

ENTANGLED STATES ARE TYPICALLY INCOMPARABLE

VISHESH JAIN, MATTHEW KWAN, AND MARCUS MICHELEN

ABSTRACT. Consider a bipartite quantum system, where Alice and Bob jointly possess a pure state $|\psi\rangle$. Using local quantum operations on their respective subsystems, and unlimited classical communication, Alice and Bob may be able to transform $|\psi\rangle$ into another state $|\phi\rangle$. Famously, *Nielsen's theorem* [Phys. Rev. Lett., 1999] provides a necessary and sufficient algebraic criterion for such a transformation to be possible (namely, the *local spectrum* of $|\phi\rangle$ should *majorise* the local spectrum of $|\psi\rangle$).

In the paper where Nielsen proved this theorem, he conjectured that in the limit of large dimensionality, for *almost all* pairs of states $|\psi\rangle, |\phi\rangle$ (according to the natural unitary invariant measure) such a transformation is not possible. That is to say, typical pairs of quantum states $|\psi\rangle, |\phi\rangle$ are entangled in fundamentally different ways, that cannot be converted to each other via local operations and classical communication.

Via Nielsen's theorem, this conjecture can be equivalently stated as a conjecture about majorisation of spectra of random matrices from the so-called *trace-normalised complex Wishart–Laguerre ensemble*. Concretely, let X and Y be independent $n \times m$ random matrices whose entries are i.i.d. standard complex Gaussians; then Nielsen's conjecture says that the probability that the spectrum of $XX^\dagger / \text{tr}(XX^\dagger)$ majorises the spectrum of $YY^\dagger / \text{tr}(YY^\dagger)$ tends to zero as both n and m grow large. We prove this conjecture, and we also confirm some related predictions of Cunden, Facchi, Florio and Gramegna [J. Phys. A., 2020; Phys. Rev. A., 2021].

1. INTRODUCTION

Suppose a bipartite pure quantum state is distributed to two parties Alice and Bob. Imagine that Alice and Bob then travel to distant labs; after this point, they may only perform quantum operations on their own subsystem, though they are free to coordinate their operations via classical communication (e.g., they can share the results of any measurements). This paradigm is known as LOCC (local operations with classical communication), and plays a fundamental role in quantum information theory. See the Nielsen–Chuang monograph [24] for a general introduction to quantum information theory, and see the survey of Chitambar, Leung, Mančinska, Ozols, and Winter [9] for a thorough survey on the LOCC paradigm.

The LOCC paradigm naturally defines a partial order (called *LOCC-convertibility* or *entanglement transformation*) on the set of bipartite pure states $\mathbb{C}^n \otimes \mathbb{C}^m$ for any dimensions $n, m \in \mathbb{N}$. Namely, for a pair of states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$, we write $|\psi\rangle \rightarrow |\phi\rangle$ (read “ $|\psi\rangle$ transforms to $|\phi\rangle$ ”) if and only if, starting from the state $|\psi\rangle$, it is possible for Alice and Bob to coordinate some sequence of local operations to guarantee that they end up with the state $|\phi\rangle$.

In 1999, Nielsen [23] found a beautiful connection between LOCC-convertibility and the mathematical theory of *majorisation*. For the precise statement, and more context, the reader should refer to [23] (and Chapter 12.5.1 of Nielsen–Chuang [24]), but as a brief summary: although Alice and Bob share a pure state $|\psi\rangle$, Alice can interpret her subsystem as a *mixed* state (a statistical ensemble of pure states in \mathbb{C}^n). This mixed state can be described by a *density operator* $\rho_\psi = \text{tr}_2(|\psi\rangle\langle\psi|)$ (a Hermitian linear operator $\mathbb{C}^n \rightarrow \mathbb{C}^n$ obtained via a *partial trace*, “tracing out” Bob's subsystem). The eigenvalues of ρ_ψ are all nonnegative real numbers, summing to 1¹. Nielsen's theorem is that LOCC-convertibility $|\psi\rangle \rightarrow |\phi\rangle$ is equivalent to the property that the spectrum of ρ_ϕ *majorises* the spectrum of ρ_ψ : the sum of the k smallest² eigenvalues of ρ_ψ should be at most the sum of the k smallest eigenvalues of ρ_ϕ , for all $k \leq n$.

Vishesh Jain was partially supported by NSF CAREER award DMS-2237646. Matthew Kwan was supported by ERC Starting Grant “RANDSTRUCT” No. 101076777. Marcus Michelen was supported in part by NSF grants DMS-2137623 and DMS-2246624.

¹We could equally well have defined these eigenvalues by tracing out Alice's subsystem, instead of Bob's subsystem. It is not hard to see that this choice only affects the number of zero eigenvalues; the nonzero eigenvalues do not actually depend on which of the two subsystems we choose to trace out.

²Note that since the eigenvalues sum to 1, this is equivalent to the condition that the sum of the k largest eigenvalues of ρ_ψ is at most the sum of the k largest eigenvalues of ρ_ϕ , for all $k \leq n$.

Nielsen’s theorem has a number of consequences. In particular, it lays bare the fact that (if $n, m \geq 3$) there are fundamentally different types of entanglement that cannot be LOCC-converted to each other: there exists a pair of states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$ such that $|\psi\rangle \not\rightarrow |\phi\rangle$ and $|\phi\rangle \not\rightarrow |\psi\rangle$. Nielsen conjectured that this phenomenon is actually *typical*: if $|\psi\rangle, |\phi\rangle$ are independent random states (both sampled from the Haar measure, with respect to unitary transformations, on the unit sphere in $\mathbb{C}^n \otimes \mathbb{C}^m$), then the probability of the event $|\psi\rangle \rightarrow |\phi\rangle$ tends to zero as $n, m \rightarrow \infty$. This conjecture has remained unproved for the last 25 years, though it is supported by strong numerical evidence (cf. the very convincing computations of Cunden, Facchi, Florio and Gramegna [11]), and there are quite convincing heuristic arguments for why it should be true (a brief heuristic argument was given by Nielsen [23], and another heuristic argument via integral geometry was given by Życzkowski and Bengtsson [36]). Also, the *infinite-dimensional* counterpart to Nielsen’s conjecture was proved by Clifton, Hepburn and Wüthrich [10].

The main purpose of this paper is to prove Nielsen’s conjecture. Due to Nielsen’s characterisation of LOCC-convertibility in terms of majorisation, we can actually completely abandon the language of quantum information theory: it was observed by Życzkowski and Sommers [37] that for a random pure state $|\psi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$, the density operator ρ_ψ can be interpreted as a random matrix sampled from the *trace-normalised complex Wishart–Laguerre ensemble*, defined as follows.

Definition. For parameters $n, m \in \mathbb{N}$, the *complex Wishart–Laguerre ensemble* is the distribution of the random Hermitian matrix GG^\dagger , where “ \dagger ” denotes Hermitian transpose, and $G \in \mathbb{C}^{n \times m}$ is an $n \times m$ matrix with independent complex standard Gaussian entries³. Moreover, by the *trace-normalised complex Wishart–Laguerre ensemble*, we mean the distribution of $M/\text{tr}(M)$, where M is sampled from the complex Wishart–Laguerre ensemble.

So, Nielsen’s conjecture is equivalent to the following theorem, which is the main result of this paper.

Theorem 1.1. *Let $A^{(1)}, A^{(2)} \in \mathbb{C}^{n \times n}$ be independent samples from the trace-normalised complex Wishart–Laguerre ensemble with parameters n, m . Let $\lambda_1^{(1)} \geq \dots \geq \lambda_n^{(1)} \geq 0$ and $\lambda_1^{(2)} \geq \dots \geq \lambda_n^{(2)} \geq 0$ be the eigenvalues of $A^{(1)}$ and $A^{(2)}$, respectively. Then*

$$\mathbb{P}[\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)} \text{ for all } k \leq n] \rightarrow 0$$

as $n, m \rightarrow \infty$.

We discuss the proof approach for [Theorem 1.1](#) in a bit more detail later in this introduction ([Section 1.3](#)), but to give a very brief flavour: due to the so-called *eigenvalue repulsion* phenomenon, the (complex) Wishart–Laguerre spectrum has only tiny fluctuations, and existing technology in random matrix theory is not precise enough to describe the distributions of quantities of the form $\lambda_k + \dots + \lambda_n$. So, instead of considering such quantities directly, we consider a sequence of convex test functions which emphasise different parts of the spectrum (increasingly “focusing in” on the “edges” of the spectrum); this defines a sequence of linear statistics with nearly independent fluctuations, for which we can derive a multivariate central limit theorem. We can then take advantage of the relationship between convexity and majorisation, together with some analytic and combinatorial estimates.

Remark. Our proof approach, as written, is not quantitative; we prove that the desired majorisation probability converges to zero without proving any explicit bounds on this probability. In principle, it may be possible to establish quantitative central limit theorems for the complex Wishart–Laguerre spectrum, which would provide an explicit bound, though it seems that substantial new ideas would be required to precisely characterise the asymptotic rate of decay of the majorisation probability (computer experiments by Cunden, Facchi, Florio and Gramegna [11] indicate that the probability has order of magnitude $n^{-\theta}$ for some constant $\theta > 0$, assuming that the ratio m/n is held constant).

Remark. We have only discussed LOCC-convertibility of *bipartite* quantum states, and it is natural to also ask about *multipartite* states with more than two subsystems. The theory of LOCC-convertibility becomes much more challenging in this general setting (in the tripartite case, a necessary and sufficient condition is known [31], but it is rather complicated). However, our proof of Nielsen’s conjecture automatically implies a similar result for any number of subsystems (e.g., if there are three subsystems, we can simply view two of the subsystems as comprising a single larger subsystem to reduce to the bipartite case; this only makes LOCC-convertibility more permissive).

In addition to [Theorem 1.1](#), we also prove some other related results, which we describe next.

³There are two slightly different conventions for the definition of “complex standard Gaussian” (differing by a factor of $\sqrt{2}$). Due to the trace-normalisation, the distinction is irrelevant, but for concreteness we use the normalisation that for complex standard Gaussian X we have $\mathbb{E}[(\Re X)^2 + (\Im X)^2] = 1$.

1.1. The uniform measure on the simplex. For any possible outcome M of the trace-normalised complex Wishart–Laguerre ensemble, the eigenvalues are nonnegative real numbers that sum to 1. That is to say, the spectrum can be interpreted as a point in the $(n - 1)$ -simplex

$$\Delta_{n-1} = \{\vec{x} \in [0, \infty)^n : x_1 + \dots + x_n = 1\},$$

and Nielsen’s conjecture can be interpreted as a conjecture about majorisation of two independent random points sampled from a particular probability distribution (defined in random-matrix-theoretic terms) on Δ_{n-1} .

Of course, there are other interesting distributions on Δ_{n-1} ; perhaps the most natural (and most tractable) such measure is the *uniform* measure on Δ_{n-1} . Cunden, Facchi, Florio and Gramegna [12] proved the natural analogue of Nielsen’s conjecture for this measure. Their proof was fundamentally non-quantitative (it critically used Kolmogorov’s zero-one law), but based on numerical experiments and heuristic comparison with persistence probabilities of random walks, they predicted that the majorisation probability should decay “polynomially fast” (specifically, they predicted that the majorisation probability should scale like $n^{-\theta}$ for some positive constant $\theta \approx 0.41$). We are able to prove a bound of this shape (though presumably not with an optimal value of θ), as follows.

Theorem 1.2. *Let $(X_k^{(1)})_{k=1}^n, (X_k^{(2)})_{k=1}^n$ be two independent uniform random vectors in Δ_{n-1} , and denote by $(\lambda_k^{(1)})_{k=1}^n, (\lambda_k^{(2)})_{k=1}^n$ their decreasing rearrangements (i.e., the results of sorting $(X_k^{(1)})_{k=1}^n, (X_k^{(2)})_{k=1}^n$ in decreasing order). Then*

$$\mathbb{P}[\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)} \text{ for all } k \leq n] = O((n/\log n)^{-1/16}).$$

As observed by Cunden, Facchi, Florio and Gramegna [12], this result actually has a quantum-information-theoretic interpretation: in the same way that Nielsen’s conjecture studies the resource of *entanglement* (i.e., the operations that can be made without introducing new entanglement into the system), [Theorem 1.2](#) can be interpreted as a theorem about the resource of *coherence*. We direct the reader to [12], and the references therein, for more details.

1.2. Approximate LOCC-convertibility. So far, we have only been concerned about *exact* LOCC-convertibility: we write $|\psi\rangle \rightarrow |\phi\rangle$ if it is possible for Alice and Bob to orchestrate a sequence of local quantum operations (with unlimited classical communication) which *guarantee*, with probability 1, that $|\psi\rangle$ will be converted to $|\phi\rangle$. It is natural to ask how the situation changes if we allow some probability of failure.

Shortly after Nielsen’s 1999 paper, Vidal [34] found a generalisation of Nielsen’s theorem along these lines. For bipartite states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$, he proved that the maximum possible success probability $\Pi(|\psi\rangle \rightarrow |\phi\rangle)$ for the task of converting $|\psi\rangle$ to $|\phi\rangle$ (in the LOCC paradigm) is exactly

$$\min_{1 \leq k \leq n} \frac{\lambda_k^{(1)} + \lambda_{k+1}^{(1)} \dots + \lambda_n^{(1)}}{\lambda_k^{(2)} + \lambda_{k+1}^{(2)} \dots + \lambda_n^{(2)}},$$

where $\lambda_1^{(1)} \geq \dots \geq \lambda_n^{(1)}$ are the eigenvalues of ρ_ψ and $\lambda_1^{(2)} \geq \dots \geq \lambda_n^{(2)}$ are the eigenvalues of ρ_ϕ .

In [11], Cunden, Facchi, Florio and Gramegna considered $\Pi(|\psi\rangle \rightarrow |\phi\rangle)$ for *random* states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$: with what probability can one successfully convert between two randomly entangled states? In their computer simulations, they found that the answer seems to heavily depend (in a somewhat surprising way) on the relationship between n and m . Indeed, suppose $n, m \rightarrow \infty$ with $m/n = c$ (for some constant $c \in (0, \infty)$). If $c = 1$, computational evidence suggests that $\Pi(|\psi\rangle \rightarrow |\phi\rangle)$ has a nontrivial limiting distribution supported on the whole interval $[0, 1]$ (with mean about 0.58). On the other hand, if $c \neq 1$ then computational evidence suggests that $\Pi(|\psi\rangle \rightarrow |\phi\rangle)$ converges to 1 in probability.

As our final result in this paper, we confirm the latter prediction.

Theorem 1.3. *Let $A^{(1)}, A^{(2)} \in \mathbb{C}^{n \times n}$ be independent samples from the trace-normalised complex Wishart–Laguerre ensemble with parameters n, m . Let $\lambda_1^{(1)} \geq \dots \geq \lambda_n^{(1)}$ and $\lambda_1^{(2)} \geq \dots \geq \lambda_n^{(2)}$ denote the respective eigenvalues, and let*

$$\Pi = \min_{1 \leq k \leq n} \frac{\lambda_k^{(1)} + \lambda_{k+1}^{(1)} \dots + \lambda_n^{(1)}}{\lambda_k^{(2)} + \lambda_{k+1}^{(2)} \dots + \lambda_n^{(2)}}.$$

Then, for any fixed $\varepsilon > 0$ and $c \neq 1$, we have

$$\mathbb{P}[\Pi < 1 - \varepsilon] \rightarrow 0$$

as $n, m \rightarrow \infty$ with $m/n \rightarrow c$.

Remark 1.4. The conclusion of [Theorem 1.3](#) cannot hold without some separation between m and n . Indeed, in the case $m = n$, Edelman [15] showed that the smallest eigenvalue λ_n of a matrix drawn from the complex Wishart–Laguerre ensemble has probability density function given by $f(\lambda) = ne^{-\lambda n/2}/2$. Using the concentration of the trace (as in the proof of [Theorem 1.3](#)) along with the observation that $\Pi \leq \lambda_n^{(1)}/\lambda_n^{(2)}$, we see that for every $\varepsilon < 1$, there exists $c_\varepsilon > 0$ such that $\mathbb{P}[\Pi < 1 - \varepsilon] > c_\varepsilon$.

Our proof of [Theorem 1.3](#) really shows that $\mathbb{P}[\Pi < 1 - \varepsilon] \rightarrow 0$ as long as $|m - n|$ grows faster than $\sqrt{n \log n}$ (see [Theorem 4.1](#) for a precise statement), and it is conceivable that existing results in random matrix theory can be used to extend the above argument (showing $\mathbb{P}[\Pi < 1 - \varepsilon] > c_\varepsilon$) to the range $|m - n| = O(\sqrt{n})$. However, this would still leave a “gap” of order $\sqrt{\log n}$. It would be interesting to investigate this further.

1.3. Discussion of proof ideas. If $(\lambda_k^{(1)})_{k=1}^n$ and $(\lambda_k^{(2)})_{k=1}^n$ are the decreasing rearrangements of two independent random vectors on the simplex, sampled according to any continuous distribution, then for every *individual* k , it is easy to see by symmetry considerations that $\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)}$ with probability exactly 1/2. So, a first instinct for how to study majorisation of $(\lambda_k^{(1)})_{k=1}^n$ and $(\lambda_k^{(2)})_{k=1}^n$ (e.g., in the settings of [Theorems 1.1](#) and [1.2](#)) is to study the dependence between the events that $\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)}$ for different k . For example, if we consider a sequence of indices $k_1 < \dots < k_\ell$ which grow very rapidly or are in some other sense “well-spaced” from each other, one might hope that the corresponding events are approximately independent from each other, meaning that one might hope for an upper bound of about $2^{-\ell}$ on the majorisation probability. This type of approach has been very successful for some similar problems about majorisation in random settings (see for example [\[21, 25, 26\]](#) for its use in the study of random integer partitions).

This is the approach we take in our proof of [Theorem 1.2](#). Indeed, we employ an *exact description* of the joint distribution of $(\lambda_k)_{k=1}^n$, in terms of independent exponential random variables. Via some simple approximation arguments, we can relate the quantities $\lambda_1^{(1)} + \dots + \lambda_k^{(1)}$ to so-called *iterated partial sums*. There is a fairly extensive literature on persistence of iterated partial sums (see for example [\[3, 13, 30, 35\]](#)); in particular we take advantage of a general estimate due to Dembo, Ding and Gao [\[13\]](#).

Unfortunately, the above strategy seems to be extremely difficult to apply in the setting of [Theorem 1.1](#). In this setting there does not seem to be any convenient exact description of the distribution of $\lambda_k + \dots + \lambda_n$, so it seems we are forced to turn to limit theorems. There is a huge body of research on central limit theorems for certain statistics of Wishart–Laguerre ensembles (see for example [\[1, 2, 4, 6, 8, 16, 17, 19, 29\]](#)), but none of this applies to quantities of the form $\lambda_k + \dots + \lambda_n$. The difficulty here is that there is a so-called *eigenvalue repulsion* phenomenon for Hermitian random matrices: the eigenvalues “want to be separated from each other”, which has the effect of “locking the spectrum in place”. Specifically (at least in the case where m and n are approximately equal), for any bounded test function f , the fluctuations of $f(\lambda_1) + \dots + f(\lambda_n)$ are tiny, of comparable size to the eigenvalues themselves (and so there is no chance of describing the fluctuations in $\lambda_k + \dots + \lambda_n$ via a central limit theorem for linear statistics of the form $f(\lambda_1) + \dots + f(\lambda_n)$).

Instead, we take advantage of the relationship between majorisation and convex functions: if $(\lambda_k^{(1)})_{k=1}^n$ majorises $(\lambda_k^{(2)})_{k=1}^n$, then for any convex function $f : \mathbb{R} \rightarrow \mathbb{R}$ we have

$$f(\lambda_1^{(1)}) + \dots + f(\lambda_n^{(1)}) \geq f(\lambda_1^{(2)}) + \dots + f(\lambda_n^{(2)}).$$

So, we choose a sequence of convex test functions f_1, \dots, f_ℓ which “emphasise distinct parts of the spectrum” and whose corresponding statistics therefore fluctuate almost independently. Specifically, for each i , we choose the function f_i to grow much more rapidly than f_1, \dots, f_{i-1} , and therefore focus much more strongly on the upper edge (the so-called *soft edge*) of the spectrum. For simplicity, we will ultimately take the functions f_j to be polynomials of rapidly growing degree. We can then apply a general multivariate central limit theorem (with an adjustment for trace-normalisation) to the ℓ different statistics of the form $f_i(\lambda_1) + \dots + f_i(\lambda_n)$, and via analysis of certain integrals we can upper-bound the covariances between these statistics. The upshot is that we are able to show that the majorisation probability is at most about $2^{-\ell}$; since we can take ℓ arbitrarily large, the desired result follows.

The proof of [Theorem 1.3](#) is completely different. Since the trace is extremely well-concentrated, it suffices to prove a version of [Theorem 1.3](#) for eigenvalues of the *non-trace-normalised* complex Wishart–Laguerre ensemble. Writing μ_1, \dots, μ_n for these non-normalised eigenvalues, our approach is to prove that each individual eigenvalue μ_k is tightly concentrated (in a multiplicative sense) around its expected value $\mathbb{E}\mu_k$. Our proof of this fact relies on two ingredients: first, by Weyl’s inequalities, $\mu_k^{1/2}$ can be

interpreted as a 1-Lipschitz function (with respect to the Euclidean norm) of a Gaussian random vector, so it follows from the Gaussian concentration inequality of Sudakov–Tsirelson and Borell that μ_k is well-concentrated in an *additive* sense. Second, since $m \geq n + C\sqrt{n \log n}$, it follows from known estimates on the smallest singular value of random matrices that $\mathbb{E}[\mu_n^{1/2}]$ is sufficiently large to convert the additive error to multiplicative error. The second ingredient is the only place where we use the separation between m and n ; as we discuss in [Remark 1.4](#), this is necessary.

2. PROOF OF NIELSEN’S CONJECTURE

In this section we present the proof of [Theorem 1.1](#) (Nielsen’s conjecture), modulo some analytic estimates that will be deferred to [Section 5](#).

[Theorem 1.1](#) concerns parameters m, n which both tend to infinity, without making any assumption about the relationship between m and n . However, it is straightforward to reduce to the case where $m \geq n$ (by the symmetry between m and n), and it is then straightforward to reduce to the case where $m/n \rightarrow c$ for some constant $c \in [1, \infty]$ (by compactness). There are certain complications in the $c = \infty$ (“imbalanced”) case that are not present in the $c < \infty$ (“approximately balanced”) case, so we choose to present these two cases separately.

2.1. Majorisation and convexity. First, as discussed in [Section 1.3](#), it is crucial that instead of working with majorisation directly, we can work with convex test functions. To this end, we need the following well-known theorem (appearing, for example, as [\[27, p. 259, Theorem B\]](#)).

Theorem 2.1. *Consider any real numbers $\lambda_1^{(1)} \geq \dots \geq \lambda_n^{(1)}$ and $\lambda_1^{(2)} \geq \dots \geq \lambda_n^{(2)}$, such that $\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)}$ for all $k \in \{1, \dots, n-1\}$ and $\lambda_1^{(1)} + \dots + \lambda_n^{(1)} = \lambda_1^{(2)} + \dots + \lambda_n^{(2)}$. Then for any convex function $f : \mathbb{R} \rightarrow \mathbb{R}$, we have*

$$f(\lambda_1^{(1)}) + \dots + f(\lambda_n^{(1)}) \leq f(\lambda_1^{(2)}) + \dots + f(\lambda_n^{(2)}).$$

In particular, to show that a sequence $(\lambda_k^{(1)})_{k=1}^n$ is *not* majorised by $(\lambda_k^{(2)})_{k=1}^n$, it suffices to exhibit a convex function f for which $\sum_{k=1}^n f(\lambda_k^{(1)}) > \sum_{k=1}^n f(\lambda_k^{(2)})$. To prove [Theorem 1.1](#), we will consider an explicit sequence of convex test functions f_i , and show that the events of the form $\sum_{k=1}^n f_i(\lambda_k^{(1)}) > \sum_{k=1}^n f_i(\lambda_k^{(2)})$ (for different i) are nearly independent from each other, meaning that it is very likely that at least one of these events holds (and therefore $(\lambda_k^{(1)})_{k=1}^n$ is likely not majorised by $(\lambda_k^{(2)})_{k=1}^n$).

2.2. A central limit theorem for linear statistics, in the approximately balanced case. The other key general ingredient we will need is a (joint) central limit theorem for linear statistics of the complex Wishart–Laguerre spectrum. To start with, we present a central limit theorem that is suitable for the $c < \infty$ case (i.e., the case where m and n are approximately balanced). The statement of our central limit theorem requires some preparation.

Definition 2.2. For continuous functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, define

$$\Gamma(f, g) = \frac{1}{\pi^2} \int_{-1}^1 \int_{-1}^1 \left(\frac{f(x) - f(y)}{x - y} \right) \left(\frac{g(x) - g(y)}{x - y} \right) \frac{1 - xy}{\sqrt{1 - x^2} \sqrt{1 - y^2}} dx dy.$$

Then, for $c \in [1, \infty)$, let $a_{\pm} = (1 \pm \sqrt{c})^2$ and $\Phi_c(x) = 2\sqrt{c}x + c + 1$, and for continuous $f, g : \mathbb{R} \rightarrow \mathbb{R}$, define

$$\gamma_c(f) = \frac{1}{2\pi} \int_{a_-}^{a_+} f(x) \frac{\sqrt{(a_+ - x)(x - a_-)}}{x} dx, \quad \Gamma_c(f, g) = \Gamma(f \circ \Phi_c, g \circ \Phi_c).$$

Also, let \mathcal{F} be the set of all continuously differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with “sub-exponential growth”, in the sense that $(\log |f(x)|)/|x| \rightarrow 0$ when $|x| \rightarrow \infty$.

Theorem 2.3. *Fix functions $f_1, \dots, f_\ell \in \mathcal{F}$ and let μ_1, \dots, μ_ℓ be the eigenvalues of a random matrix sampled from the complex Wishart–Laguerre ensemble with parameters m, n . Suppose $m, n \rightarrow \infty$ in such a way that $m/n \rightarrow c \in [1, \infty)$. For each $i \in \{1, \dots, \ell\}$ let*

$$X_i = \sum_{k=1}^n f_i(\mu_k/n).$$

Then the following two asymptotic properties hold.

(1) For each $i \in \{1, \dots, \ell\}$, we have the L^1 -convergence

$$\mathbb{E}|X_i/n - \gamma_c(f_i)| \rightarrow 0.$$

(2) We have the multivariate convergence in distribution

$$(X_1 - \mathbb{E}X_1, \dots, X_\ell - \mathbb{E}X_\ell) \xrightarrow{d} \mathcal{N}(\vec{0}, \Sigma),$$

where the limiting covariance matrix $\Sigma \in \mathbb{R}^{\ell \times \ell}$ has (i, j) -entry equal to $\Gamma_c(f_i, f_j)$.

Theorem 2.3 can be proved by combining a few different results in the literature. The asymptotic formulas for the means in (1) go back to the seminal work of Marchenko and Pastur [20], and a central limit theorem was first proved by Arharov [2]. However, Arharov did not provide formulas for the covariances; for the covariance formulas in (2), we use a result of Lytova and Pastur [19]. There are three differences between these results and **Theorem 2.3**: first, our result is valid for a larger class of test functions; second, in (1), we obtain convergence in mean (whereas previous results obtain almost-sure convergence or convergence in probability—the distinction will be important for our application), and third, our result is a multivariate central limit theorem (whereas most previous results are univariate). In **Appendix A** we show how to deduce **Theorem 2.3** from previous results (the deduction is standard, but we include this for completeness).

Remark. It is perhaps more common to study the *real* Wishart–Laguerre ensemble, which is the distribution of AA^\top when $A \in \mathbb{R}^{n \times m}$ is a random matrix with independent *real* standard Gaussian entries. The distinction between the real and complex case does not affect **Theorem 2.3(1)**, and affects the limiting covariances in **Theorem 2.3(2)** by a factor of exactly 2 (i.e., the covariances are twice as large in the real case as in the complex case); see e.g. [19, Remark 4.1].

In our proof of **Theorem 1.1**, the only reason we need **Theorem 2.3** is for the following corollary, providing a central limit theorem for sums of powers of trace-normalised eigenvalues.

Corollary 2.4. *Let μ_1, \dots, μ_n be the eigenvalues of a random matrix sampled from the complex Wishart–Laguerre ensemble with parameters $m \geq n$ and let $\lambda_1, \dots, \lambda_n$ denote the trace-normalised eigenvalues. Fix ℓ , and for each $i \in \{1, \dots, \ell\}$ let*

$$X_i = \sum_{k=1}^n (\mu_k/n)^i \quad \text{and} \quad Y_i = \sum_{k=1}^n \lambda_k^i.$$

Suppose $n, m \rightarrow \infty$ in such a way that $m/n \rightarrow c$, for some fixed $c \in [1, \infty)$. Then we have multivariate convergence in distribution

$$\left(n \left(Y_i \cdot \frac{(\mathbb{E}X_1)^i}{\mathbb{E}X_i} - 1 \right) \right)_{i \leq \ell} \xrightarrow{d} \left(\frac{Z_i}{\gamma_c(x^i)} - i \frac{Z_1}{\gamma_c(x)} \right)_{i \leq \ell}$$

where $(Z_i)_{i \leq \ell} \sim \mathcal{N}(\vec{0}, \Sigma)$ for a matrix $\Sigma \in \mathbb{R}^{\ell \times \ell}$ with (i, j) -entry $\Gamma_c(x^i, x^j)$.

(Here we are abusing notation slightly, writing x^i to indicate the function $x \mapsto x^i$).

Proof. Writing $\tilde{Z}_i = X_i - \mathbb{E}X_i$, note that by **Theorem 2.3** we have convergence in distribution $(\tilde{Z}_i)_{i \leq \ell} \rightarrow (Z_i)_{i \leq \ell}$. We may then expand

$$Y_i \cdot \frac{(\mathbb{E}X_1)^i}{\mathbb{E}X_i} = \frac{X_i}{X_1^i} \cdot \frac{(\mathbb{E}X_1)^i}{\mathbb{E}X_i} = \left(1 + \frac{\tilde{Z}_i}{\mathbb{E}X_i} \right) \left(1 + \frac{\tilde{Z}_1}{\mathbb{E}X_1} \right)^{-i} = 1 + \frac{\tilde{Z}_i}{\mathbb{E}X_i} - \frac{i\tilde{Z}_1}{\mathbb{E}X_1} + O(n^{-2}). \quad (2.1)$$

Here, asymptotics are “in probability”, treating c, i as constants: the notation $O(n^{-2})$ denotes a random variable $E_{c,i}$ such that $n^2 E_{c,i}$ is bounded in probability (as $n \rightarrow \infty$, holding c, i fixed). This estimate for $E_{c,i}$ follows from the fact that the random variables \tilde{Z}_i are bounded in probability (which is a consequence of **Theorem 2.3(2)**) and the fact that $\mathbb{E}X_i = n\gamma_c(x^i) + o(n)$ has order of magnitude n (here we used **Theorem 2.3(1)**, and the observation that $\gamma_c(x^i) \in (0, \infty)$ for all $c \in [1, \infty)$ and $i \geq 1$).

Rearranging (2.1) and using again that $\mathbb{E}X_i = n\gamma_c(x^i) + o(n)$ completes the proof. \square

Remark. Our choice of test functions $x \mapsto x^i$ is mainly to keep the proof of **Corollary 2.4** (and its counterpart for the imbalanced case, **Corollary 2.8**) as simple as possible. If one were interested in optimising the quantitative aspects of our proof strategy, one should presumably consider alternative test functions.

2.3. The covariance structure, and putting the pieces together. In order to use [Corollary 2.4](#), we will need some control over the quantities γ_c and Γ_c , which describe the limiting means and covariances of our test functions. This is the content of the next lemma, whose proof we defer to [Section 5](#). Recall the definition of a_+ from [Definition 2.2](#).

Lemma 2.5. *Fix $c \in [1, \infty)$. For $i \in \mathbb{N}$, let $h_i(x) = (x/a_+)^i$.*

- (1) *We have $\lim_{i \rightarrow \infty} i\gamma_c(h_i) = 0$.*
- (2) *The limit $\alpha_c = \lim_{i \rightarrow \infty} \Gamma_c(h_i, h_i)$ exists, and satisfies $\alpha_c > 0$.*
- (3) *We have $\lim_{A \rightarrow \infty} \lim_{i \rightarrow \infty} \Gamma_c(h_i, h_{Ai}) = 0$.*

Recall the random variables $Y_i = \sum_{k=1}^n \lambda_k^i$ from [Corollary 2.4](#), which can be interpreted as linear statistics associated with convex test functions $f_i : x \mapsto x^i$. The estimates in [Lemma 2.5](#) can be interpreted as saying that the fluctuations of Y_i and Y_j are nearly independent, as long as i is very large and j is much larger than i (specifically, (1) can be used to show that correlations due to trace-normalisation are not very impactful, and (2) and (3) can then be used together to show that the covariance between the non-trace-normalised random variables X_i and X_j is negligible compared to the variances of X_i and X_j individually).

Remark. One can compute the relevant integrals in [Lemma 2.5](#), and apply them to prove [Theorem 1.1](#), without any intuitive understanding of *why* estimates of this type should hold. However, roughly speaking, the picture to keep in mind is that if j is much larger than i , then the largest eigenvalues (and their fluctuations) contribute much more strongly to Y_j than to Y_i . That is to say, Y_j is dominated by a few very large eigenvalues, which play only a negligible role in Y_i . So, intuitively speaking, [Lemma 2.5](#) corresponds to the fact that different parts of the Wishart–Laguerre spectrum have nearly independent fluctuations, and the sizes of these fluctuations do not significantly decay as we approach the upper edge of the spectrum (the latter is important as there is a “competition” between these fluctuations and the fluctuations due to trace-normalisation). We remark that the *lower* edge of the spectrum does not enjoy these properties: since all eigenvalues are always at least zero, the fluctuations of the smallest eigenvalues are more constrained. The upper and lower edges of the spectrum are sometimes called the “soft edge” and “hard edge” for this reason.

Given the preparations in this section so far, we can now prove Nielsen’s conjecture in the case where $m/n \rightarrow c$ for $c \in [1, \infty)$.

Proof of [Theorem 1.1](#), in the case $m/n \rightarrow c \in [1, \infty)$. Let \mathcal{E}_{maj} be the event that $\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)}$ for all $k \leq n$. Recall that we wish to show that $\mathbb{P}[\mathcal{E}_{\text{maj}}] \rightarrow 0$, when $m, n \rightarrow \infty$ in such a way that $m/n \rightarrow c \in [1, \infty)$.

For $s \in \{1, 2\}$ define $Y_i^{(s)} = \sum_{j=1}^n (\lambda_j^{(s)})^i$. By convexity of the functions $x \mapsto x^i$ on $[0, \infty)$, and [Theorem 2.1](#), we see that on the event \mathcal{E}_{maj} we must have $Y_i^{(1)} \leq Y_i^{(2)}$. By [Corollary 2.4](#) this implies that for any ℓ , we have

$$\limsup_{n \rightarrow \infty} \mathbb{P}[\mathcal{E}_{\text{maj}}] \leq \mathbb{P} \left[Z_i^{(1)} \leq Z_i^{(2)} - \frac{i\gamma_c(x^i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \leq \ell \right],$$

where $(Z_i^{(1)})_{i \leq \ell}$ and $(Z_i^{(2)})_{i \leq \ell}$ are independent Gaussian vectors as defined in [Corollary 2.4](#). Since the left-hand-side is independent of ℓ , we may take $\ell \rightarrow \infty$, so that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}[\mathcal{E}_{\text{maj}}] &\leq \liminf_{\ell \rightarrow \infty} \mathbb{P} \left[Z_i^{(1)} \leq Z_i^{(2)} - \frac{i\gamma_c(x^i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \leq \ell \right] \\ &= \mathbb{P} \left[Z_i^{(1)} \leq Z_i^{(2)} - \frac{i\gamma_c(x^i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \right]. \end{aligned}$$

Define $W_i^{(s)} = Z_i^{(s)}/a_+^i$, and let $h_i(x) = (x/a_+)^i$ as in [Lemma 2.5](#). Note that the covariance between $W_i^{(s)}$ and $W_j^{(s)}$ is precisely $\Gamma_c(h_i, h_j)$, and note that

$$\mathbb{P} \left[Z_i^{(1)} \leq Z_i^{(2)} - \frac{i\gamma_c(x^i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \right] = \mathbb{P} \left[W_i^{(1)} \leq W_i^{(2)} - \frac{i\gamma_c(h_i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \right].$$

Now, for each $j, q, A \in \mathbb{N}$, set $i_j(q, A) = qA^j$. So, by [Lemma 2.5](#), as we send $A, q \rightarrow \infty$ (where $A \rightarrow \infty$ much more rapidly than $q \rightarrow \infty$), we observe:

- (1) $i_j(q, A)\gamma_c(h_{i_j(q, A)}) \rightarrow 0$;

- (2) the variances of the individual $W_{i_j(q,A)}^{(s)}$ tend to $\alpha_c > 0$;
- (3) the covariances between different $W_{i_j(q,A)}^{(s)}$ tend to zero.

So, letting $(\zeta_j^{(1)}, \zeta_j^{(2)})_{j \leq N}$ be i.i.d. (real) Gaussians with mean zero and variance $\alpha_c > 0$, we have

$$\begin{aligned} & \mathbb{P} \left[W_i^{(1)} \leq W_i^{(2)} - \frac{i\gamma_c(h_i)}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } i \right] \\ & \leq \lim_{A \rightarrow \infty} \lim_{q \rightarrow \infty} \mathbb{P} \left[W_{i_j}^{(1)} \leq W_{i_j}^{(2)} - \frac{i_j(q,A)\gamma_c(h_{i_j(q,A)})}{\gamma_c(x)} (Z_1^{(2)} - Z_1^{(1)}) \text{ for all } j \leq N \right] \\ & = \mathbb{P}[\zeta_j^{(1)} \leq \zeta_j^{(2)} \text{ for all } j \leq N] \\ & = 2^{-N}. \end{aligned}$$

Taking $N \rightarrow \infty$ completes the proof. \square

2.4. The imbalanced case. In this subsection, we present analogues of [Theorem 2.3](#), [Corollary 2.4](#), and [Lemma 2.5](#) for the “imbalanced” case where $m/n \rightarrow \infty$. We then deduce the statement of [Theorem 1.1](#) in this case, and provide the details for how to deduce the full statement of [Theorem 1.1](#) from the approximately-balanced and imbalanced cases.

The main technical difficulty in the imbalanced case is that the spectrum of the complex Wishart-Laguerre ensemble is concentrated around m (i.e., typically all eigenvalues are of the form $m - o(m)$). This means that linear statistics $Y_i = \sum_{k=1}^n \lambda_k^i$ of the type considered in [Corollary 2.4](#) have variance $o(1)$, and their limiting distribution is degenerate.

The root of the problem is that the normalisation in [Theorem 2.3](#) (considering statistics of the form $\sum_{k=1}^n f(\mu_k/n)$) is not appropriate in the imbalanced regime, and to get a sensible central limit theorem, one should renormalise with an “additive shift”: namely, one should consider statistics of the form $\sum_{k=1}^n f((\mu_k - m)/\sqrt{nm})$. The following central limit theorem is an analogue of [Theorem 2.3](#) with this renormalisation.

Definition 2.6. Recall the set of functions \mathcal{F} with “sub-exponential growth”, and the notation $\Gamma(f, g)$, from [Definition 2.2](#). For a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$, let

$$\gamma(f) = \frac{1}{\pi} \int_{-1}^1 f(x) \sqrt{1-x^2} dx.$$

Theorem 2.7. Fix functions $f_1, \dots, f_\ell \in \mathcal{F}$ and let μ_1, \dots, μ_n be the eigenvalues of a random matrix sampled from the complex Wishart-Laguerre ensemble with parameters m, n . Suppose $m, n \rightarrow \infty$ in such a way that $m/n \rightarrow \infty$. For each $i \in \{1, \dots, \ell\}$ let

$$X_i = \sum_{k=1}^n f_i \left(\frac{\mu_k - m}{2\sqrt{nm}} \right).$$

Then the following two asymptotic properties hold.

- (1) For each $i \in \{1, \dots, \ell\}$, we have the L^1 -convergence

$$\mathbb{E}|X_i/n - \gamma(f_i)| \rightarrow 0.$$

- (2) We have the multivariate convergence in distribution

$$(X_1 - \mathbb{E}X_1, \dots, X_\ell - \mathbb{E}X_\ell) \xrightarrow{d} \mathcal{N}(\vec{0}, \Sigma),$$

where $\Sigma \in \mathbb{R}^{\ell \times \ell}$ is a matrix with (i, j) -entry $\Gamma(f_i, f_j)$.

As for [Theorem 2.3](#), we can prove [Theorem 2.7](#) by combining various results in the literature (the deduction can again be found in [Appendix A](#)). Specifically, the formulas for the means in [\(1\)](#) are essentially due to Bai and Yin [\[5\]](#)⁴, and a central limit theorem in this regime was proved independently by Bao [\[6\]](#) and by Chen and Pan [\[8\]](#).

⁴In random matrix theory language, when $m/n \rightarrow c < \infty$ (as $n \rightarrow \infty$), the spectrum converges to the *Marchenko-Pastur law* with parameter c . As $c \rightarrow \infty$, it is not hard to see that the Marchenko-Pastur law (with an appropriate normalisation) converges to the *semicircular law*. Bai and Yin proved that when $m/n \rightarrow \infty$, the (suitably-normalised) spectrum converges to the semicircular law, thus showing that the $n \rightarrow \infty$ and $c \rightarrow \infty$ limits commute.

Remark. We have stated [Theorem 2.7](#) only for the case $m/n \rightarrow \infty$ (as we have already handled the case where m/n converges to a finite limit), but we remark that it would be possible to state a more general (albeit somewhat complicated) central limit theorem that holds for all $m, n \rightarrow \infty$, and use this to give a unified proof of [Theorem 1.1](#) without a case distinction.

Now, as an analogue of [Corollary 2.4](#), we next deduce a corollary of [Theorem 2.7](#) for powers of (suitably shifted) trace-normalised eigenvalues. Due to the shifting, the deduction is somewhat more complicated than for [Corollary 2.4](#).

Corollary 2.8. *Let μ_1, \dots, μ_n be the eigenvalues of a random matrix sampled from the complex Wishart–Laguerre ensemble with parameters n, m and let $\lambda_1, \dots, \lambda_n$ denote the trace-normalised eigenvalues. Fix ℓ , and for each $i \in \{1, \dots, 2\ell\}$ let*

$$X_i = \sum_{k=1}^n \left(\frac{\mu_k - m}{2\sqrt{mn}} \right)^i \quad \text{and} \quad Y_i = \sum_{k=1}^n \left(\lambda_k - \frac{1}{n} \right)^i.$$

Suppose $n, m \rightarrow \infty$ in such a way that $m/n \rightarrow \infty$. Then we have multivariate convergence in distribution

$$\left(n \left(Y_{2i} \cdot \frac{(mn/4)^i}{(\mathbb{E}X_{2i})} - 1 \right) \right)_{i \leq \ell} \xrightarrow{d} \left(\frac{Z_{2i}}{\gamma(x^{2i})} \right)_{i \leq \ell}$$

where $(Z_{2i})_{i \leq \ell} \sim \mathcal{N}(\vec{0}, \Sigma)$ for a matrix $\Sigma \in \mathbb{R}^{\ell \times \ell}$ with (i, j) -entry $\Gamma(x^{2i}, x^{2j})$.

In our proof of [Corollary 2.8](#), we will need the following basic fact.

Fact 2.9. *In the setting of [Corollary 2.8](#) we have $\mathbb{E}[\mu_1 + \dots + \mu_n] = mn$.*

Proof. Let $G \in \mathbb{R}^{n \times m}$ be a random matrix with independent standard complex Gaussian entries, in such a way that μ_1, \dots, μ_n are the eigenvalues of GG^\dagger . Then, we have

$$\mu_1 + \dots + \mu_n = \sum_{i=1}^n \sum_{j=1}^m |G_{ij}|^2,$$

and the desired result follows (recalling that the entries G_{ij} are complex standard Gaussian). \square

Proof of [Corollary 2.8](#). Throughout this proof, all implicit constants in asymptotic notation are allowed to depend on ℓ (i.e., we think of ℓ as a constant).

Writing $\tilde{Z}_i = X_i - \mathbb{E}X_i$, by [Theorem 2.7\(2\)](#) we have the convergence in distribution $(\tilde{Z}_i)_{i \leq 2\ell} \rightarrow (Z_i)_{i \leq 2\ell}$. Then, writing $T = \mu_1 + \dots + \mu_n$, note that

$$T = mn + 2\sqrt{nm}\tilde{Z}_1,$$

using [Fact 2.9](#). We may then write

$$Y_{2i} = T^{-2i} \sum_{k=1}^n \left(\mu_k - \frac{T}{n} \right)^{2i} = \left(\frac{4}{mn} \right)^i \left(1 + \frac{2\tilde{Z}_1}{\sqrt{mn}} \right)^{-2i} \sum_{k=1}^n \left(\frac{\mu_k - m}{2\sqrt{mn}} - \frac{\tilde{Z}_1}{n} \right)^{2i}. \quad (2.2)$$

Since \tilde{Z}_1 is bounded in probability by [Theorem 2.7\(2\)](#), we have that

$$\left(1 + \frac{2\tilde{Z}_1}{\sqrt{mn}} \right)^{-2i} = 1 + O\left(\frac{1}{\sqrt{mn}} \right) = 1 + o\left(\frac{1}{n} \right), \quad (2.3)$$

where asymptotics are in probability. Moreover, writing $\varepsilon = -\tilde{Z}_1/n$ we can expand

$$\begin{aligned} \sum_{k=1}^n \left(\frac{\mu_k - m}{2\sqrt{mn}} + \varepsilon \right)^{2i} &= \sum_{k=1}^n \sum_{j=0}^{2i} \binom{2i}{j} \left(\frac{\mu_k - m}{2\sqrt{mn}} \right)^{2i-j} \varepsilon^j \\ &= \sum_{j=0}^{2i} \binom{2i}{j} X_{2i-j} \varepsilon^j = \sum_{j=0}^{2i} \binom{2i}{j} (\mathbb{E}X_{2i-j} + \tilde{Z}_{2i-j}) \varepsilon^j. \end{aligned}$$

Now, recall from [Theorem 2.7\(1\)](#) that for each $i \leq 2\ell$ we have

$$\mathbb{E}X_i = n\gamma(x^i) + o(n).$$

For each $j \leq \ell$, observing that $\gamma(x^{2j}) \in (0, \infty)$ and $\gamma(x^{2j-1}) = 0$, it follows that $\mathbb{E}X_{2j} = O(n)$ and $\mathbb{E}X_{2j-1} = o(n)$. Also, by [Theorem 2.7\(2\)](#), we have that $\varepsilon = O(1/n)$ in probability. Therefore,

$$\sum_{k=1}^n \left(\frac{\mu_k - m}{2\sqrt{mn}} + \varepsilon \right)^{2i} = \mathbb{E}X_{2i} + \tilde{Z}_{2i} + o(1), \quad (2.4)$$

where the $o(1)$ term is in probability. The desired result then follows from [Equations \(2.2\) to \(2.4\)](#). \square

Finally, as an analogue of [Lemma 2.5](#), we need the following technical estimates on γ and H , which are proved in [Section 5](#).

Lemma 2.10.

- (1) The limit $\alpha = \lim_{i \rightarrow \infty} \Gamma(x^i, x^i)$ exists, and satisfies $\alpha > 0$.
- (2) We have $\lim_{A \rightarrow \infty} \lim_{i \rightarrow \infty} \Gamma(x^i, x^{Ai}) = 0$.

We can now complete the proof of [Theorem 1.1](#), supplementing the approximately balanced case in the last subsection with similar considerations in the imbalanced case (using [Corollary 2.8](#) and [Lemma 2.10](#)).

Proof of [Theorem 1.1](#). As before, let \mathcal{E}_{maj} be the event that $\lambda_k^{(1)} + \dots + \lambda_n^{(1)} \leq \lambda_k^{(2)} + \dots + \lambda_n^{(2)}$ for all $k \leq n$. We wish to show that $\mathbb{P}[\mathcal{E}_{\text{maj}}] \rightarrow 0$ as $m, n \rightarrow \infty$.

First, note that we can restrict our attention to the case where $m \geq n$. Indeed, recall that the complex Wishart–Laguerre ensemble with parameters n, m is the distribution of a random matrix of the form AA^\dagger , where $A \in \mathbb{C}^{n \times m}$ is a matrix with standard complex Gaussian entries. Switching the roles of m and n is equivalent to considering the matrix $A^\dagger A$, which has the same nonzero eigenvalues as AA^\dagger (if $n \leq m$, this switch merely introduces $m - n$ additional zero eigenvalues) and this is the only relevant information for the event \mathcal{E}_{maj} .

Second, note that if [Theorem 1.1](#) were not true, then there would be some sequence of parameter-pairs (n, m_n) , with $m_n \geq n \rightarrow \infty$, such that $\limsup_{n \rightarrow \infty} \mathbb{P}[\mathcal{E}_{\text{maj}}] > 0$. By compactness of $[1, \infty]$, there would then be an infinite subsequence of values of n along which $m_n/n \rightarrow c$ for some $c \in [1, \infty]$. So, in proving [Theorem 1.1](#) we may assume that $m_n/n \rightarrow c$ for some $c \in [1, \infty]$. In [Section 2.3](#) we have already handled the case where $c < \infty$, we may (and do) assume that $m/n \rightarrow \infty$.

For $s \in \{1, 2\}$, let $Y_i^{(s)} = \sum_{k=1}^n (\lambda_k - 1/n)^i$. By convexity of the functions $x \mapsto (x - 1/n)^{2i}$, it follows from [Corollary 2.8](#) and [Theorem 2.1](#) (as in the approximately balanced case) that

$$\limsup_{n \rightarrow \infty} \mathbb{P}[\mathcal{E}_{\text{maj}}] \leq \mathbb{P}[Z_{2i}^{(1)} \leq Z_{2i}^{(2)} \text{ for all } i],$$

where $(Z_i^{(1)})_i$ and $(Z_i^{(2)})_i$ are independent vectors distributed as in [Corollary 2.8](#).

Proceeding as in the approximately balanced case, but now using [Lemma 2.10](#), it follows that if we let $(\zeta_j^{(1)}, \zeta_j^{(2)})$ denote i.i.d. (real) Gaussians with mean zero and variance α , then

$$\mathbb{P}[Z_i^{(1)} \leq Z_i^{(2)} \text{ for all } i] \leq \lim_{N \rightarrow \infty} \mathbb{P}[\zeta_j^{(1)} \leq \zeta_j^{(2)} \text{ for all } j \leq N] = 0. \quad \square$$

3. THE UNIFORM MEASURE

In this section we prove [Theorem 1.2](#), for the uniform measure on the simplex. The starting point is that if (X_1, \dots, X_n) is uniformly distributed on the simplex, then the decreasing rearrangement $(\lambda_1, \dots, \lambda_n)$ has a *closed form description* in terms of independent random variables, as follows.

Theorem 3.1. *Let Z_1, \dots, Z_n be i.i.d. exponential random variables with mean 1. Then the normalised vector*

$$\frac{1}{Z_1 + \dots + Z_n} (Z_1, \dots, Z_n)$$

is uniformly distributed in the simplex Δ_{n-1} .

Theorem 3.2. *Let Z_1, \dots, Z_n be i.i.d. exponential random variables with mean 1, and let $Z_{(1)} < \dots < Z_{(n)}$ be their order statistics (i.e., $(Z_{(n-i+1)})_{i=1}^n$ is the decreasing rearrangement of $(Z_i)_{i=1}^n$). Then $(Z_{(i)})_{i=1}^n$ has the same distribution as*

$$\left(\frac{Z_1}{n}, \frac{Z_1}{n} + \frac{Z_2}{n-1}, \frac{Z_1}{n} + \frac{Z_2}{n-1} + \frac{Z_3}{n-2}, \dots, \frac{Z_1}{n} + \dots + \frac{Z_n}{1} \right).$$

[Theorems 3.1](#) and [3.2](#) are both classical. In particular, [Theorem 3.1](#) is a direct consequence of [[14](#), Chapter V, Theorems 2.1 and 2.2], and [Theorem 3.2](#) can be found in [[14](#), Chapter V, Section 3.3].

Now, before getting into the details, we briefly describe the plan to prove [Theorem 1.2](#). First, by a standard concentration inequality (e.g. Bernstein's inequality, see for example [[27](#), Theorem 2.8.1 and Example 2.7.8]), sums of exponential random variables do not fluctuate very much, as follows.

Theorem 3.3. *Let Z_1, \dots, Z_N be i.i.d. exponential random variables with mean 1, and let $U = a_1 Z_1 + \dots + a_N Z_N$ for some coefficients $a_1, \dots, a_N \in \mathbb{R}$. There is an absolute constant $c > 0$ such that*

$$\mathbb{P}[|U - \mathbb{E}U| \geq t] \leq \exp\left(-c \min\left(\frac{t^2}{a_1^2 + \dots + a_N^2}, \frac{t}{\max_i |a_i|}\right)\right)$$

for every $t \geq 0$.

Given [Theorem 3.3](#) (applied with $a_1 = \dots = a_n = 1$), we can approximate quantities of the form $\lambda_k + \dots + \lambda_n$ (as in the statement of [Theorem 1.2](#)) with sums of the form

$$\sum_{j=1}^{n-k+1} \sum_{i=1}^j \frac{Z_i}{n - (i-1)}$$

If $n - k$ is small, then each of the denominators $n - i + 1$ is very close to n (i.e., the denominators are all almost the same), meaning that it suffices to understand the *iterated partial sums* of the form $\sum_{j=1}^q \sum_{i=1}^j Z_i$. There is a fairly substantial literature on this topic; in particular, we take advantage of the following powerful theorem of Dembo, Ding and Gao [[13](#), Theorem 1.1], on *persistence probabilities* for iterated partial sums.

Theorem 3.4. *Let $(W_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables with zero mean and $0 < \mathbb{E}W_i^2 < \infty$. For $q \in \mathbb{N}$, let $S_q = \sum_{j=1}^q \sum_{i=1}^j W_i$. Then for any fixed $t \in \mathbb{R}$ we have*

$$\mathbb{P}[S_q < t \text{ for all } q \leq N] = \Theta_t(N^{-1/4}).$$

We now provide the full details of the proof of [Theorem 1.2](#).

Proof of [Theorem 1.2](#). Let $(Z_j)_{j=1}^n$ and $(Z'_j)_{j=1}^n$ be i.i.d. sequences of standard exponential random variables, and let $(Z_{(j)})_{j=1}^n$ and $(Z'_{(j)})_{j=1}^n$ be their order statistics. Let \mathcal{E}_{maj} be the event that

$$\frac{Z_{(1)} + \dots + Z_{(q)}}{Z_{(1)} + \dots + Z_{(n)}} \leq \frac{Z'_{(1)} + \dots + Z'_{(q)}}{Z'_{(1)} + \dots + Z'_{(n)}}$$

for all $q \leq n$. Recalling [Theorem 3.1](#), our goal is to prove that $\mathbb{P}[\mathcal{E}_{\text{maj}}] \leq O((n/\log n)^{-1/16})$.

By [Theorem 3.3](#), for a sufficiently large constant C we have $|Z_1 + \dots + Z_n - n| \leq C\sqrt{n \log n}$ with probability at most n^{-100} . So,

$$\mathbb{P}[\mathcal{E}_{\text{maj}}] \leq \mathbb{P}\left[Z_{(1)} + \dots + Z_{(q)} \leq \left(1 + 3C\sqrt{\frac{\log n}{n}}\right)(Z'_{(1)} + \dots + Z'_{(q)}) \text{ for all } q \leq n\right] + n^{-100}.$$

Recalling [Theorem 3.2](#), it suffices to prove that

$$\mathbb{P}\left[\sum_{j=1}^q \sum_{i=1}^j \frac{Z_i}{n - (i-1)} \leq \left(1 + 3C\sqrt{\frac{\log n}{n}}\right) \sum_{j=1}^q \sum_{i=1}^j \frac{Z'_i}{n - (i-1)} \text{ for all } q \leq n\right] \leq O((n/\log n)^{-1/16}). \quad (3.1)$$

Let $Q = \lfloor c(n/\log n)^{1/4} \rfloor$ for some small constant $c > 0$. By [Theorem 3.3](#), with probability at least $1 - n^{-100}$ we have

$$\sum_{j=1}^Q \sum_{i=1}^j \frac{Z'_i}{n - (i-1)} = \sum_{i=1}^Q \left(\frac{Q - (i-1)}{n - (i-1)} \cdot Z'_i\right) \leq \frac{2Q^2}{n} \leq \frac{2c^2}{\sqrt{n \log n}},$$

and similarly, with probability at least $1 - n^{-100}$ we have

$$\left| \sum_{j=1}^k \sum_{i=1}^j \frac{Z_i - Z'_i}{n - (i-1)} - \sum_{j=1}^k \sum_{i=1}^j \frac{Z_i - Z'_i}{n} \right| = \left| \sum_{i=1}^k \frac{i-1}{n} \cdot \frac{k - (i-1)}{n - (i-1)} \cdot (Z_i - Z'_i) \right| \leq \frac{1}{2n}$$

for all $k \leq Q$. Assuming that c is sufficiently small (in terms of C), it follows that the probability in (3.1) is at most

$$\begin{aligned} & \mathbb{P} \left[\sum_{j=1}^q \sum_{i=1}^j \frac{Z_i}{n-(i-1)} \leq \sum_{j=1}^q \sum_{i=1}^j \frac{Z'_i}{n-(i-1)} + 3C \sqrt{\frac{\log n}{n}} \cdot \frac{2c^2}{\sqrt{n \log n}} \text{ for all } q \leq Q \right] + n^{-100} \\ & \leq \mathbb{P} \left[\sum_{j=1}^q \sum_{i=1}^j \frac{Z_i}{n-(i-1)} \leq \sum_{j=1}^q \sum_{i=1}^j \frac{Z'_i}{n-(i-1)} + \frac{1}{2n} \text{ for all } q \leq Q \right] + n^{-100} \\ & \leq \mathbb{P} \left[\sum_{j=1}^q \sum_{i=1}^j \frac{Z_i - Z'_i}{n} \leq \frac{1}{n} \text{ for all } q \leq Q \right] + 2n^{-100} \\ & = \mathbb{P} \left[\sum_{j=1}^q \sum_{i=1}^j (Z_i - Z'_i) \leq 1 \text{ for all } q \leq Q \right] + 2n^{-100}. \end{aligned}$$

This probability is at most $O(Q^{-1/4}) = O((n/\log n)^{-1/16})$, by [Theorem 3.4](#). \square

4. APPROXIMATE LOCC-CONVERTIBILITY

In this section, we prove [Theorem 1.3](#) in the following, more precise, form.

Theorem 4.1. *Let $M^{(1)}, M^{(2)} \in \mathbb{C}^{n \times n}$ be independent samples from the trace-normalised complex Wishart–Laguerre ensemble with parameters n, m . Let $\lambda_1^{(1)} \geq \dots \geq \lambda_n^{(1)}$ and $\lambda_1^{(2)} \geq \dots \geq \lambda_n^{(2)}$ denote the respective eigenvalues, and let*

$$\Pi = \min_{1 \leq k \leq n} \frac{\lambda_k^{(1)} + \lambda_{k+1}^{(1)} \cdots + \lambda_n^{(1)}}{\lambda_k^{(2)} + \lambda_{k+1}^{(2)} \cdots + \lambda_n^{(2)}}.$$

For every $\varepsilon, \delta > 0$, there exists a constant $C(\varepsilon, \delta) > 0$ such that if $m \geq n + C(\varepsilon, \delta)\sqrt{n \log n}$ and $n \geq C(\varepsilon, \delta)$, then $\mathbb{P}[\Pi < 1 - \varepsilon] \leq \delta$.

In our proof of [Theorem 4.1](#) we need a few general tools. First, we need concentration of the trace of the complex Wishart–Laguerre ensemble.

Lemma 4.2. *There is an absolute constant c such that the following holds. Let M be a sample from the complex Wishart–Laguerre distribution with parameters n, m . Then*

$$\mathbb{P}[|\operatorname{tr}(M) - m| > t] \leq \exp\left(-c \min\left(t^2/(mn), t\right)\right).$$

Proof. As in the proof of [Fact 2.9](#), we can write

$$\operatorname{tr}(M) = \sum_{i=1}^n \sum_{j=1}^m |G_{ij}|^2,$$

where the G_{ij} are independent standard complex Gaussians. The terms $|G_{ij}|^2$ are sub-exponential (see [\[33, Lemma 2.7.6\]](#)), so the desired result follows from Bernstein’s inequality (see [\[33, Theorem 2.8.1\]](#)). \square

We also need the Gaussian concentration inequality for Lipschitz functions, due independently to Borell and Sudakov–Tsirelson (see [\[33, Definition 2.5.6 and Theorem 5.2.2\]](#)).

Theorem 4.3. *There is an absolute constant $c > 0$ such that the following holds. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be a Lipschitz function (with respect to the Euclidean metric), with Lipschitz constant at most r . Let $X_1, \dots, X_n \in \mathbb{R}$ be i.i.d. standard Gaussian random variables, and let $Y = F(X_1, \dots, X_n)$. Then for any $t \geq 0$ we have*

$$\mathbb{P}[|Y - \mathbb{E}Y| \geq t] \leq 2 \exp(-ct^2/r^2).$$

For a matrix $G \in \mathbb{C}^{n \times m}$, let $\sigma_k(G)$ denote the k -th largest singular value of G and let $\|G\|_{\mathbb{F}} = \sqrt{\sum_{ij} |G_{ij}|^2}$ denote the Frobenius norm of G . The next ingredient we will need is a version of Weyl’s inequality. The following statement is a consequence of e.g. [\[32, Exercise 1.3.22\(iv\)\]](#), using the inequality $\|H\|_{\mathbb{F}} = (\operatorname{tr}(H^\dagger H))^{1/2} = (\sum_i \sigma_i(H)^2)^{1/2} \geq \sigma_1(H)$.

Theorem 4.4. *For any matrices $G, H \in \mathbb{C}^{n \times m}$, we have*

$$|\sigma_k(G + H) - \sigma_k(G)| \leq \|H\|_{\mathbb{F}}.$$

Finally, we need a lower bound on the expected least singular value of complex Gaussian rectangular matrices. The following estimate is a direct consequence of [18, Theorem 3.4], which is an adaptation of the main result of [28] to the complex case.

Lemma 4.5. *There is an absolute constant $c > 0$ such that the following holds. Let G be an $n \times m$ matrix with independent complex standard Gaussian entries. Then, for all $1 \leq k \leq n$,*

$$\mathbb{E}[\sigma_k(G)] \geq \mathbb{E}[\sigma_n(G)] \geq c(\sqrt{m} - \sqrt{n-1}). \quad (4.1)$$

Now we are ready to prove [Theorem 4.1](#).

Proof of Theorem 4.1. Throughout this proof, we treat ε, δ as constants. In particular, implicit constants in asymptotic notation are allowed to depend on ε, δ . Our goal is to prove that $\mathbb{P}[\Pi < 1 - \varepsilon] \leq \delta$, as long as n is sufficiently large and $m \geq n + C\sqrt{n \log n}$ for large enough C .

Let $G^{(1)}, G^{(2)}$ be independent $n \times m$ matrices with independent complex standard Gaussian entries, and for $s \in \{1, 2\}$ let $M^{(s)} = G^{(s)}(G^{(s)})^\dagger \in \mathbb{C}^{n \times n}$ denote the corresponding sample from the complex Wishart–Laguerre ensemble with parameters n, m . Let $\mu_1^{(s)} \geq \dots \geq \mu_n^{(s)} \geq 0$ denote the eigenvalues of $M^{(s)}$; we may take $\lambda_1^{(s)} \geq \dots \geq \lambda_n^{(s)} \geq 0$ to be the eigenvalues of $M^{(s)} / \text{tr}(M^{(s)})$.

By [Lemma 4.2](#), with probability at least $1 - \delta/2$ we have $\text{tr}(M^{(i)}) = mn + O(\sqrt{mn})$. Therefore, it suffices to show that $\mathbb{P}[\Pi_\mu < 1 - \varepsilon] \leq \delta/2$, where

$$\Pi_\mu = \min_{1 \leq k \leq n} \frac{\mu_k^{(1)} + \mu_{k+1}^{(1)} \cdots + \mu_n^{(1)}}{\mu_k^{(2)} + \mu_{k+1}^{(2)} \cdots + \mu_n^{(2)}}.$$

We will, in fact, show the stronger statement that

$$\mathbb{P}\left[\mu_k^{(1)} \in [(1 - \varepsilon)\mu_k^{(2)}, (1 + \varepsilon)\mu_k^{(2)}] \quad \text{for all } 1 \leq k \leq n\right] \geq 1 - \delta/2.$$

By the union bound, it suffices to show that

$$\mathbb{P}\left[\mu_k^{(1)} \in [(1 - \varepsilon)\mu_k^{(2)}, (1 + \varepsilon)\mu_k^{(2)}]\right] \geq 1 - \delta/(2n) \quad \text{for all } 1 \leq k \leq n. \quad (4.2)$$

For $s \in \{1, 2\}$, let $\sigma_k^{(s)} = \sqrt{\mu_k^{(s)}}$ be the k -th largest singular value of $G^{(s)}$. We can interpret $\sigma_k^{(s)} : \mathbb{C}^{n \times m} \rightarrow \mathbb{R}$ as being a function of $2mn$ i.i.d. standard (real) Gaussian random variables (the real part and imaginary part of each entry of $G^{(s)}$), and by Weyl’s inequality ([Theorem 4.4](#)), this function is 1-Lipschitz with respect to the Euclidean metric on $\mathbb{C}^{n \times m} \cong \mathbb{R}^{2nm}$. It follows from the Gaussian concentration inequality ([Theorem 4.3](#)) that for any $1 \leq k \leq n$ (and $s \in \{1, 2\}$),

$$\mathbb{P}\left[|\sigma_k^{(s)} - \mathbb{E}[\sigma_k^{(s)}]| \geq t\right] \leq 2 \exp(-t^2/2). \quad (4.3)$$

Note that our assumption $m \geq n + C\sqrt{n \log n}$, together with [Lemma 4.5](#), yields

$$\mathbb{E}[\sigma_k^{(s)}] \geq C' \sqrt{\log n}, \quad (4.4)$$

where C' can be made arbitrary large by taking large enough C .

We now prove (4.2). Applying (4.3) with $t = 10\sqrt{\log(n/\delta)} = O(\sqrt{\log n})$, and using (4.4), we see that with probability at least $1 - \delta/(2n)$, for $s \in \{1, 2\}$ we have $\sigma_k^{(s)} = \mathbb{E}\sigma_k^{(s)} + O(\sqrt{\log n}) = \mathbb{E}\sigma_k^{(s)}(1 + O(1/C'))$. Since $\mathbb{E}\sigma_k^{(1)} = \mathbb{E}\sigma_k^{(2)}$, this implies that

$$\mu_k^{(1)} = (\sigma_k^{(1)})^2 = \left(\sigma_k^{(2)}(1 + O(1/C'))\right)^2 = \mu_k^{(2)}(1 + O(1/C')).$$

The desired result follows, assuming C' is sufficiently large (which we may ensure by taking C sufficiently large in terms of ε). \square

5. INTEGRAL ESTIMATES

In this section we prove the estimates in [Lemma 2.5](#) and [Lemma 2.10](#), via a sequence of estimates of various integrals. In this section we write $f \lesssim g$ to mean $f = O(g)$.

First, for the convenience of the reader we recall the relevant parts of [Definitions 2.2](#) and [2.6](#).

Definition 5.1. For functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, let

$$\begin{aligned}\gamma(f) &= \frac{1}{\pi} \int_{-1}^1 f(x) \sqrt{1-x^2} dx, \\ \Gamma(f, g) &= \frac{1}{\pi^2} \int_{-1}^1 \int_{-1}^1 \left(\frac{f(x) - f(y)}{x-y} \right) \left(\frac{g(x) - g(y)}{x-y} \right) \frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}} dx dy.\end{aligned}$$

Then, for $a_{\pm} = (1 \pm \sqrt{c})^2$ and $\Phi_c(x) = 2\sqrt{c}x + c + 1$, let

$$\gamma_c(f) = \frac{1}{2\pi} \int_{a_-}^{a_+} f(x) \frac{\sqrt{(a_+ - x)(x - a_-)}}{x} dx, \quad \Gamma_c(f, g) = \Gamma(f \circ \Phi_c, g \circ \Phi_c).$$

Now, to begin with, we show that the kernel appearing in the definition of Γ in fact yields a density of a probability measure on $[-1, 1]^2$.

Fact 5.2. *We have*

$$\frac{1}{\pi^2} \int_{-1}^1 \int_{-1}^1 \frac{1-xy}{\sqrt{(1-x^2)(1-y^2)}} dx dy = 1.$$

Proof. With the substitution $(s, t) = (1-x, 1-y)$ and the symmetry between s and t , we compute

$$\begin{aligned}\frac{1}{\pi^2} \int_{-1}^1 \int_{-1}^1 \frac{1-xy}{\sqrt{(1-x^2)(1-y^2)}} dx dy &= \frac{1}{\pi^2} \int_0^2 \int_0^2 \frac{s+t-ts}{\sqrt{st(2-s)(2-t)}} ds dt \\ &= \frac{2}{\pi^2} \int_0^2 \int_0^2 \frac{\sqrt{s}}{\sqrt{t(2-s)(2-t)}} ds dt - \frac{1}{\pi^2} \int_0^2 \int_0^2 \frac{\sqrt{st}}{\sqrt{(2-s)(2-t)}} ds dt = 1,\end{aligned}$$

where we used the identities

$$\int_0^2 \sqrt{\frac{s}{2-s}} ds = 2 \int_0^1 \frac{x^{1/2}}{(1-x)^{1/2}} dx = \pi, \quad \int_0^2 \frac{1}{\sqrt{s(2-s)}} ds = \int_0^1 \frac{x^{-1/2}}{(1-x)^{1/2}} = \pi$$

(these are Beta integrals; see for example [22, Eqns. (6.1.6) and (6.1.7)]). \square

Now we proceed to study the asymptotics of the quantities appearing in Lemmas 2.5 and 2.10.

Lemma 5.3. *Fix $c \in [1, \infty)$, and let $h_k(x) = (x/a_+)^k$. There is a constant $C_c \in (0, \infty)$ so that*

$$\lim_{k \rightarrow \infty} k^{3/2} \gamma_c(h_k) = C_c.$$

Proof. It will be more convenient to work with h_{k+1} . By changing variables $x = c + 1 + (2\sqrt{c})t$, we see

$$\begin{aligned}\gamma_c(h_{k+1}) &= \frac{1}{2\pi} \int_{a_-}^{a_+} \left(\frac{x}{a_+} \right)^{k+1} \frac{\sqrt{(a_+ - x)(x - a_-)}}{x} dx \\ &= \frac{2c}{\pi a_+} \int_{-1}^1 \left(1 - (1-t) \frac{2\sqrt{c}}{c+1+2\sqrt{c}} \right)^k \sqrt{(1-t)(t+1)} dt.\end{aligned}$$

Writing $\beta = \frac{2\sqrt{c}}{c+1+2\sqrt{c}}$ and changing variables by $1-t = s/k$, we have

$$\gamma_c(h_{k+1}) = \frac{4c^2 k^{-3/2}}{\pi a_+} \int_0^{2k} \left(1 - \frac{\beta s}{k} \right)^k \sqrt{s(2-s/k)} ds.$$

We note that $(1 - \beta s/k)^k \leq e^{-\beta s}$ and so by the dominated convergence theorem, we have

$$\lim_{k \rightarrow \infty} \int_0^{2k} \left(1 - \frac{\beta s}{k} \right)^k \sqrt{s(2-s/k)} ds = \sqrt{2} \int_0^\infty e^{-\beta s} \sqrt{s} ds = \frac{\sqrt{\pi}}{\sqrt{2}\beta^{3/2}}$$

(the last integral is a Gamma integral; see for example [22, Eqn. (4.1.7)]). The desired result follows. \square

For the next estimate, we need the following simple inequality.

Fact 5.4. *For $1 \geq b > a \geq 0$ and $k > 0$ we have*

$$(1-a)^k - (1-b)^k \leq e(e^{-ak} - e^{-bk})$$

Proof. Bound

$$(1-a)^k - (1-b)^k = \int_a^b k(1-\theta)^{k-1} d\theta \leq \int_a^b k e^{-\theta(k-1)} d\theta \leq e(e^{-ak} - e^{-bk}). \quad \square$$

Lemma 5.5. For each fixed $c, A \in [1, \infty)$, and $h_k(x) = (x/a_+)^k$ for $k \geq 1$, we have

$$\lim_{k \rightarrow \infty} \Gamma_c(h_k, h_{Ak}) = \frac{1}{2\pi^2} \int_0^\infty \int_0^\infty \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \frac{s+t}{\sqrt{st}} ds dt.$$

Proof. Writing $\beta = 2\sqrt{c}/a_+$ and changing variables $(x, y) = (1 - s/(\beta k), 1 - t/(\beta k))$, recalling that $\Phi_c(x) = 2\sqrt{c}x + c + 1$, we see that

$$\begin{aligned} \Gamma_c(h_k, h_{Ak}) &= \frac{1}{\pi^2} \int_{-1}^1 \int_{-1}^1 \left(\frac{(\Phi_c(x)/a_+)^k - (\Phi_c(y)/a_+)^k}{x-y} \right) \left(\frac{(\Phi_c(x)/a_+)^{Ak} - (\Phi_c(y)/a_+)^{Ak}}{x-y} \right) \\ &\quad \cdot \frac{1-xy}{\sqrt{1-x^2}\sqrt{1-y^2}} dx dy \\ &= \frac{1}{\pi^2} \int_0^{2\beta k} \int_0^{2\beta k} \left(\frac{(1-s/k)^k - (1-t/k)^k}{s-t} \right) \left(\frac{(1-s/k)^{Ak} - (1-t/k)^{Ak}}{s-t} \right) \\ &\quad \cdot \frac{s+t - st/(\beta k)}{\sqrt{st(2-s/(\beta k))(2-t/(\beta k))}} ds dt. \end{aligned}$$

We may bound the integrand using [Fact 5.4](#) by a constant multiple of

$$\left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \frac{s+t}{\sqrt{st}}$$

which is integrable (see [Lemma 5.7](#), to follow), so the dominated convergence theorem completes the proof. \square

Lemma 5.6. For $A \geq 1$, we have

$$\lim_{k \rightarrow \infty} \Gamma(x^k, x^{Ak}) = \frac{1}{\pi^2} \int_0^\infty \int_0^\infty \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \frac{s+t}{\sqrt{st}} ds dt.$$

Proof. We first claim that

$$\lim_{k \rightarrow \infty} \int_0^1 \int_{-1}^0 \left(\frac{x^k - y^k}{x-y} \right) \left(\frac{x^{Ak} - y^{Ak}}{x-y} \right) \frac{1-xy}{\sqrt{(1-x^2)(1-y^2)}} dx dy = 0. \quad (5.1)$$

To see this, note that on the set $[0, 1] \times [-1, 0]$ we may bound

$$\left(\frac{x^k - y^k}{x-y} \right) \left(\frac{x^{Ak} - y^{Ak}}{x-y} \right) \leq \begin{cases} 16 & \text{if either } x \geq 1/2 \text{ or } y \leq -1/2 \\ Ak^2 2^{-2(k-1)} & \text{if } |x| \leq 1/2 \text{ and } |y| \leq 1/2 \end{cases}$$

by the mean-value theorem. Applying the dominated convergence theorem shows (5.1). By symmetry of the integrand under $(x, y) \mapsto -(x, y)$ this shows that

$$\lim_{k \rightarrow \infty} \Gamma(h_k, h_{Ak}) = \lim_{k \rightarrow \infty} \frac{2}{\pi^2} \int_0^1 \int_0^1 \left(\frac{x^k - y^k}{x-y} \right) \left(\frac{x^{Ak} - y^{Ak}}{x-y} \right) \frac{1-xy}{\sqrt{(1-x^2)(1-y^2)}} dx dy.$$

Changing variables $(x, y) = (1 - s/k, 1 - t/k)$ and arguing as in [Lemma 5.5](#) completes the proof. \square

Lemma 5.7. For $A \geq 1$ we have

$$\int_0^\infty \int_0^\infty \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \frac{s+t}{\sqrt{st}} ds dt \lesssim A^{-1/2}(1 + \log A).$$

Proof. Since the integrand is symmetric in s and t it is sufficient to integrate over the region $s < t$, in which case we may bound the integrand by a constant times

$$\iint_{0 \leq s < t} \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \sqrt{\frac{t}{s}} ds dt.$$

We will further break up the region of integration depending on the value of $t - s$. For $t - s \leq 1$, bound the integrand by

$$e^{-s} \left| \frac{e^{-As} - e^{-At}}{s-t} \right| (1 + s^{-1/2}).$$

(Here we used that $(e^{-s} - e^{-t})/(s-t) \leq e^{-s}$, which follows from the classical inequality $1 - e^x \leq x$ with $x = s - t$). Then, write $t = s + h$, yielding

$$\iint_{0 \leq s \leq t \leq s+1} \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \sqrt{\frac{t}{s}} ds dt \leq \int_0^\infty e^{-s} (1 + s^{-1/2}) e^{-As} \int_0^1 \frac{1 - e^{-Ah}}{h} dh ds.$$

Now, observe that

$$\begin{aligned} \int_0^1 \frac{1 - e^{-Ah}}{h} dh &= \int_0^{1/A} \frac{1 - e^{-Ah}}{h} dh + \int_{1/A}^1 \frac{1 - e^{-Ah}}{h} dh \\ &\leq \int_0^{1/A} A dh + \int_{1/A}^1 h^{-1} dh \\ &= 1 + \log A, \end{aligned}$$

and

$$\begin{aligned} \int_0^\infty e^{-s} (1 + s^{-1/2}) e^{-As} ds &\leq \frac{1}{\sqrt{A}} + \sqrt{A+1} \int_{1/\sqrt{A}}^\infty ((A+1)s)^{-1/2} e^{-(A+1)s} ds \\ &= \frac{1}{\sqrt{A}} + \frac{1}{\sqrt{A+1}} \int_1^\infty q^{-1/2} e^{-q} dq \lesssim \frac{1}{\sqrt{A}}. \end{aligned} \quad (5.2)$$

We deduce that

$$\iint_{0 \leq s \leq t \leq s+1} \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \sqrt{\frac{t}{s}} ds dt \lesssim \frac{1 + \log A}{\sqrt{A}}.$$

This takes care of the region where $t - s \leq 1$; we now turn to the complementary region $t - s \geq 1$. With the substitution $q = (t - s)/s$ we first compute

$$\int_{s+1}^\infty (t-s)^{-2} \sqrt{1 + \frac{(t-s)}{s}} dt = s \int_{1/s}^\infty q^{-2} \sqrt{1+q} dq \lesssim s^{-1/2}.$$

Combining this with (5.2), we have

$$\begin{aligned} \iint_{t \geq s+1} \left(\frac{e^{-s} - e^{-t}}{s-t} \right) \left(\frac{e^{-As} - e^{-At}}{s-t} \right) \sqrt{\frac{t}{s}} ds dt &\leq \int_0^\infty e^{-s(1+A)} \int_{s+1}^\infty (t-s)^{-2} \sqrt{1 + \frac{(t-s)}{s}} dt ds \\ &\lesssim \int_0^\infty e^{-s(1+A)} s^{-1/2} ds \lesssim \frac{1}{\sqrt{A}}. \quad \square \end{aligned}$$

Proof of Lemma 2.5. Note that (1) follows from Lemma 5.3. The other two items follow by combining Lemma 5.5 with Lemma 5.7. \square

Proof of Lemma 2.10. Both items follow by combining Lemma 5.6 with the estimate from Lemma 5.7. \square

REFERENCES

- [1] G. W. Anderson and O. Zeitouni, *A CLT for a band matrix model*, Probab. Theory Related Fields **134** (2006), no. 2, 283–338.
- [2] L. V. Arharov, *Limit theorems for the characteristic roots of a sample covariance matrix*, Dokl. Akad. Nauk SSSR **199** (1971), 994–997.
- [3] F. Aurzada and S. Dereich, *Universality of the asymptotics of the one-sided exit problem for integrated processes*, Ann. Inst. Henri Poincaré Probab. Stat. **49** (2013), no. 1, 236–251.
- [4] Z. D. Bai and J. W. Silverman, *CLT for linear spectral statistics of large-dimensional sample covariance matrices*, Ann. Probab. **32** (2004), no. 1A, 553–605.
- [5] Z. D. Bai and Y. Q. Yin, *Convergence to the semicircle law*, Ann. Probab. **16** (1988), no. 2, 863–875.
- [6] Z. Bao, *On asymptotic expansion and central limit theorem of linear eigenvalue statistics for sample covariance matrices when $N/M \rightarrow 0$* , Theory Probab. Appl. **59** (2015), no. 2, 185–207.
- [7] P. Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1995, A Wiley-Interscience Publication.
- [8] B. Chen and G. Pan, *CLT for linear spectral statistics of normalized sample covariance matrices with the dimension much larger than the sample size*, Bernoulli **21** (2015), no. 2, 1089–1133.
- [9] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, *Everything you always wanted to know about LOCC (but were afraid to ask)*, Comm. Math. Phys. **328** (2014), no. 1, 303–326.
- [10] R. Clifton, B. Hepburn, and C. Wüthrich, *Generic incomparability of infinite-dimensional entangled states*, Phys. Lett. A **303** (2002), no. 2-3, 121–124.

- [11] F. D. Cunden, P. Facchi, G. Florio, and G. Gramegna, *Volume of the set of LOCC-convertible quantum states*, J. Phys. A **53** (2020), no. 17, 175303, 26.
- [12] F. D. Cunden, P. Facchi, G. Florio, and G. Gramegna, *Generic aspects of the resource theory of quantum coherence*, Phys. Rev. A **103** (2021), no. 2, Paper No. 022401, 11.
- [13] A. Dembo, J. Ding, and F. Gao, *Persistence of iterated partial sums*, Annales de l’IHP Probabilités et statistiques, vol. 49, 2013, pp. 873–884.
- [14] L. Devroye, *Non-uniform random variate generation*, Springer New York, 1986.
- [15] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM journal on matrix analysis and applications **9** (1988), no. 4, 543–560.
- [16] V. L. Girko, *Theory of stochastic canonical equations. Vols. I and II*, Mathematics and its Applications, vol. 535, Kluwer Academic Publishers, Dordrecht, 2001.
- [17] D. Jonsson, *Some limit theorems for the eigenvalues of a sample covariance matrix*, J. Multivariate Anal. **12** (1982), no. 1, 1–38.
- [18] K. Luh and S. O’Rourke, *Eigenvector delocalization for non-Hermitian random matrices and applications*, Random Structures & Algorithms **57** (2020), no. 1, 169–210.
- [19] A. Lytova and L. Pastur, *Central limit theorem for linear eigenvalue statistics of random matrices with independent entries*, The Annals of Probability (2009), 1778–1840.
- [20] V. A. Marčenko and L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices*, Mat. Sb. (N.S.) **72(114)** (1967), 507–536.
- [21] S. Melczer, M. Michelen, and S. Mukherjee, *Asymptotic bounds on graphical partitions and partition comparability*, Int. Math. Res. Not. IMRN (2021), no. 4, 2842–2860.
- [22] V. H. Moll, *Special integrals of Gradshteyn and Ryzhik—the proofs. Vol. I*, Monographs and Research Notes in Mathematics, CRC Press, Boca Raton, FL, 2015.
- [23] M. A. Nielsen, *Conditions for a class of entanglement transformations*, Physical Review Letters **83** (1999), no. 2, 436–439.
- [24] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000.
- [25] B. Pittel, *Confirming two conjectures about the integer partitions*, J. Combin. Theory Ser. A **88** (1999), no. 1, 123–135.
- [26] B. Pittel, *Asymptotic joint distribution of the extremities of a random Young diagram and enumeration of graphical partitions*, Adv. Math. **330** (2018), 280–306.
- [27] A. W. Roberts and D. E. Varberg, *Convex functions*, Pure and Applied Mathematics, Vol. 57, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1973.
- [28] M. Rudelson and R. Vershynin, *Smallest singular value of a random rectangular matrix*, Comm. Pure Appl. Math. **62** (2009), no. 12, 1707–1739.
- [29] M. Shcherbina, *Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices*, J. Math. Phys. Anal. Geom. **7** (2011), no. 2, 176–192, 197, 199.
- [30] Y. G. Sinai, *Distribution of some functionals of the integral of a random walk*, Teoret. Mat. Fiz. **90** (1992), no. 3, 323–353.
- [31] H. Tajima, *Deterministic LOCC transformation of three-qubit pure states and entanglement transfer*, Ann. Physics **329** (2013), 1–27.
- [32] T. Tao, *Topics in random matrix theory*, Graduate Studies in Mathematics, vol. 132, American Mathematical Society, Providence, RI, 2012.
- [33] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*, vol. 47, Cambridge university press, 2018.
- [34] G. Vidal, *Entanglement of pure states for a single copy*, Physical Review Letters **83** (1999), no. 5, 1046–1049.
- [35] V. Vysotsky, *On the probability that integrated random walks stay positive*, Stochastic Process. Appl. **120** (2010), no. 7, 1178–1193.
- [36] K. Życzkowski and I. Bengtsson, *Relativity of pure states entanglement*, Annals of Physics **295** (2002), no. 2, 115–135.
- [37] K. Życzkowski and H.-J. Sommers, *Induced measures in the space of mixed quantum states*, vol. 34, 2001, Quantum information and computation, pp. 7111–7125.

APPENDIX A. A GENERAL CENTRAL LIMIT THEOREM FOR LINEAR STATISTICS OF THE WISHART–LAGUERRE ENSEMBLE

In this appendix we explain how to deduce the statements of [Theorems 2.3](#) and [2.7](#) from results in [\[5, 6, 19\]](#). As in the statements of these theorems, let μ_1, \dots, μ_n be the eigenvalues of a random matrix sampled from the complex Wishart–Laguerre ensemble with parameters n, m , and assume $m/n \rightarrow c \in [1, \infty]$.

We recall the relevant statements from [\[5, 6, 19\]](#). Let $a_{\pm} = (1 \pm \sqrt{c})^2$, let \mathcal{F}_0 be the set of all differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that both f and f' have compact support⁵, and fix $g \in \mathcal{F}_0$.

⁵The different theorems we are citing here have slightly different assumptions, but this is a simple assumption which certainly suffices for each of the theorems we cite.

First, if $c < \infty$ then let $W = g(\mu_1/n) + \dots + g(\mu_n/n)$. In this case, [19, Theorem 4.1] (which is really a restatement of results in [20]) says that

$$\frac{W}{n} \xrightarrow{P} \frac{1}{2\pi} \int_{a_-}^{a_+} g(\lambda) \frac{\sqrt{(\lambda - a_-)(a_+ - \lambda)}}{\lambda} d\lambda = \gamma_c(g), \quad (\text{A.1})$$

and [19, Theorem 4.2] says that $W - \mathbb{E}W \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, where

$$\begin{aligned} \sigma^2 &= \frac{1}{2} \cdot \frac{1}{2\pi^2} \int_{a_-}^{a_+} \int_{a_-}^{a_+} \left(\frac{g(\lambda_1) - g(\lambda_2)}{\lambda_1 - \lambda_2} \right)^2 \frac{4c - (\lambda_1 - c - 1)(\lambda_2 - c - 1)}{\sqrt{4c - (\lambda_1 - c - 1)^2} \sqrt{4c - (\lambda_2 - c - 1)^2}} d\lambda_1 d\lambda_2 \\ &= \Gamma_c(g, g). \end{aligned} \quad (\text{A.2})$$

Second, if $c = \infty$, then with $X = g((\mu_1 - m)/2\sqrt{mn}) + \dots + g((\mu_n - m)/2\sqrt{mn})$, the convergence to the semicircle law proved in [5] implies that

$$\frac{X}{n} \xrightarrow{P} \frac{2}{\pi} \int_{-1}^1 g(\lambda) \sqrt{1 - \lambda^2} d\lambda = \gamma(g). \quad (\text{A.3})$$

With the slightly different parameterisation $Y = g((\mu_1 - m)/\sqrt{mn}) + \dots + g((\mu_n - m)/\sqrt{mn})$, [6, Theorem 1.2] says that $Y - \mathbb{E}Y \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, where

$$\sigma^2 = \frac{1}{2} \cdot \frac{1}{2\pi^2} \int_{-2}^2 \int_{-2}^2 \left(\frac{g(\lambda_1) - g(\lambda_2)}{\lambda_1 - \lambda_2} \right)^2 \frac{4 - \lambda_1 \lambda_2}{\sqrt{4 - \lambda_1^2} \sqrt{4 - \lambda_2^2}} d\lambda_1 d\lambda_2 = \Gamma(g, g). \quad (\text{A.4})$$

Remark. The above cited papers all primarily concern the *real* Wishart–Laguerre ensemble; we need the complex case. The same proofs apply to both cases (the only difference is that the variance in the real case is a factor of 2 larger than the variance in the complex case). See [19, Remark 4.1].

These results imply the univariate cases of [Theorems 2.3](#) and [2.7](#), assuming that $f \in \mathcal{F}_0$ and provided we weaken L^1 -convergence to convergence in probability. We now discuss how to upgrade this to the desired statements, using known results on concentration of singular values of Gaussian random matrices. Specifically, we will need the following.

Lemma A.1. *Let $\mu_1 \geq \dots \geq \mu_n$ be the eigenvalues of a random matrix sampled from the complex Wishart-Laguerre ensemble with parameters $m \geq n$. There exists an absolute constant C such that for any $t \geq 0$,*

$$\sqrt{m} - C(\sqrt{n} + t) \leq \sqrt{\mu_n} \leq \sqrt{\mu_1} \leq \sqrt{m} + C(\sqrt{n} + t)$$

except with probability at most $2 \exp(-t^2)$.

Remark. In the real case, this statement is proved in [33, Theorem 4.6.1], but the same proof extends readily to the complex case as well.

Therefore, in the case $c < \infty$, it follows from the upper bound in [Lemma A.1](#) that for all sufficiently large s (depending on c),

$$\mathbb{P} \left[\max_k \mu_k/n \geq s \right] = \mathbb{P}[\sqrt{\mu_1} \geq \sqrt{sn}] \leq \mathbb{P}[\sqrt{\mu_1} \geq \sqrt{m} + C\sqrt{n} + C\sqrt{sn/2}] \leq 2 \exp(-sn/4),$$

so that if $g \in \mathcal{F}$ (i.e. g is continuously differentiable and has sub-exponential growth) then for any $\varepsilon > 0$, we can (smoothly) truncate $g \in \mathcal{F}$ to some function $g_0 \in \mathcal{F}_0$ with compact support, in such a way that

$$\mathbb{E} \left| \left(g(\mu_1/n) + \dots + g(\mu_n/n) \right) - \left(g_0(\mu_1/n) + \dots + g_0(\mu_n/n) \right) \right| \leq \varepsilon.$$

That is to say, the effect of the truncation can be made arbitrarily small, so if the asymptotic results in [\(A.1\)](#) and [\(A.2\)](#) hold for every $g \in \mathcal{F}_0$ they must also hold for every $g \in \mathcal{F}$.

Further, by the upper bound in [Lemma A.1](#) and using that $g \in \mathcal{F}$ has sub-exponential growth, it follows that the sequence of random variables W/n is uniformly integrable. Since this sequence converges to $\gamma_c(g)$ in probability by [\(A.1\)](#), it follows by uniform integrability that it also converges in L^1 .

In the case $c = \infty$, we use both the upper and the lower bound in [Lemma A.1](#) to see that for all sufficiently large s ,

$$\begin{aligned} \mathbb{P} \left[\max_k (\mu_k - m)/\sqrt{mn} \geq s \right] &\leq \mathbb{P}[\mu_1 \geq m + s\sqrt{mn}] + \mathbb{P}[\mu_n \leq m - s\sqrt{mn}] \\ &\leq 2 \min\{\exp(-s^2n/2), \exp(-s\sqrt{mn}/2)\} \\ &\leq 2 \exp(-sn/4), \end{aligned}$$

from which it follows (as before) that we can upgrade the assumption $g \in \mathcal{F}_0$ to the assumption $g \in \mathcal{F}$ and also upgrade the convergence in (A.3) to convergence in L^1 .

Finally, we discuss how to deduce multivariate central limit theorems from the above univariate central limit theorems. Given functions $f_1, \dots, f_\ell \in \mathcal{F}$, for any fixed vector $\vec{t} = (t_1, \dots, t_\ell) \in \mathbb{R}^\ell$ we can apply the above observations to the function $f_{\vec{t}} = t_1 f_1 + \dots + t_\ell f_\ell$, to see that

$$t_1(W_1 - \mathbb{E}W_1) + \dots + t_\ell(W_\ell - \mathbb{E}W_\ell) \xrightarrow{d} \mathcal{N}(\vec{0}, \Gamma_c(f_{\vec{t}}, f_{\vec{t}})),$$

in the case $c < \infty$ and

$$t_1(Y_1 - \mathbb{E}Y_1) + \dots + t_\ell(Y_\ell - \mathbb{E}Y_\ell) \xrightarrow{d} \mathcal{N}(\vec{0}, \Gamma(f_{\vec{t}}, f_{\vec{t}}))$$

in the case $c = \infty$. The desired convergence in [Theorem 2.3\(2\)](#) and [Theorem 2.7\(2\)](#) now follows from the Cramér–Wold device (see for example [7, Theorem 29.4]).

UNIVERSITY OF ILLINOIS CHICAGO
Email address: visheshj@uic.edu

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (ISTA).
Email address: matthew.kwan@ist.ac.at

UNIVERSITY OF ILLINOIS CHICAGO
Email address: michelen@uic.edu