MAXIMUM LIKELIHOOD ESTIMATION FOR MATRIX NORMAL MODELS VIA QUIVER REPRESENTATIONS

HARM DERKSEN AND VISU MAKAM

Abstract. We study the log-likelihood function and Maximum Likelihood Estimate (MLE) for the matrix normal model for both real and complex models. We describe the exact number of samples needed to achieve (almost surely) three conditions, namely a bounded log-likelihood function, existence of MLEs, and uniqueness of MLEs. As a consequence, we observe that almost sure boundedness of log-likelihood function guarantees almost sure existence of an MLE, thereby proving a conjecture of Drton, Kuriki and Hoff. The main tools we use are from the theory of quiver representations, in particular, results of Kac, King and Schofield on canonical decomposition and stability.

Contents

1. Introduction

The following problem is fundamental in statistics in a variety of settings: Among a collection of probability distributions (a.k.a. a statistical model), find the one that best fits some empirical data. A probability distribution in the collection that maximizes the likelihood of the empirical data is called a Maximum Likelihood Estimate (MLE). Understanding the existence and uniqueness of MLEs is an important problem that is widely studied. A related problem is to understand when the likelihood function (or equivalently the log-likelihood function) is bounded above. Henceforth, we will simply say likelihood and log-likelihood functions are bounded to mean that they are bounded above.

In many settings, data is observed in two domains, and hence observations are naturally matrix-valued. For such observations, one sometimes assumes that they follow a *matrix normal distribution*. These matrix normal models have been used for various purposes in various settings, for example to EEG/MEG data [6, 11, 32, 54], Environmental data [23, 45], Netflix movie rating data [1] and facial recognition [55] to name a few. The existence and uniqueness of MLEs, and the boundedness of the

Key words and phrases. maximum likelihood estimation, matrix normal models, stability, quiver representations.

The first author was partially supported by NSF grants IIS-1837985 and DMS-2001460. The second author was partially supported by NSF grants DMS-1638352 and CCF-1900460 and the University of Melbourne.

likelihood function in matrix normal models for small sample sizes is important to understand so as to use them effectively in applications (see [21] for more motivation). This problem was studied in many manuscripts before, e.g., [21, 22, 49, 51], each making partial progress.

More recently, Amendola, Kohn, Reichenbach and Seigal [2] uncovered connections between MLEs and stability notions in invariant theory and adapted results of Bürgin and Draisma [8] to further improve the results on sample size required for the (almost sure) boundedness of the log-likelihood function. It is worthwhile to mention some interesting connections to invariant theory even though it will not be relevant to our paper. The flip-flop algorithm for computing an MLE for matrix normal models [22, 44] is very similar to the algorithm proposed by Gurvits [28] for computing capacity of completely positive operators, which is well known to be equivalent to null cone membership (a central problem in algorithmic invariant theory) for the so called left-right action. Even more curiously, the notion of geodesic convexity that has played a major role in understanding invariant theoretic algorithms in recent years [9] can already be seen in Wiesel's work [56] several years prior in the setting of the flip-flop algorithm.

Summary of our main results. In this paper, for matrix normal models (both real and complex), we will compute the *exact* number of samples needed to achieve (almost surely) three conditions, i.e., (1) a bounded log-likelihood function, (2) existence of MLEs and (3) uniqueness of MLEs, thereby completely resolving the problems. In particular, we prove a conjecture of Drton, Kuriki and Hoff [21] that almost sure boundedness of the likelihood function implies almost sure existence of an MLE. We utilize heavily the connections between invariant theory and MLEs and in particular the connection between matrix normal models and geometric invariant theory for quiver representations discovered in [2]. Our techniques, however, are significantly different from any of the previous work on these problems and rely on the algebraic aspects of theory of quiver representations. We also study a related model called the model of proportional covariance matrices and give complete answers to the aforementioned questions in that case as well.

Before we get into precise definitions and results, a few remarks on background literature. We refer to [2, 21] and references therein for more details regarding real and complex Gaussian models, matrix normal models, their MLEs and associated thresholds as well as more motivation for the problems we discuss in this paper. A detailed explanation and proofs of the connections between MLEs and invariant theory can be found in [2]. We point the reader to the book [18] as a comprehensive introductory text on quiver representations.

1.1. **Real Gaussian models.** We denote by PD_n, the cone of $n \times n$ positive definite matrices with entries in $\mathbb{R}$, the field of real numbers. For an n-dimensional Gaussian distribution with mean 0 and covariance matrix $\Sigma \in \mathrm{PD}_n$, the density function is described by

$$f_\Sigma(y) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} e^{-\frac{1}{2} y^\top \Sigma^{-1} y}.$$

The inverse of the covariance matrix, i.e., Σ^{-1} is called the concentration matrix and denoted Ψ. A subset of $\mathcal{M} \subseteq \mathrm{PD}_n$ defines a statistical model consisting of the n-dimensional Gaussian distributions with mean 0 and concentration matrix $\Psi \in \mathcal{M}$. For Gaussian models, the sample data is a tuple of vectors $Y = (Y_1, \ldots, Y_m) \in (\mathbb{R}^n)^m$, where m denotes the sample size. The likelihood function $L_Y : \mathrm{PD}_n \to \mathbb{R}$ is given by

$$L_Y(\Psi) = \prod_{i=1}^m f_{\Psi^{-1}}(Y_i) = \det\left(\frac{\Psi}{2\pi}\right)^{m/2} e^{-\frac{1}{2}\sum_{i=1}^m Y_i^\top \Psi Y_i}.$$

The log-likelihood function $l_Y : \mathrm{PD}_n \to \mathbb{R}$ (up to an additive constant) is given by

$$l_Y(\Psi) = \frac{m}{2} \log\det(\Psi) - \frac{1}{2} \operatorname{Tr}\left(\Psi \sum_{i=1}^{m} Y_i Y_i^\top\right).$$

A Maximum Likelihood Estimate (MLE) is a point $\widehat{\Psi} \in \mathcal{M}$ that maximizes the likelihood of observing the data Y, which is equivalent to maximizing the log-likelihood function l_Y. In other words, $\widehat{\Psi}$ is an MLE if $l_Y(\widehat{\Psi}) \geq l_Y(\Psi)$ for all $\Psi \in \mathcal{M}$. If the log-likelihood function is unbounded, then of course MLEs do not exist. But even when the log-likelihood function is bounded, MLEs do not necessarily exist because the supremum of the log-likelihood function may not be achieved by any particular concentration matrix. Finally, even when an MLE exists, there is no guarantee that it is unique as there may be many points in the model that achieve the maximum possible value of the log-likelihood function.

For a Gaussian model $\mathcal{M} \subseteq \mathrm{PD}_n$, we define three threshold functions as follows:

(1) We define $\mathrm{mlt}_b(\mathcal{M})$ to be the smallest integer m such that for $d \geq m$, the log-likelihood function l_Y for $Y = (Y_1, \ldots, Y_d) \in (\mathbb{R}^n)^d$ is bounded almost surely.
(2) We define $\mathrm{mlt}_e(\mathcal{M})$ to be the smallest integer m such that for $d \geq m$, an MLE exists almost surely for $Y \in (\mathbb{R}^n)^d$.
(3) We define $\mathrm{mlt}_u(\mathcal{M})$ to be the smallest integer m such that for $d \geq m$, there almost surely exists a unique MLE for $Y \in (\mathbb{R}^n)^d$.

In the above, almost surely means that the property holds away from a subset of $(\mathbb{R}^n)^d$ of Lebesgue measure zero. We will refer to $\mathrm{mlt}_b, \mathrm{mlt}_e$ and mlt_u as maximum likelihood threshold functions. By the above discussion, we observe that $\mathrm{mlt}_b \leq \mathrm{mlt}_e \leq \mathrm{mlt}_u$.

1.2. **Complex Gaussian models.** The setting of complex Gaussian models is very much analogous, with minor changes. The density function for a (circularly symmetric) complex n-dimensional Gaussian with covariance matrix $\Sigma \in \mathrm{PD}_n$ (the cone of positive definite $n \times n$ complex matrices) is given by

$$f_\Sigma(y) = \frac{1}{\det(\pi\Sigma)} e^{-y^\dagger \Sigma^{-1} y},$$

where $y^\dagger$ denotes the adjoint of y, i.e., conjugate transpose.

By an analogous computation the real case, we see that the log-likelihood function (up to an additive constant) is given by

$$l_Y(\Psi) = m \log\det(\Psi) - \operatorname{Tr}\left(\Psi \sum_{i=1}^{m} Y_i Y_i^\dagger\right),$$

where $Y_i^\dagger$ denotes the adjoint of Y_i. Observe the similarity to the log-likelihood function for real Gaussian (except for the multiplicative factor of $\frac{1}{2}$ which can be ignored for the purposes of understanding whether l_Y is bounded above, achieves its maximum, etc). The rest of the discussion follows analogously. In particular, MLEs and the maximum likelihood thresholds are defined as above.

Remark 1.1. We will reuse the same notation for real and complex models (for e.g., PD_n, l_Y, etc). It will always be clear whether we are in a real or a complex model, so there will be no scope for confusion.

1.3. **Matrix normal models and main results.** If $n = pq$, we consider the subset

$$\mathcal{M}(p,q) = \{\Psi_1 \otimes \Psi_2 \mid \Psi_1 \in \mathrm{PD}_p, \Psi_2 \in \mathrm{PD}_q\} \subseteq \mathrm{PD}_{pq},$$

where $\otimes$ denotes the Kronecker (or tensor) product of matrices. Such a statistical model is called a matrix normal model. Sometimes it is also called a Kronecker covariance model. We will consider and deal with both real and complex matrix normal models. As mentioned above, PD_n denotes positive definite real or complex matrices depending on whether we are working with real or complex matrix normal models. When we want to differentiate between the real and complex models, we will use $\mathcal{M}_{\mathbb{R}}(p,q)$ and $\mathcal{M}_{\mathbb{C}}(p,q)$ respectively. For matrix normal models, it will be convenient to interpret the data as a $p \times q$ matrix, rather than a vector of size pq, and we will do so.

Drton, Kuriki, and Hoff [21] suggest that an exact formula for maximum likelihood thresholds may be complicated because of the following behavior. For a sample size of two (i.e., $Y = (Y_1, Y_2) \in \mathrm{Mat}^2_{p,q}$, where $\mathrm{Mat}_{p,q}$ denotes the space of $p \times q$ matrices), consider the matrix normal model $\mathcal{M}(p,q)$.

- If $(p,q) = (5,4)$, then we almost surely have a unique MLE;
- If $(p,q) = (6,4)$, then we almost surely have an MLE that is not unique;
- If $(p,q) = (7,4)$, then MLEs do not exist;
- If $(p,q) = (8,4)$, then we almost surely have an MLE that is not unique.

We obtain *exact* formulas for the $\mathrm{mlt}_b, \mathrm{mlt}_e$ and mlt_u for matrix normal models $\mathcal{M}(p,q)$, both real and complex. There are some delicate differences between real matrix normal models and complex matrix normal models which we will elaborate on later, see also [2, Example 4.2]. Nevertheless, the maximum likelihood threshold functions are the same for both real and complex matrix normal models.

From the point of view of quiver representations, it is natural to fix the number of samples, and then study the boundedness of log-likelihood function, existence and uniqueness of MLEs as p and q vary. This subtle change in point of view offers a significantly different perspective from earlier work. The added advantage is that the answer comes out very crisp!

Theorem 1.2. *Suppose $K = \mathbb{R}$ or $\mathbb{C}$. Let $Y = (Y_1, \ldots, Y_m) \in \mathrm{Mat}^m_{p,q}(K)$. Let $d = \gcd(p,q)$. Then, for the matrix normal model $\mathcal{M}_K(p,q)$:*

1. *If $p^2 + q^2 - mpq < 0$, then there almost surely exists a unique MLE.*
2. *If $p^2 + q^2 - mpq \in \{0, d^2\}$, then an MLE exists almost surely. Further this MLE is almost surely unique if and only if $d = 1$.*
3. *In all other cases, the log-likelihood function is unbounded always (not just almost surely). Consequently MLEs do not exist.*

At this juncture, we invite the reader to verify that the complicated behavior in the examples mentioned above is consistent with the statement of the above theorem. We reformulate Theorem 1.2 to compute exactly $\mathrm{mlt}_b, \mathrm{mlt}_e$, and mlt_u for $\mathcal{M}(p,q)$. While not as elegant as the formulation in the above theorem, it remains fairly simple.

Theorem 1.3. *Consider the (real or complex) matrix normal model $\mathcal{M}(p,q)$. Let $\gcd(p,q) = d$, and let $r = \dfrac{p^2 + q^2 - d^2}{pq}$. Then*

1. *If $p = q = 1$, then $\mathrm{mlt}_b = \mathrm{mlt}_e = \mathrm{mlt}_u = 1$.*
2. *If $p = q > 1$, then $\mathrm{mlt}_b = \mathrm{mlt}_e = 1$ and $\mathrm{mlt}_u = 3$.*
3. *If $p \neq q$ and r is an integer, then $\mathrm{mlt}_b = \mathrm{mlt}_e = r$. If $d = 1$, then $\mathrm{mlt}_u = r$, and if $d > 1$, then $\mathrm{mlt}_u = r + 1$.*
4. *If $p \neq q$ and r is not an integer, then $\mathrm{mlt}_b = \mathrm{mlt}_e = \mathrm{mlt}_u = \lceil \frac{p^2+q^2}{pq} \rceil$.*

Now, it is a simple observation to see that the Drton-Kuriki-Hoff conjecture [21] follows immediately from the above theorems:

Corollary 1.4 (Drton-Kuriki-Hoff conjecture [21]). *For the (real or complex) matrix normal model $\mathcal{M}(p,q)$, almost sure boundedness of (log-)likelihood function implies almost sure existence of MLE. In particular, for all (p,q),*

$$\mathrm{mlt}_b(\mathcal{M}(p,q)) = \mathrm{mlt}_e(\mathcal{M}(p,q)).$$

We also consider a variant of the matrix normal model where one of the matrices is diagonal which is called the model of proportional covariance matrices (see e.g., [24, 35]). Let

$$\mathcal{N}(p,q) = \{\Psi \otimes D \mid \Psi \in \mathrm{PD}_p, D \in \mathrm{PD}_q \text{ is a diagonal matrix}\} \subseteq \mathrm{PD}_{pq}.$$

This model has also been considered before [49] in the context of maximum likelihood threshold functions. We have the following results:

Theorem 1.5. *Consider the model $\mathcal{N}(p,q)$. Let $r = p/q$.*

1. *If r is an integer, then $\mathrm{mlt}_b = \mathrm{mlt}_e = r$. If $q = 1$, then $\mathrm{mlt}_u = r$ and if $q > 1$, then $\mathrm{mlt}_u = r + 1$.*
2. *If r is not an integer, then $\mathrm{mlt}_b = \mathrm{mlt}_e = \mathrm{mlt}_u = \lceil r \rceil$.*

Once again, we see that $\mathrm{mlt}_b(\mathcal{N}(p,q)) = \mathrm{mlt}_e(\mathcal{N}(p,q))$.

1.4. **Organization.** In Section 2, we recall invariant theory and the connections to MLE for Gaussian group models. We also study stability notions when the underlying field is $\mathbb{R}$ or $\mathbb{C}$ and discuss an important result, i.e., Proposition 2.23 that allow us to transfer generic stability results from $\mathbb{C}$ (where it is easier to prove things) to $\mathbb{R}$ (which is more important for statistics). We discuss quiver representations, stability for quiver representations and canonical decompositions in Sections 3, 4 and 5 respectively. In Section 6 and Section 7, we bring together all the material we develop to prove our main results on maximum likelihood thresholds.

1.5. **Acknowledgements.** We would like to thank Carlos Améndola, Mathias Drton, Kathlén Kohn, Philipp Reichenbach, Anna Seigal for interesting discussions and comments on an earlier draft of this paper. We also thank Ronno Das and Siddharth Krishna for helpful discussions.

2. Invariant theory

Invariant theory is the study of symmetries captured by group actions. The roots of this subject can be traced back to the masters of computation in the 19th century. At the turn of the 20th century, the work of Hilbert and Weyl brought invariant theory to the forefront of mathematics, and served to establish the foundations for modern algebra and algebraic geometry.

The basic setting is as follows. Let G be a group. A representation of G is an action of G on a (finite-dimensional) vector space V (over a field K) by linear transformations. This is captured succinctly as a group homomorphism $\rho : G \to \mathrm{GL}(V)$. In particular, an element $g \in G$ acts on V by the linear transformation $\rho(g)$. We write $g \cdot v$ or gv to mean $\rho(g)v$. Throughout this paper, we will only consider the setting where G is a linear algebraic group (over the underlying field K), i.e., G is an (affine) variety, the multiplication and inverse maps are morphism of varieties, and the action is a rational action (or rational representation), i.e., $\rho : G \to \mathrm{GL}(V)$ is a morphism of algebraic groups.

The G-orbit of $v \in V$ is the set of all vectors that you can get from v by applying elements of the group, i.e.,

$$O_v := \{gv \mid g \in G\} \subseteq V.$$

We denote by $K[V]$, the ring of polynomial functions on V (a.k.a. the coordinate ring of V). A polynomial function $f \in K[V]$ is called *invariant* if $f(gv) = f(v)$ for all $g \in G$ and $v \in V$. In other words, a polynomial is called invariant if it is constant along orbits. The invariant ring is

$$K[V]^G := \{f \in K[V] \mid f(gv) = f(v) \ \forall\ g \in G, v \in V\}.$$

The invariant ring has a natural grading by degree, i.e., $K[V]^G = \oplus_{d=0}^{\infty} K[V]_d^G$ where $K[V]_d^G$ consists of all invariant polynomials that are homogeneous of degree d. For $v \in V$, we denote by $\overline{O_v}$ the closure of the orbit O_v.

Remark 2.1. To define the closure, we need to define a topology on V. In this paper, we will only use the fields $K = \mathbb{R}$ or $\mathbb{C}$. Hence, we will use the standard Euclidean topology on V for orbit closures, unless otherwise specified. This is not standard. In literature, the topology is usually taken as the Zariski topology. We will need to use the Zariski topology at times, but we will be careful in specifying it each time. For $K = \mathbb{C}$, the orbit closure w.r.t. Euclidean topology agrees with the orbit closure w.r.t. Zariski topology (in the setting of rational actions of reductive groups). We caution the reader that the interplay between the Euclidean and Zariski topology can be a bit tricky at times for $K = \mathbb{R}$.

The stabilizer of the action at a point $v \in V$ is defined to be the subgroup $G_v := \{g \in G \mid gv = v\}$. We make a few definitions.

Definition 2.2. Let $K = \mathbb{R}$ or $\mathbb{C}$, and let G be an algebraic group (over K) with a rational action on a vector space V (over K), i.e., $\rho : G \to \mathrm{GL}(V)$. Let Δ denote the kernel of the homomorphism ρ. Give V the standard Euclidean topology. Then, for $v \in V$, we say v is

- *unstable* if $0 \in \overline{O_v}$;
- *semistable* if $0 \notin \overline{O_v}$;
- *polystable* if $v \neq 0$ and O_v is closed;
- *stable* if v is polystable and the quotient G_v/Δ is finite.

We point out again that our definitions may not be quite standard because we use the Euclidean topology. However, this is the form that is most suited for our purposes. Clearly, stable $\implies$ polystable $\implies$ semistable. A point is unstable if and only if it is not semistable. Also, note for any action of G on V, there is a natural diagonal action on the direct sum V^m by $g \cdot (v_1, \ldots, v_m) = (gv_1, \ldots, gv_m)$ for all $g \in G$ and $v_i \in V$. Moreover, note that for any group action $\rho : G \to \mathrm{GL}(V)$, the notions of semistable, polystable and stable are the same whether we consider the action of G or the action of $\rho(G)$.

We make another important definition:

Definition 2.3. Let $K = \mathbb{R}$ or $\mathbb{C}$, and let G be an algebraic group (over K) with a rational action on a vector space V (over K). Then, we say V is generically G-semistable (resp. polystable, stable, unstable) if there is a non-empty Zariski-open subset $U \subseteq V$ such that every $v \in U$ is G-semistable (resp. polystable, stable, unstable).

The following notion of a null cone plays a central role in computational invariant theory.

Definition 2.4 (Null cone)**.** Let $K = \mathbb{R}$ or $\mathbb{C}$, and let G be an algebraic group (over K) with a rational action on a vector space V (over K). Then, the null cone is defined by

$$\mathcal{N}_G(V) := \{v \in V \mid 0 \in \overline{O_v}\}.$$

In other words, the null cone consists of all the unstable points in V.

2.1. **MLE for Gaussian group models and invariant theory.** In this subsection, we will briefly recall Gaussian group models and their connections to invariant theory. Suppose $K = \mathbb{R}$ or $\mathbb{C}$. For any group G acting on K^n by linear transformations (i.e., $\rho : G \to \mathrm{GL}_n$), there is a corresponding *Gaussian group model* $\mathcal{M}_G := \{\rho(g)^\dagger \rho(g) \mid g \in G\} \subseteq \mathrm{PD}_n$. Note that adjoint is the same as transpose for a matrix with real entries. The following result for Gaussian group models was proved in [2] (we state a more general, but equivalent form of their result).

Theorem 2.5 ([2]). *Let $K = \mathbb{R}$ or $\mathbb{C}$, and let V be a finite dimensional Hilbert space, i.e., a vector space with a positive definite inner product (Hermitian when $K = \mathbb{C}$). Let $\rho : G \to \mathrm{GL}(V)$ be a rational action of G on V. Suppose $\rho(G) \subseteq \mathrm{GL}(V)$ is a Zariski closed subgroup, closed under adjoints and non-zero scalar multiples. Let $G_{\mathrm{SL}} \subseteq G$ be a subgroup such that $\rho(G_{\mathrm{SL}}) = \rho(G) \cap (\mathrm{SL}(V))$ and let $Y \in V^m$ be an m-tuple of samples. Then, for the (diagonal) action of G_{SL} and the model $\mathcal{M}_G$, we have*

- *Y is semistable $\iff$ l_Y is bounded from above;*
- *Y is polystable $\iff$ an MLE exists;*
- *Y is stable $\implies$ there is a unique MLE. Further, if $K = \mathbb{C}$, the converse also holds, i.e., there is a unique MLE $\implies$ Y is stable.*

Remark 2.6. In the above result, it suffices to ask for $\rho(G_{\mathrm{SL}})$ and $\rho(G) \cap (\mathrm{SL}(V))$ to have the same identity component since the stability notions for the action of either group will be the same. Indeed, we will need this mild generalization in Theorem 2.7 and Proposition 2.9 below.

Matrix normal models are Gaussian group models. Consider the so-called *Left-Right action* of $G = \mathrm{GL}_p \times \mathrm{GL}_q$ on $V = \mathrm{Mat}_{p,q}$ given by the formula $(P, Q) \cdot Y = PYQ^{-1}$. The Gaussian group model $\mathcal{M}_G$ equals $\mathcal{M}(p, q)$. In this case, we can take G_{SL} to be the subgroup $\mathrm{SL}_p \times \mathrm{SL}_q$. This puts us squarely in the setup of semi-invariants for Kronecker quivers, which we will discuss in detail in later sections.

Theorem 2.7 ([2]). *Let $K = \mathbb{R}$ or $\mathbb{C}$. Let $Y \in \mathrm{Mat}_{p,q}^m$ be an m-tuple of matrices. Consider the left-right action of $G_{\mathrm{SL}} = \mathrm{SL}_p \times \mathrm{SL}_q$ on $\mathrm{Mat}_{p,q}^m$. Then, w.r.t. the matrix normal model $\mathcal{M}(p, q)$,*

- *Y is G_{SL}-semistable $\iff$ l_Y is bounded from above;*
- *Y is G_{SL}-polystable $\iff$ an MLE exists;*
- *Y is G_{SL}-stable $\implies$ there is a unique MLE. Further, if $K = \mathbb{C}$, the converse also holds, i.e., there is a unique MLE $\implies$ Y is G_{SL}-stable.*

The example below is a concrete illustration of the ideas in this paper that we will use to prove our main result, i.e., Theorem 1.2.

Example 2.8. We take $m = 2$, $p = 4$, and $q = 7$. For generic $(Y_1, Y_2) \in \mathrm{Mat}_{4,7}^2$, we claim (and justify below) that there is a change of basis (on the left and right) such that both Y_1 and Y_2 are simultaneously in a block form as below (where all non-starred entries are 0):

$$\left[\begin{array}{cc|cc|ccc} * & * & & & & & \\ & & * & * & & & \\ \hline & & & & * & * & * \\ & & & & * & * & * \end{array}\right]$$

So, without loss of generality, let us assume Y_1, Y_2 are in the form above.

We write $\mathrm{diag}(d_1, \ldots, d_k)$ to represent a diagonal $k \times k$ matrix with diagonal entries $d_1, \ldots, d_k$. For $t \neq 0$, consider $\lambda(t) = \mathrm{diag}(t^7, t^7, t^{-7}, t^{-7}) \in \mathrm{SL}_4$ and $\mu(t) = \mathrm{diag}(t^6, t^6, t^6, t^6, t^{-8}, t^{-8}, t^{-8}) \in \mathrm{SL}_7$. Let $g(t) = (\lambda(t), \mu(t)) \in \mathrm{SL}_4 \times \mathrm{SL}_7$. Then, one can check that $g(t) \cdot Y_i = \lambda(t) Y_i \mu(t)^{-1} = t Y_i$ follows from the pattern of zeros. Hence $\lim_{t \to 0} g(t) \cdot Y_i = 0$. This would mean that $Y = (Y_1, Y_2)$ is not semistable because the origin is a limit point of its $\mathrm{SL}_4 \times \mathrm{SL}_7$ orbit. Since generic points in $\mathrm{Mat}_{4,7}^2$ are unstable, all points in $\mathrm{Mat}_{4,7}^2$ are unstable because the null cone is closed in the Zariski topology (see Lemma 2.12 below). In other words, for the matrix normal model $\mathcal{M}(4, 7)$, MLEs do not exist if you only have two samples, which agrees with the observations in Section 1.3 due to Drton, Kuriki and Hoff.

The fact that a generic 2-tuple of 4×7 matrices can be simultaneous block form as mentioned above is a special case of the notion of canonical decomposition which we discuss in Section 5.

This particular example concerns the 2-Kronecker quiver and the dimension vector $(4,7)$. It can be deduced from [17, Section 3] that the canonical decomposition is $(4,7) = (1,2) \oplus (1,2) \oplus (2,3)$ in the notation explained in Section 5. The ability to drive the matrices Y_i to the origin in the limit using only diagonal group elements and that too of a very particular form is an exhibition of the Hilbert–Mumford criterion, see Theorem 2.13. In fact, this style of elementary argument could be used to prove Corollary 4.5.

The model of proportional covariance matrices $\mathcal{N}(p,q)$ is also a Gaussian group model. Consider the action of the group $H = GL_p \times \mathrm{T}_q$ on $V = \mathrm{Mat}_{p,q}$ given again by $(P,Q) \cdot Y = PYQ^{-1}$, where $\mathrm{T}_q \subseteq \mathrm{GL}_q$ denotes the subgroup of diagonal $q \times q$ matrices (i.e., a q-dimensional torus). It is easy to observe that the Gaussian group model $\mathcal{M}_H$ equals $\mathcal{N}(p,q)$. Further, in this case, one can take H_{SL} to be the subgroup $\mathrm{SL}_p \times \mathrm{ST}_q$ where ST_q denotes the subgroup of diagonal $q \times q$ matrices with determinant 1. We remark here that this fits the setup of semi-invariants for star quivers (details in later sections). We record this result to use in later sections.

Proposition 2.9. *Let $K = \mathbb{R}$ or $\mathbb{C}$. Let $Y \in \mathrm{Mat}_{p,q}^m$ be an m-tuple of matrices. Consider the aforementioned action of $H_{\mathrm{SL}} = \mathrm{SL}_p \times \mathrm{ST}_q$ on $\mathrm{Mat}_{p,q}^m$. Then, w.r.t. the model $\mathcal{N}(p,q)$,*

- *Y is H_{SL}-semistable $\iff$ l_Y is bounded from above;*
- *Y is H_{SL}-polystable $\iff$ an MLE exists;*
- *Y is H_{SL}-stable $\implies$ there is a unique MLE. Further, if $K = \mathbb{C}$, the converse also holds, i.e., there is a unique MLE $\implies$ Y is H_{SL}-stable.*

2.2. **Invariant theory over $\mathbb{C}$.** For this section, we take $K = \mathbb{C}$ and discuss a few notions in invariant theory. An algebraic group G (over $\mathbb{C}$) is called a reductive group if every rational representation is completely reducible, i.e., it can be decomposed into a direct sum of irreducible representations. There are other equivalent definitions of reductive groups over $\mathbb{C}$. For rational actions of reductive groups, Hilbert [29, 30] showed that the invariant ring is finitely generated.

Remark 2.10. The groups $\mathrm{GL}_n = \mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n = \mathrm{SL}_n(\mathbb{C})$, and finite groups are all reductive groups (over $\mathbb{C}$). Direct products of reductive groups are reductive, in particular $G = \mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_d}$ is reductive. For any $\sigma = (\sigma_1, \ldots, \sigma_d) \in \mathbb{Z}^d$, the subgroup $G_\sigma = \{(g_1, \ldots, g_d) \in G \mid \prod_{i=1}^d \det(g_i)^{\sigma_i} = 1\} \subseteq G$ is also a reductive group. All groups that we consider in this paper fall into the list of aforementioned examples. With reference to Theorem 2.5, note that any complex Zariski closed subgroup of GL_n that is closed under adjoints is a complex reductive group. On the other hand, if you have a Zariski closed subgroup of $\mathrm{GL}(V)$ that is reductive, then there is a choice of inner product on V such that the group is closed under adjoints, see [41]. Thus, reductivity can be thought of as a coordinate free version of a Zariski-closed subgroup of GL_n being closed under adjoints.

The first point to note about orbit closures is that since invariant polynomials are continuous (w.r.t. either Zariski or Euclidean topology), any invariant polynomial will be constant not just along orbits, but their closures as well. Hence, any invariant polynomial will not be able to distinguish two points $v, w \in V$ if their orbit closures intersect. The converse is also true for rational actions of complex reductive groups (see e.g., [31, Lemma 3.8] for a proof). We say a subset $W \subseteq V$ is G-invariant if $gW = W$ for all $g \in G$.

Theorem 2.11 (Mumford)**.** *Suppose G is a (complex) reductive group with a rational action on V. Then for $v, w \in V$,*

$$\overline{O_v} \cap \overline{O_w} \neq \emptyset \iff f(v) = f(w) \ \forall f \in \mathbb{C}[V]^G.$$

In fact a more general statement is true – if W_1, W_2 are G-invariant Zariski-closed subsets with an empty intersection, then there exists $f \in \mathbb{C}[V]^G$ such that $f(W_1) = 0$ and $f(W_2) = 1$.

While the definition of the null cone (see Definition 2.4) is analytic in nature, it happens to be an algebraic variety when we consider rational actions of reductive groups. For a collection of polynomials $\{f_i\} \subseteq \mathbb{C}[V]$, we denote by $\mathbb{V}(\{f_i\})$ the common zero locus of all the f_i's.

Lemma 2.12. *Let V be a rational representation of a (complex) reductive group G. Then*

$$\mathcal{N}_G(V) = \mathbb{V}\left(\bigcup_{d\geq 1} \mathbb{C}[V]_d^G\right).$$

Proof. This follows immediately from Theorem 2.11. □

An important result in understanding the null cone is the Hilbert–Mumford criterion which says that you can detect whether a point $v \in V$ is in the null cone using 1-parameter subgroups of G. A 1-parameter subgroup of G is simply a morphism of algebraic groups $\lambda : \mathbb{C}^* \to G$.

Theorem 2.13 (Hilbert–Mumford criterion)**.** *Let G be a (complex) reductive group with a rational action on V. Then $v \in \mathcal{N}_G(V)$ if and only if there is a 1-parameter subgroup $\lambda : \mathbb{C}^* \to G$ such that $\lim_{t\to 0} \lambda(t) \cdot v = 0$.*

Definition 2.14. Let V be a rational representation of a reductive group G. Then, we define three subsets

$$V^{ss} := \{v \in V \mid v \text{ is } G\text{-semistable}\},$$
$$V^{ps} := \{v \in V \mid v \text{ is } G\text{-polystable}\},$$
$$V^{st} := \{v \in V \mid v \text{ is } G\text{-stable}\}.$$

We call V^{ss} (resp. V^{ps}, V^{st}) the semistable (resp. polystable, stable) locus. We will write $V^{G\text{-}ss}, V^{G\text{-}ps}, V^{G\text{-}st}$ if we need to clarify the group.

Since the semistable locus is precisely the complement of the null cone, the following is immediate from Lemma 2.12:

Corollary 2.15. *Let V be a rational representation of a complex reductive group G. Then, the semistable locus V^{ss} is Zariski-open (but may be empty).*

Similar statements are true for the polystable and stable loci.

Lemma 2.16. *Let V be a rational representation of a complex reductive group G. Then, the stable locus V^{st} is Zariski open and the polystable locus V^{ps} is Zariski-constructible, i.e., it is a union of Zariski locally closed subsets.*

Proof. For a non-negative integer r, define $\mathcal{Z}_r := \{v \in V \mid \dim O_v \leq r\}$. It is easy to show that $\mathcal{Z}_r$ is Zariski closed for all r. Let $k = \dim G/\Delta$ (where Δ is the kernel of the action), then clearly $\mathcal{Z}_k = V$ since no orbit can have dimension larger than k.

If V^{st} is empty, then it is Zariski open. If V^{st} is non-empty, there is some $w \in V^{st}$. Hence G_w/Δ is finite, which means that $\dim O_w = k$ and O_w is Zariski-closed (see Remark 2.1). Hence O_w and $\mathcal{Z}_{k-1}$ are G-invariant Zariski-closed subsets with empty intersection, so by Theorem 2.11, we have $f \in \mathbb{C}[V]^G$ such that $f(O_w) = 1$ and $f(\mathcal{Z}_{k-1}) = 0$. Consider $V_f := \{v \in V | f(v) \neq 0\}$. Then, clearly $V_f \cap \mathcal{Z}_{k-1} = \emptyset$ because $f(\mathcal{Z}_{k-1}) = 0$. Take $v \in V_f$, we have $\dim O_v = k$ since $v \notin \mathcal{Z}_{k-1}$. Further, we claim O_v is closed. If not, take $v_1 \in \overline{O}_v \setminus O_v$. Then, we must have $\dim O_{v_1} < \dim O_v = k$, so $v_1 \in \mathcal{Z}_{k-1}$. However, since $v_1 \in \overline{O}_v$, we have $f(v_1) = f(v) \neq 0$, which means that $v_1 \in V_f$, which is absurd since $V_f \cap \mathcal{Z}_{k-1}$ is empty. Hence O_v is closed. In other words, $V_f \subseteq V^{st}$. To summarize, we have shown that for each $w \in V^{st}$, there is a Zariski-open subset V_f such that $w \in V_f \subseteq V^{st}$, which means that V^{st} is Zariski-open.

Let $\mathcal{C}_r := \{v \in \mathcal{Z}_r \setminus \mathcal{Z}_{r-1} \mid v \text{ is polystable}\}$. The above argument shows that $\mathcal{C}_k = V^{st}$ is Zariski-open in $\mathcal{Z}_k = V$. A similar argument shows that $\mathcal{C}_r$ is Zariski-open in $\mathcal{Z}_r$. In particular, this means that $\mathcal{C}_r$ is a Zariski locally closed subset of V. Thus $V^{ps} = \cup_{i=1}^{k} \mathcal{C}_i$ is a union of Zariski locally closed subsets, i.e., it is Zariski-constructible. □

For an algebraic group G, we denote its connected component of the identity by G^0, which is an algebraic subgroup. The following result tells us that as far as stability notions are concerned, one might as well restrict themselves to the connected component of identity.

Lemma 2.17. *Let V be a rational representation of a complex reductive group G. Let $G^0 \subseteq G$ denote its connected component of identity. Then for $v \in V$, v is G-semistable/stable/polystable if and only if it is G^0-semistable/stable/polystable.*

Proof. For semistability, observe that the Hilbert-Mumford criterion is the same whether you use G or G^0. For polystability, use the fact that the G-orbit is a finite disjoint union of G^0-orbits, each of which forms a connected component of the G-orbit. For stability, using the orbit-stabilizer theorem (i.e., dimension of stabilizer + dimension of orbit = dimension of the group), we see that $\dim(G_v) = \dim((G^0)_v)$ since the G-orbit and G^0-orbit have the same dimension. Let $\rho : G \to \mathrm{GL}(V)$ be the map that defines the action of G on V. Let Δ denote the kernel of $\rho : G \to \mathrm{GL}(V)$. Then $\Delta \cap G^0$ is the kernel of ρ restricted to G^0. Both Δ and $\Delta \cap G^0$ have the same Lie algebra, so they have the same dimension. Thus $\dim(G_v/\Delta) = 0$ if and only if $\dim((G^0)_v/\Delta \cap G^0) = 0$. Thus v is G-stable if and only if it is G^0-stable. □

2.3. **Invariant theory over $\mathbb{R}$.** Even though invariant theory is nicest when $K = \mathbb{C}$, the case of $K = \mathbb{R}$ is perhaps more important in the context of MLE and statistics in general. Hence, in this subsection, we collect some results on invariant theory over the real numbers which will help us "transfer" results from $\mathbb{C}$ to $\mathbb{R}$. We also intend that this subsection serve as a general reference for the reader who is not familiar with the intricacies of invariant theory and algebraic groups over $\mathbb{R}$. The following definitions are from [3].

For a complex (affine) variety X, we denote its coordinate ring by $\mathbb{C}[X]$. A complex (affine) variety X with an $\mathbb{R}$-structure (i.e., an $\mathbb{R}$-subalgebra $\mathbb{R}[X] \subseteq \mathbb{C}[X]$ such that $\mathbb{R}[X] \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{C}[X]$) is called an (affine) $\mathbb{R}$-variety. As a technical point, we identify a variety X with its complex points $X_{\mathbb{C}}$ (which can be viewed as algebra morphisms $\mathbb{C}[X] \to \mathbb{C}$). A morphism $f : X \to Y$ of (affine) varieties is equivalent to a map on the coordinate rings $f^* : \mathbb{C}[Y] \to \mathbb{C}[X]$. The morphism f is said to be defined over $\mathbb{R}$ if $f^*(\mathbb{R}[Y]) \subseteq \mathbb{R}[X]$. The real points $X_{\mathbb{R}}$ as the points in $X_{\mathbb{C}}$ (i.e., the algebra morphisms $\mathbb{C}[X] \to \mathbb{C}$) that are defined over $\mathbb{R}$. A complex algebraic group G is called an $\mathbb{R}$-group if it is an (affine) $\mathbb{R}$-variety such that the multiplication map and inverse map are defined over $\mathbb{R}$, and its real points $G_{\mathbb{R}}$ is an algebraic group over $\mathbb{R}$.

Remark 2.18. All varieties in this paper will be affine, so we will henceforth drop the prefix affine.

For this entire section, let G be a connected reductive $\mathbb{R}$-group. Let V be a rational representation of G that is defined over $\mathbb{R}$ – this means that V is an $\mathbb{R}$-variety, and that $\rho : G \to \mathrm{GL}(V)$ is defined over $\mathbb{R}$. In particular, we get an action of $G_{\mathbb{R}}$ on $V_{\mathbb{R}}$.

Remark 2.19. All the groups we consider in this paper will be connected reductive $\mathbb{R}$-groups, and all representations will be defined over $\mathbb{R}$. Groups such as GL_n and SL_n are connected reductive, and are naturally $\mathbb{R}$-groups, and hence so are their direct products. The group G_σ defined in Remark 2.10 is also naturally an $\mathbb{R}$-group, and connected if σ is indivisible, i.e., $\gcd(\sigma_1, \ldots, \sigma_d) = 1$.

The following proposition is a fundamental result:

Proposition 2.20. *Let G be a connected reductive $\mathbb{R}$-group. Let V be a rational representation of G that is defined over $\mathbb{R}$. Then, the invariant ring for the action of G on V is obtained by base change from the invariant ring for the action of $G_{\mathbb{R}}$ on $V_{\mathbb{R}}$, i.e.,*

$$\mathbb{R}[V]^{G_{\mathbb{R}}} \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{C}[V]^{G}.$$

In other words, $\operatorname{Spec}(\mathbb{C}[V]^G)$ *(also called the categorical quotient) is naturally an $\mathbb{R}$-variety.*

Proof. To begin with, first note that by definition $\mathbb{R}[V]$ is the same as $\mathbb{R}[V_{\mathbb{R}}]$, the coordinate ring of $V_{\mathbb{R}}$. The action $G \times V \to V$ gives a map $\mu_{\mathbb{C}} : \mathbb{C}[V] \to \mathbb{C}[G] \otimes \mathbb{C}[V]$. Similarly, we get a map $\mu_{\mathbb{R}} : \mathbb{R}[V] \to \mathbb{R}[G] \otimes \mathbb{R}[V]$, and note that $\mu_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} = \mu_{\mathbb{C}}$ is simply a restatement that the action is defined over $\mathbb{R}$.

It is easy to see that the polynomial $f \in \mathbb{C}[V]$ is G-invariant if and only if $\mu_{\mathbb{C}}(f) = 1 \otimes f$. Write $f = f_1 + if_2$ with $f_j \in \mathbb{R}[V]$. Then, we see that $\mu_{\mathbb{C}}(f) = \mu_{\mathbb{R}}(f_1) + i\mu_{\mathbb{R}}(f_2)$. Thus $\mu_{\mathbb{R}}(f_1)$ and $\mu_{\mathbb{R}}(f_2)$ are the real and imaginary parts of $\mu_{\mathbb{C}}(f)$ which are $1 \otimes f_1$ and $1 \otimes f_2$ respectively. So, for each j, we have $\mu_{\mathbb{R}}(f_j) = 1 \otimes f_j$, which means that f_j is $G_{\mathbb{R}}$-invariant. This proves $\supseteq$.

To prove $\subseteq$, it suffices to prove that $\mathbb{R}[V]^{G_{\mathbb{R}}} = \mathbb{R}[V_{\mathbb{R}}]^{G_{\mathbb{R}}} \subseteq \mathbb{C}[V]^G$. Indeed, observe that the action of G on V gives an action of G on $\mathbb{C}[V]$. A function $f \in \mathbb{C}[V]$ is G-invariant if and only if $g \cdot f = f$ for all $g \in G$ (easy to see and only uses that $\mathbb{C}$ is an infinite field). Similarly, we have an action of $G_{\mathbb{R}}$ on $\mathbb{R}[V_{\mathbb{R}}]$ and $f \in \mathbb{R}[V_{\mathbb{R}}]^{G_{\mathbb{R}}}$ if and only if $g \cdot f = f$ for all $g \in G_{\mathbb{R}}$. Now, suppose $f \in \mathbb{R}[V_{\mathbb{R}}]^{G_{\mathbb{R}}} \subseteq \mathbb{C}[V]$. Then, since $g \cdot f = f$ for all $g \in G_{\mathbb{R}}$ and $G_{\mathbb{R}}$ is Zariski-dense in G (see e.g., [48] or [3, Corollary 18.3]), we get that $g \cdot f = f$ for all $g \in G$, so $f \in \mathbb{C}[V]^G$. Thus $\mathbb{R}[V_{\mathbb{R}}]^{G_{\mathbb{R}}} \subseteq \mathbb{C}[V]^G$ as required. □

The following result is crucial for transferring our results for complex matrix normal models to real matrix normal models. We will need some results from Lie algebras. The theory of Lie algebras is well understood and we do not intend to recall the theory here. We refer the interested reader to standard references, e.g., [3, 4, 5, 53].

Proposition 2.21. *Let G be a connected reductive $\mathbb{R}$-group. Let V be a rational representation of G that is defined over $\mathbb{R}$. Let $v \in V_{\mathbb{R}}$. Then v is semistable/polystable/stable for the $G_{\mathbb{R}}$-action if and only if v is semistable/polystable/stable for the G-action.*

Proof. Let us split the argument for each notion of stability.

- **Semistability:** Suppose v is G-semistable. Then, there is a homogeneous polynomial invariant $f \in \mathbb{C}[V]^G$ such that $f(v) \neq 0$. Write $f = f_1 + if_2$ where $f_j \in \mathbb{R}[V]$ and hence in $\mathbb{R}[V]^G$ by the above proposition. Thus, $f_j(v) \neq 0$ for some j. By homogeneity, $f_j(0) = 0$. Since f_i is $G_{\mathbb{R}}$-invariant, the $G_{\mathbb{R}}$-orbit closure of v cannot contain the origin. For the converse, suppose $v \in V_{\mathbb{R}}$ is G-unstable. Then, by Theorem 2.13 there is a 1-parameter subgroup that drives v to 0 in the limit. By a result of Birkes [7, Theorem 5.2], you can choose a 1-parameter subgroup defined over $\mathbb{R}$ that drives v to 0 in the limit. Hence, v is $G_{\mathbb{R}}$-unstable.
- **Polystability:** Suppose the G-orbit of v is closed in the Euclidean topology (and hence in the Zariski topology, see Remark 2.1). Borel and Harish-Chandra [5, Proposition 2.3] show that this implies that the $G_{\mathbb{R}}$-orbit of v is closed in the Euclidean topology. Birkes showed that if the $G_{\mathbb{R}}$-orbit of v is closed in the Euclidean topology, then the G-orbit is closed in the Zariski topology (and hence the Euclidean topology), see [7, Corollary 5.3].
- **Stability:** Stability is polystability along with the fact that the stabilizer (modulo the kernel) is finite. We already know from above that v is G-polystable if and only if v is $G_{\mathbb{R}}$-polystable. So, we only have to analyze the stabilizers. We will utilize heavily the fact that the dimensions of Lie algebras reflect the dimensions of the groups themselves (in both the real and complex settings). Let Δ denote the kernel of the representation $\rho : G \to \mathrm{GL}(V)$.

Since $G_v \supseteq \Delta$, we see that G_v/Δ is finite if and only if we have an equality of Lie algebras $\mathrm{Lie}(G_v) = \mathrm{Lie}(\Delta)$. Since G_v (resp. Δ) is defined over $\mathbb{R}$ (see [53, Proposition 12.1.2, Corollary 12.1.3]), we get that $(G_v)_{\mathbb{R}}$ (resp. $\Delta_{\mathbb{R}}$) is a real manifold whose dimension equals the complex dimension of G_v (resp. Δ), see [4, Section 5.2]. In fact, more is true, the Lie algebra of $\mathrm{Lie}((G_v)_{\mathbb{R}})$ (resp. $\mathrm{Lie}(\Delta_{\mathbb{R}})$) is a real form of $\mathrm{Lie}(G_v)$ (resp. $\mathrm{Lie}(\Delta)$), see [4, Section 5.3]. In other words, we have $\mathrm{Lie}((G_v)_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C} = \mathrm{Lie}(G_v)$ and $\mathrm{Lie}(\Delta_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C} = \mathrm{Lie}(\Delta)$.

Since, we have inclusions $\Delta \subseteq G_v$ (and consequently $\Delta_{\mathbb{R}} \subseteq (G_v)_{\mathbb{R}}$), we get that $\mathrm{Lie}(G_v) = \mathrm{Lie}(\Delta)$ if and only if $\mathrm{Lie}((G_v)_{\mathbb{R}}) = \mathrm{Lie}(\Delta_{\mathbb{R}})$. It is perhaps trivial, but nonetheless necessary to observe that $(G_v)_{\mathbb{R}} = (G_{\mathbb{R}})_v$ and $\Delta_{\mathbb{R}}$ is the kernel of $\rho_{\mathbb{R}}$. Thus $\mathrm{Lie}((G_v)_{\mathbb{R}}) = \mathrm{Lie}(\Delta_{\mathbb{R}})$ if and only if both $(G_{\mathbb{R}})_v$ and $\Delta_{\mathbb{R}}$ have the same dimension if and only if $(G_{\mathbb{R}})_v/\Delta_{\mathbb{R}}$ is finite (since $(G_{\mathbb{R}})_v \supseteq \Delta_{\mathbb{R}}$, and real algebraic groups have finitely many components).

To summarize, we have that G_v/Δ is finite if and only if $(G_{\mathbb{R}})_v/\Delta_{\mathbb{R}}$ is finite, so v is G-stable if and only if it is $G_{\mathbb{R}}$-stable.

□

We write $(V^{G\text{-}ss})_{\mathbb{R}}$ to denote the real points of the set of G-semistable points of V, i.e., $(V^{G\text{-}ss})_{\mathbb{R}} = V^{G\text{-}ss} \cap V_{\mathbb{R}}$. We write $(V_{\mathbb{R}})^{G_{\mathbb{R}}\text{-}ss}$ to denote the $G_{\mathbb{R}}$-semistable points of $V_{\mathbb{R}}$. We will use similar notation for polystable and stable loci as well. The following is immediate from Proposition 2.21.

Corollary 2.22. *We have* $(V^{G\text{-}ss})_{\mathbb{R}} = (V_{\mathbb{R}})^{G_{\mathbb{R}}\text{-}ss}$, $(V^{G\text{-}ps})_{\mathbb{R}} = (V_{\mathbb{R}})^{G_{\mathbb{R}}\text{-}ps}$, *and* $(V^{G\text{-}st})_{\mathbb{R}} = (V_{\mathbb{R}})^{G_{\mathbb{R}}\text{-}st}$.

Recall the notions of generic G-semistability/polystability/stability from Definition 2.3.

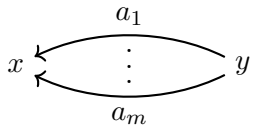
Proposition 2.23. *Let G be a connected reductive $\mathbb{R}$-group. Let V be a rational representation of G that is defined over $\mathbb{R}$. Then V is generically G-semistable (resp. G-polystable, G-stable) if and only if $V_{\mathbb{R}}$ is generically $G_{\mathbb{R}}$-semistable (resp. $G_{\mathbb{R}}$-polystable, $G_{\mathbb{R}}$-stable).*

Proof. Let $X \subseteq V$ be a Zariski-constructible subset. Then it is easy to see that X contains a dense Zariski-open subset of V if and only if $X_{\mathbb{R}} = X \cap V_{\mathbb{R}}$ contains a dense Zariski-open subset of $V_{\mathbb{R}}$. Now, the proposition follows from the fact that V^{ss}, V^{ps} and V^{st} are Zariski-constructible (by Corollary 2.15 and Lemma 2.16) along with Corollary 2.22. □

3. Quiver representations

The theory of quivers and their representations forms a rich generalization of linear algebra. Numerous applications of quivers have been discovered in various algebraic subjects ranging from cluster algebras and cluster categories [19, 38], Schubert calculus [15, 16, 47], moduli spaces, Donaldson-Thomas invariants and cohomological Hall algebras and non-commutative algebraic geometry (see [46] and references therein) and symplectic resolutions (see [27] and references therein) to name a few. More recently, the invariant theory of quivers has played an influential role in areas of theoretical computer science, notably to Geometric Complexity Theory and non-commutative identity testing [12, 13, 33, 34, 42], Brascamp–Lieb inequalities [26] and simultaneous robust subspace recovery [10].

Let K denote the ground field. The reader should keep in mind $K = \mathbb{R}$ or $\mathbb{C}$. A quiver Q is a directed acyclic graph, i.e. a set of vertices denoted Q_0 and a set of arrows Q_1. For each arrow $a \in Q_1$, we denote by ta and ha, the tail vertex and head vertex of the arrow. We will demonstrate all the basic notions and definitions in the crucial example (below) of the m-Kronecker quiver $\Theta(m)$ with two vertices x and y with m arrows $a_1, \ldots, a_m$ from y to x.



A representation V of Q is simply an assignment of a finite-dimensional vector space (over a ground field K) $V(x)$ for each $x \in Q_0$ and a linear transformation $V(a) : V(ta) \to V(ha)$ for each arrow $a \in Q_1$. A morphism of quiver representations $\phi : V \to W$ is a collection of linear maps $\phi(x) : V(x) \to W(x)$ for each $x \in Q_0$ subject to the condition that for every $a \in Q_1$, the diagram below commutes.

$$\begin{array}{ccc} V(ta) & \xrightarrow{V(a)} & V(ha) \\ \downarrow{\scriptstyle \phi(ta)} & & \downarrow{\scriptstyle \phi_{ha}} \\ W(ta) & \xrightarrow{W(a)} & W(ha) \end{array}$$

A representation V of $\Theta(m)$ is given by assigning vector spaces $V(x)$ and $V(y)$ to x and y, and m linear maps $V(a_1), \ldots, V(a_m)$ from $V(y)$ to $V(x)$. A morphism between two representations V and W of $\Theta(m)$ is two linear maps $\phi(x) : V(x) \to W(x)$ and $\phi(y) : V(y) \to W(y)$ such that $\phi(x) \circ V(a_i) = W(a_i) \circ \phi(y)$ for all $1 \leq i \leq m$.

A subrepresentation U of V is a collection of subspaces $U(x) \subseteq V(x)$ such that for every edge the linear map $U(a)$ is simply a restriction of $V(a)$. In particular, this means that the image of $U(ta)$ under $V(a)$ will need to be contained in $U(ha)$. For two representations V and W, we define their direct sum $V \oplus W$ to be the representation that assigns $V(x) \oplus W(x)$ to each vertex x and the linear map $\begin{pmatrix} V(a) & 0 \\ 0 & W(a) \end{pmatrix}$ for each arrow $a \in Q_1$. Similarly, the notion of direct summand, image, kernel, co-image, etc are all defined in the obvious way, see [18] for details. In summary, the category of quiver representations forms an abelian category.

The dimension vector of a representation V is $\underline{\dim}(V) = (\dim V(x))_{x \in Q_0}$. So, for a representation V of $\Theta(m)$, its dimension vector is $\underline{\dim}(V) = (\dim(V(x)), \dim(V(y)))$. For any representation V of a quiver Q, for each $x \in Q_0$, picking a basis for $V(x)$ identifies $V(x)$ with $K^{\dim(V(x))}$. Further, with this identification, every linear map $V(a)$ is just a matrix of size $\dim(V(ha)) \times \dim(V(ta))$. Thus, we come to the following definition. For any dimension vector $\alpha = (\alpha(x))_{x \in Q_0} \in \mathbb{N}^{Q_0}$ (where $\mathbb{N} = \{0, 1, 2, \ldots, \}$), we define the representation space

$$\mathrm{Rep}(Q, \alpha) = \bigoplus_{a \in Q_1} \mathrm{Mat}_{\alpha(ha), \alpha(ta)} .$$

Any point $V = (V(a))_{a \in Q_1} \in \mathrm{Rep}(Q, \alpha)$ can be interpreted as a representation of Q with dimension vector as follows: for each $x \in Q_0$, assign the vector space $K^{\alpha(x)}$, and for each arrow $a \in Q_1$, the matrix $V(a)$ describes a linear transformation from $K^{\alpha(ta)}$ to $K^{\alpha(ha)}$. The base change group $\mathrm{GL}(\alpha) = \prod_{x \in Q_0} \mathrm{GL}_{\alpha(x)}$ acts on $\mathrm{Rep}(Q, \alpha)$ in a natural fashion where $\mathrm{GL}_{\alpha(x)}$ acts on the vector space $K^{\alpha(x)}$ assigned to vertex x. More concretely, for $g = (g_x)_{x \in Q_0} \in \mathrm{GL}(\alpha)$ and $V = (V(a))_{a \in Q_1} \in \mathrm{Rep}(Q, \alpha)$, the point $g \cdot V \in \mathrm{Rep}(Q, \alpha)$ is defined by the formula

$$(g \cdot V)(a) = g_{ha} V(a) g_{ta}^{-1}.$$

The $\mathrm{GL}(\alpha)$ orbits in $\mathrm{Rep}(Q, \alpha)$ are in $1-1$ correspondence with isomorphism classes of α-dimensional representations.

Consider the subgroup $\mathrm{SL}(\alpha) = \prod_{x \in Q_0} \mathrm{SL}(\alpha(x)) \subseteq \mathrm{GL}(\alpha)$. Then, the invariant ring for the action of $\mathrm{SL}(\alpha)$ on $\mathrm{Rep}(Q, \alpha)$ is called the ring of semi-invariants

$$\mathrm{SI}(Q, \alpha) = K[\mathrm{Rep}(Q, \alpha)]^{\mathrm{SL}(\alpha)}.$$

For the m-Kronecker quiver $\Theta(m)$, suppose we pick a dimension vector $\alpha = (p, q)$ (we use the convention that the first entry corresponds to vertex x), then the representation space

$$\mathrm{Rep}(\Theta(m), (p, q)) = \mathrm{Mat}_{p,q}^m .$$

Now, $\mathrm{GL}(\alpha) = \mathrm{GL}_p \times \mathrm{GL}_q$, and the action is given by the formula

$$(g_1, g_2) \cdot (Y_1, \ldots, Y_m) = (g_1 Y_1 g_2^{-1}, \ldots, g_1 Y_m g_2^{-1}).$$

The orbits of this action correspond to isomorphism classes of (p, q)-dimensional representations of $\Theta(m)$. The subgroup $\mathrm{SL}(\alpha)$ is $\mathrm{SL}_p \times \mathrm{SL}_q$. First, observe that $Y = (Y_1, \ldots, Y_m)$ is semistable/polystable/stable (for the action of $\mathrm{SL}_p \times \mathrm{SL}_q$) if and only if $\lambda Y = (\lambda Y_1, \ldots, \lambda Y_m)$ is semistable/polystable/stable for $\lambda \in K^*$. This is a simple consequence of the fact that the action is by linear transformations. Thus, we see that whether $Y = (Y_1, \ldots, Y_m)$ is semistable, polystable, or stable (for the action of $\mathrm{SL}(\alpha)$) only depends on the isomorphism class of the quiver representation it defines (i.e., the $\mathrm{GL}(\alpha)$-orbit). This is the starting point of understanding the various stability notions from a representation theoretic perspective, which we will discuss in more detail in the next section.

Remark 3.1. The space $\mathrm{Rep}(Q, \alpha)$ is a representation of $\mathrm{GL}(\alpha)$ and its various subgroups such as $\mathrm{SL}(\alpha)$. At the same time, we refer to a point $V \in \mathrm{Rep}(Q, \alpha)$ also as a representation. We advise the reader to keep in mind that we think of V as a representation of the quiver Q to avoid confusion. Moreover, if $V \in \mathrm{Rep}(Q, \alpha)_{\mathbb{R}} = \bigoplus_{a \in Q_1} \mathrm{Mat}_{\alpha(ha), \alpha(ta)}(\mathbb{R})$, then it can be thought of as both a real and complex representation of Q.

3.1. **Indecomposability of modules over field extensions.** Let A be a finite-dimensional $\mathbb{R}$-algebra, and let A-mod denote the category of finite dimensional (left)-modules over A. We denote by $A_{\mathbb{C}} := A \otimes_{\mathbb{R}} \mathbb{C}$, the $\mathbb{C}$-algebra obtained by extending scalars. Let $A_{\mathbb{C}}$-mod denote the category of finite dimensional (left)-modules over $A_{\mathbb{C}}$. For any module $M \in A$-mod, let $M_{\mathbb{C}} := M \otimes_{\mathbb{R}} \mathbb{C} \in A_{\mathbb{C}}$-mod. The $A_{\mathbb{C}}$ structure on $M_{\mathbb{C}}$ is the obvious one you get by extending scalars. We can interpret $M_{\mathbb{C}}$ as an A-module: $M_{\mathbb{C}} = M \oplus iM$, and hence we have $M_{\mathbb{C}} \cong M^{\oplus 2}$ as A-modules.

The Krull-Remak-Schmidt Theorem says that modules in the category A-mod can be decomposed as a direct sum of indecomposables and this decomposition is essentially unique in the sense that any two such decompositions have the same summands (counted with multiplicities). The Krull-Remak-Schmidt theorem holds for $A_{\mathbb{C}}$-mod as well. Note that an indecomposable module is one that cannot be written as a direct sum of two or more (proper) submodules, and not to be confused with an irreducible (or simple) module, which is a module with no non-trivial submodules.

Lemma 3.2. *Let $M \in A$-mod. If $M_{\mathbb{C}} \in A_{\mathbb{C}}$-mod is indecomposable, then M is indecomposable.*

Proof. If $M = M_1 \oplus M_2$, then $M_{\mathbb{C}} = (M_1)_{\mathbb{C}} \oplus (M_2)_{\mathbb{C}}$ is a decomposition of $M_{\mathbb{C}}$ as $A_{\mathbb{C}}$-modules. □

We will need the following lemma.

Lemma 3.3. *Let $M \in A$-mod. Suppose $M_{\mathbb{C}}$ can be written as a direct sum of three or more (non-zero) submodules (as an $A_{\mathbb{C}}$-module). Then M is not indecomposable as an A-module.*

Proof. Let $M_{\mathbb{C}} = N_1 \oplus N_2 \oplus \cdots \oplus N_d$ with $d \geq 3$. Each N_i is an $A_{\mathbb{C}}$-module summand, and hence an A-module summand as well. If we further refine the N_i into a direct sum of indecomposable A-modules, we can write $M_{\mathbb{C}} = N_1' \oplus N_2' \oplus \ldots N_{d'}'$ for some $d' \geq d \geq 3$, where each N_i' is an indecomposable A-module. Suppose M is indecomposable. The module $M_{\mathbb{C}} = M \oplus iM$ as an A-module. Hence, by the Krull-Remak-Schmidt theorem, we know that any decomposition into indecomposables has to have exactly two summands (and each of which is isomorphic to M as A-modules). But this contradicts the fact that $M_{\mathbb{C}} = N_1' \oplus N_2' \oplus \ldots N_{d'}'$ is a decomposition with $d' > 2$ summands. □

3.2. **Quiver representations as modules over the path algebra.** Let K denote the ground field. For a quiver $Q = (Q_0, Q_1)$, we will define the path algebra KQ. A path p of length k is a sequence of k arrows $a_k a_{k-1} \dots a_1$ such that $t(a_{i+1}) = h(a_i)$ for $1 \leq i < k$. The head vertex of the path is $h(a_k)$ and the tail vertex is $t(a_1)$. We introduce trivial paths e_x of length zero for each $x \in Q_0$ with $h(e_x) = t(e_x) = x$.

The path algebra KQ is a K-algebra with a basis labeled by all paths in Q. The multiplication is as follows. For paths p and q, their product $p \cdot q$ is the concatenation of the two paths if $tp = hq$ and 0 otherwise. For any representation V of Q over K, we can interpret it as KQ-module $\oplus_{x \in Q_0} V(x)$. For $w \in V(x)$, and a path $p = a_k a_{k-1} a_{k-2} \dots a_1$, the action is given by $p \cdot w = V(a_k)V(a_{k-1}) \dots V(a_1)w \in V_{hp}$ if $tp = x$, and 0 otherwise. This is in fact an equivalence of categories, see [18] for details.

Of particular importance is the fact that if we take $A = \mathbb{R}Q$, then $A_{\mathbb{C}} = \mathbb{C}Q$, and hence Lemma 3.3 applies.

4. Stability notions for quiver representations

We follow the conventions from [18] for consistency. For this section, we let $K = \mathbb{C}$. Let Q be a quiver with no oriented cycles (self loops are counted as oriented cycles). Let α be a dimension vector. For any $\sigma \in \mathbb{Z}^{Q_0}$ (which we call a weight), we have a character of $\mathrm{GL}(\alpha)$ which we also denote σ by abuse of notation. The character $\sigma : \mathrm{GL}(\alpha) \to K^*$ is given by $\sigma((g_x)_{x \in Q_0}) = \prod_{x \in Q_0} \det(g_x)^{\sigma(x)}$. The ring of semi-invariants has a decomposition

$$\mathrm{SI}(Q, \alpha) = \bigoplus_{\sigma \in \mathbb{Z}^{Q_0}} \mathrm{SI}(Q, \alpha)_\sigma,$$

where $\mathrm{SI}(Q, \alpha)_\sigma = \{f \in \mathrm{SI}(Q, \alpha) \mid f(g \cdot x) = \sigma(g^{-1}) f(x) \ \forall g \in \mathrm{GL}(\alpha)\}$.

We define the effective cone of weights

$$C(Q, \alpha) := \{\sigma \in \mathbb{Z}^{Q_0} \mid \mathrm{SI}(Q, \alpha)_{m\sigma} \neq 0 \text{ for some } m \in \mathbb{Z}_{>0}\}.$$

For a weight σ and a dimension vector β, we define $\sigma(\beta) := \sum_{x \in Q_0} \sigma(x)\beta(x)$. We point out there that every $\sigma \in C(Q, \alpha)$ must satisfy $\sigma(\alpha) = 0$. For each $0 \neq \sigma \in C(Q, \alpha)$ that is indivisible (i.e., $\gcd(\sigma(x) : x \in Q_0) = 1$), we consider the subring

$$\mathrm{SI}(Q, \alpha, \sigma) := \oplus_{m=0}^{\infty} \mathrm{SI}(Q, \alpha)_{m\sigma}.$$

For a sincere dimension vector α (i.e., $\alpha(x) \neq 0 \ \forall x \in Q_0$), it turns out that this subring can also be seen as an invariant ring, i.e., $\mathrm{SI}(Q, \alpha, \sigma) = K[\mathrm{Rep}(Q, \alpha)]^{\mathrm{GL}(\alpha)_\sigma}$ where $\mathrm{GL}(\alpha)_\sigma = \{g \in \mathrm{GL}(\alpha) \mid \sigma(g) = 1\}$. Note that $\mathrm{GL}(\alpha)_\sigma$ is a reductive group. It is well-known that the associated projective variety $\mathrm{Proj}(\mathrm{SI}(Q, \alpha, \sigma))$ defines a moduli space for the α-dimensional representations of Q, see [39].

We make a definition following King [39]. We follow the convention from [18] which is consistent with our notational choices so far, but differs from King's original convention by a sign.

Definition 4.1 (King [39])**.** Let Q be a quiver with no oriented cycles, V be a representation of Q and $\sigma \in \mathbb{Z}^{Q_0}$ a weight such that $\sigma(\underline{\dim} V) = 0$.

- V is σ-semistable if $\sigma(\beta) \leq 0$ for all $\beta \in \mathbb{Z}_{\geq 0}^{Q_0}$ such that V contains a subrepresentation of dimension β.
- V is σ-stable if $\sigma(\beta) < 0$ for all $\beta \in \mathbb{Z}_{\geq 0}^{Q_0}$ (other than 0 and $\underline{\dim}(V)$) such that V contains a subrepresentation of dimension β.
- V is σ-polystable if $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ such that V_i are all σ-stable representations.

Observe here that any σ-stable representation must be indecomposable, i.e., it cannot be written as a direct sum of (proper) subrepresentations. Indeed, suppose $V = V_1 \oplus V_2$, then $0 = \sigma(\underline{\dim} V) = \sigma(\underline{\dim} V_1) + \sigma(\underline{\dim} V_2)$. Hence at least one of $\sigma(\underline{\dim} V_i) \geq 0$, and hence V cannot be σ-stable. Also observe that if V is a direct sum $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, then V is σ-semistable (or σ-polystable) if and only if all the V_i are. Thus, in order to understand whether a generic representation of dimension α is σ-semistable/polystable/stable, it is useful to understand how it decomposes as a direct sum of indecomposables, which is the topic of discussion in the next section.

We now relate σ-stability notions to $\mathrm{GL}(\alpha)_\sigma$-stability notions:

Theorem 4.2 (King [39]). *Let Q be a quiver with no oriented cycles, $\alpha \in \mathbb{Z}_{>0}^{Q_0}$ a sincere dimension vector and $0 \neq \sigma$ an indivisible weight such that $-\sigma \notin C(Q, \alpha)$. A representation $V \in \mathrm{Rep}(Q, \alpha)$ is σ-semistable (resp. σ-polystable, σ-stable) if and only if V is $\mathrm{GL}(\alpha)_\sigma$-semistable (resp. $\mathrm{GL}(\alpha)_\sigma$-polystable, $\mathrm{GL}(\alpha)_\sigma$-stable).*

King's original formulation is slightly different from the one above, but can be seen to be equivalent (details in Appendix A). Now, we proceed to discuss these stability notions for the m-Kronecker quiver.

4.1. **Stability notions for the m-Kronecker quiver.** For the m-Kronecker quiver $\Theta(m)$, let us take $\alpha = (p, q)$. Let $p' = p/\gcd(p, q)$ and $q' = q/\gcd(p, q)$. Then

$$\mathrm{SI}(\Theta(m), (p, q)) = \bigoplus_{k=0}^{\infty} \mathrm{SI}(\Theta(m), (p, q))_{(-kq', kp')}. \tag{1}$$

Indeed, observe that for $\sigma \in \mathbb{Z}^2$ to be in the cone of effective weights $C(\Theta(m), (p, q))$, we need $\sigma(p, q) = 0$. This means that any $\sigma \in C(\Theta(m), (p, q))$ must be a multiple of $(-q', p')$. As it turns out, non-trivial semi-invariants do not exist when you take a weight that is a negative scalar multiple of $(-q', p')$, which one can see directly from the fundamental theorem that describes semi-invariants of quivers in a determinantal fashion [16, 20, 52] (see also [18, Theorem 10.7.1]) because there are no paths from x to y.

Further, this means that

$$\mathbb{C}[\mathrm{Rep}(\Theta(m), (p, q))]^{\mathrm{SL}_p \times \mathrm{SL}_q} = \mathbb{C}[\mathrm{Rep}(\Theta(m), (p, q))]^{\mathrm{SL}(\alpha)} = \mathbb{C}[\mathrm{Rep}(\Theta(m), (p, q))]^{\mathrm{GL}(\alpha)_{(-q', p')}}.$$

In fact, we have the following result:

Lemma 4.3. *Let $\rho : \mathrm{GL}_p \times \mathrm{GL}_q \to \mathrm{GL}(\mathrm{Rep}(\Theta(m), (p, q))$ be the left-right action. Let $\alpha = (p, q)$ and let $p' = p/\gcd(p, q)$ and $q' = q/\gcd(p, q)$. Then*

$$\rho(\mathrm{SL}_p \times \mathrm{SL}_q) = \rho(\mathrm{GL}(\alpha)_{(-q', p')}).$$

In particular, this means that σ-semistability (resp. polystability, stability) for $\sigma = (-q', p')$ is the same as semistability (resp. polystability, stability) for the $\mathrm{SL}_p \times \mathrm{SL}_q$-action.

Proof. Since $\mathrm{SL}_p \times \mathrm{SL}_q \subseteq \mathrm{GL}(\alpha)_{(-q', p')}$, we only need to show $\supseteq$. Suppose $(g, h) \in \mathrm{GL}(\alpha)_{(-q', p')}$. This means that $\det(g)^{q'} = \det(h)^{p'}$. Note that $\rho(g, h) = (g \otimes (h^{-1})^\top)^{\oplus m}$. Without loss of generality, we can assume that $\det(g) = 1$ (otherwise, choose a λ such that $\det(\lambda g) = 1$ and replace (g, h) with $(\lambda g, \lambda h)$). Thus, we have $\det(h)^{p'} = 1$. Thus $\det(h) = e^{2\pi i n/p'}$ for some n. Now, choose an integer t such that $t \equiv 0 \bmod q'$ and $t \equiv -n \bmod p'$. Such a t exists by the Chinese remainder theorem since p' and q' are coprime. Let $\mu = e^{2\pi i t/dp'q'}$, where $d = \gcd(p, q) = p/p' = q/q'$. Then $\det(\mu g) = \mu^p = e^{2\pi i t/q'} = 1$ and $\det(\mu h) = \mu^q \cdot e^{2\pi i n/p'} = e^{2\pi i t/p'} \cdot e^{2\pi i n/p'} = 1$. Now, observe that $\rho(\mu g, \mu h) = \rho(g, h)$ and $\mu g \in \mathrm{SL}_p$ and $\mu h \in \mathrm{SL}_q$.

Further, we observe that for $\sigma = (-q', p')$, $\mathrm{GL}(\alpha)_\sigma$-semistability (resp. polystability, stability) is equivalent to σ-semistability (resp. polystability, stability) because $-\sigma \notin C(Q, \alpha)$ follows from Equation 1 above. □

From Lemma 4.3 and Theorem 4.2, we deduce:

Corollary 4.4. *Consider the $G = \mathrm{SL}_p \times \mathrm{SL}_q$-action on $\mathrm{Rep}(\Theta(m), (p,q))$, and let $\sigma = (-q', p')$, where $p' = p/\gcd(p,q)$ and $q' = q/\gcd(p,q)$. A representation $V \in \mathrm{Rep}(\Theta(m), (p,q))$ is*

- *semistable if and only if $\sigma(\beta) \leq 0$ for all dimension vectors β such that V has a subrepresentation of dimension β.*
- *stable if and only if $\sigma(\beta) < 0$ for all dimension vectors β (other than 0 and (p,q)) such that V has a subrepresentation of dimension β.*
- *polystable if and only if V is a direct sum of σ-stable representations.*

A simple corollary of the above is the following, which will be very useful.

Corollary 4.5. *Let $V \in \mathrm{Rep}(\Theta(m), (p,q))$. Suppose $V = V_1 \oplus V_2 \cdots \oplus V_k$ is the decomposition of V into indecomposables, and let $\beta_i = \underline{\dim}(V_i)$. If for some i and j, β_i and β_j are linearly independent, then V is unstable (w.r.t. $\mathrm{SL}_p \times \mathrm{SL}_q$ action).*

Proof. Let $\sigma = (-q', p')$, where $p' = p/\gcd(p,q)$ and $q' = q/\gcd(p,q)$. Then, by Lemma 4.3, if V is to be semistable, then $\sigma(\beta_i) \leq 0$ for all i and further $\sigma(\sum_i \beta_i) = \sigma((p,q)) = 0$. This means that $\sigma(\beta_i) = 0$ for all i. However, the kernel of σ is clearly 1-dimensional, so both β_i and β_j cannot be in the kernel if they are linearly independent. □

5. Canonical decomposition

For this section, we assume $K = \mathbb{C}$. Let Q be a quiver with no oriented cycles and let α be a dimension vector. Every representation $V \in \mathrm{Rep}(Q, \alpha)$ can be decomposed into a direct sum $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ where each V_i is an indecomposable representation. The Krull-Remak-Schmidt theorem tells us that the summands that occur in any such decomposition are isomorphic (up to permutation). Of course, this decomposition will be different for different choices of $V \in \mathrm{Rep}(Q, \alpha)$, but for a (non-empty) Zariski-open subset of $\mathrm{Rep}(Q, \alpha)$, the dimension vectors of the indecomposables in the decomposition will be the same. This brings us to the definition of canonical decomposition that was first defined by Kac.

Definition 5.1 (canonical decomposition [36, 37])**.** We write $\alpha = \beta_1 \oplus \cdots \oplus \beta_k$ and call it the canonical decomposition if a generic representation $V \in \mathrm{Rep}(Q, \alpha)$ decomposes as a direct sum of indecomposables whose dimension vectors are $\beta_1, \ldots, \beta_k$.

The existence and uniqueness of canonical decomposition requires a little argument and we refer the reader to [18]. To fully understand canonical decomposition, we need to recall the notion of roots. We need to define a bilinear form on $\mathbb{R}^{Q_0}$. For $\alpha, \beta \in \mathbb{R}^{Q_0}$, we define

$$\langle \alpha, \beta \rangle = \sum_{x \in Q_0} \alpha(x)\beta(x) - \sum_{a \in Q_1} \alpha(ta)\beta(ha).$$

Definition 5.2. A dimension vector α is called a root if there is an indecomposable representation of dimension α. A root is called real if $\langle \alpha, \alpha \rangle = 1$, isotropic if $\langle \alpha, \alpha \rangle = 0$ and non-isotropic imaginary if $\langle \alpha, \alpha \rangle < 0$. Further, it is called a Schur root if there exists a non-empty Zariski open subset of $\mathrm{Rep}(Q, \alpha)$ such that every representation in it is indecomposable. Note that isotropic roots are also considered imaginary roots.

For the rest of this section, we fix a quiver Q with no oriented cycles. We will recall some standard results. First two lemmas that are straightforward, see [17, 18].

Lemma 5.3. *For any Schur root α, its canonical decomposition is $\alpha = \alpha$.*

Lemma 5.4. *Suppose $\alpha = \beta_1 \oplus \beta_2 \oplus \cdots \oplus \beta_k$ is the canonical decomposition of α. Then each β_i is a Schur root.*

We now state a theorem of Schofield that will be very useful for us.

Theorem 5.5 (Schofield [50]). *Suppose $\alpha = \beta_1 \oplus \beta_2 \oplus \cdots \oplus \beta_k$ is the canonical decomposition for α. Then the canonical decomposition for $m\alpha$ is*

$$m\alpha = [m\beta_1] \oplus [m\beta_2] \oplus \cdots \oplus [m\beta_k],$$

where $[m\beta] = \beta^{\oplus m}$ if β is a real or isotropic Schur root and $[m\beta] = m\beta$ if β is a non-isotropic imaginary Schur root.

Corollary 5.6. *Suppose $\alpha = \beta_1^{\oplus m_1} \oplus \beta_2^{\oplus m_2} \oplus \cdots \oplus \beta_k^{\oplus m_k}$ is the canonical decomposition of α. For all i such that $m_i > 1$, β_i must be a real Schur root or an isotropic Schur root.*

We make a definition:

Definition 5.7. Let α be a dimension vector. Then, we call α σ-stable (resp. σ-semistable, σ-polystable) if a generic representation of dimension α is σ-stable (resp. σ-semistable, σ-polystable).

It is easy to see that in order for α to be σ-stable for any σ, it must be a Schur root. Schofield proved a result in the other direction, which will be very useful for us:

Theorem 5.8 (Schofield [50]). *Let α be a Schur root. Then there exists $0 \neq \sigma \in C(Q, \alpha)$ such that α is σ-stable.*

Corollary 5.9. *Let α be a dimension vector and σ be a weight. Suppose $\alpha = \beta_1 \oplus \beta_2 \oplus \cdots \oplus \beta_k$ is the canonical decomposition with β_i being σ-stable for all i. Then, α is σ-polystable. Moreover, α is σ-stable if and only if $k = 1$.*

Proof. We have a map $\phi : \mathrm{GL}(\alpha) \times \prod_{i=1}^{k} \mathrm{Rep}(Q, \beta_i) \to \mathrm{Rep}(Q, \alpha)$, that takes $(g, (V^{(i)})_{1 \leq i \leq k}) \mapsto g \cdot (V^{(1)} \oplus V^{(2)} \oplus \cdots \oplus V^{(k)})$. The fact that $\alpha = \beta_1 \oplus \beta_2 \oplus \cdots \oplus \beta_k$ is the canonical decomposition means that ϕ is dominant, i.e., its image $\mathrm{Im}(\phi)$ contains a (non-empty) Zariski open subset of $\mathrm{Rep}(Q, \alpha)$.

Now, the fact that each β_i is σ-stable means that there is a non-empty Zariski open subset $U_i \subseteq \mathrm{Rep}(Q, \beta_i)$ that consists of σ-stable representations. Let $U = \mathrm{GL}(\alpha) \times \prod_{i=1}^{k} U_i$. Then for any representation $V \in \phi(U) \subseteq \mathrm{Rep}(Q, \alpha)$, it decomposes as a direct sum of representations of dimension $\beta_1, \ldots, \beta_k$, each of which is σ-stable. Hence $\phi(U)$ consists of σ-polystable representations. Now, $\phi(U)$ is Zariski-dense in $\mathrm{Im}(\phi)$ which is Zariski-dense in $\mathrm{Rep}(Q, \alpha)$. Thus the Zariski-closure of $\phi(U)$ is $\mathrm{Rep}(Q, \alpha)$. Since U is constructible, its image under the map ϕ is constructible (by Chevalley's theorem on constructible sets) and hence contains a (dense, hence non-empty) Zariski-open subset of its closure, i.e., there exists a Zariski-open subset of $\mathrm{Rep}(Q, \alpha)$ that is contained in $\phi(U)$. Thus α is σ-polystable.

That α is σ-stable if and only if $k = 1$ is obvious. □

6. Matrix normal models

Let us explicitly compute the canonical decomposition for the m-Kronecker quiver $\Theta(m)$.

Proposition 6.1 (Canonical decomposition for the m-Kronecker quiver). *Consider the m-Kronecker quiver $\Theta(m)$ and let $\alpha = (p, q)$ be a dimension vector, and let $d = \gcd(p, q)$.*

1. *If $p^2 + q^2 - mpq < 0$, then α is a (non-isotropic) imaginary Schur root and its canonical decomposition is $\alpha = \alpha$.*
2. *If $p^2 + q^2 - mpq = 0$, then $\frac{\alpha}{d}$ is an isotropic Schur root and the canonical decomposition is $\alpha = (\frac{\alpha}{d})^{\oplus d}$ (note that $\frac{\alpha}{d} \in \mathbb{Z}_{\geq 0}^{Q_0}$).*
3. *If $p^2 + q^2 - mpq = d^2$, then $\frac{\alpha}{d}$ is a real Schur root and the canonical decomposition is $\alpha = (\frac{\alpha}{d})^{\oplus d}$.*

(4) *In all other cases (i.e., $p^2+q^2-mpq>0$, but not equal to d^2), the canonical decomposition has at least two linearly independent dimension vectors.*

Proof. First, observe that for a dimension vector $\gamma=(a,b)$, we have $\langle\gamma,\gamma\rangle = a^2+b^2-mab$. The set of roots for $\Theta(m)$ are precisely the dimension vectors (a,b) such that $a^2+b^2-mab\leq 1$ [36]. All real roots and non-isotropic imaginary roots are Schur [36, Theorem 4] (note that non-isotropic imaginary roots only occur for $m\geq 3$). To be precise, Kac shows that all real and non-isotropic imaginary roots occur in a canonical decomposition, and hence must be Schur. Isotropic roots only occur for $m=2$, and these are precisely (a,a). In this case, $(1,1)$ is Schur, but the rest are of course not Schur. Moreover, observe that any real root (a,b) must be indivisible as otherwise, it would not be possible for $\langle(a,b),(a,b)\rangle = a^2+b^2-mab=1$. Thus, all real and isotropic Schur roots are indivisible. Further, we can conclude that if (a,b) is indivisible, then (a,b) is a Schur root if and only if $a^2+b^2-mab\leq 1$.

Let us understand when the canonical decomposition of (p,q) has at least two linearly independent dimension vectors and when it does not. If it does not have two linearly independent dimension vectors, then all the dimension vectors in the canonical decomposition must be parallel to α, so $\alpha=\lambda_1\alpha\oplus\lambda_2\alpha\oplus\cdots\oplus\lambda_k\alpha$ is the canonical decomposition for some scalars λ_i. This means that α is a scalar (not necessarily integral) multiple of a Schur root, i.e., $\lambda_1\alpha$. So, let us now turn to understanding dimension vectors that are scalar multiples of Schur roots.

Let $p'=p/d$ and $q'=q/d$. We claim that (p,q) is a scalar (not necessarily integral) multiple of a Schur root if and only if (p',q') is a Schur root. The "if" is obvious and we have to prove "only if". So, let us assume (p,q) is a multiple of a Schur root. Suppose $p^2+q^2-mpq<0$, then clearly both (p,q) and (p',q') are non-isotropic imaginary Schur roots. If $p^2+q^2-mpq\geq 0$, then (p,q) must be a multiple of a real or isotropic Schur root, and since real/isotropic Schur roots are indivisible, that Schur root must be (p',q'). Note that as a consequence of the above arguments, we get that (p,q) is a scalar multiple of a Schur root if and only if it is an integral multiple of a Schur root.

Having proved the claim in the previous paragraph, we know that (p,q) is a multiple of a Schur root if and only if $p'^2+q'^2-mp'q'\leq 1$ or equivalently $p^2+q^2-mpq=d^2$ or ≤ 0. This is precisely the first three cases and in these cases, Theorem 5.5 tells us precisely what the canonical decomposition has to be, depending on whether the Schur root (p',q') is real, isotropic or non-isotropic imaginary. In all other cases, (p,q) is not a multiple of a Schur root and as argued above its canonical decomposition will have two linearly independent dimension vectors. □

6.1. **Maximum Likelihood thresholds for complex matrix normal models.** Let us prove Theorem 1.2 for the case of $K=\mathbb{C}$ first.

Proof of Theorem 1.2 for $K=\mathbb{C}$. For this proof, let G_{SL} denote $\mathrm{SL}_p\times\mathrm{SL}_q$. For $\sigma=(-q',p')$, we know that σ-stable/polystable/semistable is the same as G_{SL}-stable/semistable/polystable by Lemma 4.3.

For (1), (resp. (2)), observe (by Proposition 6.1) that α (resp. $\frac{\alpha}{d}$) are Schur roots and hence π-stable for some (indivisible) π by Theorem 5.8. Such a π must satisfy $\pi(p,q)=0$ and so π must be $\sigma=(-q',p')$ (naively, it could also have been $(q',-p')$, but this is not in $C(\Theta(m),(p,q))$ as remarked before). Thus, from Corollary 5.9, we deduce that α is σ-stable (resp. σ-polystable) and further than in the case of (2), α is σ-stable if and only if $d=1$. Applying Theorem 2.7, we get the required conclusion.

(4) follows immediately by combining Corollary 4.5, Proposition 6.1 and Theorem 2.7. □

6.2. **Maximum Likelihood thresholds for real matrix normal models.**

Lemma 6.2. *Let $Y\in\mathrm{Rep}(\Theta(m),(p,q))_{\mathbb{R}}=\mathrm{Mat}^m_{p,q}(\mathbb{R})$. If an MLE given Y is unique for the real matrix normal model $\mathcal{M}(p,q)$, then Y is indecomposable over $\mathbb{R}$. In other words, it is indecomposable when thought of as a representation over $\mathbb{R}$ or equivalently, an $\mathbb{R}Q$-module.*

Proof. Let $G = \mathrm{SL}_p \times \mathrm{SL}_q$, so that $G_{\mathbb{R}} = \mathrm{SL}_p(\mathbb{R}) \times \mathrm{SL}_q(\mathbb{R})$. If an MLE given Y exists, then Y is polystable. Hence, without loss of generality, assume Y is polystable. Moreover, without loss of generality, assume Y is a point in the $G_{\mathbb{R}}$-orbit with minimal norm. Let $(G_{\mathbb{R}})_Y$ denote the stabilizer at Y.

There is a constant $\lambda \in \mathbb{R} \setminus \{0\}$ such that for each $(g,h) \in (G_{\mathbb{R}})_Y$, $\lambda(g^\top g \otimes (h^{-1})(h^{-1})^\top) \in \mathrm{PD}_{pq}$ is an MLE, see [2, Proposition 5.2, Remark 5.5]. There is a very minor change in the formula because the action we use is slightly different (yet equivalent) from the one used in [2]. The two actions are related by an automorphism of SL_q, given by $h \mapsto (h^{-1})^\top$), and we modify appropriately the formula for an MLE.

Clearly λI_{pq} is an MLE, so if it is unique, then for all $(g,h) \in (G_{\mathbb{R}})_Y$, we must have $g^\top g = \alpha I_p$ and $h^\top h = \alpha I_q$ for some $0 \neq \alpha \in \mathbb{R}$. Since $g^\top g$ and $h^\top h$ are positive definite, we must have $\alpha > 0$, and since $\det(g) = \det(h) = 1$, we must have $\alpha = 1$. In other words, we must have $g^\top g = I_p$ and $h^\top h = I_q$, i.e., g and h are orthogonal matrices.

Suppose on the contrary that Y is decomposable over $\mathbb{R}$. Interpreting this as a representation over $\mathbb{R}$ for $\Theta(m)$, we assign $\mathbb{R}^p$ to the vertex x and $\mathbb{R}^q$ to the vertex y, and to each arrow a_i, we assign the linear map $Y_i : \mathbb{R}^q \to \mathbb{R}^p$. Now, Y is decomposable means that there is a decomposition $\mathbb{R}^p = W(x) \oplus Z(x)$ and $\mathbb{R}^q = W(y) \oplus Z(y)$, such that for each i, $Y_i(W(y)) \subseteq W(x)$ and $Y_i(Z(y)) \subseteq Z(x)$. Consider $\dim(W) = (\dim(W(x)), \dim(W(y)))$ and $\dim(Z) = (\dim(Z(x)), \dim(Z(y)))$. Then since Y is σ-semistable, we must have that $\dim(W) = (ap, aq)$ and $\dim(Z) = (bp, bq)$ by Corollary 4.5 (to be precise, we have to complexify everything to apply Corollary 4.5, we leave the details to the reader). Now, let $c, d \in \mathbb{R}_{>0}$ such that $c^a d^b = 1$ and $|c|, |d| \neq 1$ (for e.g., $c = 2$ and $d = 2^{-a/b}$). Let $g \in \mathrm{SL}_p(\mathbb{R})$ be the linear map that is defined by $g(v) = cv$ for $v \in W(x)$ and $g(v) = dv$ for $v \in Z(x)$, and let $h \in \mathrm{SL}_q(\mathbb{R})$ be the linear map defined by $h(v) = cv$ for $v \in W(y)$ and $h(v) = dv$ for $v \in Z(y)$. Then, it is a simple check to see that $(g,h) \in (G_{\mathbb{R}})_Y$. However, clearly g and h are not orthogonal matrices because they have eigenvalues with absolute value $\neq 1$. This contradicts uniqueness of MLE by the above discussion.

Thus, Y must be indecomposable over $\mathbb{R}$.

□

Proof of Theorem 1.2 for $K = \mathbb{R}$. Let $G_{\mathrm{SL}} = \mathrm{SL}_p(\mathbb{C}) \times \mathrm{SL}_q(\mathbb{C})$ and so $(G_{\mathrm{SL}})_{\mathbb{R}} = \mathrm{SL}_p(\mathbb{R}) \times \mathrm{SL}_q(\mathbb{R})$. By using Proposition 2.23, all the generic stability notions (Definition 2.3) carry over without any change from the case of $K = \mathbb{C}$ to the case of $K = \mathbb{R}$.

In particular, the statements regarding boundedness of log-likelihood function and existence of MLEs also carry over from $K = \mathbb{C}$ to $K = \mathbb{R}$. The only issue arises in terms of uniqueness of an MLE. Over the reals, stability implies uniqueness of MLEs, but not conversely. Thus, even when $\mathrm{Rep}(\Theta(m), (p,q))_{\mathbb{R}}$ is not generically $(G_{\mathrm{SL}})_{\mathbb{R}}$-stable, we might still have almost sure uniqueness of MLEs. So, we need to look at the cases where we have generic polystability but not generic stability. This happens exactly when $d \geq 2$ and $p^2 + q^2 - mpq$ is either 0 or d^2. This is precisely why we proved the above lemma.

Now, suppose $p^2 + q^2 - mpq = 0$ or d^2 and $d \geq 3$. Then, by Lemma 3.3, we get that a generic point in $\mathrm{Rep}(\Theta(m), (p,q))_{\mathbb{R}}$ is decomposable over $\mathbb{R}$, and by Lemma 6.2 that MLE is not unique.

Now, suppose $p^2 + q^2 - mpq = d^2$ and $d = 2$. This is precisely the case where $(p,q) = 2\beta$ where β is a real Schur root. This means that there is a unique indecomposable of dimension β and it is defined over $\mathbb{R}$ – this is because for a real Schur root, the representation space has a Zariski-dense orbit corresponding to this unique indecomposable [18, Lemma 11.1.3]. This unique indecomposable corresponds to an $\mathbb{R}Q$-module (for $Q = \Theta(m)$) that we will call W. Take a generic point $Y \in \mathrm{Rep}(\Theta(m), (p,q))_{\mathbb{R}}$. Interpret this as an $\mathbb{R}Q$-module, which we will call M. Then, by genericity, we know that $M_{\mathbb{C}} \cong W_{\mathbb{C}} \oplus W_{\mathbb{C}}$ as $\mathbb{C}Q$-modules. Hence $M_{\mathbb{C}} \cong W^{\oplus 4}$ as $\mathbb{R}Q$-modules. Thus $M^{\oplus 2} \cong W^{\oplus 4}$ as $\mathbb{R}Q$-modules. By Krull-Remak-Schmidt theorem, we get that $M \cong W^{\oplus 2}$ as

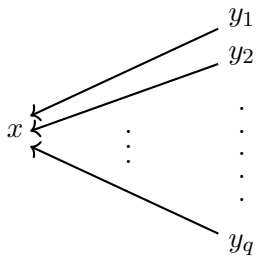
$\mathbb{R}Q$-modules. In other words, Y is decomposable over $\mathbb{R}$, and hence by Lemma 6.2, there is not a unique MLE.

Now, suppose $p^2+q^2-mpq=0$ and $d=2$. This happens precisely in the case of $(p,q)=(2,2)$ and $m=2$. This is a slightly tricky case, and it turns out that we cannot claim uniqueness or non-uniqueness of MLEs generically. In fact, the subset of points having a unique MLE is semi-algebraic and full-dimensional, but not dense. Nevertheless, it remains that we do not have almost sure uniqueness of MLEs, see [21, Section 4] (in particular Corollary 4.6) for a more thorough explanation of this behavior.

Thus every statement in the case of $K=\mathbb{C}$ transfers to the case of $K=\mathbb{R}$. □

7. Model of proportional covariance matrices

In this section, we will focus on the model of proportional covariance matrices $\mathcal{N}(p,q)$. Once again, we will first work with $K=\mathbb{C}$ (and then transfer the result to $K=\mathbb{R}$). For this case, we consider the quiver $\mathcal{B}(q,m)$ with vertices $x,y_1,\ldots,y_q$ and m arrows from each y_i to x. The quiver $\mathcal{B}(q,1)$ is pictured below.



Let us first define an operation for quivers. For any quiver $Q=(Q_0,Q_1)$, define $Q^{[m]}$ to be the following quiver: Let its vertex set be Q_0, the vertex set for Q. For each $a\in Q_1$, define m arrows $a^{[1]},\ldots,a^{[m]}$ in $Q_1^{[m]}$ such that $ta=ta^{[i]}$ and $ha=ha^{[i]}$ for all i. For example, we have $\mathcal{B}(q,m)=\mathcal{B}(q,1)^{[m]}$. For any dimension vector $\alpha\in\mathbb{Z}_{\geq 0}^{Q_0}$, we have $\mathrm{Rep}(Q^{[m]},\alpha)=\mathrm{Rep}(Q,\alpha)^{\oplus m}$. Further, the action of $\mathrm{GL}(\alpha)$ on $\mathrm{Rep}(Q^{[m]},\alpha)=\mathrm{Rep}(Q,\alpha)^{\oplus m}$ is the diagonal action obtained from the action on $\mathrm{Rep}(Q,\alpha)$. The same holds for the action of any subgroup of $\mathrm{GL}(\alpha)$.

We use the convention that in a dimension vector for $\mathcal{B}(q,m)$, the coordinates correspond to $x,y_1,\ldots,y_q$ in order. If we take the dimension vector $\alpha=(p,1,1,\ldots,1)$, then $\mathrm{Rep}(\mathcal{B}(q,1),\alpha)$ can be identified with $\mathrm{Mat}_{p,q}$. Let $\sigma=(-q',p',\ldots,p')$ where $(p',q')=\frac{1}{\gcd(p,q)}(p,q)$. Analogous to Lemma 4.3, we can prove that σ-semistability/polystability/stability coincides with $\mathrm{SL}_p\times\mathrm{ST}_q$-semistability/polystability/stability.

Since $\mathcal{B}(q,m)=\mathcal{B}(q,1)^{[m]}$, by the above discussion we conclude the following:

Lemma 7.1. *Consider the action of $H_{\mathrm{SL}}=\mathrm{SL}_p\times\mathrm{ST}_q$ on $\mathrm{Mat}_{p,q}^m=\mathrm{Rep}(\mathcal{B}(q,m),(p,1,1,\ldots,1))$. Let $\sigma=(-q',p',\ldots,p')$ be a weight for $\mathcal{B}(q,m)$, where $(p',q')=\dfrac{1}{\gcd(p,q)}(p,q)$. Suppose $Y\in\mathrm{Mat}_{p,q}^m$. Then, Y is H_{SL}-semistable/polystable/stable if and only if Y is σ-semistable/polystable/stable.*

Proof. This is analogous to Lemma 4.3. □

Proposition 7.2. *Let $Q=\mathcal{B}(q,m)$, $\alpha=(p,1,1,\ldots,1)$ and $\sigma=(-q',p',\ldots,p')$ where $(p',q')=\frac{1}{\gcd(p,q)}(p,q)$. If $mq<p$, then every $Y\in\mathrm{Rep}(Q,\alpha)=\mathrm{Mat}_{p,q}^m$ is σ-unstable. If $mq=p$, then α is σ-polystable (and σ-stable precisely when $q=1$). If $mq>p$, then α is σ-stable.*

Proof. Let $mq<p$, and let $Y\in\mathrm{Rep}(Q,\alpha)$. We claim that there is a subrepresentation of dimension $\beta=(mq,1,1,\ldots,1)$. This is because from each vertex y_i, there are m arrows, and each one of them has a 1-dimensional image. There are mq of such 1-dimensional subspaces (one for each

arrow), so there is a subspace $U \subseteq \mathbb{C}^p$ of dimension mq that contains all of these. This gives a subrepresentation of dimension β. Now $\sigma(\beta) > 0$, so α is not σ-semistable.

Now, let $mq = p$, and let $Y \in \mathrm{Rep}(Q, \alpha)$ be generic. Then, for each y_i, the images of the m arrows starting from y_i form an m-dimensional subspace of $\mathbb{C}^p$ (the vector space at the vertex x). You get one such m-dimensional subspace for each y_i (call it U_i), hence there are q of them. By genericity, we will have that $\mathbb{C}^q = \oplus_{i=1}^q U_i$. This means that placing $\mathbb{C}$ at the vertex y_i, U_i at vertex x and $\mathbb{C}^0$ at all other vertices gives a subrepresentation, and in fact a direct summand. Thus, Y is a direct sum of q indecomposables of dimensions $(m, 1, 0, \ldots, 0), (m, 0, 1, \ldots, 0), \ldots, (m, 0, \ldots, 1)$. It is straightforward to see (by genericity) that each one of these summands will be indecomposable, has no non-trivial subrepresentations, and is σ-stable. Thus, Y is σ-polystable. In fact, it is easy to see that the canonical decomposition is $\alpha = (m, 1, 0, \ldots, 0) \oplus (m, 0, 1, \ldots, 0) \oplus \cdots \oplus (m, 0, \ldots, 1)$ and that each of the dimension vectors appearing in the canonical decomposition are real Schur roots that are σ-stable.

Now, let $mq > p$, and let $Y \in \mathrm{Rep}(Q, \alpha)$ be generic. Similar arguments as above will show that any subrepresentation has a dimension vector of the form $\beta = (\min\{m(\sum \epsilon_i), p\}, \epsilon_1, \epsilon_2, \ldots, \epsilon_q)$, where $\epsilon_i \in \{0, 1\}$. For each subrepresentation, we observe that $\sigma(\beta) < 0$ unless $\beta = \alpha$, when $\sigma(\alpha) = 0$. Hence Y is σ-stable. □

Proof of Theorem 1.5. We claim the following three statements. If $mq < p$, then the log-likelihood function is unbounded. If $mq = p$, then (almost surely) an MLE exists and we have almost sure uniqueness precisely when $m = 1$. If $mq > p$, then (almost surely) we have a unique MLE. For $K = \mathbb{C}$, they follow from Proposition 7.2, Lemma 7.1 and Proposition 2.9. Transfering the result to $K = \mathbb{R}$ is analogous to Theorem 1.2. For the case $mq = p$, one has to look into the proof of Proposition 7.2 to see that the canonical decomposition of (p, q) consists of real Schur roots, so the argument parallels part (2) of Theorem 1.2.

Reformulating this in terms of maximum likelihood thresholds gives us the required conclusion. □

References

[1] G. I. Allen, R. Tibshirani, *Transposable regularized covariance models with an application to missing data imputation*, Ann. Appl. Stat. **4**, no. 2, (2010), 764–790.

[2] C. Amendola, K. Kohn, P. Reichenbach and A. Seigal, *Invariant theory and scaling algorithms for maximum likelihood estimation*, `arXiv:2003.13662`, [math.ST], 2020.

[3] A. Borel, *Linear algebraic groups*, Second edition. Graduate Texts in Mathematics, 126. Springer-Verlag, New York, 1991. xii+288 pp.

[4] A. Borel, *Lie groups and linear algebraic groups. I. Complex and real groups*, Lie groups and automorphic forms, 1–49, AMS/IP Stud. Adv. Math., 37, Amer. Math. Soc., Providence, RI, 2006.

[5] A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Annals of Mathematics, Second Series **75** (1962), 485–535.

[6] F. Bijma, J. C. De Munck, R. M. Heethaar, *The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products*, NeuroImage **27**, no. 2, (2005), 402–415.

[7] D. Birkes, *Orbits of linear algebraic groups*, Annals of Mathematics, Second Series **93** (1971), 459–475.

[8] M. Bürgin and J. Draisma, *The Hilbert null-cone on tuples of matrices and bilinear forms*, Math. Z. **254** (2006), no. 4, 785–809.

[9] P. Bürgisser, C. Franks, A. Garg, R. Oliveira, M. Walter, and A. Wigderson, *Towards a theory of non-commutative optimization: geodesic 1st and 2nd order methods for moment maps and polytopes*, FOCS 2019, 845–861.

[10] C. Chindris and D. Kline, *Simultaneous robust subspace recovery and semi-stability of quiver representations*, `arXiv:2003.02962` [math.RT], 2020.

[11] J. C. De Munck, H. M. Huizenga, L. J. Waldorp, R. M. Heethaar, *Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise*, IEEE Transactions Sign. Proc. **50**, no. 7 (2002), 1565–1572.

[12] H. Derksen and V. Makam, *Polynomial degree bounds for matrix semi-invariants*, Adv. Math. **310** (2017), 44–63.

[13] H. Derksen and V. Makam, *Generating invariant rings of quivers in arbitrary characteristic*, J. Algebra **489** (2017), 435–445.

[14] H. Derksen and V. Makam, *Degree bounds for semi-invariant rings of quivers*, J. Pure Appl. Algebra **222** (2018), no. 10, 3282–3292.

[15] H. Derksen, A. Schofield and J. Weyman, *On the number of subrepresentations of a general quiver representation*, J. Lond. Math. Soc. (2) **76** (2007), no. 1, 135–147.

[16] H. Derksen and J. Weyman, *Semi-invariants of quivers and saturation of Littlewood-Richardson co-efficients*, Journal of the American Math. Soc. **13** (2000), 467-479.

[17] H. Derksen and J. Weyman, *On the canonical decomposition of quiver representations*, Compositio Math. **133** (2002), no. 3, 245–265.

[18] H. Derksen and J. Weyman, *An introduction to quiver representations*, Graduate Studies in Mathematics **184**, American Mathematical Society, Providence, RI, 2017. x+334 pp.

[19] H. Derksen, J. Weyman and A. Zelevinsky, *Quivers with potentials and their representations II: applications to cluster algebras*, J. Amer. Math. Soc. **23** (2010), no. 3, 749–790.

[20] M. Domokos and A. N. Zubkov, *Semi-invariants of quivers as determinants*, Transformation groups **6** (2001), 9-24.

[21] M. Drton, S. Kuriki and P. Hoff, *Existence and Uniqueness of the Kronecker Covariance MLE*, `arXiv:2003.06024`, [math.ST], 2020.

[22] P. Dutilleul, *The MLE algorithm for the matrix normal distribution*, J. Statist. Comput. Simul. **64** (1999), 105–123.

[23] P. Dutilleul, B. Pinel-Alloul, *A doubly multivariate model for statistical analysis of spatio-temporal environmental data*, Environmetrics **7** (1996), 551–566.

[24] S. P. Eriksen, *Proportionality of covariance matrices*, Ann. Statist. **15** (1987), no. 2, 732–748.

[25] A. Garg, L. Gurvits, R. Oliveira and A. Widgerson, *A deterministic polynomial time algorithm for non-commutative rational identity testing*, 57th Annual IEEE Symposium on Foundations of Computer Science–FOCS 2016, 109–117, IEEE Computer Soc., Los Alamitos, CA, 2016.

[26] A. Garg, L. Gurvits, R. Oliveira and A. Widgerson, *Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via operator scaling*, Geom. Funct. Anal. **28** (2018), no. 1, 100–145.

[27] V. Ginzburg, *Lectures on Nakajima's quiver varieties*, Geometric methods in representation theory. I, 145–219, Sémin. Congr., 24-I, Soc. Math. France, Paris, 2012.

[28] L. Gurvits, *Classical complexity and quantum entanglement*, Journal of Computer and System Sciences **69** (2004), no. 3, 448–484.

[29] D. Hilbert, *Über die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), 473–534.

[30] D. Hilbert, *Über die vollen Invariantensysteme*, Math. Ann. **42** (1893), 313–370.

[31] V. Hoskins, *Geometric invariant theory and symplectic quotients*, http://userpage.fu-berlin.de/hoskins/GITnotes.pdf (2012).

[32] H. M. Huizenga, J. C. De Munck, L. J. Waldorp, R. P. P. P. Grasman, *Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model*, IEEE Trans. Biomed. Eng. **49** (2002), no. 6, 533–539.

[33] G. Ivanyos, Y. Qiao and K. V. Subrahmanyam, *Non-commutative Edmonds' problem and matrix semi-invariants*, Comput. Complexity **26** (2017), no. 3, 717–763.

[34] G. Ivanyos, Y. Qiao and K. V. Subrahmanyam, *Constructive non-commutative rank computation is in deterministic polynomial time*, Comput. Complexity **27** (2018), no. 4, 561–593.

[35] S. T. Jensen and S. Johansen, *Estimation of proportional covariances*, Statistics & probability letters **6** (1987), no. 2, 83–85.

[36] V. Kac, *Infinite root systems, representations of graphs and invariant theory*, Invent. Math. **56** (1980), no. 1, 57–92.

[37] V. Kac, *Infinite root systems, representations of graphs and invariant theory. II*, J. Algebra **78** (1982), no. 1, 141–162.

[38] B. Keller, *Cluster algebras, quiver representations and triangulated categories*, Triangulated categories, 76–160, London Math. Soc. Lecture Note Ser. bf 375, Cambridge Univ. Press, Cambridge, 2010.

[39] A. D. King, *Moduli of representations of finite-dimensional algebras*, Quart. J. Math. Oxford Ser. **45** (1994), no. 180, 515–530.

[40] Y. Matsushima, *Espaces homogènes de Stein des groupes de Lie complexes*, Nagoya Math. Journal **16** (1960), 205–218.

[41] G. D. Mostow, *Self-adjoint groups*, Annals of Mathematics **62** (1955), no. 1, 44–55.

[42] K. Mulmuley, *Geometric Complexity Theory V: Equivalence between blackbox derandomization of polynomial identity testing and derandomization of Noether's normalization lemma*, J. Amer. Math. Soc. **30** (2017), no. 1, 225–309.

[43] L. Le Bruyn and C. Procesi, *Semisimple representations of quivers*, Trans. Amer. Math. Soc **317** (1990), 585-598.
[44] N. Lu and D. L. Zimmerman, *The likelihood ratio test for a separable covariance matrix*, Stat. and Prob. Letters **73**, no. 4 (2005), 449–457.
[45] K. V. Mardia and C. R. Goodall, *Spatial-temporal analysis of multivariate environmental monitoring data*, Multivariate Environ. Stat. **6** (1993), 347–386.
[46] M. Reineke, *Moduli of representations of quivers*, Trends in representation theory of algebras and related topics, 589–637, EMS Ser. Congr. Rep., Eur. Math. Soc., Zürich, 2008.
[47] N. Ressayre, *Multiplicative formulas in Schubert calculus and quiver representation*, Indag. Math. (N.S.) **22** (2011), no. 1-2, 87–102.
[48] M. Rosenlicht, *Some rationality questions on algebraic groups*, Annali di Matematica Pura ed Applicata **43** (1957), no. 1, 25–50.
[49] B. Roś, B. Fetsje, J. C. de Munck and Mathisca C. M. de Gunst, *Existence and uniqueness of the maximum likelihood estimator for models with a Kronecker product covariance structure*, J. Multivariate Anal. **143** (2016), 345–361.
[50] A. Schofield, *General representations of quivers*, Proc. London Math. Soc. (3) **65**, (1992), no. 1, 46–64.
[51] M. S. Srivastava, T. von Rosen, D. von Rosen, *Models with a Kronecker product covariance structure: estimation and testing*, Math. Methods Statist. **17** (2008), no. 4, 357–370.
[52] A. Schofield and M. van der Bergh, *Semi-invariants of quivers for arbitrary dimension vectors*, Indag. Mathem., N.S **12** (2001), 125–138.
[53] T. A. Springer, *Linear algebraic groups*, Reprint of the 1998 second edition. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009. xvi+334 pp.
[54] B. Torrésani and E. Villaron, *Harmonic hidden Markov models for the study of EEG signals*, 18th European Signal Processing Conference, EUSIPCO-2010.
[55] Y. Zhang and J. Schneider, *Learning multiple tasks with a sparse matrix-normal penalty*, Adv. Neural Inf. Process. Syst. **23** (2010), 2550–2558.
[56] A. Wiesel, *Geodesic convexity and covariance estimation*, IEEE Trans. Signal Process. **60** (2012), no. 12, 6182–6189.

Appendix A. Equivalence of stability notions

In this appendix, we reconcile Theorem 4.2 with King's original formulation [39].

Let Q be a quiver with no oriented cycles, let α be a sincere dimension vector, i.e., $\alpha(x) \neq 0$ for all $x \in Q_0$, and let $\sigma \in C(Q,\alpha)$ be a non-zero indivisible weight. Then, it is easy to see that $\mathbb{C}[\mathrm{Rep}(Q,\alpha)]^{\mathrm{GL}(\alpha)_\sigma} = \bigoplus_{n\in\mathbb{Z}} \mathrm{SI}(Q,\alpha)_{n\sigma}$. But, in fact, $\mathbb{C}[\mathrm{Rep}(Q,\alpha)]^{\mathrm{GL}(\alpha)_\sigma} = \bigoplus_{n\in\mathbb{Z}_{\geq 0}} \mathrm{SI}(Q,\alpha)_{n\sigma}$ because we assume the quiver has no oriented cycles. This follows from the fact that the form $\langle -,-\rangle$ is non-degenerate (see [18, Definition 2.5.3]) and that $\mathrm{SI}(\mathrm{Q},\alpha)_\gamma \neq 0$ implies that $\gamma = \langle \beta, -\rangle$ for some dimension vector β (see [18, Theorem 10.7.1]). Now, since we chose $0 \neq \sigma \in \mathbb{C}(Q,\alpha)$, we know that for some $n \in \mathbb{Z}_{>0}$, $n\sigma = \langle \beta, -\rangle$ for some dimension vector $0 \neq \beta \in \mathbb{Z}_{\geq 0}^{Q_0}$. So for $m \in \mathbb{Z}_{>0}$, we get that $-m\sigma = \langle -\frac{m}{n}\beta, -\rangle$, but $-\frac{m}{n}\beta$ cannot be a dimension vector as it contains negative entries, so $\mathrm{SI}(Q,\alpha)_{-m\sigma} = 0$ for all $m \in \mathbb{Z}_{>0}$.

Let $\mathbb{C}_\sigma$ denote the 1-dimensional representation of $\mathrm{GL}(\alpha)$ corresponding to σ, i.e., $\mathbb{C}_\sigma = \mathbb{C}$ as a vector space and the linear action of $\mathrm{GL}(\alpha)$ is given by $g \cdot 1 = \sigma(g)$.

Proposition A.1. *Let Q, α, σ be as above. Then $V \in \mathrm{Rep}(Q,\alpha)$ is $\mathrm{GL}(\alpha)_\sigma$-semistable/polystable/stable if and only if $(V,1) \in \mathrm{Rep}(Q,\alpha) \oplus \mathbb{C}_\sigma$ is $\mathrm{GL}(\alpha)$-semistable/polystable/stable.*

Proof. Let z denote the coordinate of $\mathbb{C}_\sigma$ in $\mathrm{Rep}(Q,\alpha) \oplus \mathbb{C}_\sigma$. Let us split the argument for each notion of stability

- **Semistability:** Suppose V is $\mathrm{GL}(\alpha)_\sigma$-semistable. Then there exists $f \in \mathrm{SI}(Q,\alpha)_{n\sigma}$ such that $f(V) \neq 0$. This means that $\widetilde{f} = fz^n$ is $\mathrm{GL}(\alpha)$ invariant (with no constant term) and $\widetilde{f}(V,1) = f(V) \neq 0$. So $(V,1)$ is $\mathrm{GL}(\alpha)$-semistable.

  Conversely, if $\widetilde{f}(V,1) \neq 0$ for some $\widetilde{f}$ that is $\mathrm{GL}(\alpha)$-invariant and homogeneous (say of degree $m > 0$), then write $\widetilde{f} = \sum_{i=0}^{m} f_{m-i} z^i$, with f_j homogeneous of degree j for all j. Then, each $f_{m-i}z^i$ is $\mathrm{GL}(\alpha)$-invariant. For some i, we have that $f_{m-i}z^i$ does not vanish

on $(V,1)$. So, $f_{m-i} \in \mathrm{SI}(Q,\alpha)_{i\sigma}$ is homogeneous of degree $(m-i)$ such that $f_{m-i}(V) \neq 0$. If $i = m$, this means that f_0 is a constant, but $f_0 \in \mathrm{SI}(Q,\alpha)_{m\sigma}$, which is absurd because $m\sigma \neq 0$. Thus $i < m$ and so $f_{m-i} \in \mathrm{SI}(Q,\alpha)_{i\sigma} \in \mathbb{C}[\mathrm{Rep}(Q,\alpha)]^{\mathrm{GL}(\alpha)_\sigma}$ is a homogeneous polynomial of positive degree that does not vanish on V. Thus, V is $\mathrm{GL}(\alpha)_\sigma$-semistable.

- **Polystability:** Suppose $0 \neq V$ is not $\mathrm{GL}(\alpha)_\sigma$-polystable. Then, by the (generalized) Hilbert–Mumford criterion ([18, Proposition 9.6.2]) there exists a 1-parameter subgroup $\lambda : \mathbb{C}^* \to \mathrm{GL}(\alpha)_\sigma$ such that $\lim_{t\to 0} \lambda(t)V = W$ where $W \notin \mathrm{GL}(\alpha)_\sigma \cdot V$. This means that $\lim_{t\to 0} \lambda(t)(V,1) = (W,1)$. Now, we will show that $(W,1) \notin \mathrm{GL}(\alpha)\cdot(V,1)$. Otherwise, we have $(W,1) = g(V,1) = (gV, \sigma(g))$ for some $g \in \mathrm{GL}(\alpha)$. Thus $\sigma(g) = 1$, i.e., $g \in \mathrm{GL}(\alpha)_\sigma$ and $gV = W$ and hence $W \in \mathrm{GL}(\alpha)_\sigma \cdot V$, which is a contradiction. So, $(V,1)$ is not polystable. In the case that $V = 0$, note that $(0,1)$ is not even $\mathrm{GL}(\alpha)$-semistable if there exists $g \in \mathrm{GL}(\alpha)$ such that $|\sigma(g)| < 1$ (since that would mean $\lim_{k\to\infty} g^k(0,1) = (0,0)$). It is easy to construct such a $g \in \mathrm{GL}(\alpha)$ with our assumptions, i.e., Q has no oriented cycles, α is sincere and σ is non-zero.

  Conversely, suppose $(V,1)$ is not polystable. Then there is a 1-parameter subgroup λ of $\mathrm{GL}(\alpha)$ such that $\lim_{t\to 0} \lambda(t)(V,1) = (W,c)$, with $(W,c) \notin \mathrm{GL}(\alpha)(V,1)$. Suppose $c = 0$, then $(W,0)$ is easily seen to be unstable because all points are unstable for the action of $\mathrm{GL}(\alpha)$ on $\mathrm{Rep}(Q,\alpha)$ if Q has no oriented cycles (as is the case for us). This would mean that $(V,1)$ is not even $\mathrm{GL}(\alpha)$-semistable, which means that V is not $\mathrm{GL}(\alpha)_\sigma$-semistable and hence not $\mathrm{GL}(\alpha)_\sigma$-polystable. Hence w.l.o.g., assume $c \neq 0$ from now on. Now, the function $t \mapsto \sigma(\lambda(t))$ is a character of $\mathbb{C}^*$ and has to be of the form $t \mapsto t^k$ for some integer k. Since $c = \lim_{t\to 0} t^k$ is defined, we must have $k \geq 0$. If $k = 0$, we get $c = 1$ and if $k > 0$, we get $c = 0$. Since $c \neq 0$, we must have $c = 1$ and $\lambda(t) \in \mathrm{GL}(\alpha)_\sigma$. This means that $\lim_{t\to 0} \lambda(t)V = W$, so $W \in \overline{\mathrm{GL}(\alpha)_\sigma \cdot V}$. But $W \notin \mathrm{GL}(\alpha)_\sigma \cdot V$, because if it were, then $gV = W$ for some $g \in \mathrm{GL}(\alpha)_\sigma$, which means $g(V,1) = (W,1) = (W,c)$, which is a contradiction. Thus the orbit of V is not closed, and hence V is not polystable.
- **Stability:** Since V is $\mathrm{GL}(\alpha)_\sigma$-polystable if and only if $(V,1)$ is $\mathrm{GL}(\alpha)$-polystable, we only need to now understand the stabilizers. First, observe that if $\Delta \subseteq \mathrm{GL}(\alpha)_\sigma$ is the kernel for its action on $\mathrm{Rep}(Q,\alpha)$, then Δ is the kernel for the action of $\mathrm{GL}(\alpha)$ on $\mathrm{Rep}(Q,\alpha)\oplus\mathbb{C}_\sigma$. Thus, all we need to do is to show that the two stabilizers, i.e., $\mathrm{GL}(\alpha)_{(V,1)}$ and $(\mathrm{GL}(\alpha)_\sigma)_V$, have the same dimension. In fact they are both the same. Indeed $g \in \mathrm{GL}(\alpha)_{(V,1)}$ if and only if $g(V,1) = (gV,\sigma(g)) = (V,1)$ if and only if $\sigma(g) = 1$ and $gV = V$ if and only if $g \in (\mathrm{GL}(\alpha)_\sigma)_V$.

□

Now, we bridge the gap between the result stated in [39] and Theorem 4.2. In Theorem 4.2, we have σ such that $-\sigma \notin C(Q,\alpha)$. We have two cases

- **Case 1:** $\sigma \in C(Q,\alpha)$ – In this case, King [39] showed that σ-semistability/polystability/stability for $V \in \mathrm{Rep}(Q,\alpha)$ was the same as the $\mathrm{GL}(\alpha)$-semistability/polystability/stability of $(V,1) \in \mathrm{Rep}(Q,\alpha)\oplus\mathbb{C}_\sigma$, which is equivalent to $\mathrm{GL}(\alpha)_\sigma$-semistability/polystability/stability of V as we have shown above.
- **Case 2:** $\sigma \notin C(Q,\alpha)$ – In this case, we see that $\mathbb{C}[\mathrm{Rep}(Q,\alpha)]^{\mathrm{GL}(\alpha)_\sigma} = \mathbb{C}$ since we also assume $-\sigma \notin C(Q,\alpha)$. Thus, every $V \in \mathrm{Rep}(Q,\alpha)$ is $\mathrm{GL}(\alpha)_\sigma$-unstable. On the other hand, King [39] showed that σ-semistability of $V \in \mathrm{Rep}(Q,\alpha)$ is equivalent to the existence of $f \in \mathrm{SI}(Q,\alpha)_{d\sigma}$ for some $d \in \mathbb{Z}_{\geq 1}$ such that $f(V) \neq 0$. But as $\mathrm{SI}(Q,\alpha)_{d\sigma} = 0$ for all $d \in \mathbb{Z}_{\geq 1}$ (because $\sigma \notin C(Q,\alpha)$), we see that every $V \in \mathrm{Rep}(Q,\alpha)$ is σ-unstable (i.e., not σ-semistable). Thus, in this case, Theorem 4.2 is clearly true.

Finally, we remark that the hypothesis on Q, α and σ cannot be entirely removed. For example, if you take $\sigma = 0$, then for $V = 0$, it is easy to see that $(V, 1)$ is $\mathrm{GL}(\alpha)$-semistable, but V is not $\mathrm{GL}(\alpha)_\sigma$-semistable.