The Large-N Limits of Brownian Motions on $\mathbb{GL}_N$

Todd Kemp*
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112
tkemp@math.ucsd.edu

January 23, 2015

Abstract

We introduce a two-parameter family of diffusion processes $(B^N_{r,s}(t))_{t\geq 0}$, $r, s > 0$, on the general linear group $\mathbb{GL}_N$ that are Brownian motions with respect to certain natural metrics on the group. At the same time, we introduce a two-parameter family of free Itô processes $(b_{r,s}(t))_{t\geq 0}$ in a faithful, tracial W^*-probability space, and we prove that the process $(B^N_{r,s}(t))_{t\geq 0}$ converges to $(b_{r,s}(t))_{t\geq 0}$ in noncommutative distribution as $N \to \infty$ for each $r, s > 0$. The processes $(b_{r,s}(t))_{t\geq 0}$ interpolate between the free unitary Brownian motion when $(r, s) = (1, 0)$, and the free multiplicative Brownian motion when $r = s = \frac{1}{2}$; we thus resolve the open problem of convergence of the Brownian motion on $\mathbb{GL}_N$ posed by Philippe Biane in 1997.

Contents

*Supported by NSF CAREER Award DMS-1254807

1 Introduction

1.1 Main Theorems and Discussion

Let $\mathbb{M}_N$ denote the space of $N \times N$ complex matrices, and let $\mathbb{GL}_N$ denoted the Lie group of invertible matrices in $\mathbb{M}_N$. In this paper, we will address the behavior of Brownian motion on this group as $N \to \infty$. In fact, we introduce a two-parameter family $B^N_{r,s}$ of diffusion processes that are all left-invariant Brownian motions with respect to a family of metrics on $\mathbb{GL}_N$ (achieved by scaling the inner product by independent factors $r, s > 0$ on the real and imaginary parts of the Lie algebra); see Definitions 1.3 and 1.5. The canonical Brownian motion on $\mathbb{GL}_N$ coincides with $B^N_{\frac{1}{2},\frac{1}{2}}$, while the degenerate case $B^N_{1,0}$ is Brownian motion on the unitary group $\mathbb{U}_N$.

Our main concern is with the large-N limit of the finite-dimensional (noncommutative) distributions of these (r, s)-Brownian motions. To be precise: the classical distribution of the stochastic process $B^N_{r,s}$ is a measure on paths taking values in $\mathbb{GL}_N$, and it is very difficult to make sense of a large-N limit of such objects (though attempts have been made in the analogous case of $\mathbb{U}_N$-valued processes, cf. [11, 12]). Motivated instead by random matrix theory and free probability, we study statistics of the process that live in an N-independent space. For a single random matrix ensemble $X = X^N$ taking values in the normal matrices in $\mathbb{M}_N$, the standard object of study is the empirical spectral distribution: the random probability measure on $\mathbb{C}$ that places equal weights at the eigenvalues of the matrix. This measure is captured by its (random) trace moments: random variables of the form $\{\mathrm{tr}(X^k X^{*m})\colon k, m \in \mathbb{N}\}$, where $\mathrm{tr} = \frac{1}{N}\mathrm{Tr}$ is the normalized trace on $\mathbb{M}_N$ and X^* is the adjoint (conjugate transpose) of X. For a collection $X_1, \ldots, X_n$ of random matrix ensembles that do not generally commute, the natural analog is the *noncommutative distribution*: the collection of all random variables $\mathrm{tr}(f(X_1, X_1^*, \ldots, X_n, X_n^*))$ for all noncommutative polynomials f in the matrices and their adjoints.

The main theorem of this paper is the identification of the large-N limit of the noncommutative distribution of any finite collection of instances of the Brownian motion $B^N_{r,s}(t_1), \ldots, B^N_{r,s}(t_n)$. In the limit, one does not find the distribution of a diffusion, or any (classical) Itô process at all. Rather, the limit is a *free* Itô process $b_{r,s}$, which we refer to as a free multiplicative (r, s)-Brownian motion; see Definition 1.6. (Sections 2.2 and 2.3 give brief recollections of the basics of free probability and free stochastic analysis, with some references to more in depth treatments.) That is: the limit of $B^N_{r,s}$ is identified as a one-parameter family of operators $\{b_{r,s}(t)\}_{t\geq 0}$ in a tracial noncommutative probability space $(\mathscr{A}, \tau)$, whose finite-dimensional noncommutative distributions are precisely the large-N limits of those of $B^N_{r,s}$. As is standard in noncommutative probability, we refer to this as *convergence of the process* (as this is the strongest notion of convergence that makes sense for noncommutative stochastic processes whose distributions are not measures on a fixed path space).

Theorem 1.1. *For $r, s > 0$, let $B^N_{r,s}$ be an (r, s)-Brownian motion on $\mathbb{GL}_N$, and let $b_{r,s}$ be a free multiplicative (r, s)-Brownian motion. Then $(B^N_{r,s}(t))_{t\geq 0}$ on $\mathbb{GL}_N$ converges, as a noncommutative stochastic process, to $(b_{r,s}(t))_{t\geq 0}$ as $N \to \infty$. More precisely: if $n \in \mathbb{N}$ and f is any noncommutative polynomial in $2n$ indeterminates, then for any $t_1, \ldots, t_n \geq 0$,*

$$\begin{aligned} &\mathbb{E}\mathrm{tr}\left[f(B^N_{r,s}(t_1), B^N_{r,s}(t_1)^*, \ldots, B^N_{r,s}(t_n), B^N_{r,s}(t_n)^*)\right] \\ &\qquad = \tau[f(b_{r,s}(t_1), b_{r,s}(t_1)^*, \ldots, b_{r,s}(t_n), b_{r,s}(t_n)^*)] + O\left(\frac{1}{N^2}\right). \end{aligned} \tag{1.1}$$

Theorem 1.1 is proved in Section 5. It answers a question left by Biane in [2]. The main result of that paper was the analogous statement of Theorem 1.1 for the canonical Brownian motion U^N on the unitary group $\mathbb{U}_N$. In that paper, Biane introduced the free unitary Brownian motion for the first time; it has now become a standard tool used in free probability theory (see [7], [28], and many others). Biane's proof had two main steps: first establishing the convergence for a fixed time $t > 0$, and then using group properties resepcted by the process and its limit (namely independent multiplicative increments) together with complementary asymptotic freeness results to extend the convergence to all finite-dimensional distributions. Since a single instance $U^N(t)$

of the unitary Brownian motion is a unitary (hence normal) matrix, the spectral theorem is available. Thus the noncommutative distribution becomes the empirical spectral distribution ν_t, whose limit is then computed via a careful analysis of the characters of irreducible representation of $\mathbb{U}_N$. See Definition 1.7 for a closed formula for the limiting moments of this distribution; these moments will come into play in the present analysis as well.

Our proof is similarly broken into two parts, first establishing the convergence for a fixed t, and then leveraging more general asymptotic freeness results (cf. Section 2.3) to extend to convergence of the process. For a single t, the story is quite different. The process $B_{r,s}^N$ is almost surely never normal (cf. Proposition 4.15), and so the empirical spectral distribution of $B_{r,s}^N(t)$ has no simple connection to the noncommutative distribution of the process (it is not even a continuous function of the moments in the limit). In [2, p. 19], Biane states "It is very likely that the process $(\Gamma_t)_{t\in\mathbb{R}_+}$" (which is our process $B_{r,s}^N$ with $r = s = \frac{1}{2}$) "is the limit in distribution… of the Brownian motion with values in $\mathbb{GL}_N$… but we have not proved this." He goes on to list a partial result, showing that the convergence holds for a single time $t \geq 0$ for the self-adjoint process $\Gamma_t^*\Gamma_t$, which, he states, can be computed following the same general outline as his analysis of the heat kernel on $\mathbb{U}_N$ but using the spherical functions for the pair $(\mathbb{GL}_N, \mathbb{U}_N)$ in the place of the characters of $\mathbb{U}_N$. It is possible that a more involved representation theoretic approach like this might yield a proof of our Theorem 1.1 for a single time t, but such a proof has not appeared in the literature in the 17 years since this question was posed.

Our approach is more geometric, using a structure theorem for the Laplacian on $\mathbb{GL}_N$ proved in the author's earlier joint paper [9], and associated concentration of heat kernel measure results from that paper. (A similar approach was indepednently developed by Guillaume Cébron in [5, Theorem 4.6], and is used there to give a somewhat different proof of the special case of Theorem 1.1 for a single time $t \geq 0$ and for the canonical case $r = s = \frac{1}{2}$.) These ideas go back to earlier papers by Eric Rains [22] and Ambar Sengupta [25]. To give a little more detail presently: for a single time $t \geq 0$, we compare the left- and right-hand sides of (1.1) using stochastic calculus. Each can be represented as a stochastic integral involving noncommutative polynomials of lower order (thanks to the linearity of the diffusion and drift coefficients in (1.5)), and the proof proceeds by a careful induction using the following key concentration of measure result, which is another main theorem of the present paper.

Theorem 1.2. *Let $n \in \mathbb{N}$, $t_1, \ldots, t_n \geq 0$, and let $B_{r,s}^{1,N}(t_1), \ldots, B_{r,s}^{n,N}(t_n)$ be independent copies of the Brownian motion $B_{r,s}^N(\cdot)$ at these times. These operators possess a limit joint distribution, and, for any noncommutative polynomials in $2n$ indeterminates, there is a constant $C = C(r, s, t_1, \ldots, t_n, f, g)$ such that*

$$\mathrm{Cov}\left[\mathrm{tr}(f(B_{r,s}^{1,N}(t_1), \ldots, B_{r,s}^{1,N}(t_n)^*)), \mathrm{tr}(g(B_{r,s}^{1,N}(t_1), \ldots, B_{r,s}^{n,N}(t_n)^*))\right] \leq \frac{C}{N^2}. \tag{1.2}$$

Theorem 1.2 is proved in Section 3. It is a multivariate extension of the technology in [9, Sections 3 & 4].

1.2 Definitions, Notation, and Auxiliary Results

To motivate our interest in the diffusions $B_{r,s}^N$ (formally defined below), let us briefly discuss the complexity of the Lie group $\mathbb{GL}_N$ as compared to $\mathbb{U}_N$. The Lie algebra $\mathrm{Lie}(\mathbb{GL}_N) = \mathfrak{gl}_N = \mathbb{M}_N$ possesses no $\mathrm{Ad}(\mathbb{GL}_N)$-invariant inner product (as it is not of compact type, cf. [15]). However, $\mathbb{GL}_N$ is the complexification of $\mathbb{U}_N$, which in particular gives the decomposition of its Lie algebra $\mathfrak{gl}_N = \mathfrak{u}_N \oplus i\mathfrak{u}_N$, where $\mathfrak{u}_N = \mathrm{Lie}(\mathbb{U}_N)$ consists of skew-Hermitian matrices.

The standard inner product on $\mathfrak{u}_N$ is $\langle \xi, \eta \rangle = -\mathrm{Tr}(\xi\eta)$, which is $\mathrm{Ad}(\mathbb{U}_N)$-invariant. It extends to the real Hilbert-Schmidt inner product $\langle \xi, \eta \rangle = \Re\mathrm{Tr}(\xi\eta^*)$ on $\mathfrak{gl}_N$ which, while not $\mathrm{Ad}(\mathbb{GL}_N)$-invariant, remains invariant under conjugation by $\mathbb{U}_N$. It is surely not the only such $\mathrm{Ad}(\mathbb{U}_N)$-invariant inner product on $\mathfrak{gl}_N$. The simplest generalization is given by scaling the inner product independently on the two part of the decomposition $\mathfrak{gl}_N = \mathfrak{u}_N \oplus i\mathfrak{u}_N$. This is closely related to the two-parameter Segal–Bargmann transform of Driver and Hall (cf. [8]) whose large-N limit is the topic of [9], and is the real motivation for the present discussion. Let us define the inner products now.

Definition 1.3. *Let $r, s > 0$. Define the real inner product $\langle\cdot,\cdot\rangle_{r,s}$ on $\mathfrak{gl}_N$ by*

$$\langle \xi_1 + i\eta_1, \xi_2 + i\eta_2\rangle_{r,s}^N = -\frac{1}{r}N\mathrm{Tr}(\xi_1\xi_2) - \frac{1}{s}N\mathrm{Tr}(\eta_1\eta_2), \qquad \xi_1,\xi_2,\eta_1,\eta_2 \in \mathfrak{u}_N. \tag{1.3}$$

That is: $\langle\cdot,\cdot\rangle_{r,s}^N$ makes $\mathfrak{u}_N$ and $i\mathfrak{u}_N$ orthogonal, and its restrictions to these two orthocomplementary subspaces are positive scalar multiples of the Hilbert-Schmidt inner product.

Remark 1.4. The inner product $\langle\cdot,\cdot\rangle_{r,s}^N$ may alternatively be written in the form

$$\langle A, B\rangle_{r,s}^N = \frac{1}{2}\left(\frac{1}{s}+\frac{1}{r}\right)N\Re\mathrm{Tr}(AB^*) + \frac{1}{2}\left(\frac{1}{s}-\frac{1}{r}\right)N\Re\mathrm{Tr}(AB).$$

We scale with $N\mathrm{Tr}$ in order to produce a meaningful limit as $N\to\infty$. That is must scale opposite to the trace $\mathrm{tr} = \frac{1}{N}\mathrm{Tr}$ used to define the limit moments is a consequence of the fact that, in general, the Laplacian scales opposite to the metric.

Any real inner product on $\mathfrak{gl}_N$ gives rise to a left-invariant Riemannian metric on $\mathbb{GL}_N$, and hence to a left-invariant Laplacian and associated diffusion process: the Brownian motion.

Definition 1.5. *Let $r, s > 0$. Let $\Delta_{r,s}^N$ denote the Laplace-Beltrami operator on $\mathbb{GL}_N$ associated to the left-invariant Riemannian metric induced by the inner product $\langle\cdot,\cdot\rangle_{r,s}^N$ (cf. (2.8)). The diffusion process $B_{r,s}^N(t)$ on $\mathbb{GL}_N$, started at $B_{r,s}^N(0) = I_N$, with generator $\frac{1}{2}\Delta_{r,s}^N$, is called an (r,s)*-**Brownian motion** *on $\mathbb{GL}_N$. Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ from which the random matrices $B_{r,s}^N(t)$ are sampled, and denote $\mathbb{E} = \int_\Omega \cdot\, d\mathbb{P}$.*

Theorem 1.1 characterizes the large-N limit of $B_{r,s}^N(t)$ as a noncommutative stochastic process. To do so, we introduce the following free stochastic processes (for a discussion of free stochastic calculus, see Section 2.2).

Definition 1.6. *Fix $r, s \geq 0$. Let $(\mathscr{A}, \tau)$ be a W^*-probability space that contains two freely independent free semicircular Brownian motions x and y. For $t \geq 0$, let*

$$w_{r,s}(t) = i\sqrt{r}\,x(t) + \sqrt{s}\,y(t). \tag{1.4}$$

The **free multiplicative Brownian motion** *of parameters r, s, denoted $b_{r,s}$, is the unique solution to the following free stochastic differential equation:*

$$db_{r,s}(t) = b_{r,s}(t)\,dw_{r,s}(t) - \frac{1}{2}(r-s)b_{r,s}(t)\,dt, \quad b_{r,s}(0) = 1. \tag{1.5}$$

Note that, when $r = s = \frac{1}{2}$, $w_{r,s}(t) = \frac{1}{\sqrt{2}}(ix(t)+y(t))$ which is the variance-normalized circular Brownian motion (the large-N limit of the appropriately scaled Brownian motion on the Lie algebra $\mathbb{M}_N$, cf. [26]), more commonly denoted Z_t. In this case, setting $\Gamma_t = b_{\frac{1}{2},\frac{1}{2}}(t)$, (1.5) reduces to

$$d\Gamma_t = \Gamma_t\,dZ_t, \qquad \Gamma_0 = 1, \tag{1.6}$$

which is the left-invariant version of the free multiplicative Brownian motion referenced in [2, 3]. We may also consider the degenerate case $(r,s) = (1,0)$. Let $u(t) = b_{1,0}(t)$; then (1.5) becomes

$$du(t) = iu(t)\,dx(t) - \frac{1}{2}u(t)\,dt, \qquad u(0) = 1 \tag{1.7}$$

which is the free SDE for the (left) *free unitary Brownian motion*, introduced in [2].

In order to prove Theorem 1.1, we need to describe more concretely the noncommutative distribution of $b_{r,s}$; to that end, we introduce the following indispensable constants, which are (extensions) of the moment of the free unitary Brownian motion.

Definition 1.7. *For each $t \in \mathbb{R}$, there exists a unique probability measure ν_t on $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ with the following properties. For $t > 0$, ν_t is supported in the unit circle $\mathbb{U}$; for $t < 0$, ν_t is compactly supported in $\mathbb{R}_+ = (0, \infty)$; and $\nu_0 = \delta_1$. In all cases, ν_t is determined by its moments: $\nu_0(t) \equiv 1$ and, for $n \in \mathbb{Z} \setminus \{0\}$,*

$$\nu_n(t) \equiv \int_{\mathbb{C}^*} u^n \, \nu_t(du) = e^{-\frac{|n|}{2}t} \sum_{k=0}^{|n|-1} \frac{(-t)^k}{k!} |n|^{k-1} \binom{|n|}{k+1}. \tag{1.8}$$

The existence of the measure was proved in [1] for $t < 0$ and in [2] for $t > 0$. In the latter case, it is the a.s. limit of the empirical spectral distribution of the free unitary Brownian motion; in the former case, it has a similar description in terms of a positive free diffusion process sometimes called multiplicative Brownian motion.

The two parts of the proof of Theorem 1.1 (convergence for a single t, and then asymptotic freeness of increments to extend to multiple t) rely on the following auxiliary results. They are fairly straightforward computations using the Itô formula, and their proofs are outlined in Section 4.

Proposition 1.8. *Let $r, s, t \geq 0$ and $n \in \mathbb{N}$. Then*

$$\tau\left[b_{r,s}(t)^n\right] = \tau\left[b_{r,s}(t)^{*n}\right] = \nu_n((r-s)t), \tag{1.9}$$

$$\tau\left[(b_{r,s}(t)b_{r,s}(t)^*)^n\right] = \nu_n(-4st), \tag{1.10}$$

$$\tau\left[b_{r,s}(t)^2 b_{r,s}(t)^{*2}\right] = e^{4st} + 4st(1+st)e^{(3s-r)t}. \tag{1.11}$$

Equations (1.9) and (1.10) were proved in the author's paper [16, Theorems 1.3 & 1.5]. They are included here to show how they can be derived more directly from the limit process $b_{r,s}(t)$; the final steps of the calculations are in Corollaries 4.5 and 4.8. Equation (1.11) is needed in the proof of Proposition 1.10 below. In particular, comparing (1.10) and (1.11) shows that $b_{r,s}(r)$ is never normal if $r, s, t > 0$, as holds for finite N as well.

Remark 1.9. Proposition 4.15 below shows the unsurprising fact that $B_{r,s}^N$ is non normal for all $t > 0$ (with probability 1), since the submanifold of normal matrices is of codimension > 1 and is therefore a polar set for the diffusion $B_{r,s}^N$. One might hope to be able to prove this directly using Itô's formula, but the best one can do in that framework is a calculation akin to (1.10) and (1.11) which shows the weaker statement that for each *fixed* $t > 0$, $B_{r,s}^N(t)$ is a.s. non-normal.

Finally, we will use the fact that the free stochastic process $(b_{r,s}(t))_{t \geq 0}$ inherits all of the invariant properties from $B_{r,s}^N(t)$ that qualify it as a Brownian motion.

Proposition 1.10. *For $r, s > 0$ and $N \in \mathbb{N}$, the $\mathbb{GL}_N$ Brownian motion $(B_{r,s}^N(t))_{t \geq 0}$ has independent, stationary multiplicative increments. If $N \geq 2$, then, with probability 1, $B_{r,s}^N(t)$ is not a normal matrix for any $t > 0$.*

For $r, s \geq 0$, the free multiplicative Brownian motion $(b_{r,s}(t))_{t \geq 0}$ is invertible for all $t \geq 0$, and has freely independent, stationary multiplicative increments. If $s > 0$, then $b_{r,s}(t)$ is not a normal operator for any $t > 0$.

Remark 1.11. A simple time change argument shows that if $s = 0$, then $b_{r,0}$ is unitary, and $u(t) \equiv b_{r,0}(t/r)$ is a free unitary Brownian motion for any $r > 0$. The same applies to $B_{r,0}^N$, which we define (in this degenerate case) as the solution to the SDE (2.11) below.

2 Background

In this section, we briefly outline the technology needed to prove the results in this paper: stochastic calculus for matrix-valued Itô processes (particularly for invertible random matrices), the corresponding stochastic calculus in the free probability setting, and the notion of *asymptotic freeness* that ties the two together.

2.1 Stochastic Calculus on $\mathbb{GL}_N$

Let G be a connected Lie group, with Lie algebra $\mathfrak{g}$. For $\xi \in \mathfrak{g}$, the associated left-invariant vector field on G is denoted ∂_ξ:

$$(\partial_\xi f)(g) = \left.\frac{d}{dt}\right|_{t=0} f(g\exp(t\xi)), \qquad f \in C^\infty(G). \tag{2.1}$$

Let $\langle\cdot,\cdot\rangle$ be a real inner product on $\mathfrak{g}$, and let β be an orthonormal basis for $(\mathfrak{g}, \langle\cdot,\cdot\rangle)$. Then the Laplace-Beltrami operator on G for the Riemannian metric induced by $\langle\cdot,\cdot\rangle$ is

$$\Delta_G = \sum_{\xi\in\beta} \partial_\xi^2, \tag{2.2}$$

which does not depend on the particular orthonormal basis used.

If $G \subset \mathbb{M}_N$ is a linear Lie group, then the Brownian motion on G (the diffusion process with generator $\frac{1}{2}\Delta_G$) may be constructed as the solution to a stochastic differential equation. Fix an orthonormal basis β for $\mathfrak{g}$, and let $W(t)$ denote the following Wiener process in $\mathfrak{g}$:

$$W(t) = \sum_{\xi\in\beta} W_\xi(t)\,\xi,$$

where $\{W_\xi : \xi \in \beta\}$ are i.i.d. standard $\mathbb{R}$-valued Brownian motions. Then the Brownian motion $B(t)$ is determined by the Stratonovich SDE

$$dB(t) = B(t) \circ dW(t), \qquad W(0) = I_N. \tag{2.3}$$

While convenient for proving geometric invariance, the Stratonovich form is less well-adapted to computation. We can convert (2.3) to Itô form. The result, due to McKean [18, p. 116], is

$$dB(t) = B(t)\,dW(t) + \frac{1}{2}B(t)\left(\sum_{\xi\in\beta}\xi^2\right)dt, \qquad B(0) = I_N. \tag{2.4}$$

See, also, [13].

Let us specialize to the case of interest, with $G = \mathbb{GL}_N$ and $\mathfrak{gl}_N$ equipped with an $\mathrm{Ad}_{\mathbb{U}_N}$-invariant inner product $\langle\cdot,\cdot\rangle_{r,s}^N$ of (1.3). To clarify: let $\langle\cdot,\cdot\rangle_{\mathfrak{u}_N}$ denote the following real inner product on $\mathfrak{u}_N$:

$$\langle\xi,\eta\rangle_{\mathfrak{u}_N} = -N\mathrm{Tr}(\xi\eta). \tag{2.5}$$

Then the inner product $\langle\cdot,\cdot\rangle_{r,s}^N$ on $\mathfrak{gl}_N = \mathfrak{u}_N \oplus i\mathfrak{u}_N$ is given by

$$\langle\xi_1 + i\eta_1, \xi_2 + i\eta_2\rangle_{r,s}^N = \frac{1}{r}\langle\xi_1,\xi_2\rangle_{\mathfrak{u}_N} + \frac{1}{s}\langle\eta_1,\eta_2\rangle_{\mathfrak{u}_N}. \tag{2.6}$$

It is straightforward to check that, if β_N is an orthonormal basis for $\mathfrak{u}_N$ with respect to $\langle\cdot,\cdot\rangle_{\mathfrak{u}_N}$, then

$$\beta_{r,s}^N = \left\{\sqrt{r}\xi : \xi \in \beta_N\right\} \cup \left\{\sqrt{s}i\xi : \xi \in \beta_N\right\} \tag{2.7}$$

is an orthonormal basis for $\mathfrak{gl}_N$ with respect to $\langle\cdot,\cdot\rangle_{r,s}^N$. Equation (2.2) and a straightforward application of the chain rule in (2.1) then shows that the Laplace-Beltrami operator is

$$\Delta_{r,s}^N = \sum_{\xi\in\beta}(r\partial_\xi^2 + s\partial_{i\xi}^2). \tag{2.8}$$

Remark 2.1. In [9, 16], we used the elliptic operator

$$A^N_{s,t} = \left(s - \frac{t}{2}\right) \sum_{\xi\in\beta_N} \partial_\xi^2 + \frac{t}{2} \sum_{\xi\in\beta_N} \partial_{i\xi}^2 = \Delta^N_{s-t/2,t/2}.$$

The linear change of parameters was convenient for our discussion of the two-parameter Segal–Bargmann transform, and so all of the theorems in [16] are stated using this language as well.

We will have frequent use for the following "magic formula"; it was stated and proved as [9, Proposition 3.1], but it surely goes back further (for example to the work of Sengupta [25], and Gordina [14] where it was used in the context of infinite dimensional orthogonal groups). If β_N is an orthonormal basis of $\mathfrak{u}_N$, then

$$\sum_{\xi\in\beta_N} \xi A \xi = -\mathrm{tr}(A) I_N, \qquad A \in \mathbb{M}_N. \tag{2.9}$$

In particular, taking $A = I_N$ yields

$$\sum_{\xi\in\beta_N} \xi^2 = -I_N. \tag{2.10}$$

Combining this with (2.7) gives

$$\sum_{\xi\in\beta^N_{r,s}} \xi^2 = -(r-s)I_N,$$

and so, by (2.4), the $\mathbb{U}_N$-invariant Brownian motion $B^N_{r,s}(t)$ is determined by the Itô SDE

$$dB^N_{r,s}(t) = B^N_{r,s}(t)\, dW^N_{r,s}(t) - \frac{1}{2}(r-s) B^N_{r,s}(t)\, dt, \tag{2.11}$$

where $W^N_{r,s}(t) = \sum_{\xi\in\beta^N_{r,s}} W_\xi(t)\,\xi$. It will be convenient to express this Itô process in a slightly different form. Let us choose the following orthonormal basis β_N for $\mathfrak{u}_N$:

$$\beta_N = \left\{ \frac{1}{\sqrt{N}} E_{jj}, \frac{1}{\sqrt{2N}}(E_{jk} - E_{kj}), \frac{1}{\sqrt{2N}} i(E_{jk} + E_{jk})\colon 1 \le j < k \le N \right\}, \tag{2.12}$$

where E_{jk} is the matrix unit with a 1 in the (j,k)-entry and 0 elsewhere. Then it is strightforward to check that

$$W^N_{r,s}(t) = \sqrt{r} \sum_{\xi\in\beta_N} B_\xi(t)\,\xi + i\sqrt{s} \sum_{\xi\in\beta_N} B_{i\xi}(t)\,\xi = \sqrt{r}\, i X^N(t) + \sqrt{s}\, Y^N(t),$$

where $X^N(t)$ and $Y^N(t)$ are independent GUE_N *Brownian motions*. That is: the matrices $X^N(t), Y^N(t)$ are Hermitian, the diagonal entries are $\mathbb{R}$-valued Brownian motions of variance t/N, and the entries $[X^N(t)]_{jk}$ and $[Y^N(t)]_{jk}$ with $1 \le j < k \le N$ are complex Brownian motions of total variance t/N (i.e. $\frac{1}{\sqrt{2}}(B(t) + iB'(t))$ where $B(t), B'(t)$ are independent $\mathbb{R}$-valued Brownian motions of variance t/N). This is a convenient representation, due to the following easily-verified stochastic calculus rules that apply to matrix stochastic integrals with respect to (linear combinations of) $X^N(t)$ and $Y^N(t)$.

Lemma 2.2. *Let $\Theta(t), \Theta_1(t), \Theta_2(t)$ be $\mathbb{M}_N$-valued stochastic processes that are adapted to the filtration $\mathscr{F}_t$ of $X^N(t)$ and $Y^N(t)$, with all entries in $L^2(\Omega, \mathscr{F}, \mathbb{P})$. Then, for any $t \ge 0$, the following hold:*

$$\mathbb{E}\left(\int_0^t \Theta_1(s)\, dX^N(s)\, \Theta_2(s)\right) = \mathbb{E}\left(\int_0^t \Theta_1(s)\, dY^N(s)\, \Theta_2(s)\right) = 0 \tag{2.13}$$

$$\int_0^t dX^N(s)\, \Theta(t)\, dX^N(s) = \int_0^t dY^N(s)\, \Theta(t)\, dY^N(s) = \int_0^t \mathrm{tr}(\Theta(s))\, ds \cdot I_N \tag{2.14}$$

$$\int_0^t dX^N(t)\, \Theta(s)\, dY^N(s) = \int_0^t dY^N(s)\, \Theta(s)\, dX^N(s) = 0. \tag{2.15}$$

Moreover, let $\Theta_1(t)$ and $\Theta_2(t)$ be $\mathbb{M}_N$-valued Itô processes: solutions to SDEs of the form

$$d\Theta(t) = f_1(\Theta(t))\,dX^N(t)\,f_2(\Theta(t)) + g_1(\Theta(t))\,dY^N(t)\,g_2(\Theta(t)) + h(\Theta(t))\,dt, \tag{2.16}$$

for Lipschitz functions $f_1, f_2, g_1, g_2, h\colon \mathbb{M}_N \to \mathbb{M}_N$. Then the following Itô product rules hold:

$$d(\Theta_1(t)\Theta_2(t)) = d\Theta_1(t)\cdot\Theta_2(t) + \Theta_1(t)\cdot d\Theta_2(t) + d\Theta_1(t)\cdot d\Theta_2(t) \tag{2.17}$$

$$\Theta_1(t)\,dX^N(t)\,\Theta_2(t)\,dt = \Theta_1(t)\,dY^N(t)\,\Theta_2(t)\,dt = 0. \tag{2.18}$$

Lemma 2.2 is straightforward to verify from the standard Itô calculus for vector-valued processes. Note, for example, that (2.14) is a consequence of the magic formula (2.9).

Remark 2.3. The global Lipschitz assumption on the drift and diffusion coefficient functions in (2.16) guarantee the existence of a unique solution for all time by the standard theory, cf. [10]. In all the examples considered presently, these functions will be first-order polynomials in the matrix entries.

2.2 Free Stochastic Calculus

For an introduction to noncommutative probability theory, and free probability in particular, we refer the reader to [21]. We assume familiarity with noncommutative probability spaces and W^*-probability spaces. The reader is directed to [17, Sections 1.1–1.3] for a quick introduction to free additive (semicircular) Brownian motion, and to [7, Section 1.3] for a brief introduction to free unitaty Brownian motion. Also, we give a brief discussion of free independence at the beginning of Section 2.3 below.

Let $(\mathscr{A},\tau)$ be a faithful, tracial W^*-probability space. To fix notation, for $a \in \mathscr{A}$ denote its noncommutative distribution as φ_a. I.e. letting $\mathbb{C}\langle X, X^*\rangle$ denote the noncommutative polynomials in two variables, $\varphi_a\colon \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ is the linear functional

$$\varphi_a(f) = \tau(f(a,a^*)), \qquad f \in \mathbb{C}\langle X, X^*\rangle.$$

A *free semicircular Brownian motion* $x(t)$ is a self-adjoint stochastic process $(x(t))_{t\ge 0}$ in $\mathscr{A}$ such that $x(0)=0$, $\mathrm{Var}(x(1)) = 1$, and the additive increments of x are stationary and freely independent: for $0 \le t_1 < t_2 < \infty$, $\varphi_{x(t_2)-x(t_1)} = \varphi_{x(t_2-t_1)}$, and $x(t_2)-x(t_1)$ is freely independent from the W^*-subalgebra $\mathscr{A} \supset \mathscr{A}_{t_1} \equiv W^*\{x(t)\colon 0 \le t \le t_1\}$. Since $x(t)$ is a bounded self-adjoint operator, its distribution is given by a compactly-supported probability measure on $\mathbb{R}$; the freeness of increments and stationarity then implies that $\varphi_{x(t_2)-x(t_1)}$ is the *semicircle law*: setting $t = t_2 - t_1$,

$$\tau[(x(t_2)-x(t_1))^n] = \int_{-2\sqrt{t}}^{2\sqrt{t}} s^n\,\frac{1}{2\pi t}\sqrt{4t-s^2}\,ds, \qquad n \in \mathbb{N}.$$

In [26], it was proven that, if $X^N(t)$ is a GUE_N Brownian motion, then the *process* $(X^N(t))_{t\ge0}$ converges to a free semicircular Brownian motion: for any n and any $t_1, t_2, \ldots, t_n \ge 0$, and any noncommutative polynomial $f \in \mathbb{C}\langle X_1, \ldots, X_n\rangle$,

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}[f(X^N(t_1),\ldots,X^N(t_n))] = \tau[f(x(t_1),\ldots,x(t_n))].$$

Appealing to Lemma 2.2, this paves the way to *free stochastic differential equations*.

Let $x(t)$ and $y(t)$ be two freely independent free semicircular Brownian motions in a W^*-probability space $(\mathscr{A},\tau)$, and let $\mathscr{A}_t = W^*\{x(s), y(s)\colon 0 \le s \le t\}$. Let $\theta(t), \theta_1(t), \theta_2(t)$ be processes that are adapted to the filtration $\mathscr{A}_t$. The *free Itô integral*

$$\int_0^t \theta_1(s)\,dx(s)\,\theta_2(s) \tag{2.19}$$

is defined in precisely the same manner as Itô integrals of real-valued processes with respect to real Brownian motion: as $L^2(\mathscr{A}_t,\tau)$-limits of sums $\sum_j \theta_1(t_j)(x(t_j)-x(t_{j-1}))\theta_2(t_j)$ over partitions $\{0=t_0\le\cdots\le t_n=t\}$ as the partition width $\sup_j |t_j-t_{j-1}|$ tends to 0. Standard Picard iteration techniques show that, if f_1,f_2,g_1,g_2,h are Lipschitz functions then the integral equation

$$b(t)=1+\int_0^t f_1(b(s))\,dx(s)\,f_2(b(s))+\int_0^t g_1(b(s))\,dy(s)\,g_2(b(s))+\int_0^t h(b(s))\,ds, \tag{2.20}$$

has a unique adapted solution $b(t)\in\mathscr{A}_t$ satisfying $b(0)=1$.

Remark 2.4. We are over-simplifying here: (2.19) should really be a sum of such terms (or a limit thereof) representing the stochastic integral of a *biprocess*. It is only possible to make sense of Lipschitz functional calculus for self-adjoint (or at least normal) biprocesses; in the simplified form of (2.19), this would require $\theta_1=\theta_2$. Otherwise, we are restricted to polynomial functions f_1,f_2,g_1,g_2,h, and the (global) Lipschitz requirement then limits the theory to first-order polynomials. Fortunately, that is sufficient for the present purposes (cf. (1.5)). The question of extending a more general theory of existence of solutions to free stochastic differential equations involving non-self-adjoint biprocesses is an active area of current research.

As usual, we use differential notation to express (2.20) in the form

$$db(t)=f_1(b(t))\,dx(t)\,f_2(b(t))+g_1(b(t))\,dy(t)\,g_2(b(t))+h(b(t))\,dt,\qquad b(0)=1. \tag{2.21}$$

We refer to (2.21) as a *free stochastic differential equation*. Solutions of such equations are called *free Itô processes*. The matrix stochasic calculus of Lemma 2.2 has a precise analogue for free Itô processes.

Lemma 2.5. *Let $(\mathscr{A},\tau)$ be a W^*-probability space containing two freely independent free semicircular Brownian motions $x(t)$ and $y(t)$, adapted to the filtration $\{\mathscr{A}_t\}_{t\ge0}$. Let $\theta(t),\theta_1(t),\theta_2(t)$ be processes adapted to $\mathscr{A}_t$. Then, for any $t\ge0$, the following hold:*

$$\tau\left(\int_0^t \theta_1(s)\,dx(s)\,\theta_2(s)\right)=\tau\left(\int_0^t \theta_1(s)\,dy(s)\,\theta_2(s)\right)=0 \tag{2.22}$$

$$\int_0^t dx(s)\,\theta(s)\,dx(s)=\int_0^t dy(s)\,\theta(s)\,dy(s)=\int_0^t \tau(\theta(s))\,ds \tag{2.23}$$

$$\int_0^t dx(s)\,\theta(s)\,dy(s)=\int_0^t dy(s)\,\theta(s)\,dx(s)=0. \tag{2.24}$$

Moreover, if $\theta_1(t)$ and $\theta_2(t)$ are free Itô processes, then the following Itô product rules hold:

$$d(\theta_1(t)\theta_2(t))=d\theta_1(t)\cdot\theta_2(t)+\theta_1(t)\cdot d\theta_2(t)+d\theta_1(t)\cdot d\theta_2(t) \tag{2.25}$$

$$\theta_1(t)\,dx(t)\,\theta_2(t)\,dt=\theta_1(t)\,dy(t)\,\theta_2(t)\,dt=0. \tag{2.26}$$

For a proof of Lemma 2.5, see [4].

2.3 Asymptotic Freeness

Definition 2.6. *Let $(\mathscr{A},\tau)$ be a noncommutative probability space. Unital $*$-subalgebras $\mathscr{A}_1,\ldots,\mathscr{A}_m\subset\mathscr{A}$ are called* **free** *with respect to τ if, given any $n\in\mathbb{N}$ and $k_1,\ldots,k_n\in\{1,\ldots,m\}$ such that $k_{j-1}\ne k_j$ for $1<j\le n$, and any elements $a_j\in\mathscr{A}_{k_j}$ with $\tau(a_j)=0$ for $1\le j\le n$, it follows that $\tau(a_1\cdots a_n)=0$. Random variables $a_1,\ldots,a_m$ are said to be* **freely independent** *if the unital $*$-algebras $\mathscr{A}_j=\langle a_j,a_j^*\rangle\subset\mathscr{A}$ they generate are free.*

Free independence is a $*$-moment factorization property. By centering $a_i - \tau(a_i)1_{\mathscr{A}} \in \mathscr{A}_i$, the freeness rule allows (inductively) any moment $\tau(a_{k_1}^{\varepsilon_1}\cdots a_{k_n}^{\varepsilon_n})$ to be decomposed as a polynomial in moments $\tau(a_i^{\varepsilon})$ in the variables separately. For example, if a, b are freely independent then $\tau(a^\varepsilon b^\delta) = \tau(a^\varepsilon)\tau(b^\delta)$, while

$$\tau(a^{\varepsilon_1}b^{\delta_1}a^{\varepsilon_2}b^{\delta_2}) = \tau(a^{\varepsilon_1})\tau(a^{\varepsilon_2})\tau(b^{\delta_1}b^{\delta_2}) + \tau(a^{\varepsilon_1}a^{\varepsilon_2})\tau(b^{\delta_1})\tau(b^{\delta_2}) - \tau(a^{\varepsilon_1})\tau(a^{\varepsilon_2})\tau(b^{\delta_1})\tau(b^{\delta_2}),$$

for any $\varepsilon, \varepsilon_1, \varepsilon_2, \delta, \delta_1, \delta_2 \in \{1, *\}$. In general, if $a_1, \ldots, a_n$ are freely independent, then their noncommutative joint distribution $\varphi_{a_1,\ldots,a_n}$ (a linear functional on $\mathbb{C}\langle X_1, \ldots, X_n, X_1^*, \ldots, X_n^*\rangle$) is determined by the individual distributions $\varphi_{a_1}, \ldots, \varphi_{a_n}$ (linear functionals on $\mathbb{C}\langle X, X^*\rangle$).

Let $L^{\infty-}(\Omega, \mathscr{F}, \mathbb{P}) = \bigcap_{p>1} L^p(\Omega, \mathscr{F}, \mathbb{P})$, and let $\mathbb{M}_N \otimes L^{\infty-}$ denote the algebra of $N \times N$ matrices with entries in $L^{\infty-}(\Omega, \mathscr{F}, \mathbb{P})$. There are scant few non-trivial instances of free independence in the noncommutative probability space $(\mathbb{M}_N \otimes L^{\infty-}, \mathbb{E}\mathrm{tr})$. However, *asymptotic freeness* abounds.

Definition 2.7. *Let $n \in \mathbb{N}$. For each $N \in \mathbb{N}$, let $A_1^N, \ldots, A_n^N$ be random matrices in $\mathbb{M}_N \otimes L^{\infty-}$. Say that $(A_1^N, \ldots, A_n^N)$ are* **asymptotically free** *if there is a noncommutative probability space $(\mathscr{A}, \tau)$ containing freely independent random variables $a_1, \ldots, a_n$ such that $(A_1^N, \ldots, A_n^N)$ converges in noncommutative distribution to $(a_1, \ldots, a_n)$.*

The general mantra for producing asymptotically free random matrices is as follows.

> If $A_1^N, \ldots, A_n^N$ are random matrices whose distribution is invariant under unitary conjugation, and possess a joint limit distribution, then they are asymptotically free.

The first result in this direction was proved in [26], where the matrices A_j^N were taken to have the form $A_j^N = U_j^N D_j^N (U_j^N)^*$ where $U_1^N, \ldots, U_n^N$ are independent Haar-distributed unitaries, and D_j^N are deterministic diagonal matrices with uniform bounds on their trace moments. This was later improved to include all deterministic matrices (with uniform bounds on their operator norms) in [27]; see, also, [6, 30] for related results. We will use the following form of the mantra, which is a weak form of [20, Theorem 1].

Theorem 2.8 (Mingo, Śniady, Speicher, 2007)**.** *Let $A_1^N, \ldots, A_n^N$ be independent random matrices in $\mathbb{M}_N \otimes L^{\infty-}$, with the following properties.*

(1) The joint law of $A_1^N, \ldots, A_n^N$ is invariant under conjugation by unitary matrices in $\mathbb{U}_N$.

(2) There is a joint limit distribution: for any noncommutative polynomial $f \in \mathbb{C}\langle X_1, \ldots, X_n, X_1^*, \ldots, X_n^*\rangle$, $\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(f(A_1^N, \ldots, A_n^N, (A_1^N)^*, \ldots, (A_n^N)^*))$ exists.

(3) The fluctuations are $O(1/N^2)$: for any noncommutative polynomials f, g as in (2), there is a constant $C = C(f, g)$ so that

$$\mathrm{Cov}\left[\mathrm{tr}\big(f(A_1^N, \ldots, A_n^N, (A_1^N)^*, \ldots, (A_n^N)^*)\big), \mathrm{tr}\big(g(A_1^N, \ldots, A_n^N, (A_1^N)^*, \ldots, (A_n^N)^*)\big)\right] \le \frac{C}{N^2}.$$

Then $A_1^N, \ldots, A_n^N$ are asymptotically free.

Remark 2.9. [20, Theorem 1] has a much stronger assumption than (3): it also assumes that the classical cumulants k_r in normalized traces of noncommutative polynomials are $o(1/N^r)$ for all $r > 2$, thus producing a so-called *second-order limit distribution*. However, this stronger assumption is used only to produce a stronger conclusion: that the matrices are asymptotically free *of second-order*. Following the proof, it is relatively easy to see that Theorem 2.8 is proved along the way, at least in the case $n = 2$. To go from 2 to general finite n can be achieved by induction together with the associativity of freeness; cf. [29, Proposition 2.5.5(iii)]. See, also, [19] where this is proved more explicitly in the harder case of real random matrices (where $\mathbb{U}_N$-invariance is replaced with $\mathbb{O}_N$-invariance).

3 Heat Kernels on $\mathbb{GL}_N^n$

Here we generalize the technology we developed in [9, Sections 3.4 & 4.1] to independent products of heat kernel measures on $\mathbb{GL}_N$.

3.1 Laplacians on $\mathbb{GL}_N^n$

Let $n, N \in \mathbb{N}$. Then $\mathbb{GL}_N^n = \mathbb{GL}_N \times \cdots \times \mathbb{GL}_N$ is a Lie group of real dimension $2nN^2$. Its Lie algebra is $\mathfrak{gl}_N^n = \mathfrak{gl}_N \oplus \cdots \oplus \mathfrak{gl}_N$. For $\xi \in \mathfrak{gl}_N$, and $1 \le j \le n$, let ξ_j denote the vector $(0, \ldots, 0, \xi, 0, \ldots, 0) \in \mathfrak{gl}_N^n$ (with ξ in the jth component). The Lie product on $\mathfrak{gl}_N^n$ is then determined by $[\xi_j, \eta_k] = \delta_{jk}(\xi_j\eta_k - \eta_k\xi_j)$ for $1 \le j, k \le n$. In particular, if $j \ne k$ and $\xi, \eta \in \mathfrak{gl}_N$, then the left-invariant derivations ∂_{ξ_j} and ∂_{η_k} on $C^\infty(\mathbb{GL}_N^n)$ commute. To be clear, note that, for $f \in C^\infty(\mathbb{GL}_N^n)$,

$$(\partial_{\xi_j} f)(A_1, \ldots, A_n) = \left.\frac{d}{dt}\right|_{t=0} f(A_1, \ldots, A_{j-1}, A_j e^{t\xi}, A_{j+1}, \ldots, A_n). \tag{3.1}$$

Let $\beta_{r,s}^N$ denote an orthonormal basis for $\mathfrak{gl}_N$ (with respect to $\langle \cdot, \cdot \rangle_{r,s}^N$, as in (2.7)). For $1 \le j \le n$, define

$$\Delta_{r,s}^{j,N} = \sum_{\xi \in \beta_{r,s}^N} \partial_{\xi_j}^2. \tag{3.2}$$

Note that $\Delta_{r,s}^{j,N}$ and $\Delta_{r,s}^{k,N}$ commute for all j, k. Now, fix $t_1, \ldots, t_n > 0$. Then the operator

$$t_1 \Delta_{r,s}^{1,N} + \cdots + t_n \Delta_{r,s}^{n,N}$$

is elliptic, essentially self-adjoint on $C_c^\infty(\mathbb{GL}_N^n)$, and non-positive. We may therefore use the spectral theorem to define the bounded operator

$$e^{\frac{1}{2}(t_1 \Delta_{r,s}^{1,N} + \cdots + t_n \Delta_{r,s}^{n,N})} = e^{\frac{1}{2} t_1 \Delta_{r,s}^{1,N}} \cdots e^{\frac{1}{2} t_n \Delta_{r,s}^{n,N}}.$$

Define the **heat kernel measure** $\mu_{r,s;t_1,\ldots,t_n}^{n,N}$ on $\mathbb{GL}_N^n$ by

$$\int_{\mathbb{GL}_N^n} f \, d\mu_{r,s;t_1,\ldots,t_n}^{n,N} = \left(e^{\frac{1}{2}(t_1 \Delta_{r,s}^{1,N} + \cdots + t_n \Delta_{r,s}^{n,N})} f \right)(I_N^n), \qquad f \in C_c(\mathbb{GL}_N^n), \tag{3.3}$$

where $I_N^n = (I_N, \ldots, I_N) \in \mathbb{GL}_N^n$. In particular, let $K_1, \ldots, K_n \subset \mathbb{GL}_N$ be compact sets; by approximating $\mathbb{1}_{K_1 \times \cdots \times K_n}$ with a continuous function, we see that

$$\mu_{r,s;t_1,\ldots,t_n}^{n,N}(K_1 \times \cdots \times K_n) = \left(e^{\frac{1}{2} t_1 \Delta_{r,s}^N} \mathbb{1}_{K_1} \right)(I_N) \cdots \left(e^{\frac{1}{2} t_n \Delta_{r,s}^N} \mathbb{1}_{K_n} \right)(I_N) = \mu_{r,s;t_1}^{1,N}(K_1) \cdots \mu_{r,s;t_n}^{1,N}(K_n).$$

Since $\mu_{r,s;t}^{1,N}$ is the heat kernel measure on $\mathbb{GL}_N$ corresponding to $\Delta_{r,s}^N$, it is the distribution of the Brownian motion $B_{r,s}^N(t)$, and so we have shown the following.

Lemma 3.1. *Let $(B_{r,s}^{1,N}(t))_{t \ge 0}, \ldots, (B_{r,s}^{n,N}(t))_{t \ge 0}$ be n independent (r,s)-Brownian motions on $\mathbb{GL}_N$. Then the joint law of the random vector $(B_{r,s}^{1,N}(t_1), \ldots, B_{r,s}^{n,N}(t_n))$ is $\mu_{r,s;t_1,\ldots,t_n}^{n,N}$.*

3.2 Multivariate Trace Polynomials

Let J be an index set (for our purposes in this section, we will usually take $J = \{1, \ldots, n\}$ for some $n \in \mathbb{N}$). Let $\mathscr{E}_J$ denote the set of all nonempty words in $J \times \{1, *\}$, $\mathscr{E}_J = \bigcup_{n\in\mathbb{N}} (J \times \{1, *\})^n$. Let $\mathbf{v}_J = \{v_\varepsilon : \varepsilon \in \mathscr{E}_J\}$ be commuting variables, and let

$$\mathscr{P}(J) = \mathbb{C}[\mathbf{v}_J]$$

be the algebra of (commutative) polynomials in the variables $\mathbf{v}_J$. That is: as a $\mathbb{C}$-vector space, $\mathscr{P}(J)$ has as its standard basis 1 together with the monomials

$$v_{\varepsilon^{(1)}} \cdots v_{\varepsilon^{(k)}}, \quad k \in \mathbb{N}, \quad \varepsilon^{(1)}, \ldots, \varepsilon^{(k)} \in \mathscr{E}_J, \tag{3.4}$$

and the (commutative) product on $\mathscr{P}(J)$ is the standard polynomial product.

We may identify monomials in $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$ with the variables v_ε, via

$$\Upsilon(X_{j_1}^{\varepsilon_1} \cdots X_{j_k}^{\varepsilon_k}) = v_{((j_1,\varepsilon_1),\ldots,(j_k,\varepsilon_k))}.$$

Extending linearly, $\Upsilon : \mathbb{C}\langle X_j, X_j^* : j \in J\rangle \hookrightarrow \mathscr{P}(J)$ is a linear inclusion, identifying $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$ with the *linear* polynomials in $\mathscr{P}(J)$. The algebra $\mathscr{P}(J)$ is the "universal enveloping algebra" of $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$, in the following sense: any linear functional φ on $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$ extends (via Υ) uniquely to an algebra homomorphism $\widetilde{\varphi} : \mathscr{P}(J) \to \mathbb{C}$. Conversely, any algebra homomorphism $\mathscr{P}(J) \to \mathbb{C}$ is determined by its restriction to $\Upsilon(\mathbb{C}\langle X_j, X_j^* : j \in J\rangle)$, which intertwines a unique linear functional on $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$. Hence, the noncommutative distribution $\varphi_{\{a_j : j\in J\}}$ of J random variables can be equivalently represented as an algebra homomorphism $\mathscr{P}(J) \to \mathbb{C}$.

Definition 3.2. *For a monomial (3.4), the* **trace degree** *is defined to be*

$$\deg(v_{\varepsilon^{(1)}} \cdots v_{\varepsilon^{(k)}}) = |\varepsilon^{(1)}| + \cdots + |\varepsilon^{(k)}|,$$

where $|\varepsilon| = n$ if $\varepsilon \in (J \times \{1, *\})^n$. More generally, if $P \in \mathscr{P}(J)$, then $\deg(P)$ is the maximal trace degree of the monomial terms in P. Define $\deg(0) = 0$. Note that $\deg(PQ) = \deg(P) + \deg(Q)$, and $\deg(P + Q) \le \max\{\deg(P), \deg(Q)\}$ for $P, Q \in \mathscr{P}(J)$. For $d \in \mathbb{N}$, denote by $\mathscr{P}_d(J)$ the subspace

$$\mathscr{P}_d(J) = \{P \in \mathscr{P}(J) : \deg(P) \le d\}.$$

Note that $\mathscr{P}_d(J)$ is finite dimensional (if J is finite), and $\mathscr{P}(J) = \bigcup_{d\ge 1} \mathscr{P}_d(J)$.

We now introduce a kind of functional calculus for $\mathscr{P}(J)$.

Definition 3.3. *Let $(\mathscr{A}, \tau)$ be a noncommutative probability space. Let J be an index set, and let $\{a_j : j \in J\}$ be specified elements in $\mathscr{A}$. For $n \in \mathbb{N}$, and $(J \times \{1, *\})^n \ni \varepsilon = ((j_1, \varepsilon_1), \ldots, (j_n, \varepsilon_n))$, define*

$$a^\varepsilon \equiv a_{j_1}^{\varepsilon_1} \cdots a_{j_n}^{\varepsilon_n}.$$

We define for each $P \in \mathscr{P}(J)$ a complex number $P_\tau(a_j : j \in J)$ as follows: for $\varepsilon \in \mathscr{E}_J$, $[v_\varepsilon]_\tau(a_j : j \in J) = \tau(a^\varepsilon)$; and, in general, the map $P \mapsto P_\tau(a_j : j \in J)$ is an algebra homomorphism from $\mathscr{P}(J)$ to $\mathbb{C}$.

In other words: P_τ is the unique algebra homomorphism extending (via Υ) the linear functional $\varphi_{\{a_j : j\in J\}}$ on $\mathbb{C}\langle X_j, X_j^* : j \in J\rangle$ (i.e. the noncommutative distribution of $\{a_j : j \in J\}$).

Example 3.4. Let $J = \{1, 2\}$, and consider $\mathscr{P}(J) \ni P = v_{(1,*),(2,1),(1,1)} - 2v_{(2,1)}^2$, which has trace degree 3; then

$$P_\tau(a_1, a_2) = \tau(a_1^* a_2 a_1) - 2\left(\tau(a_2)\right)^2.$$

We generally refer to the functions $\{P_\tau : P \in \mathscr{P}(J)\}$ as (multivariate) **trace polynomials**.

Notation 3.5. *For $N \in \mathbb{N}$, in the noncommutative probability space $(\mathbb{M}_N, \mathrm{tr})$, we denote the evaluation map $P \mapsto P_{\mathrm{tr}}$ of Definition 3.3 as $P \mapsto P_N$. Thus, if $A_1, \ldots, A_n \in \mathbb{M}_N \otimes L^{\infty-}$, and P is as in Example 3.4, then $P_N(A_1, \ldots, A_n) = \mathrm{tr}(A_1^* A_2 A_1) - 2\left(\mathrm{tr}(A_2)\right)^2$, which is a random variable, to be clear.*

3.3 Intertwining Formula

The following "magic formulas" appeared as [9, Proposition 1]; note that (2.10) is a special case of (3.5).

Proposition 3.6. *Let β_N be an orthonormal basis for $\mathfrak{u}_N$ with respect to the inner product (2.5). Then for any $A \in \mathbb{M}_N$*

$$\sum_{\xi\in\beta_N} \xi A\xi = -\mathrm{tr}(A)I_N, \tag{3.5}$$

$$\sum_{\xi\in\beta_N} \mathrm{tr}(A\xi)\xi = -\frac{1}{N^2}A. \tag{3.6}$$

For the remainder of this section, we usually suppress the indices r, s for notational convenience; so, for example, $\Delta^{j,N} \equiv \Delta^{j,N}_{r,s}$ for $1 \le j \le n$. Let $J = \{1, \ldots, n\}$ throughout.

Theorem 3.7. *Let $j \in J$. There are collections $\left\{Q^j_\varepsilon \colon \varepsilon \in \mathscr{E}_J\right\}$ and $\left\{R^j_{\varepsilon,\delta} \colon \varepsilon, \delta \in \mathscr{E}_J\right\}$ in $\mathscr{P}(J)$ with the following properties.*

(1) For each $\varepsilon \in \mathscr{E}_J$, Q^j_ε is a finite sum of monomials of homogeneous trace degree $|\varepsilon|$ such that

$$\Delta^{j,N}\left([v_\varepsilon]_N\right) = [Q^j_\varepsilon]_N.$$

(2) For each $\varepsilon, \delta \in \mathscr{E}_J$, $R^j_{\varepsilon,\delta}$ is a finite sum of monomials of homogeneous trace degree $|\varepsilon| + |\delta|$ such that

$$r\sum_{\xi\in\beta_N}(\partial_{\xi_j}[v_\varepsilon]_N)(\partial_{\xi_j}[v_\delta]_N) + s\sum_{\xi\in\beta_N}(\partial_{i\xi_j}[v_\varepsilon]_N)(\partial_{i\xi_j}[v_\delta]_N) = \frac{1}{N^2}[R^j_{\varepsilon,\delta}]_N,$$

for any orthonormal basis β_N of $\mathfrak{u}_N$.

Please note that Q^j_ε and $R^j_{\varepsilon,\delta}$ *do not depend on N*. The $1/N^2$ in (2) comes from the magic formula (3.6), as we will see in the proof.

Proof. Fix $\mathscr{E}_J \ni \varepsilon = ((j_1, \varepsilon_1), \ldots, (j_m, \varepsilon_m))$; then $[v_\varepsilon]_N(A_1, \ldots, A_n) = \mathrm{tr}(A^{\varepsilon_1}_{j_1} \cdots A^{\varepsilon_m}_{j_m})$. Applying the product rule, for any $\xi \in \beta_N$ we have

$$\partial^2_{\xi_j}\left([v_\varepsilon]_N\right) = \sum_{k=1}^m \delta_{j,j_k}\mathrm{tr}(A^{\varepsilon_1}_{j_1}\cdots(A_{j_k}\xi^2)^{\varepsilon_k}\cdots A^{\varepsilon_m}_{j_m}) \tag{3.7}$$

$$+ 2\sum_{1\le k<\ell\le m}\delta_{j,j_k}\delta_{j,j_\ell}\mathrm{tr}(A^{\varepsilon_1}_{j_1}\cdots(A_{j_k}\xi)^{\varepsilon_k}\cdots(A_{j_\ell}\xi)^{\varepsilon_\ell}\cdots A^{\varepsilon_m}_{j_m}). \tag{3.8}$$

Similarly, $\partial^2_{i\xi_j}$ is given by the same formula but possibly with some minus signs in some of the terms (depending on $\varepsilon_k, \varepsilon_\ell$). For convenience, let $\beta^+_N = \beta_N$ and $\beta^-_N = i\beta_N$. Magic formula (2.10) gives $\sum_{\xi\in\beta^\pm_N}\xi^2 = \mp I_N$, and so summing over $\beta^\pm_N$ we have, for each k,

$$\sum_{\xi\in\beta^\pm_N}\mathrm{tr}(A^{\varepsilon_1}_{j_1}\cdots(A_{j_k}\xi^2)^{\varepsilon_k}\cdots A^{\varepsilon_m}_{j_m}) = \pm[v_\varepsilon]_N,$$

where the $\pm$ on the left and right do not necessarily match (we will not keep careful track of signs through this proof). Thus, (3.7) summed over $\beta^\pm_N$ gives some integer multiple $n^\pm_j(\varepsilon)$ of $[v_\varepsilon]_N$. Summing the terms in (3.8) over $\xi \in \beta^\pm_N$, using (3.5), yields

$$\sum_{\xi\in\beta^\pm_N}\mathrm{tr}(A^{\varepsilon_1}_{j_1}\cdots(A_{j_k}\xi)^{\varepsilon_k}\cdots(A_{j_\ell}\xi)^{\varepsilon_\ell}\cdots A^{\varepsilon_m}_{j_m}) = \pm[v_{\varepsilon_{k,\ell}}]_N[v_{\varepsilon'_{k,\ell}}]_N,$$

where $\varepsilon_{k,\ell}$ is a substring of ε (running between index k or $k+1$ and index $\ell-1$ or ℓ, depending on $\varepsilon_k, \varepsilon_\ell$) and $\varepsilon'_{k,\ell}$ is the concatenation of the two remaining substrings of ε when $\varepsilon_{k,\ell}$ is removed. Hence, define

$$Q^{j,\pm}_\varepsilon = n_\pm(\varepsilon) v_\varepsilon + 2 \sum_{1\le k<\ell\le m} \pm \delta_{j,j_k}\delta_{j,j_\ell} v_{\varepsilon_{k,\ell}} v_{\varepsilon'_{k,\ell}}.$$

Note that $|\varepsilon| = |\varepsilon_{k,\ell}| + |\varepsilon'_{k,\ell}|$ for each k,ℓ; so $Q^{j,\pm}_\varepsilon$ are homogeneous of trace degree $|\varepsilon|$. The above argument shows that

$$\sum_{\xi\in\beta_N^\pm} \partial^2_{\xi_j}[v_\varepsilon]_N = [Q^{j,\pm}_\varepsilon]_N,$$

and so setting $Q^j_\varepsilon = rQ^{j,+}_\varepsilon + sQ^{j,-}_\varepsilon$ completes item (1) of the theorem.

For item (2), fix $\mathscr{E}_J \ni \delta = ((h_1,\delta_1),\ldots,(h_p,\delta_p))$; then $[v_\delta]_N(A_1,\ldots,A_n) = \mathrm{tr}(A^{\delta_1}_{h_1}\cdots A^{\delta_p}_{h_p})$. Thus, for $\xi\in\beta_N$,

$$(\partial_\xi[v_\varepsilon]_N)(\partial_\xi[v_\delta]_N) = \sum_{k=1}^m\sum_{\ell=1}^p \delta_{j,j_k}\delta_{j,h_\ell}\mathrm{tr}(A^{\varepsilon_1}_{j_1}\cdots(A_{j_k}\xi)^{\varepsilon_k}\cdots A^{\varepsilon_m}_{j_m})\mathrm{tr}(A^{\delta_1}_{h_1}\cdots(A_{h_\ell}\xi)^{\delta_\ell}\cdots A^{\delta_p}_{h_p}). \tag{3.9}$$

(To be clear: the terms $\delta_{j,j_k}\delta_{j,h_\ell}$ are indicator functions, not related to the string $\delta\in\mathscr{E}_J$.) Taking $\partial_{i\xi_j}$ instead yields the same formula, possibly with some minus signs inside the sum (depending on ε_k and δ_ℓ). We can write each term in (3.9) in the form

$$\pm\mathrm{tr}(\xi A^{\varepsilon^{(k)}})\mathrm{tr}(\xi A^{\delta^{(\ell)}})$$

where $\varepsilon^{(k)}$ and $\delta^{(\ell)}$ are certain cyclic permutations of ε and δ. Using (3.6), summing over $\xi\in\beta^\pm_N$ then yields

$$\frac{1}{N^2}\sum_{k=1}^m\sum_{\ell=1}^p \pm\delta_{j,j_k}\delta_{j,h_\ell}[v_{\varepsilon^{(k)}\delta^{(\ell)}}]_N,$$

where $\varepsilon^{(k)}\delta^{(\ell)}$ denotes the concatenation; in particular, $|\varepsilon^{(k)}\delta^{(\ell)}| = |\varepsilon|+|\delta|$. Thus, setting

$$R^{j,\pm}_{\varepsilon,\delta} = \sum_{k=1}^m\sum_{\ell=1}^p \pm\delta_{j,j_k}\delta_{j,h_\ell} v_{\varepsilon^{(k)}\delta^{(\ell)}}$$

(where the $\pm$ on the two sides do not necessarily match), we have shown that

$$\sum_{\xi\in\beta_N^\pm}(\partial_{\xi_j}[v_\varepsilon]_N)(\partial_{\xi_j}[v_\delta]_N) = \frac{1}{N^2}[R^{j,\pm}_{\varepsilon,\delta}]_N.$$

Set $R^j_{\varepsilon,\delta} \equiv rR^{j,+}_{\varepsilon,\delta} + sR^{j,-}_{\varepsilon,\delta}$; then $R^j_{\varepsilon,\delta}$ has homogeneous trace degree $|\varepsilon|+|\delta|$, and so satisfies item (2), concluding the proof of the theorem. □

Theorem 3.8 (Intertwining Formula)**.** *For $j\in J$, let $\left\{Q^j_\varepsilon\colon \varepsilon\in\mathscr{E}_J\right\}$ and $\left\{R^j_{\varepsilon,\delta}\colon \varepsilon,\delta\in\mathscr{E}_J\right\}$ be the collections in $\mathscr{P}(J)$ given in Theorem 3.7. Define the following operators on $\mathscr{P}(J)$:*

$$\mathcal{D}^j_{r,s} = \sum_{\varepsilon\in\mathscr{E}_J} Q^j_\varepsilon\frac{\partial}{\partial v_\varepsilon} \qquad \text{and} \qquad \mathcal{L}^j_{r,s} = \sum_{\varepsilon,\delta\in\mathscr{E}_J} R^j_{\varepsilon,\delta}\frac{\partial^2}{\partial v_\varepsilon\partial v_\delta}. \tag{3.10}$$

Then $\mathcal{D}^j_{r,s}$ and $\mathcal{L}^j_{r,s}$ preserve trace degree (when $(r,s)\neq(0,0)$), and, for all $P\in\mathscr{P}(J)$,

$$\Delta^{j,N}_{r,s}([P]_N) = \left[\left(\mathcal{D}^j_{r,s} + \frac{1}{N^2}\mathcal{L}^j_{r,s}\right)P\right]_N \tag{3.11}$$

Proof. The proof is almost identical to the proof of [9, Theorem 3.26]; we repeat it here. Let $\mathbf{V}_N \colon \mathbb{GL}_N^n \to \mathbb{M}_N^{\mathscr{E}_J}$ be the map

$$(\mathbf{V}_N(A_1,\ldots,A_n))\,((j_1,\varepsilon_1),\ldots,(j_m,\varepsilon_m)) = \mathrm{tr}(A_{j_1}^{\varepsilon_1}\cdots A_{j_m}^{\varepsilon_m}).$$

Then, by definition, $[P]_N = P\circ\mathbf{V}_N$. By the chain rule, if $\xi\in\mathfrak{gl}_N$ then

$$\begin{aligned}\partial_{\xi_j}^2 P_N = \partial_{\xi_j}^2(P\circ\mathbf{V}_N) &= \sum_{\varepsilon\in\mathscr{E}}\partial_{\xi_j}\left[\left(\frac{\partial P}{\partial v_\varepsilon}\right)(\mathbf{V}_N)\cdot\partial_{\xi_j}[v_\varepsilon]_N\right]\\ &= \sum_{\varepsilon\in\mathscr{E}}\left(\frac{\partial P}{\partial v_\varepsilon}\right)(\mathbf{V}_N)\cdot\partial_{\xi_j}^2([v_\varepsilon]_N) + \sum_{\varepsilon,\delta\in\mathscr{E}}\left(\frac{\partial^2 P}{\partial v_\varepsilon\partial v_\delta}\right)(\mathbf{V}_N)\cdot\left(\partial_{\xi_j}[v_\varepsilon]_N\right)\left(\partial_{\xi_j}[v_\delta]_N\right)\end{aligned}$$

from which it follows that

$$\begin{aligned}\Delta_{r,s}^{j,N}P_N &= \sum_{\varepsilon\in\mathscr{E}}\left(\frac{\partial P}{\partial v_\varepsilon}\right)(\mathbf{V}_N)\cdot\Delta_{r,s}^{j,N}([v_\varepsilon])\\ &\quad+\sum_{\varepsilon,\delta\in\mathscr{E}}\left(\frac{\partial^2 P}{\partial v_\varepsilon\partial v_\delta}\right)(\mathbf{V}_N)\cdot\left[r\sum_{\xi\in\beta_N}(\partial_{\xi_j}[v_\varepsilon])(\partial_{\xi_j}[v_\delta]) + s\sum_{\xi\in i\beta_N}(\partial_{\xi_j}[v_\varepsilon])(\partial_{\xi_j}[v_\delta]_N)\right].\end{aligned}$$

Combining this equation with the results of Theorem 3.7 completes the proof. □

This prompts us to define the following operators.

Definition 3.9. *Let* $\mathbf{t} = (t_1,\ldots,t_n)$ *for some* $t_1,\ldots,t_n > 0$. *Define*

$$\mathcal{D}_{r,s}^{\mathbf{t}} = \frac{1}{2}\sum_{j=1}^n t_j\mathcal{D}_{r,s}^j, \qquad \mathcal{L}_{r,s}^{\mathbf{t}} = \frac{1}{2}\sum_{j=1}^n t_j\mathcal{L}_{r,s}^j.$$

Corollary 3.10. *For any* $\mathbf{t} = (t_1,\ldots,t_n)\in\mathbb{R}_+^n$, *and* $d\in\mathbb{N}$, $\mathcal{D}_{r,s}^{\mathbf{t}}$ *and* $\mathcal{L}_{r,s}^{\mathbf{t}}$ *preserve the finite dimensional space* $\mathscr{P}_d(J)$, *and*

$$e^{\frac{1}{2}(t_1\Delta_{r,s}^{1,N}+\cdots+t_n\Delta_{r,s}^{n,N})}P_N = [e^{\mathcal{D}_{r,s}^{\mathbf{t}}+\frac{1}{N^2}\mathcal{L}_{r,s}^{\mathbf{t}}}P]_N, \qquad P\in\mathscr{P}_d(J).$$

In particular, $e^{\mathcal{D}_{r,s}^{\mathbf{t}}+\frac{1}{N^2}\mathcal{L}_{r,s}^{\mathbf{t}}}$ *and* $e^{\mathcal{D}_{r,s}^{\mathbf{t}}}$ *are well-defined operators on the space* $\mathscr{P}(J)$.

Proof. Since $\mathcal{D}_{r,s}^j$ and $\mathcal{L}_{r,s}^j$ preserve trace degree, the corollary follows by expanding the exponentials as power series of operators acting on the finite dimensional spaces $\mathscr{P}_d(J)$ and $[\mathscr{P}_d(J)]_N$. □

Remark 3.11. Since $\Delta_{r,s}^{j,N}$ commute for $1\le j\le n$, it is natural to expect the same holds for the intertwining operators $\mathcal{D}_{r,s}^j$ and $\mathcal{L}_{r,s}^j$. This is true, and follows easily from examining the explicit form of the coefficients of these operators given in Theorem 3.7. One must be careful about drawing such conclusions in general, however; the map $P\mapsto P_N$ is generally not one-to-one, due to the Cayley-Hamilton Theorem. It is *asymptotically* one-to-one, in the sense that its restriction to $\mathscr{P}_d(J)$ is one-to-one for all sufficiently large N (depending on d), and this can be used to prove this commutation result. Note, however, that $[\mathcal{D}_{r,s}^j,\mathcal{L}_{r,s}^j]\neq 0$ in general.

3.4 Concentration of Measure

We restate a general linear algebra result here, given as [9, Lemma 4.1].

Lemma 3.12. *Let V be a finite dimensional normed $\mathbb{C}$-space and supposed that D and L are two operators on V. Then there exists a constant $C = C(D, L, \|\cdot\|_V) < \infty$ such that, for any linear functional $\psi \in V^*$,*

$$\left|\psi(e^{D+\epsilon L}x) - \psi(e^D x)\right| \le C\|\psi\|_{V^*}\|x\|_V|\epsilon|, \quad x \in V,\ |\epsilon| \le 1, \tag{3.12}$$

where $\|\cdot\|_{V^*}$ is the dual norm on V^*.

Coupled with Corollary 3.10, this gives the following.

Proposition 3.13. *Let $P \in \mathscr{P}(J)$. Let $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}_+^n$. Then there is a constant $C = C(r, s, \mathbf{t}, P)$ so that, for all $N \in \mathbb{N}$,*

$$\left|\int_{\mathbb{GL}_N^n} P_N \, d\mu_{r,s;\mathbf{t}}^{n,N} - \left(e^{\mathcal{D}_{r,s}^{\mathbf{t}}} P\right)(\mathbf{1})\right| \le \frac{C}{N^2},$$

where, for $Q \in \mathscr{P}(J)$, $Q(\mathbf{1})$ is the complex number given by evaluating all variables of Q at 1.

Proof. Let $d = \deg(P)$; then $P \in \mathscr{P}_d(J)$. By definition (3.3),

$$\int_{\mathbb{GL}_N^n} P_N \, d\mu_{r,s;\mathbf{t}}^{n,N} = \left(e^{\frac{1}{2}(t_1\Delta_{r,s}^{1,N} + \cdots + t_n\Delta_{r,s}^{n,N})} P_N\right)(I_N^n).$$

(To be clear: the function P_N is not compactly-supported, so this does not fall strictly into the purview of (3.3); that the formula extends to such trace polynomials follows from Langland's Theorem; cf. [24, Theorem 2.1 (p. 152)]. See [9, Appendix A] for a concise sketch of the proof.) From Corollary 3.10, therefore

$$\int_{\mathbb{GL}_N^n} P_N \, d\mu_{r,s;\mathbf{t}}^{n,N} = \left[e^{\mathcal{D}_{r,s}^{\mathbf{t}} + \frac{1}{N^2}\mathcal{L}_{r,s}^{\mathbf{t}}} P\right]_N (I_N^n) = \left(e^{\mathcal{D}_{r,s}^{\mathbf{t}} + \frac{1}{N^2}\mathcal{L}_{r,s}^{\mathbf{t}}} P\right)(\mathbf{1}).$$

Note that $\psi_{\mathbf{1}}(P) = P(\mathbf{1})$ is a linear functional on the finite dimensional space $\mathscr{P}_d(J)$; thus the result follows from (3.12) by choosing any norm $\|\cdot\|_{\mathscr{P}_d(J)}$ on $V = \mathscr{P}_d(J)$, and setting

$$C(r, s, \mathbf{t}, P) = C(\mathcal{D}_{r,s}^{\mathbf{t}}, \mathcal{L}_{r,s}^{\mathbf{t}}, \|\cdot\|_{\mathscr{P}_d(J)})\|\psi_{\mathbf{1}}\|_{\mathscr{P}_d(J)^*}\|P\|_{\mathscr{P}_d(J)},$$

thus concluding the proof. □

We now come to the main theorems of this section.

Theorem 3.14. *Let $(B_{r,s}^{1,N}(t))_{t\ge0}, \ldots, (B_{r,s}^{n,N}(t))_{t\ge0}$ be independent Brownian motions on $\mathbb{GL}_N$. Then these matrix processes have a joint limit distribution: for any $m \in \mathbb{N}$, $j_1, \ldots, j_m \in \{1, \ldots, n\}$, $t_1, \ldots, t_n \ge 0$ and $\varepsilon_1, \ldots, \varepsilon_m \in \{1, *\}$,*

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{r,s}^{j_1,N}(t_{j_1})^{\varepsilon_1} \cdots B_{r,s}^{j_m,N}(t_{j_m})^{\varepsilon_m}) \qquad \textit{exists.}$$

Proof. Let $\mathbf{t} = (t_1, \ldots, t_n)$. The given expected trace is computed in terms of the joint law $\mu_{r,s;\mathbf{t}}^{n,N}$ of the independent Brownian random matrices as

$$\mathbb{E}\mathrm{tr}(B_{r,s}^{j_1,N}(t_{j_1})^{\varepsilon_1} \cdots B_{r,s}^{j_m,N}(t_{j_m})^{\varepsilon_m}) = \int_{\mathbb{GL}_N^n} \mathrm{tr}(A_{j_1}^{\varepsilon_1} \cdots A_{j_m}^{\varepsilon_m})\, \mu_{r,s;\mathbf{t}}^{n,N}(dA_1 \cdots dA_n) = \int_{\mathbb{GL}_N^n} [v_\varepsilon]_N \, d\mu_{r,s;\mathbf{t}}^{n,N}$$

where $\varepsilon = ((j_1, \varepsilon_1), \ldots, (j_m, \varepsilon_m))$. Proposition 3.13 thus shows that the limit as $N \to \infty$ exists, and is equal to $\left(e^{\mathcal{D}_{r,s}^{\mathbf{t}}} v_\varepsilon\right)(\mathbf{1})$. □

Remark 3.15. In light of Theorem 1.1, we can identify the joint limit in Theorem 3.14 as the increments $b_{r,s}(t_1), b_{r,s}(t_1)^{-1}b_{r,s}(t_2'), \ldots, b_{r,s}(t_{n-1}')^{-1}b_{r,s}(t_n')$ where $t_j = t_1 + \cdots + t_j$ for $1 \le j \le n$.

Theorem 3.16. *Let $P, Q \in \mathscr{P}(J)$, and let $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}_+^n$. There is a constant $C_2 = C_2(r, s, \mathbf{t}, P, Q)$ such that*

$$\left|\operatorname{Cov}_{\mu^{n,N}_{r,s;\mathbf{t}}}(P_N, Q_N)\right| \le \frac{C_2}{N^2}. \tag{3.13}$$

Theorem 3.16 is a generalization of [16, Proposition 4.13], and the proof is very similar. First, we need a lemma on intertwining complex conjugation, which is elementary to prove and left to the reader; cf. [16, Lemma 3.11].

Lemma 3.17. *Given $\varepsilon \in \mathscr{E}_J$, define $\varepsilon^* \in \mathscr{E}_J$ by $((j_1, \varepsilon_1), \ldots, (j_n, \varepsilon_n))^* = ((j_n, \varepsilon_n^*), \ldots, (j_1, \varepsilon_1^*))$ where $1^* = *$ and $*^* = 1$. Define $\mathcal{C}\colon \mathscr{P}(J) \to \mathscr{P}(J)$ to be the conjugate linear homomorphism satisfying $\mathcal{C}(v_\varepsilon) = v_{\varepsilon^*}$ for all $\varepsilon \in \mathscr{E}_J$. Then for all $N \in \mathbb{N}$*

$$\overline{P_N} = [\mathcal{C}(P)]_N, \qquad P \in \mathscr{P}(J). \tag{3.14}$$

Proof of Theorem 3.16. The covariance of $\mathbb{C}$-valued random variables F, G is $\operatorname{Cov}(F, G) = \mathbb{E}(F\overline{G}) - \mathbb{E}(F)\mathbb{E}(\overline{G})$. Define $\mathcal{D}^{\mathbf{t},N}_{r,s} = \mathcal{D}^{\mathbf{t}}_{r,s} + \frac{1}{N^2}\mathcal{L}^{\mathbf{t}}_{r,s}$. From Lemma 3.17, we may write $P_N\overline{Q_N} = [PQ^*]_N$, and so, from (3.3) and Corollary 3.10, we have

$$\mathbb{E}_{\mu^{n,N}_{r,s;\mathbf{t}}}(P_N\overline{Q}_N) = \left(e^{\mathcal{D}^{\mathbf{t},N}_{r,s}}(PQ^*)\right)(\mathbf{1}). \tag{3.15}$$

Similarly,

$$\mathbb{E}_{\mu^{n,N}_{r,s;\mathbf{t}}}(P_N) \cdot \mathbb{E}_{\mu^{n,N}_{r,s;\mathbf{t}}}(\overline{Q_N}) = \left(e^{\mathcal{D}^{\mathbf{t},N}_{r,s}} P\right)(\mathbf{1}) \cdot \left(e^{\mathcal{D}^{\mathbf{t},N}_{r,s}} Q^*\right)(\mathbf{1}). \tag{3.16}$$

Now, set

$$\Psi_1^N \equiv \big(e^{-\mathcal{D}^{\mathbf{t},N}_{r,s}} P\big)(\mathbf{1}), \quad \Psi_*^N \equiv \big(e^{-\mathcal{D}^{\mathbf{t},N}_{r,s}} Q^*\big)(\mathbf{1}), \quad \Psi_{1,*}^N \equiv \big(e^{-\mathcal{D}^{\mathbf{t},N}_{r,s}}(PQ^*)\big)(\mathbf{1}), \tag{3.17}$$

$$\Psi_1 \equiv \big(e^{-\mathcal{D}^{\mathbf{t}}_{r,s}} P\big)(\mathbf{1}), \quad \Psi_* \equiv \big(e^{-\mathcal{D}^{\mathbf{t}}_{r,s}} Q^*\big)(\mathbf{1}), \quad \Psi_{1,*} \equiv \big(e^{-\mathcal{D}^{\mathbf{t}}_{r,s}}(PQ^*)\big)(\mathbf{1}). \tag{3.18}$$

Thus, (3.15) and (3.16) show that

$$\operatorname{Cov}_{\mu^{n,N}_{r,s;\mathbf{t}}}(P_N, Q_N) = \Psi_{1,*}^N - \Psi_1^N\Psi_*^N. \tag{3.19}$$

We estimate this as follows. First

$$|\Psi_{1,*}^N - \Psi_1^N\Psi_*^N| \le |\Psi_{1,*}^N - \Psi_{1,*}| + |\Psi_{1,*} - \Psi_1\Psi_*| + |\Psi_1\Psi_* - \Psi_1^N\Psi_*^N|. \tag{3.20}$$

Referring to (3.18), note that $\mathcal{D}^{\mathbf{t}}_{r,s}$ is a first-order differential operator; it follows that $e^{\mathcal{D}^{\mathbf{t}}_{r,s}}$ is an algebra homomorphism, and so the second term in (3.20) is 0. The first term is bounded by $\frac{1}{N^2} \cdot C(r, s, \mathbf{t}, PQ^*)$ by Proposition 3.13. For the third term, we add and subtract $\Psi_1^N\Psi_*$ to make the additional estimate

$$\begin{aligned}
|\Psi_1\Psi_* - \Psi_1^N\Psi_*^N| &\le |\Psi_*||\Psi_1 - \Psi_1^N| + |\Psi_1^N||\Psi_* - \Psi_*^N| \\
&\le |\Psi_*||\Psi_1 - \Psi_1^N| + \big(|\Psi_1| + |\Psi_1^N - \Psi_1|\big)|\Psi_* - \Psi_*^N| \\
&\le \frac{1}{N^2} \cdot |\Psi_*| C(r, s, \mathbf{t}, P) + \left(|\Psi_1| + \frac{1}{N^2} \cdot C(r, s, \mathbf{t}, P)\right) \cdot \frac{1}{N^2} \cdot C(r, s, \mathbf{t}, Q^*) \\
&= \frac{1}{N^2} \cdot (|\Psi_*| C(r, s, \mathbf{t}, P) + |\Psi_1| C(r, s, \mathbf{t}, Q^*)) + \frac{1}{N^4} \cdot C(r, s, \mathbf{t}, P) C(r, s, \mathbf{t}, Q^*).
\end{aligned} \tag{3.21}$$

Combining (3.21) with (3.19) – (3.20) and the following discussion shows that the constant

$$C_2(r, s, \mathbf{t}, P, Q) = C(r, s, \mathbf{t}, PQ^*) + C(r, s, \mathbf{t}, P) C(r, s, \mathbf{t}, Q^*) + |\Psi_*| C(r, s, \mathbf{t}, P) + |\Psi_1| C(r, s, \mathbf{t}, Q^*) \tag{3.22}$$

verifies (3.13), proving the proposition. □

This brings us to the proof of Theorem 1.2. For convenience, we restate that the desired estimate is

$$\mathrm{Cov}\big[\mathrm{tr}(f(B^{1,N}_{r,s}(t_1),\ldots,B^{n,N}_{r,s}(t_n)^*)),\mathrm{tr}(g(B^{1,N}_{r,s}(t_1),\ldots,B^{n,N}_{r,s}(t_n)^*))\big] \le \frac{C_2}{N^2}, \tag{3.23}$$

for any $f,g\in\mathbb{C}\langle X_1,\ldots,X_n,X_1^*,\ldots,X_n^*\rangle$, for some constant $C_2=C_2(r,s,\mathbf{t},f,g)$; here $B^{1,N}_{r,s}(\cdot),\ldots,B^{n,N}_{r,s}(\cdot)$ are independent (r,s)-Brownian motions on $\mathbb{GL}_N$.

Proof of Theorem 1.2. Setting $\mathbf{t}=(t_1,\ldots,t_n)$, the covariance in (3.23) is precisely

$$\mathrm{Cov}_{\mu^{n,N}_{r,s;\mathbf{t}}}\left([\Upsilon(f)]_N,[\Upsilon(g)]_N\right)$$

and so the result follows immediately from Theorem 3.16. □

Theorem 1.2, in the special case $f=g$, implies that the convergence to the joint limit distribution in Theorem 3.14 is, in fact, almost sure.

Corollary 3.18. *Let $(B^{1,N}_{r,s}(t))_{t\ge 0},\ldots,(B^{n,N}_{r,s}(t))_{t\ge 0}$ be independent Brownian motions on $\mathbb{GL}_N$. Then, for any $t_1,\ldots,t_n\ge 0$ and any $f\in\mathbb{C}\langle X_1,\ldots,X_n,X_1^*,\ldots,X_n^*\rangle$, the random variable $\mathrm{tr}(f(B^{1,N}_{r,s}(t_n),\ldots,B^{n,N}_{r,s}(t_n)^*)$ converges to its mean almost surely.*

This follows immediately from the $O(1/N^2)$ covariance estimate of Theorem 1.2, together with Chebyshev's inequality and the Borel-Cantelli lemma.

Finally, we note that we have proven asymptotic freeness of independent (r,s)-Brownian motions.

Corollary 3.19. *Let $t_1,\ldots,t_n>0$ and let $B^{1,N}_{r,s}(t_1),\ldots,B^{n,N}_{r,s}(t_n)$ be independent random matrices sampled from (r,s)-Brownian motion. Then these random matrices are asymptotically free.*

Proof. Summarizing Theorem 2.8: to verify that a collection of random matrix ensembles is asymptotically free, it sufficies to show that the collection possesses a limit distribution (which we verified in this case in Theorem 3.14) whose fluctuations are $O(1/N^2)$ (which we refified in Theorem 1.2), and that the joint distribution of the matrices for each fixed in N is invariant under $\mathbb{U}_N$-conjugation. This last properly holds trivially in our case, as the heat kernel is $\mathbb{U}_N$-invariant (since the inner product is). Hence, independent (r,s)-Brownian motion samples verify all the conditions of Theorem 2.8, concluding the proof. □

4 Invariance Properties and Moments of the (r,s)-Brownian Motions

In this section, we compute the relevant moments of the free multiplicative (r,s)-Brownian motion summarized in Proposition 1.8, and prove the basic invariance properties of both $B^N_{r,s}$ and $b_{r,s}$ needed to extend our main Theorem 1.1 from a single time to multiple times, summarized in Proposition 1.10.

4.1 Moment Calculations

We begin by reiterating the following differential characterization of the constants $\nu_n(t)$ from (1.8).

Lemma 4.1. *Let $\{\nu_n\colon n\ge 0\}$ be the functions in (1.8), and let $\varrho_n(t)=e^{\frac{n}{2}t}\nu_n(t)$. The functions ϱ_n are uniquely determined by the initial conditions $\varrho_n(0)=\nu_n(0)=1$ for all n, $\varrho_1(t)\equiv 1$, and the following system of coupled ODEs for $n\ge 2$:*

$$\varrho_n'(t)=-\sum_{k=1}^{n-1}k\varrho_k(t)\varrho_{n-k}(t).$$

Indeed, in [2], this connection was the key step in identifying the distribution of a free unitary Brownian motion as the limit distribution (at each fixed time t) of a Brownian motion U_t^N on $\mathbb{U}_N$. It is also independently proved in [9, Lemma 5.4, Eq. (5.23)].

Lemma 4.2. *Let $b_{r,s}(t)$ be defined by (1.5); for short, let $b = b_{r,s}(t)$. Set $a = a_{r,s}(t) = e^{\frac{1}{2}(r-s)t}b$. Then*

$$da = a\,dw, \tag{4.1}$$

where $w = w_{r,s}(t)$ of (1.4).

Proof. Since $t \mapsto e^{\frac{1}{2}(r-s)t}$ is a free Itô process with $de^{\frac{1}{2}(r-s)t} = \frac{1}{2}(r-s)e^{\frac{1}{2}(r-s)t}\,dt$, (2.25) shows that

$$da = de^{\frac{1}{2}(r-s)t}\cdot b + e^{\frac{1}{2}(r-s)t}\cdot db + de^{\frac{1}{2}(r-s)t}\cdot db.$$

The last term is 0, while the first two simplify to

$$da = \frac{1}{2}(r-s)e^{\frac{1}{2}(r-s)t}b\,dt + e^{\frac{1}{2}(r-s)t}(b\,dw - \frac{1}{2}(r-s)b\,dt) = a\,dw,$$

by (1.5). □

We also record the following Itô formula for $dw_{r,s}(t)$ products.

Lemma 4.3. *Let $t \geq 0$ and let $\varepsilon, \delta \in \{1, *\}$. For any adapted process $\theta = \theta(t)$,*

$$dw^\varepsilon\,\theta\,dw^\delta = (s \pm r)\tau(\theta)\,dt, \tag{4.2}$$

where the sign is $-$ if $\varepsilon = \delta$ and $+$ if $\varepsilon \neq \delta$.

Lemma 4.3 is an immediate computation from (2.23) – (2.26).

We use (4.1) to give a recursive formula for the powers of $a_{r,s}(t)$.

Proposition 4.4. *For $n \in \mathbb{N}^*$,*

$$d(a^n) = \sum_{k=1}^{n} a^k\,dw\,a^{n-k} + (s-r)\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1} ka^k\tau(a^{n-k})\,dt. \tag{4.3}$$

Proof. When $n = 1$, (4.3) reduces to (4.1). We proceed by induction, supposing that (4.3) has been verified up to level n. Then, using the Itô product rule (2.25), together with (4.1) and (4.3), gives

$$\begin{aligned} d(a^{n+1}) &= d(a\cdot a^n) = da\cdot a^n + a\cdot d(a^n) + da\cdot d(a^n) \\ &= a\,dw\,a^n + \sum_{k=1}^{n} a^{k+1}\,dw\,a^{n-k} + (s-r)\sum_{k=1}^{n-1} ka^{k+1}\tau(a^{n-k})\,dt + \sum_{k=1}^{n} a\,dw\,a^k\,dw\,a^{n-k}. \end{aligned}$$

The first two terms combine, reindexing $\ell = k+1$, to give $\sum_{\ell=1}^{n+1} a^\ell\,dw\,a^{n+1-\ell}$. From (4.2), the last terms are

$$(s-r)\sum_{k=1}^{n}\tau(a^k)a^{n+1-k}\,dt$$

which, when combined with the penultimate terms, yields (4.3) at level $n+1$. This concludes the inductive proof. □

Corollary 4.5. *The moments of $a = a_{r,s}(t)$ are $\tau(a^n) = \varrho_n((r-s)t)$; consequently, the moments of $b = b_{r,s}(t)$ are $\tau(b^n) = \nu_n((r-s)t)$, verifying (1.9).*

Proof. Since $a(0) = b(0) = 1$, $\tau(a(0)^n) = 1 = \varrho_n(0)$. Taking the trace of (4.3) and using (2.22), we have

$$d\tau(a^n) = (s-r)\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1} k\tau(a^k)\tau(a^{n-k})\,dt. \tag{4.4}$$

Thus $\frac{d}{dt}\tau(a) = 0 = \varrho_1'((r-s)t)$. If $s = r$, (4.4) asserts that $\tau(a^n) = \tau(a(0)^n) = 1 = \varrho_n(0\cdot t)$ for all n. On the other hand, if $s \neq r$, let $\tilde{\varrho}_n(t) = \tau(a_{r,s}(t/(r-s))^n)$; then the chain rule applied to (4.4) shows that

$$\tilde{\varrho}_n'(t) = -\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1} k\tilde{\varrho}_k(t)\tilde{\varrho}_{n-k}(t).$$

By Lemma 4.1, it follows that $\tilde{\varrho}_n(t) = \varrho_n(t)$ for all $n, t \geq 0$. Hence, $\tau(a_{r,s}(t)^n) = \varrho_n((r-s)t) = e^{\frac{n}{2}(r-s)t}\nu_n((r-s)t)$, as claimed. As defined in Lemma 4.2, we therefore have

$$\tau(b^n) = \tau[(e^{-\frac{1}{2}(r-s)t}a)^n] = e^{-\frac{n}{2}(r-s)t}\varrho_n((r-s)t) = \nu_n((r-s)t),$$

verifying (1.9), and concluding the proof. □

We now turn to the moments of $b_{r,s}(t)b_{r,s}(t)^*$. A different exponential scaling from Lemma 4.2 is in order here.

Lemma 4.6. *Let $c_{r,s}(t) = e^{-st}b_{r,s}(t)$; for short, let $c = c_{r,s}(t)$. Then*

$$d(cc^*) = 2\sqrt{s}\,c\,dy\,c^*, \tag{4.5}$$

where $y = y(t)$.

Proof. First note that $cc^* = e^{-2st}bb^*$. As in Lemma 4.2, we have

$$d(cc^*) = -2s\,cc^*\,dt + e^{-2st}d(bb^*). \tag{4.6}$$

By the Itô product rule (2.25) and (1.5),

$$\begin{aligned} d(bb^*) &= db\cdot b^* + b\cdot db^* + db\cdot db^* \\ &= b\,dw\,b^* - \frac{1}{2}(r-s)bb^*\,dt + b\,dw^*\,b^* - \frac{1}{2}(r-s)bb^*\,dt + b\,dw\,dw^*\,b^* \\ &= b(dw + dw^*)b^* - (r-s)bb^*\,dt + (r+s)bb^*\,dt \end{aligned}$$

where the last equality follows from Lemma 4.3. Note that $dw + dw^* = 2\sqrt{s}\,dy$, and so this simplifies to $d(bb^*) = 2\sqrt{s}b\,dy\,b^* + 2s\,bb^*\,dt$. Combining this with (4.6) yields the result. □

Proposition 4.7. *For $n \in \mathbb{N}^*$,*

$$d[(cc^*)^n] = 2\sqrt{s}\sum_{k=1}^{n}(cc^*)^{k-1}c\,dy\,c^*(cc^*)^{n-k} + 4s\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1}k(cc^*)^k\tau[(cc^*)^{n-k}]\,dt. \tag{4.7}$$

Proof. When $n = 1$, (4.7) reduces to (4.6), so we proceed by induction: suppose that (4.7) has been verified up to level n. Then we use the Itô product formula (2.25), together with (4.6) and (4.7), to compute

$$\begin{aligned} d[(cc^*)^{n+1}] &= d(cc^*)\cdot(cc^*)^n + cc^*\cdot d[(cc^*)^n] + d(cc^*)\cdot d[(cc^*)^n] \\ &= 2\sqrt{s}\,c\,dy\,c^*(cc^*)^n + 2\sqrt{s}\sum_{k=1}^{n}(cc^*)^k c\,dy\,c^*(cc^*)^{n-k} + 4s\sum_{k=1}^{n-1}k(cc^*)^{k+1}\tau[(cc^*)^{n-k}]\,dt \\ &\qquad + 4s\sum_{k=1}^{n} c\,dy\,c^*(cc^*)^{k-1}c\,dy\,c^*(cc^*)^{n-k}. \end{aligned}$$

Reindexing $\ell = k+1$, the first two terms combine to give $2\sqrt{s}\sum_{\ell=1}^{n+1}(cc^*)^{\ell-1}c\,dy\,c^*(cc^*)^{n+1-\ell}$. In the last term, we use (2.23) to yield

$$dy\,c^*(cc^*)^{k-1}c\,dy = \tau(c^*(cc^*)^{k-1}c)\,dt = \tau[(cc^*)^k]\,dt.$$

Hence, reindexing $j = n+1-k$, the final sum is

$$4s\sum_{k=1}^{n}\tau[(cc^*)^k](cc^*)^{n+1-k}\,dt = 4s\sum_{j=1}^{n}(cc^*)^j\tau[(cc^*)^{n+1-j}].$$

Also reindexing the penultimate sum with $\ell = k+1$, the last two sums combine to give

$$4s\sum_{\ell=2}^{n}(\ell-1)(cc^*)^\ell\tau[(cc^*)^{n+1-\ell}]\,dt + 4s\sum_{j=1}^{n}(cc^*)^j\tau[(cc^*)^{n+1-j}].$$

Note that the first sum could just as well be started at $\ell = 1$ (since that term is 0), and these two combine to give the second term in (4.7), concluding the inductive proof. □

Corollary 4.8. *The moments of cc^* are $\tau[(cc^*)^n] = \varrho_n(-4st)$; consequently, the moments of bb^* are $\tau[(bb^*)^n] = \nu_n(-4st)$, verifying (1.11).*

Proof. Since $b(0) = 1$, $\tau[(cc^*(0))^n] = 1 = \varrho_n(0)$ for all n. Taking the trace of (4.7), we have

$$d\tau[(cc^*)^n] = 4s\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1}k\tau[(cc^*)^k]\tau[(cc^*)^{n-k}]\,dt. \tag{4.8}$$

Thus $\frac{d}{dt}\tau(cc^*) = 0 = \varrho_1'(-4st)$. If $s = 0$, (4.8) asserts that $\tau[(cc^*)^n] = \tau[(cc^*(0))^n] = 1 = \varrho_n(0\cdot t)$ for all n. If $s \neq 0$, let $\hat{\varrho}_n(t) = \tau[((cc^*)(-t/4s))^n]$; then the chain rule applied to (4.8) shows that

$$\hat{\varrho}_n'(t) = -\mathbb{1}_{n\geq 2}\sum_{k=1}^{n-1}k\hat{\varrho}_k(t)\hat{\varrho}_{n-k}(t).$$

By Lemma 4.1, it follows that $\hat{\varrho}_n(t) = \varrho_n(t)$ for all $n, t \geq 0$. Hence,

$$\tau[(cc^*)^n)] = \varrho_n(-4st) = e^{\frac{n}{2}(-4s)t}\nu_n(-4st),$$

as claimed. As defined in Lemma 4.6, we therefore have

$$\tau[(bb^*)^n] = \tau[(e^{2st}cc^*)^n] = e^{-2nst}\varrho_n(-4st) = \nu_n(-4st),$$

verifying (1.10), and concluding the proof. □

Finally, we calculate $\tau(b^2b^{*2})$. We need the following cubic moment as part of the recursive computation.

Lemma 4.9. *Let $a = e^{\frac{1}{2}(r-s)t}b$ as in Lemma 4.2. Then*

$$\tau(a^2a^*) = (1+2st)e^{(s+r)t}. \tag{4.9}$$

Proof. From the Itô product rule (2.25), we have

$$d(a^2a^*) = da\cdot aa^* + a\cdot da\cdot da^* + a^2da^* + (da)^2\cdot a^* + da\cdot a\cdot da^* + a\cdot da\cdot da^*.$$

Lemma 4.2 asserts that $da = a\,dw$. To compute $d\tau(a^2a^*)$, we can ignore the first three terms that have trace 0 by (2.22); the last three terms become

$$a\,dw\,a\,dw\,a^* + a\,dw\,a\,dw^*\,a^* + a^2\,dw\,dw^*\,a^* = (s-r)\tau(a)aa^*\,dt + (s+r)\tau(a)aa^*\,dt + (s+r)a^2a^*\,dt$$

by Lemma 4.3. Taking traces, we therefore have

$$d\tau(a^2a^*) = 2s\tau(a)\tau(aa^*)\,dt + (s+r)\tau(a^2a^*)\,dt. \tag{4.10}$$

In Corollary 4.5, we computed that $\tau(a) = \varrho_1((r-s)t) = e^{\frac{1}{2}(r-s)t}\nu_1((r-s)t)$, which, referring to (1.8), is equal to 1. Similarly, in Corollary 4.8, we calculated that $\tau(bb^*) = \nu_1(-4st) = e^{2st}$, and so $\tau(aa^*) = e^{(r-s)t}\tau(bb^*) = e^{(r+s)t}$. Hence, (4.10) reduces to the ODE

$$\frac{d}{dt}\tau(a^2a^*) = 2se^{(r+s)t} + (s+r)\tau(a^2a^*), \qquad \tau(a^2a^*(0)) = 1.$$

It is simple to verify that (4.9) is the unique solution of this ODE. □

Remark 4.10. As a sanity check, note that in the case $(r,s) = (1,0)$ (4.9) shows that $\tau(b^2b^*) = e^{-\frac{3}{2}t}\tau(a^2a^*) = e^{-t/2}$. As pointed out in (1.7), $b_{1,0}(t) = u(t)$ is a free unitary Brownian motion, and so $\tau(b^2b^*) = \tau(b)$ in this case; thus, we have consistency with (1.8).

Proposition 4.11. *Let $a = e^{\frac{1}{2}(r-s)t}b$ as in Lemma 4.2. Then*

$$\tau(a^2a^{*2}) = 4st(1+st)e^{(s+r)t} + e^{2(s+r)t} \tag{4.11}$$

and thus (1.11) holds true.

Proof. Expanding, once again, using the Itô product rule (2.25), we have

$$\begin{aligned} d(a^2a^{*2}) &= da\cdot aa^{*2} + a\cdot da\cdot a^{*2} + a^2\cdot da^*\cdot a^* + a^2a^*\cdot da^* && (4.12)\\ &\quad + (da)^2\cdot a^{*2} + da\cdot a\cdot da^*\cdot a^* + da\cdot aa^*\cdot da^* && (4.13)\\ &\quad + a\cdot da\cdot da^*\cdot a^* + a\cdot da\cdot a^*\cdot da^* + a^2\cdot(da^*)^2. && (4.14)\end{aligned}$$

The terms in (4.12) all have trace 0. We simplify the terms in (4.13) and (4.14) using $da = a\,dw$ and Lemma 4.3 as follows:

$$\begin{aligned}(4.13) &= a\,dw\,a\,dw\,a^{*2} + a\,dw\,a\,dw^*\,a^{*2} + a\,dw\,aa^*\,dw^*\,a^*\\ &= (s-r)\tau(a)aa^{*2}\,dt + (s+r)\tau(a)aa^{*2}\,dt + (s+r)\tau(aa^*)aa^*\,dt,\end{aligned}$$

and

$$\begin{aligned}(4.14) &= a^2\,dw\,dw^*\,a^{*2} + a^2\,dw\,a^*\,dw^*\,a^* + a^2\,dw^*\,a^*\,dw^*\,a^*\\ &= (s+r)a^2a^{*2}\,dt + (s+r)\tau(a^*)a^2a^*\,dt + (s-r)\tau(a^*)a^2a^*\,dt.\end{aligned}$$

Taking traces, and using the fact (from Lemma 4.9) that $\tau(a^*)\tau(a^2a^*)$ is real, this yields

$$d\tau(a^2a^{*2}) = 2s\tau(a)\tau(aa^{*2})\,dt + (s+r)[\tau(aa^*)]^2\,dt + (s+r)\tau(a^2a^{*2})\,dt + 2s\tau(a^*)\tau(a^2a^*)\,dt.$$

Using (4.9), together with (1.10) and the fact (pointed out in the proof of Lemma 4.9) that $\tau(a) = 1$, gives

$$\frac{d}{dt}\tau(a^2a^{*2}) = 4s(1+2st)e^{(s+r)t} + (s+r)e^{2(s+r)t} + (s+r)\tau(a^2a^{*2}). \tag{4.15}$$

It is easy to verify that (4.11) is the unique solution to this ODE with initial condition 1. Substituting $b = e^{\frac{1}{2}(s-r)t}a$ then yields (1.11). ☐

Remark 4.12. Again, as a sanity check, (1.11) reduces to $\tau(b^2b^{*2}) = 1$ when $s = 0$; this is consistent with the fact that b is unitary in this case.

4.2 Invariance Properties

Proposition 1.10 summarizes the main properties of both the matrix Brownian motions $B^N_{r,s}(t)$ on $\mathbb{GL}_N$ and its limit $(b_{r,s}(t))_{t\geq 0}$. We will prove these properties separately for finite N versus the limit, although in many cases the proofs are extremely similar.

We begin by noting that the invertibility of $B^N_{r,s}(t)$ follows from the SDE (2.11).

Proposition 4.13. *The diffusion $B^N_{r,s}(t)$ is invertible for all $t \geq 0$ (with probability 1); the inverse $B^N_{r,s}(t)^{-1}$ is a* right-*invariant version of an (r,s)-Brownian motion.*

Proof. Fix a Brownian motion $W^N_{r,s}(t) = \sqrt{r}\,iX^N(t) + \sqrt{s}\,Y^N(t)$ on $\mathfrak{gl}_N$, so that $B^N_{r,s}(t)$ is the solution of (2.11) with respect to $W^N_{r,s}(t)$. Then define $A^N_{r,s}(t)$ to be the solution to

$$dA^N_{r,s}(t) = -dW^N_{r,s}(t)\,A^N_{r,s}(t) - \frac{1}{2}(r-s)A^N_{r,s}(t)\,dt. \tag{4.16}$$

Note that $-X^N(t)$ and $-Y^N(t)$ are also independent GUE_N Brownian motions, so $A^N_{r,s}(t)$ is a right-invariant version of $B^N_{s,t}(t)$. (Indeed, the reader can readily check that, if ∂_ξ is replaced with the right-invariant derivative $\frac{d}{dt}f(\exp(-t\xi)g)$, thus defining a right-invariant Laplacian, the associated Brownian motion satisfies (4.16).) To simplify notation, let $W = W^N_{r,s}(t)$, $B = B^N_{r,s}(t)$, and $A = A^N_{r,s}(t)$. Using the Itô product rule (2.17), we have

$$\begin{aligned} d(BA) &= dB\cdot A + B\cdot dA + dB\cdot dA \\ &= B\,dW\,A - \frac{1}{2}(r-s)BA\,dt - B\,dW\,A - \frac{1}{2}(r-s)BA\,dt - B\,(dW)^2\,A. \end{aligned}$$

From (2.14) – (2.18), we compute exactly as in Lemma 4.3 that $(dW)^2 = (s-r)I_N\,dt$. This shows that $d(BA) = 0$. Since $B^N_{r,s}(0) = A^N_{r,s}(0) = I_N$, it follows that $BA = I_N$, so $A^N_{r,s}(t) = B^N_{r,s}(t)^{-1}$, as claimed. ☐

Proposition 4.14. *The multiplicative increments of $(B^N_{r,s}(t))_{t\geq 0}$ are independent and stationary.*

Proof. Let $0 \leq t_1 < t_2 < \infty$, and let $\mathscr{F}_{t_1}$ denote the σ-field generated by $\{X^N(t), Y^N(t)\}_{0\leq t\leq t_1}$. From the defining SDE (2.11), we have

$$B^N_{r,s}(t_2) - B^N_{r,s}(t_1) = \int_{t_1}^{t_2} B^N_{r,s}(t)\,dW^N_{r,s}(t) - \frac{1}{2}(r-s)\int_{t_1}^{t_2} B^N_{r,s}(t)\,dt,$$

or, in other words,

$$B_{r,s}^N(t_1)^{-1}B_{r,s}^N(t_2) = I_N + \int_{t_1}^{t_2} B_{r,s}^N(t_1)^{-1}B_{r,s}^N(t)\,dW_{r,s}^N(t) - \frac{1}{2}(r-s)\int_{t_1}^{t_2} B_{r,s}^N(t_1)^{-1}B_{r,s}^N(t)\,dt. \tag{4.17}$$

This shows that the process $C^N(t) = B_{r,s}^N(t_1)^{-1}B_{r,s}^N(t)$ for $t \ge t_1$ satisfies the SDE

$$dC^N(t) = C^N(t)\,d(W_{r,s}^N(t) - W_{r,s}^N(t_1)) - \frac{1}{2}(r-s)C^N(t)\,dt.$$

Note that $W_{r,s}^N(t) - W_{r,s}^N(t_1) = \sqrt{r}\,i(X^N(t) - X^N(t_1)) + \sqrt{s}\,(Y^N(t) - Y^N(t_1))$. Since $(X^N(t) - X^N(t_1))_{t\ge t_1}$ and $(Y^N(t) - Y^N(t_1))_{t\ge t_1}$ are independent GUE_N Brownian motions, and since $C_{t_1}^N = I_N$, it follows that $(C^N(t))_{t\ge t_1}$ is a version of $(B_{r,s}^N(t))_{t\ge 0}$. This shows, in particular, that the multiplicative increments are stationary. Moreover, (4.17) shows that $B_{r,s}^N(t_1)^{-1}B_{r,s}^N(t_2)$ is measurable with respect to the σ-field generated by the increments $(W_{r,s}^N(t) - W_{r,s}^N(t_1))_{t_1\le t\le t_2}$, which is independent from $\mathscr{F}_{t_1}$ (since the additive increments of $X^N(t)$ and $Y^N(t)$ are independent). Since all the random matrices $B_{r,s}^N(t')$ with $t' \le t_1$ are $\mathscr{F}_{t_1}$-measurable, it follows that $(B_{r,s}^N(t))_{t\ge 0}$ has independent multiplicative increments, as claimed. □

Proposition 4.15. *For $r, s > 0$ and $N \ge 2$, with probability 1, $B_{r,s}^N(t)$ is non-normal for all $t > 0$.*

Proof. Let $\mathbb{M}_N^{\mathrm{nor}}$ denote the set of normal matrices. Let $\mathbb{D}_N$ denote the $2N$ (real) dimensional space of diagonal matrices in $\mathbb{M}_N$, and $\mathbb{T}_N \subset \mathbb{U}_N$ the N (real) dimensional maximal torus of diagonal unitary matrices. The map $\Phi\colon \mathbb{D}_N \times \mathbb{U}_N \to \mathbb{M}_N^{\mathrm{nor}}$ given by $\Phi(D,U) = UDU^*$ is smooth, and (by the spectral theorem) surjective. Since $\Phi(D,U) = \Phi(D,TU)$ for any $T \in \mathbb{T}_N$, the map descends to a smooth surjection $\widetilde{\Phi}\colon \mathbb{D}_N \times \mathbb{U}_N/\mathbb{T}_N \to \mathbb{M}_N^{\mathrm{nor}}$. It follows that

$$\dim_{\mathbb{R}}(\mathbb{M}_N^{\mathrm{nor}}) \le \dim_{\mathbb{R}}(\mathbb{D}_N) + \dim_{\mathbb{R}}(\mathbb{U}_N/\mathbb{T}_N) = 2N + N^2 - N = N^2 + N.$$

Thus, as a submanifold of $\mathbb{M}_N$ (which has real dimension $2N^2$), $\mathrm{codim}_{\mathbb{R}}(\mathbb{M}_N^{\mathrm{nor}}) \ge 2N^2 - (N^2+N) = N^2 - N$. This is ≥ 2 for $N \ge 2$.

The manifold $\mathbb{GL}_N$ is an open dense subset of $\mathbb{M}_N$, and the generator $\Delta_{r,s}^N$ is easily seen to be a non-degenerate elliptic operator on $C^\infty(\mathbb{M}_N)$. Thus, by the main theorem of [23], $\mathbb{M}_N^{\mathrm{nor}}$ is a polar set for the diffusion generated by $\frac{1}{2}\Delta_{r,s}^N$; i.e. the hitting time of $\mathbb{M}_N^{\mathrm{nor}}$ for $(B_{r,s}^N(t))_{t>0}$ is $+\infty$ almost surely. This concludes the proof. □

Remark 4.16. If D is in the open dense subset of $\mathbb{D}_N$ with all eigenvalues distinct, then the stabilizer of D in $\mathbb{U}_N$ is exactly equal to $\mathbb{T}_N$; thus the map $\widetilde{\Phi}$ above is generically a local diffeomorphism. It follows that $\dim_{\mathbb{R}}(\mathbb{M}_N^{\mathrm{nor}}) = N^2 + N$.

Now we turn to the similar properties of the free Itô process $b_{r,s}$. In many cases the proofs are nearly identical to the above ones, in which case we only highlight the necessary differences.

Proposition 4.17. *For all $r, s, t \ge 0$, the free multiplicative (r,s)-Brownian motion $b_{r,s}(t)$ is invertible; the inverse $a_{r,s}(t) = b_{r,s}(t)^{-1}$ satisfies the free SDE*

$$da_{r,s}(t) = -dw_{r,s}(t)\,a_{r,s}(t) - \frac{1}{2}(r-s)\,a_{r,s}(t)\,dt. \tag{4.18}$$

Proof. The proof proceeds very similarly to the proof of Proposition 4.13: using (2.23) – (2.26) instead of (2.14) – (2.18), we compute that $d(b_{r,s}(t)a_{r,s}(t)) = 0$, which shows, since $b_{r,s}(0) = a_{r,s}(0) = 1$, that $b_{r,s}(t)a_{r,s}(t) = 1$.

In this infinite-dimensional setting, we must also verify that $a_{r,s}(t)b_{r,s}(t) = 1$. To that end, to simplify notation, let $a_t = a_{r,s}(t)$, $b_t = b_{r,s}(t)$, and $w_t = w_{r,s}(t)$. Then we have

$$\begin{aligned} d(a_t b_t) &= da_t \cdot b_t + a_t \cdot db_t + da_t \cdot db_t \\ &= -dw_t\, a_t b_t - \frac{1}{2}(r-s)a_t b_t\, dt + a_t b_t\, dw_t - \frac{1}{2}(r-s)a_t b_t\, dt - dw_t\, a_t b_t\, dw_t \\ &= [a_t b_t, dw_t] - (r-s)a_t b_t\, dt - dw_t\, a_t b_t\, dw_t. \end{aligned}$$

From Lemma 4.3,

$$dw_t\, a_t b_t\, dw_t = (s-r)\tau(a_t b_t).$$

Thus, $a_t b_t$ satisfies the free SDE

$$d(a_t b_t) = [a_t b_t, dw_t] + (r-s)[a_t b_t - \tau(a_t b_t)],$$

with initial condition $a_0 b_0 = 1$. Notice that the free SDE $d\theta_t = [\theta_t, dw_t] + (r-s)[\theta_t - \tau(\theta_t)]$ holds true for any constant process θ_t; thus, with initial condition $\theta_0 = 1$ uniquely determining the solution, we see that $a_t b_t = 1$ as well. □

Proposition 4.18. *The multiplicative increments of $(b_{r,s}(t))_{t\geq 0}$ are freely independent and stationary.*

The proof of Proposition 4.18 is virtually identical to the proof of Proposition 4.14; one need only replace the σ-fields $\mathscr{F}_t$ with the von Neumann algebras $\mathscr{A}_t = W^*\{x(t'), y(t') \colon 0 \leq t' \leq t\}$.

Proposition 4.19. *For $r \geq 0$ and $s > 0$, $b_{r,s}(t)$ is non-normal for all $t > 0$.*

Proof. Let $b_t = b_{r,s}(t)$; we compute that

$$[b_t, b_t^*]^2 = (b_t b_t^*)^2 - b_t (b_t^*)^2 b_t - b_t^* b_t^2 b_t^* + (b_t^* b_t)^2,$$

and so

$$\tau\left([b_t, b_t^*]^2\right) = 2\tau[(b_t b_t^*)^2] - 2\tau[b_t^2 (b_t^*)^2].$$

We now use (1.8), (1.10), and (1.11) to expand this:

$$\begin{aligned} \tau[(b_t b_t^*)^2] - \tau[b_t^2 (b_t^*)^2] &= \nu_2(-4st) - (e^{4st} + 4st(1+st)e^{(3s-r)t}) \\ &= e^{4st}(1+4st) - (e^{4st} + 4st(1+st)e^{(3s-r)t}) \\ &= 4ste^{3st}[e^{st} - (1+st)e^{-rt}]. \end{aligned}$$

Since $r \geq 0$, $e^{-rt} \leq 1$, and since $st > 0$, $e^{st} > 1 + st$. It follows that $\tau([b_t, b_t^*]^2) > 0$ for $t > 0$, proving that b_t is not normal. □

5 Convergence of the Brownian Motions

This final section is devoted to the proof of Theorem 1.1: that the process $(B_{r,s}^N(t))_{t\geq 0}$ converges in noncommutative distribution to the process $(b_{r,s}(t))_{t\geq 0}$. We first show the convergence of the random matrices $B_{r,s}^N(t)$ for each fixed $t \geq 0$; the multi-time statement then follows from asymptotic freeness considerations.

5.1 Convergence for a Fixed t

We begin by noting the single-t version of Theorem 1.2, which was proved in [16, Proposition 4.13]. For any $r,s>0$ and $t\geq 0$, and any noncommutative polynomials $f,g\in\mathbb{C}\langle X,X^*\rangle$, there is a constant $C_{r,s}(t,f,g)$ such that

$$\mathrm{Cov}\left[\mathrm{tr}\big(f(B_{r,s}^N(t),B_{r,s}^N(t)^*)\big),\mathrm{tr}\big(g(B_{r,s}^N(t),B_{r,s}^N(t)^*)\big)\right]\leq\frac{C_{r,s}(t,f,g)}{N^2},\tag{5.1}$$

where $C_{r,s}(t,f,g)$ depends continuously on t.

We now proceed to prove the fixed-t case of Theorem 1.1. The idea is to compare the SDE for $B_{r,s}^N(t)$ to the free SDE for $b_{r,s}(t)$, and inductively show that traces of $*$-moments differ by $O(1/N^2)$, using (5.1).

Theorem 5.1. *Let $r,s,t\geq 0$. Let $n\in\mathbb{N}$ and let $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)\in\{1,*\}^n$. Then there is a constant $C'_{r,s}(t,\varepsilon)$ that depends continuously on r,s,t so that*

$$\left|\mathbb{E}\mathrm{tr}\big(B_{r,s}^N(t)^{\varepsilon_1}\cdots B_{r,s}^N(t)^{\varepsilon_n}\big)-\tau(b_{r,s}(t)^{\varepsilon_1}\cdots b_{r,s}(t)^{\varepsilon_n})\right|\leq\frac{C'_{r,s}(t,\varepsilon)}{N^2}.\tag{5.2}$$

Remark 5.2. We remark again that this result was proved, in the special case $r=s$, in [5] (using different techniques). In fact, Cébron's method could well be adapted to give an alternate proof of this result that does not rely explicitly on an inductive analysis of stochastic differential equations, although in some sense the central idea is the same.

Proof. In the case $n=0$, (5.2) holds true vacuously with $C'_{r,s}(t,\varnothing)=0$. When $n=1$, as computed in (1.9) we have $\tau(b_{r,s}(t)^{\varepsilon_1})=\nu_1((r-s)t)$, and so (5.2) follows immediately from [16, Theorem 1.3]. From here, we proceed by induction: assume that (5.2) has been verified up to, but not including, level n.

Fix $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)\in\{1,*\}^n$. Let $A_{r,s}^N(t)=e^{\frac{1}{2}(r-s)t}B_{r,s}^N(t)$, so that, following precisely the proof of Lemma 4.2 but using (2.17) instead of (2.25), we have

$$dA_{r,s}^N(t)=A_{r,s}^N(t)\,dW_{r,s}^N(t).\tag{5.3}$$

For convenience, denote $A=A_{r,s}^N(t)$, and denote $A^\varepsilon=A^{\varepsilon_1}\cdots A^{\varepsilon_n}$. Then, using the Itô product rule (2.17), we have

$$d(A^\varepsilon)=\sum_{j=1}^n A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}\cdot dA^{\varepsilon_j}\cdot A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_n}\tag{5.4}$$

$$+\sum_{1\leq j<k\leq n}A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}\cdot dA^{\varepsilon_j}\cdot A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}\cdot dA^{\varepsilon_k}\cdot A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n}.\tag{5.5}$$

From (2.15) and (5.3), the terms in (5.5) become

$$\begin{aligned}&A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}\cdot dA^{\varepsilon_j}\cdot A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}\cdot dA^{\varepsilon_k}\cdot A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n}\\ =\;&A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}A^{\varepsilon'_j}\,dW^{\varepsilon_j}\,A^{\varepsilon''_j}A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}A^{\varepsilon'_k}\,dW^{\varepsilon_k}\,A^{\varepsilon''_k}A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n}\end{aligned}$$

where $W=W_{r,s}^N(t)$, and $1'=1$, $1''=*'=0$, and $*''=*$. As in Lemma 4.3, (2.14) – (2.18) show that, for any adapted process Θ, and any $\varepsilon,\delta\in\{1,*\}$,

$$dW^\varepsilon\,\Theta\,dW^\delta=(s\pm r)\,\mathrm{tr}(\Theta)\,dt\tag{5.6}$$

where the sign is $-$ if $\varepsilon=\delta$ and $+$ if $\varepsilon\neq\delta$. Hence, the terms in (5.5) become

$$(s\pm r)\mathrm{tr}(A^{\varepsilon''_j}A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}A^{\varepsilon'_k})A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}A^{\varepsilon'_j}A^{\varepsilon''_k}A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n}.$$

Now, note that the expected value of all the terms in (5.4) is 0 by (2.13) and (5.3). Therefore, taking $\mathbb{E}\mathrm{tr}$ in (5.4) and (5.5), we have

$$\frac{d}{dt}\mathbb{E}\mathrm{tr}(A^\varepsilon) = \sum_{1\le j<k\le n} (s\pm r)\mathbb{E}\left[\mathrm{tr}(A^{\varepsilon_j''}A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}A^{\varepsilon_k'})\mathrm{tr}(A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}A^{\varepsilon_j'}A^{\varepsilon_k''}A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n})\right].$$

It is possible for one of the two trace terms to be trivial, in two special cases.

- If $j=1$ and $k=n$, and if $\varepsilon_1 = *$ and $\varepsilon_n = 1$, then the first trace term is equal to $\mathrm{tr}(A^\varepsilon)$, while the second one is just $\mathrm{tr}(I_N)=1$.
- For $1\le j<n$, if $k=j+1$, and $\varepsilon_j = 1$ while $\varepsilon_k = *$, then the second trace term is equal to $\mathrm{tr}(A^\varepsilon)$, while the first one is just $\mathrm{tr}(I_N)=1$.

In all other (ε, j, k) configurations, each trace term involves a non-trivial string of length $< n$. Note that, in both these exceptional cases, the two exponents must be different, and so the factor in front is $s+r$. We separate out these cases as follows:

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\mathrm{tr}(A^\varepsilon) &= (s+r)\mathbb{1}_{(\varepsilon_1,\varepsilon_n)=(*,1)}\mathbb{E}\mathrm{tr}(A^\varepsilon) + (s+r)\sum_{j=1}^{n-1}\mathbb{1}_{(\varepsilon_j,\varepsilon_{j+1})=(1,*)}\mathbb{E}\mathrm{tr}(A^\varepsilon) \\ &\quad + \widetilde{\sum_{1\le j<k\le n}} (s\pm r)\mathbb{E}\left[\mathrm{tr}(A^{\varepsilon_j''}A^{\varepsilon_{j+1}}\cdots A^{\varepsilon_{k-1}}A^{\varepsilon_k'})\mathrm{tr}(A^{\varepsilon_1}\cdots A^{\varepsilon_{j-1}}A^{\varepsilon_j'}A^{\varepsilon_k''}A^{\varepsilon_{k+1}}\cdots A^{\varepsilon_n})\right],\end{aligned}$$

where $\widetilde{\sum}$ indicates that the sum excludes the at-most-n terms accounted for in the special cases. Define

$$\kappa(\varepsilon) = \mathbb{1}_{(\varepsilon_1,\varepsilon_n)=(*,1)} + \sum_{j=1}^{n-1}\mathbb{1}_{(\varepsilon_j,\varepsilon_{j+1})=(1,*)},$$

and let

$$\varepsilon^1_{j,k} = (\varepsilon_j'', \ldots, \varepsilon_k'), \qquad \varepsilon^2_{j,k} = (\varepsilon_1, \ldots, \varepsilon_j', \varepsilon_k'', \ldots, \varepsilon_n).$$

Thus we have shown that $\mathbb{E}\mathrm{tr}(A^\varepsilon)$ satisfies the ODE

$$\frac{d}{dt}\mathbb{E}\mathrm{tr}(A^\varepsilon) = \kappa(\varepsilon)(s+r)\mathbb{E}\mathrm{tr}(A^\varepsilon) + \widetilde{\sum_{1\le j<k\le n}} (s\pm r)\mathbb{E}\left[\mathrm{tr}(A^{\varepsilon^1_{j,k}})\mathrm{tr}(A^{\varepsilon^2_{j,k}})\right], \tag{5.7}$$

where all the terms in the sum are expectations of products of traces of words in A and A^* of length *strictly less* than n. Since $A(0)=I_N$, the unique solution of this ODE (in terms of these functions in the sum) is

$$\mathbb{E}\mathrm{tr}(A^\varepsilon_T) = e^{\kappa(\varepsilon)(s+r)T} + \widetilde{\sum_{1\le j<k\le n}} (s\pm r)\int_0^T e^{\kappa(\varepsilon)(s+r)(T-t)}\,\mathbb{E}\left[\mathrm{tr}(A_t^{\varepsilon^1_{j,k}})\mathrm{tr}(A_t^{\varepsilon^2_{j,k}})\right]\,dt \tag{5.8}$$

where we have written $A_t = A^N_{r,s}(t)$ to emphasize the different times of evaluation. Now returning to $B_t = B^N_{r,s}(t) = e^{-\frac{1}{2}(r-s)t}A_t$, and noting that the total length of the two strings $\varepsilon^1_{j,k}$ and $\varepsilon^2_{j,k}$ is n, the same as the length of ε, this gives

$$\begin{aligned}\mathbb{E}\mathrm{tr}(B^\varepsilon_T) &= e^{[\kappa(\varepsilon)(s+r)-\frac{n}{2}(r-s)]T} \\ &\quad + \widetilde{\sum_{1\le j<k\le n}} (s\pm r)\int_0^T e^{[\kappa(\varepsilon)(s+r)-\frac{n}{2}(r-s)](T-t)}e^{\frac{n}{2}(r-s)t}\,\mathbb{E}\left[\mathrm{tr}(B_t^{\varepsilon^1_{j,k}})\mathrm{tr}(B_t^{\varepsilon^2_{j,k}})\right]\,dt.\end{aligned} \tag{5.9}$$

Now, repeating this derivation line-by-line, we find that, setting $b_t = b_{r,s}(t)$,

$$\tau(b_T^\varepsilon) = e^{[\kappa(\varepsilon)(s+r)-\frac{n}{2}(r-s)]T} + \widetilde{\sum_{1\le j<k\le n}} (s\pm r)\int_0^T e^{[\kappa(\varepsilon)(s+r)-\frac{n}{2}(r-s)](T-t)} e^{\frac{n}{2}(r-s)t}\tau(b_t^{\varepsilon_{j,k}^1})\tau(b_t^{\varepsilon_{j,k}^2})\,dt. \tag{5.10}$$

The principal difference is that, when applying the free Itô product rule (2.25), the trace τ factors through completely, while in the matrix Itô product rule (2.17), only the trace tr factors through, while the expectation $\mathbb{E}$ does not. Thus, the desired quantity (on the left-hand-side of (5.2)) at time T is equal to

$$\widetilde{\sum_{1\le j<k\le n}} (s\pm r)\int_0^T e^{[\kappa(\varepsilon)(s+r)-\frac{n}{2}(r-s)](T-t)} e^{\frac{n}{2}(r-s)t}\left(\mathbb{E}\left[\mathrm{tr}(B_t^{\varepsilon_{j,k}^1})\mathrm{tr}(B_t^{\varepsilon_{j,k}^2})\right] - \tau(b_t^{\varepsilon_{j,k}^1})\tau(b_t^{\varepsilon_{j,k}^2})\right)dt. \tag{5.11}$$

Again to simplify notation, fix j,k in the sum and let $B_\ell = B_t^{\varepsilon_{j,k}^\ell}$ and $b_\ell = b_t^{\varepsilon_{j,k}^\ell}$ for $\ell = 1,2$. Then we expand the difference as

$$\mathbb{E}[\mathrm{tr}(B_1)\mathrm{tr}(B_2)] - \tau(b_1)\tau(b_2) = \mathrm{Cov}[\mathrm{tr}(B_1),\mathrm{tr}(B_2)] + \mathbb{E}\mathrm{tr}(B_1)\mathbb{E}\mathrm{tr}(B_2) - \tau(b_1)\tau(b_2), \tag{5.12}$$

and the last two terms may be written (by adding and subtracting $\tau(b_1)\mathbb{E}\mathrm{tr}(B_2)$) as

$$\mathbb{E}\mathrm{tr}(B_1)\mathbb{E}\mathrm{tr}(B_2) - \tau(b_1)\tau(b_2) = \mathbb{E}\mathrm{tr}(B_2)\cdot[\mathbb{E}\mathrm{tr}(B_1)-\tau(b_1)] + \tau(b_1)\cdot[\mathbb{E}\mathrm{tr}(B_2)-\tau(b_2)]. \tag{5.13}$$

We now appeal to the inductive hypothesis. By construction, all the terms in the sum $\widetilde{\sum}$ have both strings $\varepsilon_{j,k}^1$ and $\varepsilon_{j,k}^2$ of length *strictly* $< n$. As such, the inductive hypothesis yields that $|\mathbb{E}\mathrm{tr}(B_\ell) - \tau(b_\ell)| \le C_\ell(t)/N^2$ for constants $C_\ell(t)$ that depend continuously on t (and all of the hidden parameters r,s,ε). It follows, in particular, that the constants $\mathbb{E}\mathrm{tr}(B_2)$ are uniformly bounded in N and $t\in[0,T]$. Thus, the terms in (5.13) are bounded by $C(t)/N^2$ for some constant $C(t)$ that is uniformly bounded in $t\in[0,T]$. By (5.1), the covariance term in (5.12) is also bounded by $C'(t)/N^2$ for such a constant $C'(t)$. Integrating $C(t)+C'(t)$ times the relevant exponentials, summed over j,k, in (5.11) now shows that the whole expression is $\le C''(T)/N^2$ for some constant $C''(T)$ that depends continuously on T. This concludes the proof. □

Remark 5.3. In [16, Theorem 1.6], the author showed that there exists a linear functional $\varphi_{r,s}^t\colon \mathbb{C}\langle X,X^*\rangle\to\mathbb{C}$ so that (5.2) holds with $\varphi_{r,s}^t(X^{\varepsilon_1}\cdots X^{\varepsilon_n})$ in place of $\tau(b_{r,s}(t)^{\varepsilon_1}\cdots b_{r,s}(t)^{\varepsilon_n})$; the upshot of the present theorem is to identify this linear functional as the noncommutative distribution of $b_{r,s}(t)$. In particular, it lives in a faithful, normal, tracial W^*-probability space, which could not be easily proved using the techniques in [16].

5.2 Asymptotic Freeness and Convergence of the Process

In this final section, we use the freeness of the increments of $b_{r,s}(t)$ and the asymptotic freeness of the increments of $B_{r,s}^N(t)$, together with Theorem 5.1, to prove Theorem 1.1. We begin with some preliminary lemmas.

Lemma 5.4. *Let $\varepsilon_1,\ldots,\varepsilon_n\in\{1,*\}$, and let $f\in\mathbb{C}\langle X_1,\ldots,X_n\rangle$ be a noncommutative polynomial. Given any permutation $\sigma\in\Sigma_n$, there is a noncommutative polynomial $g\in\mathbb{C}\langle X_1,\ldots,X_n,X_1^*,\ldots,X_n^*\rangle$ with the following property. If $b_1,\ldots,b_n$ are invertible random variables in a noncommutative probability space, and $a_1=b_1, a_2=b_1^{-1}b_2,\ldots,a_n=b_{n-1}^{-1}b_n$ are the corresponding multiplicative increments, then*

$$f(b_{\sigma(1)}^{\varepsilon_1},\ldots,b_{\sigma(n)}^{\varepsilon_n}) = g(a_1,\ldots,a_n,a_1^*,\ldots,a_n^*).$$

Proof. For $1 \leq j \leq n$, write

$$b_j = b_1(b_1^{-1}b_2)\cdots(b_{j-1}^{-1}b_j) = a_1a_2\cdots a_j. \tag{5.14}$$

Let $f_\sigma(X_1,\ldots,X_n) = f(X_{\sigma(1)},\ldots,X_{\sigma(n)})$; then

$$f(b_{\sigma(1)}^{\varepsilon_1},\cdots,b_{\sigma(n)}^{\varepsilon_n}) = f_\sigma(b_1^{\varepsilon_{\sigma^{-1}(1)}},\ldots,b_n^{\varepsilon_{\sigma^{-1}(n)}}).$$

In each variable, expand the term $b_j^{\varepsilon_{\sigma^{-1}(j)}}$ using (5.14) (to the $\varepsilon_{\sigma^{-1}(j)}$ power); this yields the polynomial g. □

The next lemma uses the language of Section 3.2 to give a more precise formulation of how free independence reduces the calculation of joint moments to separate moments.

Lemma 5.5. *Given any $n \in \mathbb{N}$ and any noncommutative polynomial $g \in \mathbb{C}\langle X_1,\ldots,X_n,X_1^*,\ldots,X_n^*\rangle$, there is an integer $m \in \mathbb{N}$ and a collection $\{P^{j,k}\colon 1 \leq j \leq n, 1 \leq k \leq m\}$ of elements of $\mathscr{P}$ with the property that, if $(\mathscr{A},\tau)$ is a noncommutative probability space, and $a_1,\ldots,a_n \in \mathscr{A}$ are freely independent, then*

$$\tau(g(a_1,\ldots,a_n,a_1^*,\ldots,a_n^*)) = \sum_{k=1}^m P_\tau^{1,k}(a_1)\cdots P_\tau^{n,k}(a_n). \tag{5.15}$$

Here $\mathscr{P}$ denotes the polynomial space $\mathscr{P}(J)$ with the index set J a singleton. The proof of Lemma 5.5 is contained in the proof of [21, Lemma 5.13]. The idea is to center the variables and proceed inductively. The exact machinery of how $P^{j,k}$ are computed from g is the business of the rich theory of free cumulants, which is the primary topic of [21].

Now, suppose $A_1^N,\ldots,A_n^N$ are $N\times N$ random matrices that are asymptotically free; cf. Definition 2.7. This means precisely that $(A_1^N,\ldots,A_n^N) \to (a_1,\ldots,a_n)$ in noncommutative distribution, for some freely independent collection $a_1,\ldots,a_n$ in a noncommutative probability space $(\mathscr{A},\tau)$. In other words, for any noncommutative polynomial $g \in \mathbb{C}\langle X_1,\ldots,X_n,X_1^*,\ldots,X_n^*\rangle$,

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\mathrm{tr}\big(g(A_1^N,\ldots,A_n^N,(A_1^N)^*,\ldots,(A_n^N)^*)\big) &= \tau(g(a_1,\ldots,a_n,a_1^*,\ldots,a_n^*))\\ &= \sum_{k=1}^m P_\tau^{1,k}(a_1)\cdots P_\tau^{n,k}(a_n)\end{aligned}$$

where the second equality is Lemma 5.5. Note that $P_\tau^{j,k}(a)$ is a polynomial in the trace moments of a,a^*, and by assumption of convergence of the joint distribution, we also therefore have $(P_{\mathbb{E}\mathrm{tr}}^{j,k}(A_j^N)) \to P_\tau^{j,k}(a_j)$ as $N\to\infty$. Hence, we can alternatively state asymptotic freeness as

$$\lim_{N\to\infty}\mathbb{E}\mathrm{tr}\big(g(A_1^N,\ldots,A_n^N,(A_1^N)^*,\ldots,(A_n^N)^*)\big) = \lim_{N\to\infty}\sum_{k=1}^m P_{\mathbb{E}\mathrm{tr}}^{1,k}(A_1^N)\cdots P_{\mathbb{E}\mathrm{tr}}^{n,k}(A_n^N). \tag{5.16}$$

We now stand ready to prove Theorem 1.1.

Proof of Theorem 1.1. For convenience, denote $B_{r,s}^N(t) = B_t$ and $b_{r,s}(t) = b_t$. Fix $t_1,\ldots,t_n \geq 0$ and $\varepsilon_1,\ldots,\varepsilon_n \in \{1,*\}$. Fix a permutation $\sigma\in\Sigma_n$ such that $t_{\sigma(1)} \leq \cdots \leq t_{\sigma(n)}$ and let $t_j' = t_{\sigma(j)}$. Let

$$A_1 = B_{t_1'},\ A_2 = B_{t_1'}^{-1}B_{t_2},\ \ldots,\ A_n = B_{t_{n-1}'}^{-1}B_{t_n}$$

be the increments for the partition $t_1' \leq \cdots \leq t_n'$. Using Lemma 5.4, we can write

$$\mathbb{E}\mathrm{tr}\big(B_{t_1}^{\varepsilon_1}\cdots B_{t_n}^{\varepsilon_n}\big) = \mathbb{E}\mathrm{tr}(g(A_1,\ldots,A_n,A_1^*,\ldots,A_n^*)) \tag{5.17}$$

where $g \in \mathbb{C}\langle X_1, \ldots, X_n, X_1^*, \ldots, X_n^*\rangle$ is determined by σ and $\varepsilon_1, \ldots, \varepsilon_n$.

By Proposition 4.14, the increments A_j are independent; moreover, their stationarity means that A_j has the same distribution as $B_{\Delta t'_j}$ where $\Delta t'_1 = t'_1$ and $\Delta t'_j = t'_j - t'_{j-1}$ for $1 < j \le n$. Thus, by Corollary 3.19, $A_1, \ldots, A_n$ are asymptotically free. In addition, the equality of distributions means that all $*$-moments of A_j are equal to the same $*$-moments of $B_{\Delta t'_j}$. Thus, combining (5.16) and (5.17), we have

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{t_1}^{\varepsilon_1} \cdots B_{t_n}^{\varepsilon_n}) = \lim_{N\to\infty} \sum_{k=1}^{m} P^{1,k}_{\mathbb{E}\mathrm{tr}}(B_{\Delta t'_1}) \cdots P^{n,k}_{\mathbb{E}\mathrm{tr}}(B_{\Delta t'_n}).$$

From Theorem 5.1, we therefore have

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{t_1}^{\varepsilon_1} \cdots B_{t_n}^{\varepsilon_n}) = \sum_{k=1}^{m} P^{1,k}_{\tau}(b_{\Delta t'_1}) \cdots P^{n,k}_{\tau}(b_{\Delta t'_n}).$$

Now, by Proposition 4.18, the increments $b_{\Delta t'_j}$ are freely independent and stationary; so letting

$$a_1 = b_{t'_1},\ a_2 = b_{t'_1}^{-1} b_{t'_2},\ \ldots,\ a_n = b_{t'_{n-1}}^{-1} b_{t'_n}$$

we see that $\{b_{\Delta t'_1}, \ldots, b_{\Delta t'_n}\}$ have the same joint distribution as $\{a_1, \ldots, a_n\}$. Thus

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{t_1}^{\varepsilon_1} \cdots B_{t_n}^{\varepsilon_n}) = \sum_{k=1}^{m} P^{1,k}_{\tau}(b_{\Delta t'_1}) \cdots P^{n,k}_{\tau}(b_{\Delta t'_n}) = \sum_{k=1}^{m} P^{1,k}_{\tau}(a_1) \cdots P^{n,k}_{\tau}(a_n),$$

and by the definition (5.15) of $P^{j,k}$, this yields

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{t_1}^{\varepsilon_1} \cdots B_{t_n}^{\varepsilon_n}) = \tau(g(a_1, \ldots, a_n, a_1^*, \ldots, a_n^*)).$$

Finally, by the definition (5.17) of g, we conclude that

$$\lim_{N\to\infty} \mathbb{E}\mathrm{tr}(B_{t_1}^{\varepsilon_1} \cdots B_{t_n}^{\varepsilon_n}) = \tau(b_{t_1}^{\varepsilon_1} \cdots b_{t_n}^{\varepsilon_n}),$$

concluding the proof. □

We conclude by giving an extension of Theorem 1.1: it follows essentially immediately that any collection of independent (r,s)-Brownian motions converges in finite-dimensional distributions to a collection of freely independent free multiplicative (r,s)-Brownian motions. Moreover, not only do moments converge, but all trace polynomials converge (at rate $\frac{1}{N^2}$).

Corollary 5.6. *Fix a time index set J and an integer m. Let $\{(B^{N,k}_{r,s}(t))_{t\ge 0}\}_{1\le k\le m}$ be a finite family of independent (r,s)-Brownian motions on $\mathbb{GL}_N$, and let $\{(b^k_{r,s}(t))_{t\ge 0}\}_{1\le k\le m}$ be a finite family of freely independent free multiplicative (r,s)-Brownian motions in noncommutative probability space $(\mathscr{A}, \tau)$. Fix a collection of times $\mathbf{T} = (t_j)_{j\in K}$, and set $B^{N,k}_{r,s}(\mathbf{T}) = \{B^{N,k}_{s,t}(t_j)\}_{j\in J}$ and $b^k_{r,s}(\mathbf{T}) = \{b^k_{r,s}(t_j)\}_{j\in J}$ Then for any trace polynomial $P \in \mathscr{P}(J^m)$,*

$$P_N(B^{N,1}_{r,s}(\mathbf{T}), \ldots, B^{N,m}_{r,s}(\mathbf{T})) = P_\tau(b^1_{r,s}(\mathbf{T}), \ldots, b^m_{r,s}(\mathbf{T})) + O\left(\frac{1}{N^2}\right).$$

The proof is a straightforward extension of the above techniques. A remark about the precise $O(1/N^2)$ statement: in Theorem 5.1, we prove the special case of Theorem 1.1 (and Corollary 5.6) when $t_1 = \cdots = t_n = t$. In this case, we have the quantitative bound that the difference between the moments in (1.1) is $O(\frac{1}{N^2})$. That this extends to the general case can be seen easily by tracking through the proof of Theorem 1.1 (beginning on page 29) and using the fact that all moments of the Brownian motion are bounded uniformly in N.

Acknowledgments

The author wishes to thank Philippe Biane, Bruce Driver, Brian Hall, Pat Fitzsimmons, and Jamie Mingo for useful conversations. He further wishes to express significant thanks to an anonymous referee whose extremely careful and constructively critical reading of an earlier version led to a much improved manuscript.

References

[1] BERCOVICI, H., AND VOICULESCU, D. Lévy-Hinčin type theorems for multiplicative and additive free convolution. *Pacific J. Math. 153*, 2 (1992), 217–248.

[2] BIANE, P. Free Brownian motion, free stochastic calculus and random matrices. In *Free probability theory (Waterloo, ON, 1995)*, vol. 12 of *Fields Inst. Commun.* Amer. Math. Soc., Providence, RI, 1997, pp. 1–19.

[3] BIANE, P. Segal-Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems. *J. Funct. Anal. 144*, 1 (1997), 232–286.

[4] BIANE, P., AND SPEICHER, R. Stochastic calculus with respect to free Brownian motion and analysis on Wigner space. *Probab. Theory Related Fields 112*, 3 (1998), 373–409.

[5] CÉBRON, G. Free convolution operators and free Hall transform. *J. Funct. Anal. 265*, 11 (2013), 2645–2708.

[6] COLLINS, B. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. *Int. Math. Res. Not.*, 17 (2003), 953–982.

[7] COLLINS, B., KEMP, T. Liberation of Projections. *J. Funct. anal. 266*, 4 (2014), 1988–2052.

[8] DRIVER, B. K., HALL, B. C. Yang-Mills theory and the Segal–Bargmann transform. *Comm. Math. Phys. 201*, 2 (1999), 249–290.

[9] DRIVER, B. K., HALL, B. C., AND KEMP, T. The large-N limit of the Segal-Bargmann transform on $\mathbb{U}_N$. *J. Funct. Anal. 265*, 11 (2013), 2585–2644.

[10] DURRETT, RICHARD Stochastic calculus. A practical introduction. *Probability and Stochastics Series.* CRC Press, Boca Raton, FL, 1996.

[11] GORDINA, M. Heat kernel analysis and Cameron-Mactin subgroup for infinite dimensional groups. *J. Funct. Anal. 171*, 1 (2000), 192–232.

[12] GORDINA, M. Holomorphic functions and the heat kernel measure on an infinite-dimensional complex orthogonal group. *Potential Anal. 12*, 4 (2000), 325–357.

[13] GORDINA, M. Quasi-invariance for the pinned Brownian motion on a Lie group. *Stochastic Process. Appl. 104*, 2 (2003), 243–257.

[14] GORDINA, M. Stochastic differential equations on noncommutative L^2. *Finite and infinite dimensional analysis in honor of Leonard Gross (New Orleans, LA, 2001)*, *Contemp. Math. 317*, (2003) 87–98.

[15] HALL, B. Harmonic analysis with respect to heat kernel measure. *Bull. Amer. Math. Soc. (N.S.) 38*, 1 (2001), 43–78.

[16] KEMP, T. Heat kernel empirical measures on $\mathbb{U}_N$ and $\mathbb{GL}_N$. *J. Theor. Probab.* (2014). To appear.

[17] Kemp, T., Nourdin, I., Peccati, G., and Speicher, R. Wigner chaos and the fourth moment. *Ann. Probab. 40*, 4 (2012), 1577–1635.

[18] McKean, Jr., H. P. *Stochastic integrals*. Probability and Mathematical Statistics, No. 5. Academic Press, New York, 1969.

[19] Mingo, J., and Popa, M. Real second order freeness and Haar orthogonal matrices. *J. Math. Phys. 54*, 051701 (2013).

[20] Mingo, J. A., Śniady, P., and Speicher, R. Second order freeness and fluctuations of random matrices. II. Unitary random matrices. *Adv. Math. 209*, 1 (2007), 212–240.

[21] Nica, A., and Speicher, R. *Lectures on the combinatorics of free probability*, vol. 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[22] Rains, E. M. Combinatorial properties of Brownian motion on the compact classical groups. *J. Theoret. Probab. 10*, 3 (1997), 659–679.

[23] Ramasubramanian, S. Hitting of submanifolds by diffusions. *Probab. Theory Related Fields 78*, 1 (1988), 149–163.

[24] Robinson, D. W. *Elliptic operators and Lie groups*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1991. Oxford Science Publications.

[25] Sengupta, A. N. Traces in two-dimensional QCD: the large-N limit. *Traces in number theory, geometry and quantum fields*, 193–212, *Aspects Math., E38, Friedr. Vieweg, Wiesbaden*, (2008)

[26] Voiculescu, D. Limit laws for random matrices and free products. *Invent. Math. 104*, 1 (1991), 201–220.

[27] Voiculescu, D. A strengthened asymptotic freeness result for random matrices with applications to free entropy. *Internat. Math. Res. Notices*, 1 (1998), 41–63.

[28] Voiculescu, D. The analogues of entropy and of Fisher's information measure in free probability theory. VI. Liberation and mutual free information. Adv. Math. 146, 2 (1999), 101–166.

[29] Voiculescu, D. V., Dykema, K. J., and Nica, A. *Free random variables*, vol. 1 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1992. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups.

[30] Xu, F. A random matrix model from two-dimensional Yang-Mills theory. *Comm. Math. Phys. 190*, 2 (1997), 287–307.