelihood function  $\ell_Y$  is not bounded from above if and only if there exist subspaces  $V_1 \subseteq \mathbb{R}^{m_1}$  and  $V_2 \subseteq \mathbb{R}^{m_2}$  with  $m_1 \dim V_2 > m_2 \dim V_1$  such that  $Y_i V_2 \subseteq V_1$  for all  $i = 1, \dots, n$ .*

*Proof.* The log-likelihood function  $\ell_Y$  is bounded from above if and only if  $Y$  is not in the null cone, by Theorem 4.1(b). A description of the null cone for complex tuples in  $(\mathbb{C}^{m_1 \times m_2})^n$  under  $\mathrm{SL}_{m_1}(\mathbb{C}) \times \mathrm{SL}_{m_2}(\mathbb{C})$  is given in [BD06]. There, the authors show that the null cone consists of tuples for which there exist subspaces  $W_1 \subseteq \mathbb{C}^{m_1}$  and  $W_2 \subseteq \mathbb{C}^{m_2}$  with  $m_1 \dim_{\mathbb{C}} W_2 > m_2 \dim_{\mathbb{C}} W_1$  such that  $Y_i W_2 \subseteq W_1$  for all  $i = 1, \dots, n$ . This is the same condition as appears in the statement, but in the complex case. In the rest of the proof, we see that the condition also holds for real tuples under the left-right action of  $\mathrm{SL}_{m_1}(\mathbb{R}) \times \mathrm{SL}_{m_2}(\mathbb{R})$ . The following argument is thanks to Jan Draisma.

Given the existence of the real subspaces  $V_1$  and  $V_2$  in the statement, we show null cone membership of the matrix tuple  $Y$  using the same one-parameter subgroups as in the proof of [BD06, Theorem 2.1]. Conversely, if  $Y$  is in the null cone, we know that there exist complex subspaces  $W_1 \subseteq \mathbb{C}^{m_1}$  and  $W_2 \subseteq \mathbb{C}^{m_2}$  with  $m_1 \dim_{\mathbb{C}} W_2 > m_2 \dim_{\mathbb{C}} W_1$  and  $Y_i W_2 \subseteq W_1$  for all  $i = 1, \dots, n$ . We use the complex subspaces  $W_j$  to construct real subspaces. Let  $V_j$  be the intersection of  $W_j$  with  $\mathbb{R}^{m_j}$ , and let  $V'_j$  be the image of  $W_j$  under the map that sends a complex vector to its real part. Since  $iV_j$  is the kernel of that map, where  $i^2 = -1$ , we have

$$2 \dim_{\mathbb{C}} W_j = \dim_{\mathbb{R}} V_j + \dim_{\mathbb{R}} V'_j.$$

In particular, we either have  $m_1 \dim V_2 > m_2 \dim V_1$  or  $m_1 \dim V'_2 > m_2 \dim V'_1$ . Since both inclusions  $Y_i V_2 \subseteq V_1$  and  $Y_i V'_2 \subseteq V'_1$  hold for all  $i = 1, \dots, n$ , either  $(V_1, V_2)$  or  $(V'_1, V'_2)$  are real subspaces as in the statement.  $\square$

We now come to statistical implications of Theorem 4.4.

**Corollary 4.5.** *If  $n < \frac{m_1}{m_2}$ , then the log-likelihood function  $\ell_Y$  is unbounded from above for every tuple of samples  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$ . In particular,  $\mathrm{mlt}_b(m_1, m_2) \geq \lceil \frac{m_1}{m_2} \rceil$ .*

*Proof.* For any one-dimensional subspace  $V_2 \subseteq \mathbb{R}^{m_2}$ , the dimension of  $V_1 := \sum_{i=1}^n Y_i V_2$  is at most  $n$ . If  $n < \frac{m_1}{m_2}$ , Theorem 4.4 implies that the log-likelihood  $\ell_Y$  is unbounded.  $\square$

The result in this corollary also follows from [DKH20, Lemma 1.2]. We now characterize when the null cone fills the space of matrix tuples, which extends [BD06, Proposition 2.4] from the space of complex matrix tuples to real matrix tuples. For this, we begin by defining the cut-and-paste rank from [BD06, Definition 2.2] over the real numbers.

**Definition 4.6.** The *cut-and-paste rank*  $\mathrm{cp}^{(n)}(a, b, c, d)$  of a tuple of positive integers  $a, b, c, d$  and  $n$  is the maximum rank of the  $ab \times cd$  matrix  $\sum_{i=1}^n X_i \otimes Y_i$ , as  $X_i$  and  $Y_i$  range over real matrices of sizes  $c \times a$  and  $d \times b$  respectively.

*Remark 4.7.* Analogously to Definition 4.6 one can define  $\mathrm{cp}_{\mathbb{C}}^{(n)}(a, b, c, d)$  by letting the  $X_i$  and  $Y_i$  range over complex matrices, see [BD06, Definition 2.2]. The real and complex

ranks agree, as follows. The condition for the rank of the matrix  $\sum_{i=1}^n X_i \otimes Y_i$  to drop is given by minors. Thus,  $\text{cp}_{\mathbb{C}}^{(n)}(a, b, c, d)$  is witnessed on a Zariski-open subset of  $W := (\mathbb{C}^{c \times a})^n \times (\mathbb{C}^{d \times b})^n$  and hence witnessed by some element in  $(\mathbb{R}^{c \times a})^n \times (\mathbb{R}^{d \times b})^n$ , as the latter is Zariski-dense in  $W$ .

We use the cut-and-paste rank to give a necessary and sufficient condition for the null cone under left-right action to fill the space of matrix tuples  $(\mathbb{R}^{m_1 \times m_2})^n$ , i.e. for the log-likelihood to be always unbounded from above. As above, we take  $m_1 \geq m_2$ . Moreover, since we saw in Corollary 4.5 that the likelihood is unbounded for  $m_2 n < m_1$ , it suffices to restrict to the range  $m_2 \leq m_1 \leq n m_2$ .

**Theorem 4.8.** *Let  $0 < m_2 \leq m_1 \leq n m_2$ . The log-likelihood  $\ell_Y$  is unbounded from above for every tuple of samples  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  if and only if there exists  $k \in \{1, \dots, m_2\}$  such that  $l = \lceil \frac{m_1}{m_2} k \rceil - 1$  satisfies both*

$$\begin{aligned} m_1 - l &\leq n(m_2 - k) \quad \text{and} \\ \text{cp}^{(n)}(a, b, c, d) &= cd, \quad \text{where } (a, b, c, d) = (m_2 - k, k, m_1 - l, nk - l). \end{aligned}$$

*Proof.* Let  $\mathcal{N}_{\mathbb{K}}$  be the null cone under the left-right action of  $\text{SL}_{m_1}(\mathbb{K}) \times \text{SL}_{m_2}(\mathbb{K})$  on  $(\mathbb{K}^{m_1 \times m_2})^n$ , where  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ . We note that  $\mathcal{N}_{\mathbb{C}}$  is Zariski closed and that  $(\mathbb{R}^{m_1 \times m_2})^n$  is Zariski-dense in  $(\mathbb{C}^{m_1 \times m_2})^n$ . Thus,  $\mathcal{N}_{\mathbb{R}}$  fills the space  $(\mathbb{R}^{m_1 \times m_2})^n$  if and only if  $\mathcal{N}_{\mathbb{C}}$  fills the space  $(\mathbb{C}^{m_1 \times m_2})^n$ , by Proposition 2.3. It therefore suffices to characterize when  $\mathcal{N}_{\mathbb{C}} = (\mathbb{C}^{m_1 \times m_2})^n$ . For this, define for natural numbers  $k$  and  $l$

$$Q_{k,l} := \left\{ (Y_1, \dots, Y_n) \in (\mathbb{C}^{m_1 \times m_2})^n \mid \exists V \subseteq \mathbb{C}^{m_2} : \dim_{\mathbb{C}} V = k, \dim_{\mathbb{C}} \left( \sum_{i=1}^n Y_i V \right) \leq l \right\}.$$

The null cone  $\mathcal{N}_{\mathbb{C}}$  is the union of the  $Q_{k,l}$  over  $1 \leq k \leq m_2$  and  $0 \leq l < \frac{m_1}{m_2} k$ , by [BD06, Theorem 2.1], which is the complex analogue of Theorem 4.4. We observe that the algebraic sets  $Q_{k,l}$  get larger as  $l$  increases. Hence, it suffices to consider if any of the  $Q_{k,l}$  fills  $(\mathbb{C}^{m_1 \times m_2})^n$  as  $k$  ranges over  $1 \leq k \leq m_2$ , where the corresponding  $l$  is the largest integer strictly smaller than  $\frac{m_1}{m_2} k$ , i.e.  $l = \lceil \frac{m_1}{m_2} k \rceil - 1$ .

The assumption  $m_1 \leq n m_2$  yields  $l < n k$ . Therefore, [BD06, Proposition 2.4] shows that

$$\dim_{\mathbb{C}} Q_{k,l} = n m_1 m_2 - \left( (m_1 - l)(k n - l) - \text{cp}_{\mathbb{C}}^{(n)}(a, b, \tilde{c}, d) \right),$$

where  $a = m_2 - k$ ,  $b = k$ ,  $\tilde{c} = \min\{m_1 - l, n(m_2 - k)\}$  and  $d = k n - l$ . By Remark 4.7,  $\text{cp}_{\mathbb{C}}^{(n)}(a, b, \tilde{c}, d) = \text{cp}^{(n)}(a, b, \tilde{c}, d)$ . Thus,  $Q_{k,l}$  equals  $(\mathbb{C}^{m_1 \times m_2})^n$  if and only if

$$\text{cp}^{(n)}(a, b, \tilde{c}, d) = (m_1 - l)(k n - l).$$

Finally, the latter equation is equivalent to

$$m_1 - l \leq n(m_2 - k) \quad \text{and} \quad \text{cp}^{(n)}(a, b, \tilde{c}, d) = \tilde{c} d,$$

since  $\tilde{c} = \min\{m_1 - l, n(m_2 - k)\}$ ,  $d = k n - l \geq 1$  and  $\text{cp}^{(n)}(a, b, \tilde{c}, d) \leq \tilde{c} d$ .  $\square$

In principle, Theorem 4.8 solves the problem of determining the maximum likelihood threshold  $\text{mlt}_b$ , although in terms of the cut-and-paste rank. Hence, this gives statistical

motivation for better understanding the cut-and-paste rank, e.g. by obtaining a general closed formula.

We use the above theorem to give a new upper bound for  $\text{mlt}_b$ .

**Corollary 4.9.** *Let  $0 < m_2 \leq m_1$ . If*

$$(10) \quad n > \max_{1 \leq k \leq m_2} \left( \frac{l}{k} + \frac{m_2 - k}{m_1 - l} \right), \quad \text{where } l = \left\lceil \frac{m_1}{m_2} k \right\rceil - 1,$$

*the log-likelihood  $\ell_Y$  for a generic matrix tuple  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  is bounded from above.*

*In other words,  $\text{mlt}_b \leq \left\lceil \max_{1 \leq k \leq m_2} \left( \frac{l}{k} + \frac{m_2 - k}{m_1 - l} \right) \right\rceil + 1$ .*

*Proof.* First, we observe that (10) with  $k = m_2$  yields  $n > \frac{m_1 - 1}{m_2}$ . The latter is equivalent to  $nm_2 \geq m_1$ , so we are in the setting of Theorem 4.8. Using the notation in that theorem, we see that (10) is equivalent to every  $k \in \{1, \dots, m_2\}$  satisfying  $cd > ab$ . In particular, for every such  $k$  we have  $\text{cp}^{(n)}(a, b, c, d) \leq ab < cd$ , so by Theorem 4.8 the log-likelihood  $\ell_Y$  cannot be unbounded from above for every tuple  $Y$ .  $\square$

Two simpler upper bounds, which are known in the statistics literature [DKH20, Proposition 1.3, Theorem 1.4], are obtained as follows.

**Corollary 4.10.** *If  $n \geq \frac{m_1}{m_2} + \frac{m_2}{m_1}$ , then the log-likelihood  $\ell_Y$  for a generic matrix tuple  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  is bounded from above. In other words,  $\text{mlt}_b \leq \lceil \frac{m_1}{m_2} + \frac{m_2}{m_1} \rceil$ .*

*Proof.* For every  $k \in \{1, \dots, m_2\}$  we have  $l < \frac{m_1 k}{m_2}$ , which implies that

$$\frac{m_1}{m_2} + \frac{m_2}{m_1} > \frac{l}{k} + \frac{m_2 - k}{m_1 - l}.$$

Thus, the assertion follows from Corollary 4.9.  $\square$

**Corollary 4.11.** *Let  $m_2$  divide  $m_1$ . The log-likelihood  $\ell_Y$  for a generic matrix tuple  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  is bounded from above if and only if  $n \geq \frac{m_1}{m_2}$ . In other words,  $\text{mlt}_b = \frac{m_1}{m_2}$ .*

*Proof.* If  $n < \frac{m_1}{m_2}$ , the log-likelihood is always unbounded from above by Corollary 4.5. So we write  $m_1 = \gamma m_2$  and assume  $n \geq \gamma$ . For every  $k \in \{1, \dots, m_2\}$ , using the notation from Theorem 4.8, we see that  $l = \gamma k - 1$  and  $a < c$ . If  $n > \gamma$ , we also have that  $b < d$ , so  $\text{cp}^{(n)}(a, b, c, d) \leq ab < cd$ . If  $n = \gamma$ , then  $m_1 - l > n(m_2 - k)$ . In either case, one of the two conditions in Theorem 4.8 is not satisfied, so  $\ell_Y$  is generically bounded from above.  $\square$

In Table 1 we list the maximum likelihood threshold  $\text{mlt}_b$  for boundedness of the log-likelihood for small values of  $m_1, m_2$ , and compare with the bounds discussed above. We observe that there are cases where our upper bound

$$\alpha = \left\lceil \max_{1 \leq k \leq m_2} \left( \frac{l}{k} + \frac{m_2 - k}{m_1 - l} \right) \right\rceil + 1, \quad \text{where } l = \left\lceil \frac{m_1}{m_2} k \right\rceil - 1,$$

is strictly better than the simple upper bound  $U = \lceil \frac{m_1}{m_2} + \frac{m_2}{m_1} \rceil$ , e.g. when  $(m_1, m_2) = (3, 2)$ . In most cases our bound  $\alpha$  matches the lower bound  $L = \lceil \frac{m_1}{m_2} \rceil$ , so that we can determine  $\text{mlt}_b$ . In addition, when  $m_2 | m_1$ , one can use Corollary 4.11 to determine  $\text{mlt}_b$ .

TABLE 1. Bounds for the maximum likelihood threshold  $\text{mlt}_b$ .  $L = \lceil \frac{m_1}{m_2} \rceil$  is the lower-bound from Corollary 4.5,  $U = \lceil \frac{m_1}{m_2} + \frac{m_2}{m_1} \rceil$  is the upper bound from Corollary 4.10, and  $\alpha$  is our new upper bound from Corollary 4.9.

$m_1$	$m_2$	$L$	$\text{mlt}_b$	$\alpha$	$U$
2	2	1	1	1	2
3	2	2	2	2	3
3	3	1	1	2	2
4	2	2	2	2	3
4	3	2	2	2	3
4	4	1	1	2	2
5	2	3	3	3	3
5	3	2	3	3	3
5	4	2	2	2	3
5	5	1	1	2	2
6	2	3	3	3	4
6	3	2	2	2	3
6	4	2	2	2	3
6	5	2	2	2	3
6	6	1	1	2	2

$m_1$	$m_2$	$L$	$\text{mlt}_b$	$\alpha$	$U$
7	2	4	4	4	4
7	3	3	3	3	3
7	4	2	3	3	3
7	5	2	3	3	3
7	6	2	2	2	3
7	7	1	1	2	2
8	2	4	4	4	5
8	3	3	3	3	4
8	4	2	2	3	3
8	5	2	3	3	3
8	6	2	2	2	3
8	7	2	2	2	3
8	8	1	1	2	2
9	2	5	5	5	5
9	3	3	3	3	4

$m_1$	$m_2$	$L$	$\text{mlt}_b$	$\alpha$	$U$
9	4	3	3	3	3
9	5	2	3	3	3
9	6	2	2	2	3
9	7	2	3	3	3
9	8	2	2	2	3
9	9	1	1	2	2
10	2	5	5	5	6
10	3	4	4	4	4
10	4	3	3	3	3
10	5	2	2	3	3
10	6	2	3	3	3
10	7	2	3	3	3
10	8	2	2	2	3
10	9	2	2	2	3
10	10	1	1	2	2

even if the bounds  $L$  and  $\alpha$  do not coincide, such as in  $(m_1, m_2) = (8, 4)$  or in the square cases  $m_1 = m_2$ . The rest of the values of  $\text{mlt}_b$  can be filled from [DKH20, Table 1]. We highlight the case  $(m_1, m_2) = (8, 3)$ : the maximum likelihood threshold  $\text{mlt}_b = 3$  was computed in [DKH20] via Gröbner bases, but it is not covered by the general bounds in [DKH20]. Nevertheless, our bound  $\alpha$  determines this case.

**4.3. Uniqueness of the MLE via stability.** We compare conditions for stability with conditions for the uniqueness of the MLE. We saw in Example 4.2 that stability of a real matrix tuple  $Y$  under left-right action of  $\text{SL}_{m_1}(\mathbb{R}) \times \text{SL}_{m_2}(\mathbb{R})$  is not equivalent to uniqueness of the MLE given  $Y$ . However, such an equivalence holds for complex Gaussian models, by Theorem 3.15. Matrix normal models over the complex numbers are induced by the left-right action of  $\text{SL}_{m_1}(\mathbb{C}) \times \text{SL}_{m_2}(\mathbb{C})$  on  $(\mathbb{C}^{m_1 \times m_2})^n$ . Hence we obtain conditions for the uniqueness of the MLE given  $Y \in (\mathbb{C}^{m_1 \times m_2})^n$  from characterizing the stability of  $Y$  under the left-right action by  $\text{SL}_{m_1}(\mathbb{C}) \times \text{SL}_{m_2}(\mathbb{C})$ . Characterizing this stability is a special case of the setting studied in [Kin94]. From this, we obtain the following theorem, which we prove in Appendix A.

**Theorem 4.12.** *Consider the left-right action of  $\text{SL}_{m_1}(\mathbb{C}) \times \text{SL}_{m_2}(\mathbb{C})$  on  $(\mathbb{C}^{m_1 \times m_2})^n$ , and a tuple  $Y \in (\mathbb{C}^{m_1 \times m_2})^n$  of  $n$  samples from a complex matrix normal model. The following are equivalent:*

- (a) *the complex MLE given  $Y$  exists uniquely;*
- (b) *the matrix tuple  $Y$  is stable;*
- (c) *the matrix  $(Y_1 | \dots | Y_n) \in \mathbb{C}^{m_1 \times nm_2}$  has rank  $m_1$ , and  $m_2 \dim V_1 > m_1 \dim V_2$  holds for all subspaces  $V_1 \subseteq \mathbb{C}^{m_1}$ ,  $\{0\} \subsetneq V_2 \subsetneq \mathbb{C}^{m_2}$  that satisfy  $Y_i V_2 \subseteq V_1$  for all  $i = 1, \dots, n$ .*

We note the similarity with the conditions that characterize semistability in Theorem 4.4. However, while Theorem 4.4 holds both over  $\mathbb{R}$  and  $\mathbb{C}$ , the same cannot be true for Theorem 4.12. In fact, the real analog of Theorem 4.12(c) is shown to characterize uniqueness of the MLE in [DKH20, Theorems 3.1(ii) and 3.3(ii)], which is not equivalent to stability by Example 4.2.

**4.4. The moment map.** In this section we recall the condition for the moment map for the action of  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  to vanish at a matrix tuple. By Kempf Ness, Theorem 2.2(a), this gives the condition to be at a point of minimal norm in the orbit.

The tangent space of  $\mathrm{SL}_{m_i}$  at the identity matrix consists of all matrices with trace zero. The moment map at  $Y \in (\mathbb{R}^{m_1 \times m_2})^n$  is the differential of  $(g_1, g_2) \mapsto \|(g_1, g_2) \cdot Y\|^2$  at the pair  $(I_{m_1}, I_{m_2})$  of identity matrices, i.e.

$$(11) \quad \begin{aligned} \mu(Y) : T_{I_{m_1}} \mathrm{SL}_{m_1} \times T_{I_{m_2}} \mathrm{SL}_{m_2} &\longrightarrow \mathbb{R} \\ (\dot{g}_1, \dot{g}_2) &\longmapsto 2 \sum_{i=1}^n \mathrm{tr}((\dot{g}_1 Y_i + Y_i \dot{g}_2^\top) Y_i^\top). \end{aligned}$$

**Theorem 4.13** (Kempf-Ness theorem for  $\mathrm{SL} \times \mathrm{SL}$  action). *Consider the left-right action of  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$  on the space of matrix tuples  $(\mathbb{R}^{m_1 \times m_2})^n$ . A matrix tuple is semistable (resp. polystable) if and only if there is a non-zero matrix tuple  $Y$  in its orbit closure (resp. orbit) where the moment map  $\mu$  vanishes, i.e.*

$$\exists c_1, c_2 > 0 : \sum_{i=1}^n Y_i Y_i^\top = c_1 I_{m_1} \text{ and } \sum_{i=1}^n Y_i^\top Y_i = c_2 I_{m_2}.$$

*Proof.* This follows from rewriting (11) as

$$(\dot{g}_1, \dot{g}_2) \longmapsto 2 \mathrm{tr} \left( \dot{g}_1 \sum_{i=1}^n Y_i Y_i^\top \right) + 2 \mathrm{tr} \left( \dot{g}_2^\top \sum_{i=1}^n Y_i^\top Y_i \right). \quad \square$$

**4.5. Scaling algorithms for the MLE.** In this section, we describe algorithmic consequences of the connection between invariant theory and maximum likelihood estimation. We present an algorithm for ML estimation that is well-known in statistics, and connect it to an algorithm in invariant theory; see the left hand side of Figure 1. The connection allows us to give a complexity analysis of the statistics algorithm. The algorithm in statistics is the flip-flop algorithm, which involves the group  $\mathrm{GL}_{m_1} \times \mathrm{GL}_{m_2}$ , while the invariant theory algorithm is operator scaling for the left-right action of  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$ . We begin by recalling these algorithms.

**4.5.1. Operator scaling and the flip-flop algorithm.** Operator scaling, see the top left in Figure 1, solves the norm minimization problem for the left-right action of  $\mathrm{SL}_{m_1}(\mathbb{C}) \times \mathrm{SL}_{m_2}(\mathbb{C})$  on the space of matrix tuples  $(\mathbb{C}^{m_1 \times m_2})^n$ . From an invariant theory perspective, operator scaling was first studied in [Gur04], and [GGOW16] showed that it yields a polynomial time algorithm for null cone membership. The method was generalized to tuples of tensors in [BGO<sup>+</sup>17, Algorithm 1].

The *flip-flop algorithm* [Dut99, LZ05], see the bottom left of Figure 1, is an alternating maximization procedure to find an MLE in a matrix normal model. It can be thought of as a Gaussian version of IPS for matrix normal models, since one alternately updates

the estimates in each marginal. If we consider  $\Psi_2$  to be fixed, the log-likelihood in (7) becomes, up to constants,

$$\ell_Y(\Psi_1) = m_2 \left[ \log \det(\Psi_1) - \text{tr} \left( \Psi_1 \cdot \frac{1}{nm_2} \sum_{i=1}^n Y_i \Psi_2 Y_i^\top \right) \right].$$

Maximizing the log-likelihood with respect to  $\Psi_1$  reduces to the case of a standard multivariate Gaussian model as in (2). The unique maximizer over the positive definite cone is the inverse, if it exists, of the matrix  $\frac{1}{nm_2} \sum_{i=1}^n Y_i \Psi_2 Y_i^\top$ . In the same way, we can fix  $\Psi_1$  and maximize the log-likelihood with respect to  $\Psi_2$ . Iterating these two steps gives the algorithm.

---

**Algorithm 4.1** Flip-flop

---

**Input:**  $Y_1, \dots, Y_n \in \mathbb{R}^{m_1 \times m_2}$ ,  $N \in \mathbb{Z}_{>0}$ .

**Output:** an approximation of an MLE, if it exists.

- 1: Initialize  $\Psi_2 := I_{m_2}$ .
- 2: **for**  $k = 1$  **to**  $N$  **do**
- 3:   the following pair of updates

$$(12) \quad \begin{aligned} \Psi_1 &:= \left( \frac{1}{nm_2} \sum_{i=1}^n Y_i \Psi_2 Y_i^\top \right)^{-1} \\ \Psi_2 &:= \left( \frac{1}{nm_1} \sum_{i=1}^n Y_i^\top \Psi_1 Y_i \right)^{-1}. \end{aligned}$$

- 4: **end for**
  - 5: **return**  $\Psi_1 \otimes \Psi_2$ .
- 

We now compare operator scaling with the flip-flop algorithm. The scaling algorithm in [BGO<sup>+</sup>17, Algorithm 1] gives, when specializing from tensors to matrices, the same procedure as Algorithm 4.1, up to scaling with different constants in the update steps (12). In [BGO<sup>+</sup>17, Algorithm 1], the matrices  $\Psi_1$  and  $\Psi_2$  in (12) are restricted to have determinant one, in order to stay in the  $\text{SL}_{m_1} \times \text{SL}_{m_2}$  orbit of  $Y$ . In comparison, Algorithm 4.1 has constants chosen to minimize the outer infimum in (9).

Although the algorithm in [BGO<sup>+</sup>17] is defined over the complex numbers, when restricting to real inputs operator scaling only involves computations over the reals. This allows the computation of MLEs (if they exist) in the real matrix normal model via (9), since the capacity of a real matrix tuple is the same under the action of  $\text{SL}_{m_1}(\mathbb{R}) \times \text{SL}_{m_2}(\mathbb{R})$  as under the action of  $\text{SL}_{m_1}(\mathbb{C}) \times \text{SL}_{m_2}(\mathbb{C})$ , see Proposition 2.3.

**4.5.2. Convergence.** In [BGO<sup>+</sup>17], the authors give conditions for being in the null cone, based on the convergence of their Algorithm 1. Specializing to a matrix tuple, to connect to the flip-flop algorithm, their results combine with ours to show the following. If an update step cannot be computed because one of the matrices in (12) cannot be inverted, then the matrix tuple  $Y$  is unstable under the action of  $\text{SL}_{m_1}(\mathbb{C}) \times \text{SL}_{m_2}(\mathbb{C})$ , and therefore also under the real action of  $\text{SL}_{m_1}(\mathbb{R}) \times \text{SL}_{m_2}(\mathbb{R})$ , by Proposition 2.3. This implies that the log-likelihood  $\ell_Y$  is unbounded, by Theorem 4.1(a). Otherwise, the sequence of terms

$(\Psi_1^{1/2}, \Psi_2^{1/2}) \cdot Y$  converges, possibly to infinity. We now consider the possible cases that can arise in this limit, by comparing to operator scaling, using the fact that the constants in the flip-flop algorithm minimize the outer infimum in (9).

If the sequence  $(\Psi_1^{1/2}, \Psi_2^{1/2}) \cdot Y$  converges to zero or infinity, then the log-likelihood  $\ell_Y$  is unbounded. Otherwise, the sequence converges to a matrix tuple of positive norm in the orbit closure, where the moment map (11) vanishes, and  $Y$  is semistable. Here, two further possibilities can arise. The first possibility occurs when the matrix tuple  $Y$  is polystable. Then the minimal norm is attained at an element of the group  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$ , and the flip-flop algorithm converges to an MLE; see (9). The second possibility occurs when  $Y$  is semistable but not polystable. Then, the flip-flop algorithm diverges by the following remark.

*Remark 4.14.* If the matrix tuple  $Y$  is semistable but not polystable under the left-right action of  $\mathrm{SL}_{m_1} \times \mathrm{SL}_{m_2}$ , then the likelihood  $L_Y$  (equivalently the log-likelihood  $\ell_Y$ ) is bounded from above, but does not attain its supremum. In this case, any sequence  $\Psi_N := (\Psi_{1,N} \otimes \Psi_{2,N})$  of concentration matrices with

$$\lim_{N \rightarrow \infty} L_Y(\Psi_{1,N} \otimes \Psi_{2,N}) = \sup L_Y > 0$$

diverges. Indeed, otherwise the limit  $\Psi_\infty$  would be rank-deficient, as the matrix normal model is closed in  $\mathrm{PD}_{m_1 m_2}$ . Then  $\det(\Psi_\infty) = 0$  yields the contradiction  $\sup L_Y = L_Y(\Psi_\infty) = 0$ .

**4.5.3. Complexity.** We use known results to derive a complexity analysis for the flip-flop algorithm. In [BGO<sup>+</sup>17], the authors prove convergence of their Algorithm 1, which solves the null cone membership problem up to an approximation parameter  $\varepsilon > 0$ . For tuples of tensors, choosing  $\varepsilon$  exponentially small in the dimension of the tensor space yields a deterministic test for null cone membership with exponential running time, see [BGO<sup>+</sup>17, Theorem 3.8]. When specializing to tuples of matrices, i.e. to operator scaling, it suffices to choose  $\varepsilon$  polynomially small. Thus for operator scaling, [BGO<sup>+</sup>17, Algorithm 1] recovers the polynomial time algorithm for the null cone membership problem from [GGOW16]. We adapt [BGO<sup>+</sup>17, Theorem 1.1] to our notation to derive the following.

**Theorem 4.15.** *Given  $\varepsilon > 0$  and a matrix tuple  $Y \in (\mathbb{Z}^{m_1 \times m_2})^n$  with matrix entries of bit size bounded by  $b$ , after a number of steps that is polynomial in  $(nm_1 m_2, b, 1/\varepsilon)$ , the flip-flop algorithm either identifies that the log-likelihood  $\ell_Y$  is unbounded or finds  $(\Psi_1, \Psi_2) \in \mathrm{PD}_{m_1} \times \mathrm{PD}_{m_2}$  such that the matrix tuple  $(\Psi_1^{1/2}, \Psi_2^{1/2}) \cdot Y$  is  $\varepsilon$ -close to a matrix tuple where the moment map (11) vanishes.*

In the case where the log-likelihood  $\ell_Y$  is bounded, taking the limit  $\varepsilon \rightarrow 0$  in Theorem 4.15 gives rise to two possibilities. Either the MLE exists and is the limit of the  $\Psi_1 \otimes \Psi_2$  as  $\varepsilon \rightarrow 0$ , or the sequence  $\Psi_1 \otimes \Psi_2$  diverges as  $\varepsilon \rightarrow 0$ , by Remark 4.14. Because of this divergence, there is no meaningful notion of approximate MLE in the latter scenario.

**4.5.4. Outlook.** We briefly comment on extensions of the above to general groups, see the right hand side of Figure 1. In its full generality, the algorithm in [BGO<sup>+</sup>17] is



an alternating minimization procedure to find the capacity of a tuple of  $d$ -dimensional tensors of format  $m_1 \times \dots \times m_d$  under the action of  $\mathrm{SL}_{m_1} \times \dots \times \mathrm{SL}_{m_d}$ . It can therefore be used for ML estimation in (real and complex) tensor normal models. More generally, the algorithms in [BFG<sup>+</sup>19] can be used for geodesically convex algorithms for maximum likelihood estimation in complex Gaussian group models as in Theorem 3.15. Many scaling algorithms are designed to optimize over the complex orbit, but often each update is defined over  $\mathbb{R}$  if the input is real, and hence they can also be used for real Gaussian group models.

## 5. TRANSITIVE DAGS

In this section we study graphical models that fit into the Gaussian group model framework. We study MLE existence via a corresponding null cone problem. We focus on directed graphs, although our results also cover undirected graphical models, as explained in Remark 5.9.

Let  $\mathcal{G}$  be a directed acyclic graph (DAG) with  $m$  nodes. We denote an edge from  $j$  to  $i$  by  $j \rightarrow i$ ; otherwise, if there is no such edge, we write  $j \not\rightarrow i$ . We note that edges  $i \rightarrow i$  do not appear in a DAG, because they give cycles of length one. Consider the statistical model represented by the linear structural equation

$$Y = \Lambda Y + \varepsilon,$$

where  $Y \in \mathbb{R}^m$ , the matrix  $\Lambda \in \mathbb{R}^{m \times m}$  satisfies  $\Lambda_{ij} = 0$  for  $j \not\rightarrow i$  in  $\mathcal{G}$ , and  $\varepsilon \sim N(0, \Omega)$  with  $\Omega \in \mathbb{R}^{m \times m}$  diagonal and positive definite. The model expresses each coordinate  $Y_i$  as a linear combination of all  $Y_j$  such that  $j \rightarrow i$ , up to Gaussian error. Solving for  $Y$ , we have

$$Y = (I - \Lambda)^{-1} \varepsilon,$$

where the acyclicity of  $\mathcal{G}$  implies that  $(I - \Lambda)$  is invertible. We see that  $Y$  is Gaussian with covariance matrix and concentration matrix

$$(13) \quad \Sigma = (I - \Lambda)^{-1} \Omega (I - \Lambda)^{-\top}, \quad \Psi = (I - \Lambda)^{\top} \Omega^{-1} (I - \Lambda).$$

The Gaussian graphical model  $\mathcal{M}_{\mathcal{G}}^{\rightarrow}$  consists of the set of concentration matrices  $\Psi$  of the form in (13), for  $\Lambda$  and  $\Omega$  defined in terms of  $\mathcal{G}$  as above.

We now put these models in the context of Gaussian group models. Given a DAG  $\mathcal{G}$ , we define the set of matrices

$$(14) \quad G(\mathcal{G}) = \{g \in \mathrm{GL}_m \mid g_{ij} = 0 \text{ for } i \neq j \text{ with } j \not\rightarrow i \text{ in } \mathcal{G}\}.$$

We have a *transitive* DAG (TDAG)  $\mathcal{G}$  if  $k \rightarrow j$  and  $j \rightarrow i$  in  $\mathcal{G}$  imply  $k \rightarrow i$  in  $\mathcal{G}$ .

**Proposition 5.1.** *The set of matrices  $G(\mathcal{G})$  is a group if and only if  $\mathcal{G}$  is a TDAG. In this case, the Gaussian graphical model given by  $\mathcal{G}$  is the Gaussian group model given by  $G(\mathcal{G})$ :*

$$\mathcal{M}_{\mathcal{G}}^{\rightarrow} = \mathcal{M}_{G(\mathcal{G})}.$$

*Proof.* If  $\mathcal{G}$  is not a TDAG, then there exist pairwise distinct indices  $i, j, k$  such that  $j \rightarrow i$  and  $k \rightarrow j$  but  $k \not\rightarrow i$ . Take the elementary matrices  $g = E_{ij}$  (with ones on the diagonal and at the  $(i, j)$  entry, and zero elsewhere) and  $h = E_{jk}$ . We see that  $g, h \in G(\mathcal{G})$ , but  $gh \notin G(\mathcal{G})$  since  $(gh)_{ik} = 1$ , hence  $G(\mathcal{G})$  is not a group.

Conversely, we assume that  $\mathcal{G}$  is a TDAG. Any invertible diagonal matrix, in particular the identity  $I$ , is in  $G(\mathcal{G})$ . Suppose  $g, h \in G(\mathcal{G})$  and that  $(gh)_{ik} \neq 0$  for  $i \neq k$ . This means that there must exist some index  $j$  such that  $g_{ij}h_{jk} \neq 0$ . In particular,  $g_{ij} \neq 0$  and  $h_{jk} \neq 0$ , so that we have either  $j \rightarrow i$  or  $j = i$ , and either  $k \rightarrow j$  or  $k = j$ . In all of these cases, we have  $k \rightarrow i$ , since  $\mathcal{G}$  is a TDAG. Therefore  $gh \in G(\mathcal{G})$ , as required for  $G(\mathcal{G})$  to be a group. Now if  $g \in G(\mathcal{G})$  we show that  $g^{-1} \in G(\mathcal{G})$ . We can write  $g = D(I - N)$ , where  $D$  is diagonal with same diagonal entries as  $g$  and  $N$  is nilpotent with same zero pattern (outside of the diagonal) as  $g$ . In fact, since the TDAG  $\mathcal{G}$  does not contain any path of length  $m$ , we have  $N^m = 0$ . Then

$$g^{-1} = (I + N + N^2 + \dots + N^{m-1})D^{-1} \in G(\mathcal{G}),$$

since  $\text{supp}(N^j) \subseteq \text{supp}(N)$  for  $j \geq 1$ , as  $\mathcal{G}$  is a TDAG. We have shown that  $G(\mathcal{G})$  is a group. The equality of models follows from reparametrizing  $(I - \Lambda)^\top \Omega^{-1} (I - \Lambda)$  by  $g^\top g$ , where  $g = \Omega^{-\frac{1}{2}}(I - \Lambda) \in G(\mathcal{G})$ .  $\square$

**Example 5.2.** Let  $\mathcal{G}$  be the TDAG  $1 \leftarrow 3 \rightarrow 2$ . The corresponding group  $G(\mathcal{G}) \subseteq \text{GL}_3$  consists of invertible matrices  $g$  of the form

$$g = \begin{bmatrix} * & 0 & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix}.$$

By Proposition 5.1, we have that the Gaussian graphical model  $\mathcal{M}_{\mathcal{G}}^{\rightarrow}$  is a 5-dimensional linear slice of the cone of symmetric positive definite  $3 \times 3$  matrices:

$$\mathcal{M}_{\mathcal{G}}^{\rightarrow} = \{g^\top g \mid g \in G(\mathcal{G})\} = \{\Psi \in \text{PD}_3 \mid \psi_{12} = \psi_{21} = 0\}. \quad \diamond$$

The group  $G(\mathcal{G})$  associated to a TDAG  $\mathcal{G}$  is Zariski closed and closed under non-zero scalar multiples, but not self-adjoint. Hence we are not in the setting of Theorem 3.10. However, we can apply Theorem 3.6 to derive our main result of this section. Since the group  $G(\mathcal{G})$  contains orthogonal matrices of determinant  $-1$  (e.g. the diagonal matrix whose first entry is  $-1$  and all other entries are 1), Theorem 3.6 holds for  $G(\mathcal{G})_{\text{SL}}^+$  by Remark 3.7.

We characterize boundedness of the likelihood and MLE existence, in terms of the stability of a tuple of samples. When the MLE exists generically (i.e., when the number of samples is at least the maximum likelihood threshold), it is known to be generically unique [Lau96, Section 5.4.1]. We show that the log-likelihood given  $Y$  is bounded from above if and only if the MLE given  $Y$  exists, by ruling out the possibility that a tuple can be semistable but not polystable. We provide an exact condition for the MLE given  $Y$  to exist, based on linear dependence of the rows of  $Y$ . A *parent* of a node  $i$  is a node  $j$  with edge  $j \rightarrow i$  in  $\mathcal{G}$ .

**Theorem 5.3.** *Consider a TDAG  $\mathcal{G}$  and a tuple of  $n$  samples  $Y \in \mathbb{R}^{m \times n}$ . If some row of  $Y$ , corresponding to node  $i$ , is a linear combination of the rows corresponding to the parents of  $i$ , then  $Y$  is unstable under the action by  $G(\mathcal{G})_{\text{SL}}^+$ , and the likelihood is unbounded from above. Otherwise,  $Y$  is polystable and the MLE exists.*

*Remark 5.4.* If  $Y$  has a row of zeros, it is unstable and the likelihood is unbounded from above. This satisfies the criterion in the above theorem, because a row of zeros at row  $i$  is

interpreted as a trivial linear combination, independently of whether node  $i$  has parents in  $\mathcal{G}$ .

*Proof of Theorem 5.3.* Without loss of generality, we label the nodes of  $\mathcal{G}$  such that  $j \rightarrow i$  implies  $j < i$ . Suppose the  $i$ th node of  $\mathcal{G}$  has parents the first  $s$  nodes, and that the  $i$ th row of  $Y$  is a linear combination of the first  $s$  rows,

$$r_i = \lambda_1 r_1 + \cdots + \lambda_s r_s.$$

We show that  $Y$  is unstable under  $G(\mathcal{G})_{\text{SL}}^+$ . Let  $\varepsilon > 0$  and consider the matrix  $g_\varepsilon$ , which is equal to  $\varepsilon I$  except for the  $i$ th row, which equals

$$(g_\varepsilon)_{ik} = \begin{cases} -\varepsilon^{-(m-1)} \lambda_k & k = 1, \dots, s \\ \varepsilon^{-(m-1)} & k = i \\ 0 & \text{otherwise.} \end{cases}$$

We have that  $g_\varepsilon \in G(\mathcal{G})_{\text{SL}}^+$ , since  $\det(g_\varepsilon) = 1$  and there are non-zero off-diagonal entries only when  $j \rightarrow i$ . Moreover, the  $i$ th row of  $g_\varepsilon Y$  is the zero vector. Letting  $\varepsilon \rightarrow 0$  we have that  $g_\varepsilon Y \rightarrow 0$ , so we conclude that  $Y$  is unstable. The log-likelihood is unbounded from above, by Theorem 3.6.

For the second claim, let  $Y$  be such that no row is a linear combination of the rows corresponding to its parents. We show by induction on  $m$  that  $Y$  is polystable. This implies that the MLE given  $Y$  exists, by Theorem 3.6. If  $m = 1$ , then  $G(\mathcal{G})_{\text{SL}}^+ = \{1\}$  and  $Y$  is a single non-zero row, so the statement holds. Now for the induction step,  $m > 1$ , we assume the claim holds for TDAGs with  $m - 1$  nodes.

We prove that the orbit  $G(\mathcal{G})_{\text{SL}}^+ \cdot Y$  is closed and hence  $Y$  is polystable. For this, let  $Y_0$  be an element of the orbit closure of  $Y$ . Then there exists  $g_\varepsilon \in G(\mathcal{G})_{\text{SL}}^+$  with  $g_\varepsilon Y \rightarrow Y_0$  as  $\varepsilon \rightarrow 0$ . We may assume without loss of generality that  $\alpha_\varepsilon := (g_\varepsilon)_{mm} > 0$ , by using an appropriate subsequence of the sequence  $(g_\varepsilon)$  and multiplying the last row and another row of both  $g_\varepsilon$  and  $Y_0$  by  $-1$  if needed. Let  $g'_\varepsilon$  be obtained from  $g_\varepsilon$  by dropping the last row and column and multiplying by  $\alpha_\varepsilon^{1/m-1}$ . Then  $g'_\varepsilon \in G(\mathcal{G}')_{\text{SL}}^+$ , where the TDAG  $\mathcal{G}'$  is obtained from  $\mathcal{G}$  by removing the last node (and all edges pointing to it). Similarly, let  $Y'$  and  $Y'_0$  be obtained from  $Y$  and  $Y_0$ , respectively, by dropping the last row. Since  $g_\varepsilon Y \rightarrow Y_0$ , we have that

$$(15) \quad \alpha_\varepsilon^{-1/m-1} g'_\varepsilon Y' \rightarrow Y'_0 \text{ as } \varepsilon \rightarrow 0.$$

Since no row of  $Y$  is a linear combination of the rows corresponding to its parents, the same is true of  $Y'$ , and we apply the induction hypothesis to see that  $Y'$  is polystable. We will use this to construct a group element that sends  $Y$  to  $Y_0$ .

Without loss of generality, assume  $m - s, \dots, m - 1$  are the parents of the last node  $m$ . Then the last row of  $g_\varepsilon$  is  $[0, \dots, 0, \beta_{s\varepsilon}, \dots, \beta_{1\varepsilon}, \alpha_\varepsilon]$  and therefore the last row of  $g_\varepsilon Y$  is

$$\beta_{s\varepsilon} r_{m-s} + \cdots + \beta_{1\varepsilon} r_{m-1} + \alpha_\varepsilon r_m.$$

Now, let  $t \leq s$  be the dimension of the vector space spanned by  $r_{m-s}, \dots, r_{m-1}$  and assume, without loss of generality, that the rows  $r_{m-t}, \dots, r_{m-1}$  are linearly independent. Then we can rewrite the last row of  $g_\varepsilon Y$  as

$$(16) \quad \gamma_{t\varepsilon} r_{m-t} + \cdots + \gamma_{1\varepsilon} r_{m-1} + \alpha_\varepsilon r_m$$

for some  $\gamma_{i\varepsilon} \in \mathbb{R}$ . Since  $r_m$  is not a linear combination of its parents, the rows  $r_{m-t}, \dots, r_m$  are linearly independent, i.e. the matrix  $M \in \mathbb{R}^{(t+1) \times n}$  formed by these rows has rank  $t+1$ . Thus, any standard basis vector in  $\mathbb{R}^{t+1}$  can be expressed as a linear combination of the columns of  $M$ . Applying these linear combinations to (16), which is the last row of  $g_\varepsilon Y$  and converges to the last row of  $Y_0$ , we conclude convergence of each  $\gamma_{i\varepsilon}$  ( $1 \leq i \leq t$ ) and of  $\alpha_\varepsilon$  as  $\varepsilon \rightarrow 0$ . We denote the corresponding limits by  $\gamma_{i0} \in \mathbb{R}$  and  $\alpha_0 \geq 0$  respectively.

If  $\alpha_0 = 0$ , we get from (15) that  $g'_\varepsilon Y' = \alpha_\varepsilon^{1/m-1} (\alpha_\varepsilon^{-1/m-1} g'_\varepsilon Y') \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . So  $Y'$  is unstable, in particular not polystable, which contradicts the induction hypothesis.

Therefore,  $\alpha_0 > 0$  and we have  $g'_\varepsilon Y' = \alpha_\varepsilon^{1/m-1} (\alpha_\varepsilon^{-1/m-1} g'_\varepsilon Y') \rightarrow \alpha_0^{1/m-1} Y'_0$  as  $\varepsilon \rightarrow 0$ . Applying the induction hypothesis to  $Y'$ , we obtain that  $\alpha_0^{1/m-1} Y'_0$  lies in the orbit of  $Y'$  under the action by  $G(\mathcal{G}')_{\text{SL}}^+$ . This means there exists  $h' \in G(\mathcal{G}')_{\text{SL}}^+$  such that  $\alpha_0^{1/m-1} Y'_0 = h' Y'$  and therefore

$$h := \begin{bmatrix} \alpha_0^{-1/m-1} h' & 0 \\ 0 \cdots 0 \gamma_{t0} \cdots \gamma_{10} & \alpha_0 \end{bmatrix} \in G(\mathcal{G})_{\text{SL}}^+$$

satisfies  $hY = Y_0$  as desired.  $\square$

Our approach characterizes MLE existence for any tuple  $Y$ , not just generic existence. We derive an immediate corollary for generic tuples, regarding the maximum likelihood thresholds  $\text{mlt}$  and  $\text{mlt}_b$  defined in Section 2.1. This is known for general DAGs in the graphical models literature, see [Lau96, Section 5.4.1] and [DFKP19, Theorem 1]. The *in-degree* of a DAG  $\mathcal{G}$  is the maximum number of parents of any node in  $\mathcal{G}$ .

**Corollary 5.5.** *For the model  $\mathcal{M}_{\mathcal{G}}^{\rightarrow}$  of a TDAG  $\mathcal{G}$ , we have*

$$\text{mlt}_b(\mathcal{G}) = \text{mlt}(\mathcal{G}) = \text{in-degree}(\mathcal{G}) + 1.$$

*Proof.* The equivalence of the two maximum likelihood thresholds follows from Theorem 5.3, where we also see that for the MLE to exist generically we need that every row in a generic matrix of samples  $Y \in \mathbb{R}^{m \times n}$  is not a linear combination of its parent rows. Generic linear independence is guaranteed if and only if the number of columns  $n$  is at least the number of rows involved in a node plus its parents.  $\square$

**Example 5.6.** Let  $\mathcal{G}$  be the TDAG  $1 \leftarrow 3 \rightarrow 2$  from Example 5.2. We apply Theorem 5.3 to show when the MLE given a sample matrix  $Y \in \mathbb{R}^{3 \times n}$  exists. Node 3 has no parents, while nodes 1 and 2 both have the node 3 as their parent. Hence the log-likelihood  $\ell_Y$  is unbounded from above if the first or second row is a scalar multiple of the third row, or if the third row is zero, and otherwise the MLE given  $Y$  exists.

When  $n = 1$ , the first and second rows are always scalar multiples of the third row, hence the null cone fills the space, and the log-likelihood is always unbounded from above. With  $n = 2$  samples, the null cone has two components, with vanishing ideal

$$\langle y_{11}y_{32} - y_{12}y_{31} \rangle \cap \langle y_{21}y_{32} - y_{22}y_{31} \rangle.$$

For generic  $Y \in \mathbb{R}^{3 \times 2}$ , these equations do not vanish and the MLE given  $Y$  exists. As in Corollary 5.5, the maximum likelihood threshold is  $\text{mlt}(\mathcal{G}) = \text{mlt}_b(\mathcal{G}) = 2$ .  $\diamond$

In the previous example the null cone is Zariski closed, but this is not always the case. We now give a precise criterion for when this happens. An *unshielded collider* of a directed graph  $\mathcal{G}$  is a subgraph  $j \rightarrow i \leftarrow k$  with no edge between  $j$  and  $k$ .

**Corollary 5.7.** *Let  $\mathcal{G}$  be a TDAG, and consider the action of  $G(\mathcal{G})_{\text{SL}}^+$  on tuples of  $n$  samples. The irreducible components of the Zariski closure of the null cone are determinantal varieties: each component is defined by the maximal minors of the submatrix whose rows are a childless node and its parents. For  $n \geq \text{mlt}(\mathcal{G})$ , the null cone is Zariski closed if and only if  $\mathcal{G}$  has no unshielded colliders.*

*Proof.* By Theorem 5.3, the null cone is the union

$$(17) \quad \bigcup_{i=1}^m \mathcal{L}(i),$$

where  $\mathcal{L}(i)$  consists of all  $m \times n$  matrices whose  $i$ th row is a linear combination of rows corresponding to the parents of node  $i$ . Since the closure of a finite union is the union of the closures, the Zariski closure of (17) is a union of determinantal varieties  $\overline{\mathcal{L}(i)}^Z$ , each given by the maximal minors of the submatrix formed by node  $i$  and its parents. If node  $i$  has a child  $c$ , then  $\overline{\mathcal{L}(i)}^Z \subset \overline{\mathcal{L}(c)}^Z$ , because of the transitivity of  $\mathcal{G}$ . The first part of the assertion follows.

For the second part, we assume without loss of generality that the labels are ordered such that  $j \rightarrow i$  implies  $j < i$ . We start by assuming that  $\mathcal{G}$  has no unshielded colliders. Let  $Y$  be a matrix in the Zariski closure of the null cone, i.e. there is some node  $i$  with parents  $p_1 < \dots < p_s$  such that the corresponding  $s+1$  rows  $r_i, r_{p_1}, \dots, r_{p_s}$  of  $Y$  are linearly dependent. So there is a nontrivial linear combination  $\lambda_1 r_{p_1} + \dots + \lambda_s r_{p_s} + \lambda_{s+1} r_i = 0$ . We pick the largest index  $\ell$  such that  $\lambda_\ell \neq 0$ . If  $\ell = s+1$ , the  $i$ th row is a linear combination of its parents, and  $Y$  is in the null cone. Otherwise, the row  $r_{p_\ell}$  is a linear combination of  $r_{p_1}, \dots, r_{p_{\ell-1}}$ . We claim that these are all parents of  $p_\ell$ , and therefore that  $Y$  is in the null cone. Indeed, if some  $p_j$  for  $1 \leq j \leq \ell-1$  was not a parent of  $p_\ell$ , we would have the unshielded collider  $p_j \rightarrow i \leftarrow p_\ell$ .

Conversely, we assume that some node  $i$  has two parents  $j < k$  that are not connected. If  $i$  has several such pairs of parents, we consider a pair  $(j, k)$  such that  $k$  is minimal. This assures that every parent  $p$  of  $k$  must also be a parent of  $j$ . Indeed, by transitivity of the DAG  $\mathcal{G}$ , we have that  $p \rightarrow i$  and that  $j \not\rightarrow p$  (since  $j \not\rightarrow k$ ). Moreover, by minimality of  $k$ , it cannot be that there is no edge between  $p$  and  $j$ , so  $p \rightarrow j$ .

We will now construct a matrix  $Y$  which is not in the null cone but in its Zariski closure. We assign the rows  $1, \dots, m$  in order, according to the following rules. Each row, except for  $k$ , is assigned so that it is linearly independent of its parents. We note that this is possible due to  $n \geq \text{mlt}(\mathcal{G})$ . In particular, the  $j$ th row is assigned such that it is linearly independent of its parents, which include the parents of  $k$  as observed above. We pick the  $k$ th row equal to the  $j$ th row. Since now the parents  $j$  and  $k$  of  $i$  are linearly dependent, we see that the matrix  $Y$  is in the Zariski closure of the null cone. However, by our construction, no node in  $\mathcal{G}$  is a linear combination of its parents, so  $Y$  does not lie in the null cone.  $\square$

**Example 5.8.** Let  $\mathcal{G}$  be the TDAG  $1 \rightarrow 3 \leftarrow 2$ , with an unshielded collider. The corresponding group  $G(\mathcal{G})$  consists of invertible matrices

$$g = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ * & * & * \end{bmatrix}.$$

This is the transpose of the group in Examples 5.2 and 5.6, but we observe differences between the two models. Since node 3 has the nodes 1 and 2 as parents, Corollary 5.5 tells us that  $\text{mlt}(\mathcal{G}) = \text{mlt}_b(\mathcal{G}) = 2 + 1 = 3$  (as opposed to  $\text{mlt} = 2$  in Example 5.6).

The null cone is not Zariski closed for  $n \geq 3$ , by Corollary 5.7. Note that the Zariski closure of the null cone when  $n = 3$  is generated by the single equation  $\det(Y)$ . We see that the null cone is also not closed for  $n = 2$ , using Theorem 5.3. Here, row 3 is generically a linear combination of rows 1 and 2, and hence the Zariski closure of the null cone fills the space of tuples. However, for special choices of tuple  $Y$ , the MLE does exist. For example, let

$$Y = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Rows 1 and 2 are non-zero, and row 3 is not a linear combination of rows 1 and 2, hence the MLE given  $Y$  exists. Since  $Y$  is of minimal norm in its orbit, one MLE is  $2I_3$ , where  $\lambda = 2$  minimizes  $\frac{3}{2}\lambda - 3\log(\lambda)$ , see Proposition 3.4. In fact, there are infinitely many MLEs, as follows. For any  $g$  in the stabilizer of  $Y$  the vector  $g \cdot Y$  is also of minimal norm in the orbit. Then  $\lambda g^\top g$  is also an MLE given  $Y$ , where  $\lambda = 2$  as before. The stabilizer is

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t & -t & 1 \end{bmatrix} : t \in \mathbb{R} \right\} \quad \text{thus} \quad 2I_3 + 2t \begin{bmatrix} t & -t & 1 \\ -t & t & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad t \in \mathbb{R} \quad \text{are also MLEs.}$$

In fact, we can verify that these are all MLEs using Proposition 3.4.  $\diamond$

We describe the implications of the above results for undirected Gaussian graphical models, i.e. those coming from graphs with undirected edges, see [Sul18, Chapter 13]. A Gaussian graphical model on an undirected graph  $\mathcal{G}$  is given by all concentration matrices  $\Psi$  such that  $\psi_{ij} = 0$  whenever the edge  $i - j$  is missing from  $\mathcal{G}$ . A natural question is to determine which undirected Gaussian graphical models are Gaussian group models, i.e. of the form  $\mathcal{M}_G$  for some group  $G \subseteq \text{GL}_m$ . For instance, note that the undirected model corresponding to  $1 - 3 - 2$  is the same as the directed model from Example 5.2. We argue that any undirected model that is a Gaussian group model is covered by our study of TDAGs.

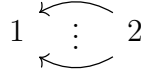
We first note that the directed model of any TDAG without unshielded colliders equals the undirected model of its underlying undirected graph, see e.g. [AMP97, Proposition 4.1]. Conversely, a necessary condition for an undirected graphical model to be a Gaussian group model can be obtained from [LM07, Theorem 2.2]: an undirected Gaussian graphical model is a transformation family if and only if the graph  $\mathcal{G}$  has neither 4-cycles nor 4-chains as induced subgraphs. There are two consequences of these conditions. One is that there is a way to *direct* the edges in  $\mathcal{G}$  so that there are no unshielded colliders.

The other consequence is that this can be done in such a way so that the undirected model coincides with the directed model  $\mathcal{M}_{\mathcal{G}}^{\rightarrow}$ , and the directed graph must be a TDAG, see page 7 of the supplementary material of [DKZ13]. In summary, we have the following equivalence.

*Remark 5.9.* The undirected graphical models that are Gaussian group models are the TDAG models without unshielded colliders. They are exactly those models whose sets of tuples of  $n$  samples with unbounded likelihood are Zariski closed for all  $n$ , by Corollary 5.7.

## APPENDIX A. CONNECTIONS TO REPRESENTATIONS OF QUIVERS

In this appendix we explain how to deduce Theorem 4.12 from the general setting given in [Kin94] in terms of representations of quivers. Our setting uses the Kronecker quiver  $Q$  with two vertices and  $n$  arrows:



An element  $Y$  in  $V := (\mathbb{C}^{m_1 \times m_2})^n$  is a finite dimensional representation of  $Q$  with dimension vector  $\alpha = (m_1, m_2)$ . We denote such a representation by  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$ . This identifies  $V$  with the space  $\mathcal{R}(Q, \alpha)$  from [Kin94], and the left-right action of  $G := \mathrm{GL}_{m_1}(\mathbb{C}) \times \mathrm{GL}_{m_2}(\mathbb{C})$  on  $V$  by  $(g_1, g_2) \cdot (Y_i)_i = (g_1 Y_i g_2^{-1})_i$  is the  $\mathrm{GL}(\alpha)$  action on  $\mathcal{R}(Q, \alpha)$ . We note that, to be consistent with [Kin94], this is a different left-right action to that we are considering elsewhere in the paper, because we take  $g_2^{-1}$  rather than  $g_2^T$ . This change preserves all stability notions, since the two actions translate to each other via the isomorphism  $(g_1, g_2) \mapsto (g_1, g_2^{-T})$ .

We consider two group actions, closely related to the action of  $G$  on  $V$ . First, we restrict to the subgroup  $H := \mathrm{SL}_{m_1}(\mathbb{C}) \times \mathrm{SL}_{m_2}(\mathbb{C})$ . Second, the group  $G$  acts on  $V \times \mathbb{C}$  by

$$g \cdot (X, z) := (g \cdot X, \chi_\theta^{-1}(g)z), \quad \text{where} \quad \chi_\theta^{-1}(g) = [\det(g_1)]^{-m_2} [\det(g_2)]^{m_1},$$

for  $\theta := (m_2, -m_1)$ . These two actions are related as follows.

**Lemma A.1.** *Consider a matrix tuple  $Y \in V = (\mathbb{C}^{m_1 \times m_2})^n$ , a scalar  $z \in \mathbb{C}^\times$  and set  $\hat{Y} := (Y, 1) \in V \times \mathbb{C}$ . Then*

- (a)  $(X, z) \in G \cdot \hat{Y} \iff z^{\frac{1}{m_1 m_2}} X \in H \cdot Y$
- (b)  $(X, z) \in \overline{G \cdot \hat{Y}} \iff z^{\frac{1}{m_1 m_2}} X \in \overline{H \cdot Y}$
- (c)  $(\exists X \in V : (X, 0) \in \overline{G \cdot \hat{Y}}) \iff 0 \in \overline{H \cdot Y}$ .

*Proof.* To prove (a), let  $g \in G$  such that  $(X, z) = g \cdot \hat{Y}$ . Then  $[\det(g_1)]^{-m_2} [\det(g_2)]^{m_1} = z$  and  $g \cdot Y = X$ . Setting

$$h := \left( [\det(g_1)]^{-\frac{1}{m_1}} g_1, [\det(g_2)]^{-\frac{1}{m_2}} g_2 \right) \in H$$

we obtain  $h \cdot Y = z^{\frac{1}{m_1 m_2}} X$ . Conversely, given the latter equation for some  $h = (h_1, h_2) \in H$  and defining  $g := (z^{-\frac{1}{m_1 m_2}} h_1, h_2)$  yields  $g \cdot \hat{Y} = (X, z)$ . Now, part (b) follows from taking a sequence in the respective orbit, that tends to a point in the orbit closure, and applying (a) to this sequence.

For  $Y = 0$  we have  $(0, 0) \in \overline{G \cdot \hat{Y}}$  and  $0 \in \overline{H \cdot Y}$ . Thus, we assume  $Y \neq 0$ . First, let  $X \in V$  and let  $g^{(k)} \in G$  be a sequence such that  $g^{(k)} \cdot \hat{Y} = (g^{(k)} \cdot Y, \chi_\theta^{-1}(g^{(k)}))$  tends to  $(X, 0)$  for  $k \rightarrow \infty$ . Since  $\chi_\theta^{-1}(g^{(k)}) \neq 0$  for all  $k$ , we can apply (a) to obtain

$$Y_k := [\chi_\theta^{-1}(g^{(k)})]^{-\frac{1}{m_1 m_2}} g^{(k)} \cdot Y \in H \cdot Y$$

for all  $k$ . With  $g^{(k)} \cdot \hat{Y} \rightarrow (X, 0)$  for  $k \rightarrow \infty$  we conclude that the sequence  $Y_k$  tends to  $0 \in V$ . On the other hand, assume there are  $Y_k \in H \cdot Y$  with  $Y_k \rightarrow 0$  for  $k \rightarrow \infty$ . As  $Y \neq 0$  we have  $Y_k \neq 0$  and hence  $c_k := \|Y_k\|^{-\frac{m_1 m_2}{2}} \neq 0$  for all  $k$ . Thus, setting

$$X_k := c_k^{-\frac{1}{m_1 m_2}} Y_k$$

and applying (a) gives  $(X_k, c_k) \in G \cdot \hat{Y}$ . The latter sequence tends to  $(0, 0) \in V \times \mathbb{C}$  by the choice of  $c_k$ .  $\square$

With the help of Lemma A.1 we can prove Theorem 4.12.

*Proof of Theorem 4.12.* The equivalence of (a) and (b) is Theorem 3.15. In the following, we prove the equivalence of (b) and (c).

Recall that  $\theta = (m_2, -m_1)$ . By [Kin94, Proposition 3.1] the matrix tuple  $Y = (Y_1, \dots, Y_n)$  is  $\chi_\theta$ -stable if and only if the representation  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  is  $\theta$ -stable. First, we show that the former is equivalent to being stable under the action of  $H$ . Then we rephrase the latter as the shrunk subspace condition Theorem 4.12(c).

We set  $\Delta := \{(tI_{m_1}, tI_{m_2}) \mid t \in \mathbb{C}^\times\}$  and let  $G_{\hat{Y}}$  denote the  $G$ -stabilizer of  $\hat{Y} = (Y, 1)$ . The tuple  $Y$  is  $\chi_\theta$ -stable if and only if the orbit  $G \cdot \hat{Y}$  is closed and the group  $G_{\hat{Y}}/\Delta$  is finite, by [Kin94, Lemma 2.2]. We show equivalence to  $Y$  being stable under the  $H$ -action, i.e. that  $Y \neq 0$ , with  $H \cdot Y$  closed and finite stabilizer  $H_Y$ . Note that  $Y = 0$  is not  $\chi_\theta$ -stable, hence we can assume  $Y \neq 0$  in the following.

Firstly, the group  $G_{\hat{Y}}/\Delta$  is finite if and only if  $H_Y$  is finite, since the group morphism

$$\varphi: G_{\hat{Y}} \rightarrow H_Y, \quad (g_1, g_2) \mapsto \left( \det(g_1)^{-\frac{1}{m_1}} g_1, \det(g_2)^{-\frac{1}{m_2}} g_2 \right)$$

induces an isomorphism  $G_{\hat{Y}}/\Delta \cong H_Y$ . Note that  $\varphi$  is well-defined, since  $\chi_\theta^{-1}(g_1, g_2) = 1$  yields  $\det(g_1)^{-\frac{1}{m_1}} = \det(g_2)^{-\frac{1}{m_2}}$ .

Secondly, we show for  $Y \neq 0$  that  $G \cdot \hat{Y}$  is closed if and only if  $H \cdot Y$  is closed using Lemma A.1. Assume that the orbit  $G \cdot \hat{Y}$  is closed and take  $X \in \overline{H \cdot Y}$ . Then  $(X, 1) \in \overline{G \cdot \hat{Y}} = G \cdot \hat{Y}$  using Lemma A.1(b), and hence  $X \in H \cdot Y$  by Lemma A.1(a). Conversely,  $H \cdot Y$  being closed with  $Y \neq 0$  implies that  $0 \notin \overline{H \cdot Y}$ . Thus, Lemma A.1(c) yields  $\overline{G \cdot \hat{Y}} \cap (V \times \{0\}) = \emptyset$ . Hence any element  $(X, z) \in \overline{G \cdot \hat{Y}}$  must satisfy  $z \in \mathbb{C}^\times$  and we conclude that  $(X, z) \in G \cdot \hat{Y}$  using Lemma A.1.

For  $\theta$ -stability,  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  is viewed as an element of the category of finite dimensional representations of the Kronecker quiver  $Q$ . We note that  $\langle \theta, (m_1, m_2) \rangle = 0$  is satisfied by our choice  $\theta = (m_2, -m_1)$ . We specialize [Kin94, Definition 1.1] to our representation  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$ . The representation is  $\theta$ -semistable if and only if for all subrepresentations of  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$ , i.e. all subspaces  $V_1 \subseteq \mathbb{C}^{m_1}$ ,  $V_2 \subseteq \mathbb{C}^{m_2}$  such that



$Y_i V_2 \subseteq V_1$  for all  $i$ , we have

$$(18) \quad \langle \theta, (\dim V_1, \dim V_2) \rangle = m_2 \dim V_1 - m_1 \dim V_2 \geq 0.$$

The representation  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  is  $\theta$ -stable if and only if in addition, the inequality in (18) is strict for all non-zero proper subrepresentations. Here, non-zero means  $V_1 \neq 0$  or  $V_2 \neq 0$ , while proper means  $V_1 \subsetneq \mathbb{C}^{m_1}$  or  $V_2 \subsetneq \mathbb{C}^{m_2}$ . Since  $V_1 \neq 0$  and  $V_2 = 0$  gives strict inequality in (18), it is enough to consider  $V_2 \neq 0$ . On the other hand, strict inequality in (18) holds for all proper subrepresentations satisfying  $V_1 \subsetneq \mathbb{C}^{m_1}$  and  $V_2 = \mathbb{C}^{m_2}$  if and only if there is *no* proper subrepresentation of this form, i.e. if and only if  $\text{rank}(Y_1, \dots, Y_n) = m_1$ . Hence, by requiring the latter condition we can restrict to the case  $V_2 \subsetneq \mathbb{C}^{m_2}$ . All together, we rephrased the  $\theta$ -stability of  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  as in the statement.  $\square$

*Remark A.2.* Proposition 3.1 in [Kin94] also provides an alternative proof of the complex analog of Theorem 4.4, i.e. [BD06, Proposition 2.1]. It states that  $Y$  is  $\chi_\theta$ -semistable if and only if  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  is  $\theta$ -semistable. The former holds if and only if

$$(V \times \{0\}) \cap \overline{G \cdot \hat{Y}} \neq \emptyset,$$

i.e. if and only if  $Y$  is semistable under the action of  $H$ , by Lemma A.1. On the other hand, the proof of Theorem 4.12 shows that  $(\mathbb{C}^{m_1}, \mathbb{C}^{m_2}; Y)$  is  $\theta$ -semistable if and only if (18) holds for all subspaces  $V_1 \subseteq \mathbb{C}^{m_1}$ ,  $V_2 \subseteq \mathbb{C}^{m_2}$  satisfying  $Y_i V_2 \subseteq V_1$  for all  $i = 1, \dots, n$ .

**Acknowledgements.** We are grateful to Peter Bürgisser, Mathias Drton, Bernd Sturmfels, and Michael Walter for fruitful discussions. We also thank Jan Draisma, Nikolay Nikolov, Panagiotis Papazoglou, Piotr Zwiernik and the anonymous referees for useful hints and suggestions. CA was partially supported by the Deutsche Forschungsgemeinschaft (DFG) in the context of the Emmy Noether junior research group KR 4512/1-1. KK was partially supported by the Knut and Alice Wallenberg Foundation within their WASP (Wallenberg AI, Autonomous Systems and Software Program) AI/Math initiative. Research of PR is funded by the European Research Council (ERC) under the European’s Horizon 2020 research and innovation programme (grant agreement no. 787840).

## REFERENCES

- [AHSE95] H.H. Andersen, M. Højbjerg, D. Sørensen, and P.S. Eriksen. *Linear and graphical models: for the multivariate complex normal distribution*, volume 101. Springer Science & Business Media, 1995.
- [AKRS20] Carlos Améndola, Kathlén Kohn, Philipp Reichenbach, and Anna Seigal. Toric invariant theory for maximum likelihood estimation in log-linear models. arXiv:2012.07793, 2020.
- [AMP97] S.A. Andersson, D. Madigan, and M.D. Perlman. On the Markov equivalence of chain graphs, undirected graphs, and acyclic digraphs. *Scandinavian Journal of Statistics*, 24(1):81–102, 1997.
- [AZGL<sup>+</sup>18] Z. Allen-Zhu, A. Garg, Y. Li, R. Oliveira, and A. Wigderson. Operator scaling via geodesically convex optimization, invariant theory and polynomial identity testing. In *STOC’18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 172–181. ACM, New York, 2018.
- [BD06] M. Bürgin and J. Draisma. The Hilbert null-cone on tuples of matrices and bilinear forms. *Mathematische Zeitschrift*, 254(4):785–809, 2006.

- [BFG<sup>+</sup>19] P. Bürgisser, C. Franks, A. Garg, R. Oliveira, M. Walter, and A. Wigderson. Towards a theory of non-commutative optimization: geodesic first and second order methods for moment maps and polytopes. *arXiv:1910.12375*, 2019.
- [BGO<sup>+</sup>17] P. Bürgisser, A. Garg, R. Oliveira, M. Walter, and A. Wigderson. Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory. *arXiv:1711.08039*, 2017.
- [BL17] C. Böhm and R.A. Lafuente. Real geometric invariant theory. *arXiv:1701.00643*, 2017.
- [BNBJJ82] O. Barndorff-Nielsen, P. Blaesild, J. Ledet Jensen, and B. Jørgensen. Exponential transformation models. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 379(1776):41–65, 1982.
- [Cra98] Erhard Cramer. Conditional iterative proportional fitting for gaussian distributions. *Journal of multivariate analysis*, 65(2):261–276, 1998.
- [DFKP19] M. Drton, C. Fox, A. Käußl, and G. Pouliot. The maximum likelihood threshold of a path diagram. *The Annals of Statistics*, 47(3):1536–1553, 2019.
- [DKH20] M. Drton, S. Kuriki, and P. Hoff. Existence and Uniqueness of the Kronecker Covariance MLE. *arXiv:2003.06024*, 2020.
- [DKZ13] J. Draisma, S. Kuhnt, and P. Zwiernik. Groups acting on Gaussian graphical models. *The Annals of Statistics*, 41(4):1944–1969, 2013.
- [DM17] H. Derksen and V. Makam. Polynomial degree bounds for matrix semi-invariants. *Advances in Mathematics*, 310:44–63, 2017.
- [Dut99] Pierre Dutilleul. The MLE algorithm for the matrix normal distribution. *Journal of statistical computation and simulation*, 64(2):105–123, 1999.
- [Fis34] Ronald Aylmer Fisher. Two new properties of mathematical likelihood. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 144(852):285–307, 1934.
- [FM81] Stephen E Fienberg and Michael M Meyer. Iterative proportional fitting. Technical report, Carnegie-Mellon University, Pittsburgh PA, Dept. of Statistics, 1981.
- [GGOW16] A. Garg, L. Gurvits, R. Oliveira, and A. Wigderson. A deterministic polynomial time algorithm for non-commutative rational identity testing. In *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016*, pages 109–117. IEEE Computer Soc., Los Alamitos, CA, 2016.
- [Goo63] N.R. Goodman. Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction). *The Annals of Mathematical Statistics*, 34(1):152–177, 1963.
- [Gur04] L. Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, 2004.
- [Hil93] D. Hilbert. Über die vollen Invariantensysteme. *Mathematische Annalen*, 42(3):313–373, 1893.
- [IQS18] G. Ivanyos, Y. Qiao, and K.V. Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. *Computational Complexity*, 27(4):561–593, 2018.
- [Kin94] A.D. King. Moduli of representations of finite-dimensional algebras. *The Quarterly Journal of Mathematics. Oxford. Second Series*, 45(180):515–530, 1994.
- [KN79] G. Kempf and L. Ness. The length of vectors in representation spaces. In *Algebraic geometry (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978)*, volume 732 of *Lecture Notes in Math.*, pages 233–243. Springer, Berlin, 1979.
- [Lau96] S. Lauritzen. *Graphical models*, volume 17. Clarendon Press, 1996.
- [LM07] G. Letac and H. Massam. Wishart distributions for decomposable graphs. *The Annals of Statistics*, 35(3):1278–1323, 2007.
- [LZ05] N. Lu and D.L. Zimmerman. The likelihood ratio test for a separable covariance matrix. *Statistics & probability letters*, 73(4):449–457, 2005.
- [MBT14] A.T. Mai, F. Bastin, and M. Toulouse. On optimization algorithms for maximum likelihood estimation. CIRRELT 2014-64, 2014.

- [Mil17] J.S. Milne. *Algebraic groups*, volume 170 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017.
- [Mos55] G.D. Mostow. Self-adjoint groups. *Annals of Mathematics. Second Series*, 62:44–55, 1955.
- [Myu03] I.J. Myung. Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1):90–100, 2003.
- [RS90] R.W. Richardson and P.J. Slodowy. Minimum vectors for real reductive algebraic groups. *Journal of the London Mathematical Society. Second Series*, 42(3):409–429, 1990.
- [Sin64] R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The Annals of Mathematical Statistics*, 35(2):876–879, 1964.
- [Sul18] S. Sullivant. *Algebraic Statistics*, volume 194 of *Graduate Studies in Mathematics*. AMS, 2018.
- [Wal17] N.R. Wallach. *Geometric invariant theory: Over the real and complex numbers*. Universitext. Springer, 2017.
- [Woo56] R.A. Wooding. The multivariate distribution of complex normal variables. *Biometrika*, 43(1/2):212–215, 1956.

**Authors’ addresses:**

Technische Universität München, Germany,  
KTH Royal Institute of Technology, Sweden,  
Technische Universität Berlin, Germany,  
University of Oxford, United Kingdom,

carlos.amendola@tum.de  
kathlen@kth.se  
reichenbach@tu-berlin.de  
seigal@maths.ox.ac.uk