

Matrix Dufresne Identities

B. Rider¹ and B. Valkó²

¹Department of Mathematics, Temple University, Philadelphia, PA 19122-6094, USA and ²Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706-1325, USA

Correspondence to be sent to: brian.rider@temple.edu

We prove a version of the classical Dufresne identity for matrix processes. More specifically, we show that the inverse Wishart laws on the space of positive definite $r \times r$ matrices can be realized by $\int_0^\infty M_s M_s^T ds$ in which $t \mapsto M_t$ is a drifted Brownian motion on $GL_r(\mathbb{R})$. This solves a problem in the study of spiked random matrix ensembles which served as the original motivation for this result. Various known extensions of the Dufresne identity (and their applications) are also shown to have analogs in this setting. In particular, we identify matrix-valued diffusions built from M_t which generalize in a natural way the scalar processes figuring into the geometric Lévy and Pitman theorems of Matsumoto and Yor.

1 Introduction

For $t \mapsto b_t$ a standard Brownian motion denote the associated geometric Brownian motion with drift, along with its (square) running integral, by

$$m_t = m_t^{(\mu)} = e^{b_t + \mu t}, \quad a_t^{(\mu)} = \int_0^t m_s^2 ds. \quad (1)$$

We will use the convention that $\mu > 0$, with the choice of sign in the superscript of $a_t^{(\pm\mu)}$ reserved to produce an integral of $(m_t)^2 = (m_t^{(\pm\mu)})^2$ either converging or diverging (almost

Received October 1, 2014; Revised April 6, 2015; Accepted April 15, 2015
Communicated by Prof. Alexei Borodin

surely) as $t \rightarrow \infty$. In certain situations (if it does not cause confusion), we will not denote the dependence on μ explicitly.

The functional $a_t^{(\mu)}$ arises in a number of contexts, including mathematical finance, diffusions in random environment, Brownian motion on hyperbolic spaces and continuum models of 1 + 1 dimensional polymers (see [23] and references therein). Connected to the valuation of a certain perpetuity, Dufresne [13] established the fundamental identity in law,

$$a_\infty^{(-\mu)} \stackrel{(law)}{=} \frac{1}{2\xi}, \tag{2}$$

in which ξ has the Gamma(μ) distribution, with density function $\frac{1}{\Gamma(\mu)} x^{\mu-1} e^{-x}$ on the positive half-line.

The Dufresne result provides us one possible starting point to what is a vast collection of beautiful distributional identities for integrated geometric Brownian motion, much of which was pioneered by the work of Matsumoto and Yor. For instance, the following process-level version of (2) was proved in [22]:

$$\left\{ \frac{1}{a_t^{(-\mu)}}, t > 0 \right\} \stackrel{(law)}{=} \left\{ \frac{1}{a_t^{(\mu)}} + \frac{1}{\tilde{a}_\infty^{(-\mu)}}, t > 0 \right\}. \tag{3}$$

Here, $\tilde{a}_\infty^{(-\mu)}$ is a copy of $a_\infty^{(-\mu)}$, independent of the original Brownian motion b_t . It is important to note that Dufresne himself had earlier established (3) at fixed times in [14]. Further afield extensions include “geometric” versions of Lévy’s $M - X$ and Pitman’s $2M - X$ theorem [21], a Brownian Burke property [28], and the integrability of the O’Connell–Yor polymer model [27].

Motivated by a problem in random matrix theory, one of the authors and Ramírez were led to a conjectured Dufresne-type identity for matrix processes [30]. Here, we prove that conjecture, and begin a program to extend the various results connected to the Dufresne identity to these matrix processes. In Section 1.1, we state our matrix analogs of (2) and (3). In Section 1.2, we introduce matrix diffusions which provide a possible generalization of those appearing in the just alluded to geometric Lévy and Pitman theorems, and discuss their asymptotics and intertwining properties. Section 1.3 states a partial Burke-type property for our matrix process. Finally, in Section 1.4, we go back and describe the motivating spiked random matrix connection, and Section 1.5 discusses some open problems and related results in the literature.

1.1 Dufresne for matrix processes

The natural matrix extension $M_t = M_t^{(\mu)}$ of the geometric Brownian motion which arises in [30] is defined by the $r \times r$ matrix Itô equation

$$dM_t = M_t dB_t + \left(\frac{1}{2} + \mu\right) M_t dt, \quad M_0 = I, \quad t \geq 0, \tag{4}$$

where $t \mapsto B_t$ is the matrix-valued Brownian motion comprised independent standard Brownian motions $\{b_{ij}(t)\}_{1 \leq i, j \leq r}$. Certainly, this coincides with m_t when $r = 1$. Note that M_t is rotational invariant: if O is a fixed orthogonal matrix then $OM_t O^T$ has the same law as a process as M_t .

As we will point out below in Section 2, M_t is almost surely invertible for all time and for any $s > 0$, the process $t \rightarrow M_s^{-1} M_{t+s}$, $t \geq 0$ has the same law as M_t , $t \geq 0$ and is independent of $\{M_r, 0 \leq r \leq s\}$. Using the independent multiplicative increment property, it is easy to extend M_t for all $t \in \mathbb{R}$. Either version of the process may be referred to as the Brownian motion (with drift μ) on the general linear group GL_r .

Along with M_t , we also define the additive functional $A_t^{(\mu)} = \int_0^t M_s M_s^T ds$ which is the matrix analog of the running integral $a_t^{(\mu)}$ from (1). Our basic matrix Dufresne identity is the following.

Theorem 1. If $2\mu > r - 1$, the $r \times r$ random matrix

$$A_\infty^{(-\mu)} = \int_0^\infty M_s M_s^T ds \tag{5}$$

has the standard inverse Wishart distribution with parameter 2μ . □

As Lemma 11 below shows, $\lim_{t \rightarrow \infty} \frac{1}{t} \log \|M_t^{(-\mu)}\| = -\mu + \frac{r-1}{2}$ with any matrix norm $\|\cdot\|$. The condition $2\mu > r - 1$ ensures that $A_\infty^{(-\mu)}$ is almost surely finite. That condition is also necessary for the nondegeneracy of underlying Wishart distribution.

Recall that the standard $r \times r$ (real) Wishart distribution with parameter $p > r - 1$ is the law on the cone of symmetric positive definitive matrices \mathcal{P} prescribed by

$$\gamma_p(dX) = \frac{1}{\Gamma_r(p/2)} (\det X)^{\frac{p-r-1}{2}} e^{-\frac{1}{2} \text{tr} X} \mathbf{1}_{\mathcal{P}}(X) dX. \tag{6}$$

Since all matrix variables will reside in GL_r , we often omit the dependence on r from the notation for the various distributions as well as their support, e.g., \mathcal{P} . Here, $\Gamma_r(p/2)$ the

multivariate gamma function $\Gamma_r(p/2) = \pi^{\frac{r(r-1)}{4}} \prod_{k=1}^r \Gamma(\frac{p-k+1}{2})$. When p is also an integer, γ_p can be realized by the random sample covariance matrix GG^T for G an $r \times p$ matrix with independent standard normal entries. In either case, it is the natural multivariate generalization of the gamma distribution. In symbols, then Theorem 1 reads $A_\infty^{(-\mu)(\text{law})} \equiv \gamma_{2\mu}^{-1}$, the latter having density proportional to $(\det X)^{-\mu - \frac{r+1}{2}} e^{-\frac{1}{2}\text{tr}X^{-1}}$ on \mathcal{P} . There are of course complex and quaternion Wishart distributions. Corollary 10 provides a version of Theorem 1 for these settings.

The original Dufresne identity (2) has a number of different proofs, not all of which appear extendable beyond the scalar case. To highlight those ideas that do carry over to the matrix case, we give two different proofs of Theorem 1. Both appear in Section 2. The first uses a time inversion strategy employed by several authors. The second mimics an argument of Baudoin–O’Connell [3] which is likely the most succinct proof of the 1D identity and which we briefly summarize now.

Let $y_t = y_0 e^{2b_t - 2\mu t}$, which is to say that y_t is m_t^2 with a variable starting point and the convergent choice of the sign of μ . The observation in [3] is that

$$u(y) = E[e^{-\frac{1}{2} \int_0^\infty y_t dt} | y_0 = y] = E[e^{-\frac{1}{2} \int_0^\infty y_t dt} | y_0 = 1] = E[e^{-\frac{1}{2} y a_\infty^{(-\mu)} }],$$

is on the one hand the Laplace transform of the desired distribution, and on the other hand, courtesy of Feynman–Kac, a solution of

$$2y \frac{d}{dy} y \frac{d}{dy} u(y) - 2\mu y \frac{d}{dy} u(y) - \frac{1}{2} y u(y) = 0, \quad u(0) = 1. \tag{7}$$

The unique bounded solution of (7) is then shown to be $u(y) = \frac{2^{1-\mu}}{\Gamma(\mu)} y^{\mu/2} K_\mu(\sqrt{y})$, where

$$K_s(a) = \frac{1}{2} \int_0^\infty x^{s-1} e^{-\frac{1}{2} a(x + \frac{1}{x})} dx \tag{8}$$

is the Macdonald function (or modified Bessel function of the second kind). After a change of variables, u is recognized as the Laplace transform of the (scaled) inverse gamma distribution.

Bessel functions of a matrix argument first appear in the 1955 work of Herz [16], and we introduce what is effectively his “ K -Bessel” function:

$$K_r(s|A, B) = \frac{1}{2} \int_{\mathcal{P}} (\det X)^{s - \frac{r+1}{2}} e^{-\frac{1}{2} \text{tr}AX - \frac{1}{2} \text{tr}BX^{-1}} dX, \tag{9}$$

for $A, B \in \mathcal{P}$. Note this reproduces the regular Macdonald function in the form $K_1(s|a, b) = (ab)^{s/2} K_s(\sqrt{ab})$ upon setting $r = 1$ and $A, B = a, b \in \mathbb{R}_+$. Both functions are well defined for all $s \in \mathbb{C}$. It is also clear that, up to a normalization, $K_r(-\mu|A, I)$ is the

Laplace transform (in the variable A) of the $\gamma_{2\mu}^{-1}$ distribution. Herz actually denotes what is effectively this function by B_r . We follow more closely the notation of Terras [34, Section 4.2.2], where this is referred to as the K -Bessel function of the second kind. We choose a slightly different normalization here (as in [17]) by introducing the extra $\frac{1}{2}$ constants in the exponential term to better align with the standard ($r = 1$) Macdonald function.

Picking up on the basic idea in [3], we set $Y_t = M_t M_t^T$ with $M_t = M_t^{(-\mu)}$ and $2\mu > r - 1$, so that $A_\infty^{(-\mu)} = \int_0^\infty Y_t dt$.

Theorem 2. The process $t \mapsto Y_t \in \mathcal{P}$ is Markovian with generator,

$$G_Y = 2\text{tr} \left(Y \frac{\partial}{\partial Y} \right)^2 - 2\mu \text{tr} \left(Y \frac{\partial}{\partial Y} \right), \tag{10}$$

expressed here through the matrix-valued operator $[\frac{\partial}{\partial Y}]_{ij} = (\frac{1}{2} + \frac{1}{2}\delta_{i,j}) \frac{\partial}{\partial Y_{ij}}$.

Furthermore, for $2\mu > r - 1$ the unique bounded solution of

$$G_Y U(Y) - \frac{1}{2}(\text{tr} Y)U(Y) = 0, \quad U(0) = 1, \quad Y \in \mathcal{P} \tag{11}$$

is the normalized K -Bessel function $U(Y) = \frac{K_r(-\mu|Y,I)}{2^{\mu-1} \Gamma_r(\mu)}$. □

Theorem 1 then follows from considerations similar to those above:

$$\mathbb{E}[e^{-\frac{1}{2}\text{tr}(Y A_\infty^{(-\mu)})}] = \mathbb{E}[e^{-\frac{1}{2}\text{tr}(Y \int_0^\infty Y_t dt)} | Y_0 = I] = \mathbb{E}[e^{-\frac{1}{2}\text{tr}(\int_0^\infty Y_t dt)} | Y_0 = Y] = U(Y).$$

The middle equality uses that Y_t started from $Y \in \mathcal{P}$ is equal in law to $\sqrt{Y}Y_t\sqrt{Y}^T$, with now Y_t started from the identity, along with the trace being cyclic.

The theory of matrix Bessel functions has been developed considerably since [16], in part due to applications to multivariate statistics as well as to the harmonic analysis of symmetric spaces. See for example [25] (particularly Chapter 7) and [34], respectively. Both references include a number of differential operator characterizations of various matrix Bessel functions. Still, the present characterization of $K_r(\cdot|A, I)$ appears new despite the obvious similarities of (7) and (11).

Remark 3. The process Y_t (modulo drift) was previously studied in [26] as one of two canonical “Brownian motions of ellipsoids”. Its Markov property, along with that of its joint process of eigenvalues, was already remarked upon there. Because of the rotational invariance of M_t , the function $U(Y)$ is actually determined by the eigenvalues

$\Lambda = \Lambda(Y) = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ alone. The eigenvalue process has generator

$$G_\Lambda = \sum_{k=1}^r \left(2\lambda_k \frac{\partial}{\partial \lambda_k} \lambda_k \frac{\partial}{\partial \lambda_k} - (r-1+2\mu)\lambda_k \frac{\partial}{\partial \lambda_k} \right) + \sum_{k < \ell} \frac{1}{\lambda_k - \lambda_\ell} \left(2\lambda_k^2 \frac{\partial}{\partial \lambda_k} - 2\lambda_\ell^2 \frac{\partial}{\partial \lambda_\ell} \right),$$

and thus (11) can be expressed instead by $(G_\Lambda - \frac{1}{2} \sum_{k=1}^r \lambda_k)U(\Lambda) = 0$. □

Last, we also have the exact matrix analog of the process-level Dufresne identity (3).

Theorem 4. There is the following identity in distribution:

$$\left\{ \left(A_t^{(\mu)} \right)^{-1}, t \geq 0 \right\} \stackrel{(\text{law})}{=} \left\{ \left(A_t^{(-\mu)} \right)^{-1} - \left(A_\infty^{(-\mu)} \right)^{-1}, t \geq 0 \right\}. \tag{12}$$

Again, it is assumed that $2\mu > r - 1$. □

Theorem 4 is proved in Section 3 by an enlargement of filtration argument, similar to the proof in [22] for the scalar case. Note the slightly different, but equivalent, presentation of the identity compared with (3). In this form, (12) can actually be made an almost sure identity by an appropriate construction of the underlying matrix Brownian motions (see Proposition 15).

1.2 Geometric Lévy and Pitman theorems

Connected to their study of the functional $a_t^{(\mu)}$, Matsumoto–Yor introduced the pair of processes,

$$x_t = m_t^{-2} \int_0^t m_s^2 ds, \quad z_t = m_t^{-1} \int_0^t m_s^2 ds, \tag{13}$$

(with $m_t = m_t^{(\mu)}$), both of which turn out to be diffusions [19–21]. For x_t the Markov property is immediate. An application of Itô’s formula produces the following simple sde for x_t for any $\mu \in \mathbb{R}$:

$$dx_t = 2x_t db_t + dt + (2 - 2\mu)x_t dt. \tag{14}$$

Plainly, the same procedure applied to z_t cannot produce a closed equation. Nonetheless, z_t is a Markov process (for any $\mu \in \mathbb{R}$) with law described by,

$$dz_t = z_t d\bar{b}_t + \left(\frac{1}{2} - \mu \right) z_t dt + \frac{K_{\mu+1}}{K_\mu} \left(\frac{1}{z_t} \right) dt. \tag{15}$$

Here, K_μ is the Macdonald function (8) and \bar{b}_t is a new Brownian motion (the subtlety is explained momentarily). An important property of z_t is its invariance under the transformation $\mu \mapsto -\mu$ which follows from identity $K_{\mu-1}(a) = K_{\mu+1}(a) - (2\mu/a)K_\mu(a)$ along with the more transparent fact $K_\mu(a) = K_{-\mu}(a)$.

The interest in x_t and z_t is that they encode generalizations of the classical $M - X$ theorem of Lévy, as well as the $2M - X$ theorem of Pitman, as was discovered by Matsumoto–Yor [19]. In particular, rescaling time by c^2 and taking μ into γ/c yields

$$x_{c^2 t}^{\gamma/c(\text{law})} \stackrel{\text{def}}{=} c^2 \int_0^t e^{c(2b_s^\gamma - 2b_t^\gamma)} ds, \quad z_{c^2 t}^{\gamma/c(\text{law})} \stackrel{\text{def}}{=} c^2 \int_0^t e^{c(2b_s^\gamma - b_t^\gamma)} ds,$$

in which b_t^γ is shorthand for the drifted Brownian motion. Simple Laplace asymptotics yield that, as $c \rightarrow \infty$, $\frac{1}{2c} \log \int_0^t e^{c(2b_s^\gamma - 2b_t^\gamma)} ds$, and $\frac{1}{c} \log \int_0^t e^{c(2b_s^\gamma - b_t^\gamma)} ds$ converge pathwise to $\max_{s < t} (b_s^\gamma - b_t^\gamma)$ and $\max_{s < t} (2b_s^\gamma - b_t^\gamma)$. Working on the sde side (that is, with (14) and (15)), shows that $\frac{1}{2c} \log x_{c^2 t}^{\gamma/c}$ and $\frac{1}{c} \log z_{c^2 t}^{\gamma/c}$ have limiting processes that are equivalent in law to the diffusions with respective generators

$$G_x = \frac{1}{2} \frac{d^2}{dx^2} - \gamma \operatorname{sgn}(x) \frac{d}{dx}, \quad G_z = \frac{1}{2} \frac{d^2}{dz^2} + \gamma \cot(\gamma z) \frac{d}{dz}. \tag{16}$$

The former is understood to be equipped with a Neumann boundary condition at the origin: the limiting x -process is reflected at the origin, whereas the z -process has an entrance boundary at that point.

Letting $\gamma \downarrow 0$, from the processes (16), we recover the reflected Brownian motion and the 3D Bessel process occurring in the celebrated results of Lévy and Pitman identifying the distributions of the processes $t \rightarrow \max_{s < t} (b_s - b_t)$ and $t \rightarrow \max_{s < t} (2b_s - b_t)$. Taking the point of view that ‘ $\exp \int \log$ ’ has replaced the running maximum, that the Brownian functionals in (13) are diffusions is now referred to as the “geometric” $M - X$ or $2M - X$ theorem.

Similar to the original Dufresne identity (2), there are various ways to identify the law of z_t with (15). The one relevant here is again due to Matsumoto–Yor [21], and rests on properties of the generalized inverse Gaussian (GIG) distribution. The GIG is a three-parameter distribution on the positive half-line with density proportional to $x^{p-1} e^{-\frac{1}{2}ax - \frac{1}{2}bx^{-1}}$, with arbitrary p and $a, b > 0$. As such, it is intimately connected to the Macdonald function K_p : the ratio appearing in the drift (15) being the mean of a GIG with parameters $(\mu, 1/z_t, 1/z_t)$. What is proved in [21] is that the law of m_t conditional on the field $\{z_s, s \leq t\}$ is exactly this GIG, and the closed equation (15) is produced by a projection (and so the indicated \bar{b}_t is measurable with respect to $\sigma(z_s, s \leq t)$).

By analogy with (13), we introduce

$$X_t = M_t^{-1} \left(\int_0^t M_s M_s^T ds \right) M_t^{-T}, \quad Z_t = M_t^{-1} \int_0^t M_s M_s^T ds, \tag{17}$$

for our matrix process $t \mapsto M_t$ as defined originally in (4). Note that from now on we use the shorthanded $M^{-T} = [M^T]^{-1}$. That X_t is Markovian is again straightforward using Itô's formula. For Z_t , as can be anticipated at this point, there is a matrix GIG distribution on \mathcal{P} , more or less defined by having normalizer given by the K -Bessel function (9). That is, it is the law

$$\eta_{p,A,B}(dX) = c(\det X)^{p-\frac{1}{2}(r+1)} e^{-\frac{1}{2}(\text{tr}AX + \text{tr}BX^{-1})} \mathbf{1}_{\mathcal{P}}(X) dX, \tag{18}$$

with $A, B \in P$ and $c = \frac{1}{2} K_r(p|A, B)^{-1}$.

Theorem 5. For all μ , the process X_t is the diffusion defined by the Itô equation

$$dX_t = Idt + (1 - 2\mu)X_t dt + \text{tr}X_t Idt - dB_t X_t - X_t dB_t^T, \tag{19}$$

run on the same Brownian motion B_t is as M_t (4). For $2\mu > r - 1$ the process X_t is reversible with respect to the invariant distribution $\gamma_{2\mu}^{-1}$.

If $|\mu| > \frac{r-1}{2}$ the process Z_t is also a diffusion. It satisfies

$$dZ_t = d\bar{B}_t Z_t + (\frac{1}{2} - \mu)Z_t dt + \kappa_\mu(I, (Z_t Z_t^T)^{-1})Z_t dt, \tag{20}$$

where now \bar{B}_t is a matrix-valued Brownian motion adapted to $\sigma(Z_s, s \leq t)$ and $\kappa_p(A, B)$ denotes the mean of the $\eta_{p,A,B}$ distribution (18). In addition, the law of $t \mapsto Z_t$ is unchanged by taking μ into $-\mu$. □

While both Z_t and the right-hand side of (20) are sensible for all μ , our method uses Theorem 1 as input and so requires the same condition. One assumes this gap might be filled by other means.

In general, the mean of a matrix GIG does not appear to have a particularly nice expression. Although if A and B are diagonal, one can see that $\kappa_p(A, B)$ is also diagonal. By bringing in a (well-known) generalization of the K -Bessel functions introduced thus far, one can get a reasonable handle on these diagonal components. See the proof of Theorem 7 for both points. Note that the invariance of Z_t under the map $\mu \mapsto -\mu$ implies the identity $\kappa_\mu(I, A) = 2\mu I + \kappa_{-\mu}(I, A)$, for $|\mu| > \frac{r-1}{2}$. (A standard analytic continuation argument extends the identity to all $\mu \in \mathbb{R}$.) A direct verification of this identity seems laborious (and nontrivial).

The key to (20) is that the conditional distribution of M_t given $\{Z_s, s \leq t, Z_t = Z\}$ is $Z^T \mathcal{E}$ for $\mathcal{E} \sim \eta_{\mu, I, (ZZ^T)^{-1}}$. This hinges on a characterization of the matrix GIG law due to Bernadac [4], which in turn builds on earlier work of Letac–Wesolowski [17]. An immediate consequence of this is the following.

Corollary 6. Let $|\mu| > \frac{r-1}{2}$. The laws of M_t and Z_t intertwine. Denote by T_t^M and T_t^Z the corresponding semigroups and define the Markov kernel Λ as

$$\Lambda h(Z) = \int_{\mathcal{P}} h(Z^T X) \eta_{\mu, I, (ZZ^T)^{-1}}(dX), \tag{21}$$

for all suitable test functions $h: \text{GL}_r^+ \mapsto \mathbb{R}$. (Here, GL_r^+ is the space of invertible $r \times r$ matrices with a positive determinant.) Then, it holds that $\Lambda T_t^M = T_t^Z \Lambda$. Since $X_t = Z_t M_t^{-T}$, it follows that X_t also intertwines with Z_t . In this case,

$$\tilde{\Lambda} h(Z) = \int_{\mathcal{P}} h(X^{-1}) \eta_{\mu, I, (ZZ^T)^{-1}}(dX), \tag{22}$$

defines the corresponding kernel for which $\tilde{\Lambda} T_t^X = T_t^Z \tilde{\Lambda}$. □

This intertwining has had far reaching implications in the scalar case. A remaining question here is whether the matrix processes contain either an $M - X$ or $2M - X$ type theorem. We show this occurs at the level of the eigenvalues (or singular values) of X_t and Z_t , each of which comprise their own Markov process.

Theorem 7. Denote by $x_r^\mu(t) \leq \dots \leq x_1^\mu(t)$ the eigenvalues of X_t . Denote the ordered singular values of Z_t similarly by $z_r^\mu(t)$. Speeding up time by a factor of c^2 and rescaling μ as in $\mu = \frac{r-1}{2} + \gamma/c$ for a fixed $\gamma > 0$, we have that

$$\lim_{c \rightarrow \infty} \frac{1}{2c} \log x_1^{\gamma/c}(c^2 t) \Rightarrow |b_t^{-\gamma \text{sgn}(\cdot)}|,$$

and

$$\lim_{c \rightarrow \infty} \frac{1}{c} \log z_r^{\gamma/c}(c^2 t) \Rightarrow b_t^{\gamma \coth(\gamma \cdot)}.$$

The notations indicate a reflected Brownian motion with constant drift $-\gamma$, and a Brownian motion with variable drift $\gamma \coth(\gamma \cdot)$, respectively. In both cases, the convergence takes place in the usual Skorohod topology. □

In either case, the reminder of the spectrum has a relatively trivial limit in the chosen scaling. For X_t , each of the similarly scaled lower eigenvalues converge to the

zero process. For Z_t , the larger singular values escape to infinity at increasing exponential rates. Note as well that Theorem 7 provides an analog for just half of the Matsumoto–Yor result—one might like at the same time to have pathwise identities by applying some sort of Laplace asymptotics to the definitions (17).

The proofs of Theorems 5 and 7, along with that of Corollary 6, are found in Section 4. Along the way, we will also show that the generator of the process $(Z_t^T Z_t)^{-1}$ can be expressed as the Doob transform of the generator of the process $Y_t = M_t M_t^T$ with an additional killing term via the function $U(Y)$ (defined in Theorem 2).

1.3 Burke properties

O’Connell–Yor [28] proved the following “Brownian Burke property”. Let b_t and c_t be independent Brownian motions, and set

$$r_t = \log \int_{-\infty}^t e^{b_{(s,t)} + c_{(s,t)} - \mu(t-s)} ds, \tag{23}$$

where $b_{(s,t)} = b_t - b_s$ and $c_{(s,t)} = c_t - c_s$. Then,

$$\left. \begin{matrix} b_t + r_0 - r_t, t \in \mathbb{R} \\ c_t + r_0 - r_t, t \in \mathbb{R} \end{matrix} \right\} \stackrel{\text{(law)}}{=} \{b_t, t \in \mathbb{R}\} \text{ and are independent.} \tag{24}$$

The analogy with the classical Burke property is made by considering $t \mapsto r_t$ as a generalized queue, where “log ∫ exp” again replaces “sup”, with $t \mapsto b_t$ and $t \mapsto \mu t - c_t$ the respective arrival and departure processes. This result is key in the construction of the semi-directed Brownian polymer (also introduced in [28]) which is now understood to be a member of the KPZ universality class [7]. A similar Burke-type property lies behind the integrability of Seppäläinen’s log-gamma polymer [32] which has also subsequently been shown to have Tracy–Widom fluctuations [8].

The above scheme constructs two new independent Brownian motions from two independent input Brownian motions. As a preliminary step, a similar statement is established in [28] that shows that

$$\{b_t^\mu + \alpha_t - \alpha_0, t \geq 0\} \stackrel{\text{(law)}}{=} \{b_t^\mu, t \geq 0\}, \quad \text{where } \alpha_t = \log \int_{-\infty}^t e^{2b_s^\mu - 2b_t^\mu} ds. \tag{25}$$

Here, we are reusing notation from before, b_t^μ denoting a Brownian motion with drift μ . Also, for fixed $\tau \in \mathbb{R}$, the field generated by the new drifted Brownian motion $b_t^\mu + \alpha_t - \alpha_0$ over $t \leq \tau$ is independent of $\{\alpha_t, t \geq \tau\}$.

The following provides a matrix extension of (25). It requires the full line version of the process M_t , the details of which are again described at the beginning of Section 2.

Theorem 8. Fix $2\mu > r - 1$ and let $M_t = M_t^{(\mu)}$ be the solution of (4) extended over $t \in (-\infty, \infty)$. Then,

$$\left(\int_{-\infty}^0 M_s M_s^T ds \right)^{-1} M_t \left(\int_{-\infty}^t M_t^{-1} M_s M_s^T M_t^{-T} ds \right) \stackrel{\text{(law)}}{=} M_t, \tag{26}$$

as processes for $t \geq 0$. In addition, for any fixed $\tau > 0$, the process defined by the left-hand side of (26) up to time τ is independent of $t \mapsto M_t^{-1} \left(\int_{-\infty}^t M_s M_s^T ds \right) M_t^{-T}$ for $t \geq \tau$. \square

Note that $M_t^{-1} \left(\int_{-\infty}^t M_s M_s^T ds \right) M_t^{-T}$ reduces to $e^{\alpha t}$ (and so (26) reduces precisely to (25)) for $r = 1$.

The next result is a matrix version of (24), the Brownian Burke property of O’Connell–Yor.

Theorem 9. Let B_t and C_t be independent two-sided matrix Brownian motions and set $2\mu > r - 1$. Consider the strong solution of

$$dH_t = H_t(dB_t + dC_t) + (2\mu + 1)H_t dt, \quad H_0 = I,$$

extended to the whole line, again using the (multiplicative) independence stationary increment property as described in Section 2. Now, define the processes $A_{(-\infty, t)} = \int_{-\infty}^t H_u H_u^T du$ and

$$F_t = B_t + 2\mu It - \frac{1}{2} \int_0^t H_s^T A_{(-\infty, s)}^{-1} H_s ds, \quad G_t = C_t + 2\mu It - \frac{1}{2} \int_0^t H_s^T A_{(-\infty, s)}^{-1} H_s ds.$$

Then, $(F_t, G_t, t \geq 0) \stackrel{\text{(law)}}{=} (B_t, C_t, t \geq 0)$. \square

Here, the analogy to the 1D case is not as immediate, but the process $t \mapsto 2\mu t I - \frac{1}{2} \int_0^t H_s^T A_{(-\infty, s)}^{-1} H_s ds$ can be seen to correspond to $r_0 - r_t$ by first differentiating and then integrating back up in the definition (23). Note as well that $H_{t/2} \stackrel{\text{(law)}}{=} M_t^{(\mu)}$.

Theorems 8 and 9 are proved in Section 3.

1.4 Connection to spiked random matrices

An important problem in mathematical statistics is to describe the law of the largest eigenvalue of sample covariance (or Wishart) matrices of the form $G \Sigma G^\dagger$. In the basic setting, G is $p \times q$ and comprised independent unit Gaussians in $\mathbb{F} = \mathbb{R}, \mathbb{C}$, or \mathbb{H} and \dagger is the associated conjugate transpose. One is typically interested in the limit as p and q tend to ∞ with Σ some deterministic sequence of symmetric population matrices.

When $\Sigma = I$, this so-called soft-edge limit is well known to be given by the $\beta = 1, 2$ or 4 Tracy–Widom laws (for the case of \mathbb{R}, \mathbb{C} , or \mathbb{H} entries respectively). Moving toward the more general problem, the spiked ensembles in which $\Sigma = \Sigma_r \oplus I_{q-r}$ and r remains fixed as p and q grow have generated considerable interest.

Using the determinantal framework at $\beta = 2$, [2] proved there exists a phase transition. Below, criticality one sees Tracy–Widom in the limit, above criticality there are Gaussian effects (the limit given by the law of the largest eigenvalue of a finite rank GUE), with a new one parameter family of spiked soft-edge laws in the crossover regime. Subsequent analytic work was carried at $\beta = 1$ and $\beta = 4$ by [24, 35], among others.

In another direction, Bloemendal and Virág [5, 6] proved that the $\beta = 1, 2$, or 4 soft-edge spiked laws can be characterized in a unified way through the eigenvalue problem for the (random) operator H acting on functions $f \in L^2([0, \infty), \mathbb{F}^r)$ defined by

$$H = -\frac{d^2}{dt^2} + rt + \sqrt{2}\mathcal{B}'_t, \quad f'(0) = Cf(0).$$

Here, \mathcal{B}_t is the standard \mathbb{F} -invariant Brownian motion, that is, for $U \in U_r(\mathbb{F}) = \{V \in \mathbb{F}^{r \times r} : VV^\dagger = I\}$ it holds that $U\mathcal{B}_tU^\dagger \stackrel{\text{(law)}}{=} \mathcal{B}_t$, and C is the scaling limit of the matrix Σ_r . At $r = 1$, the result holds for all $\beta > 0$. In that case, the noise term reduces to $\frac{2}{\sqrt{\beta}}b'_x$, and H is recognized as the stochastic Airy operator from [31], but with a Robin (rather than Dirichlet) boundary condition.

Ramírez and Rider [30] asked whether one could similarly spike the hard edge, or smallest eigenvalue laws for general β (though see [12] for earlier work specific to $\beta = 2$). This regime is defined by setting $q = p + a$ for $a > -1$ remaining fixed as $p \rightarrow \infty$. A primary motivation was to confirm that the resulting spiked hard-edge laws recover all known spiked soft edge laws via the familiar hard-to-soft transition.

To describe the spiked hard-edge operator, we set $\mu = \frac{a+r}{2}$ and introduce

$$dM_t = M_t dW_t + \left(\frac{1}{2\beta} - \mu\right) M_t dt, \quad \mathcal{M}_t = M_t M_t^\dagger, \tag{27}$$

with the usual $M_0 = I$. Here, $t \mapsto W_t$ is again an $r \times r$ matrix of independent Brownian motions, but now with each off-diagonal entry a unit \mathbb{F} -valued Brownian motion and each diagonal entry a real Brownian motion with mean zero and mean-square β^{-1} . The relevant result from Ramírez and Rider [30] is that, under the condition that $(p\Sigma_r)^{-1} \rightarrow C^{-1}$ in norm, the limiting smallest eigenvalues of $pG\Sigma G^\dagger$ are described by the eigenvalue problem for

$$Gf(t) = \int_0^\infty \left(\int_0^{t \wedge s} e^{ru} \mathcal{M}_u^{-1} du \right) \mathcal{M}_s f(s) ds + C^{-1} \int_0^\infty \mathcal{M}_s f(s) ds, \tag{28}$$

acting on $f \in L^2[\mathcal{M}]$, the space of functions $f: \mathbb{R}_+ \mapsto \mathbb{F}^r$ with $\int_0^\infty (f^\dagger \mathcal{M}_t f)(t) dt < \infty$. The operator G is positive compact and actually describes the limiting *inverse* Wishart eigenvalues, that is, the limiting Wishart eigenvalues are the spectral points λ for the problem $\lambda Gf = f$.

As in the soft edge case, when $r = 1$, the result is valid for all $\beta > 0$. Also, when $r = 1$, both \mathcal{M}_t and $e^{rt} \mathcal{M}_t^{-1}$ reduce to geometric Brownian motions and G has the interpretation of the Green’s function for a Brownian motion in a Brownian potential. Again at $r = 1$ with $C = c \in \mathbb{R}$, the limit $c \rightarrow \infty$ recovers the basic β hard edge operator introduced in [29].

While there is no critical point at the hard edge, the supercritical regime refers to choosing $C = cI$ and taking $c \rightarrow 0$ in (28). At one level, the outcome is easy to describe: cG_{cI} converges (almost surely in operator norm) to the finite rank operator defined by integration against \mathcal{M}_t . By analogy with the $r \times r$ Gaussian invariant ensemble supercritical limit at the spiked soft edge, the obvious conjecture was that $\text{spec}(\int_0^\infty \mathcal{M}_t dt)$ should be described by the inverse Wishart law(s). Of course, for $r = 1$, the conjecture was known to be correct due to the original Dufresne identity.

Corollary 10. With the appropriate choice of $t \mapsto W_t$, the eigenvalues of the $r \times r$ random matrix $\int_0^\infty \mathcal{M}_t dt$ have joint law given by the eigenvalues of the inverse \mathbb{F} -Wishart distribution with parameter 2μ . Using isotropic matrix Brownian motions and replacing the underlying process in (27) by

$$dM_t = M_t dB_t + \left(\frac{1}{\beta} - \frac{1}{2} - \mu \right) M_t dt, \quad M_0 = I, \tag{29}$$

with B_t now comprised completely of independent unit \mathbb{F} -valued Brownian motions, the full \mathbb{F} -Wishart distribution is recovered by $\int_0^\infty M_t M_t^\dagger dt$. In both cases, the natural condition on μ remains $2\mu > r - 1$. □

Of course, when $\beta = 1$, the Equations (27) and (29) agree and the above is a repeat of Theorem 1. While the structured noise in (27) is what arises in the spiked random matrix problem, we mention the result for (29) for $\beta = 2$ and 4, as it seems a more natural construction and readily produces the full matrix law. The proof of Corollary 10 is sketched alongside the proof of Theorem 1 in Section 2.

1.5 Further questions

The most obvious question is whether exists a (solvable) polymer model in matrix variables. The semi-directed Brownian polymer (or O’Connell–Yor polymer) alluded to above,

can be defined by the partition function

$$\mathcal{Z}_{n,t} = \int_{0 < s_1 < \dots < s_{n-1} < t} ds_1 \cdots ds_{n-1} \exp \left(b_{(0,s_1)}^{(1)} + b_{(s_1,s_2)}^{(2)} + \cdots + b_{(s_{n-1},t)}^{(n)} \right),$$

where $(b^{(1)}, \dots, b^{(n)})$ is a standard Brownian on \mathbb{R}^n . The stationary version defined earlier in [28] has the first level not started at zero, but instead distributed over the negative half-line by the measure $e^{b_t - \mu t}$. That partition function was in fact arrived at, and aspects of its law understood, by iterating the Brownian Burke property described in Section 1.3 (recall (23) and (24)). What is missing in our case is a matrix Burke property that can be iterated through the noncommutativity in the same fashion.

In [27], O’Connell shows that the $t \mapsto \mathcal{Z}_{n,t}$ process has the same law as the top component of a diffusion on \mathbb{R}^n whose generator is a conjugation of the quantum Toda Hamiltonian. Remarkably, when $n=2$, this result is exactly the Matsumoto-Yor $2M - X$ theorem (z_t is $\mathcal{Z}_{2,t}$ up to a change of variables). In both cases, there is an intertwining (between $\mathcal{Z}_{n,t}$ or z_t and the driving \mathbb{R}^n or \mathbb{R} Brownian motion) which provides a fairly explicit formula for the Laplace transform of the “ z ” processes. While we have an analogous intertwining (Corollary 6), the semigroup of Brownian motion on GL_r does not have a sufficiently concrete expression to afford a better characterization of the law of Z_t . Potentially, one might be able to bypass the intertwining, and find some description of the joint law (M_t, A_t) , and so Z_t , by more direct means (again, there are several such routes at $r=1$ [20]).

One might also consider various parts of the above program for different groups. In this general spirit, but from different directions, we point out the very recent papers of Chhaibi [11] and Bougerol [9]. In the second reference, geometric considerations lead to a process similar in structure to our Z_t , but constructed from a Brownian motion on the group of complex lower triangular matrices with positive diagonal. The singular values of this object are then shown to be Markov with generator given by a conjugation (by a polynomial function in copies of the Macdonald function and their derivatives) of quantum Toda on a Weyl chamber. Any direct link to the formulas derived here—to the generator of Z_t (Theorem 5) or that for its singular values (see Lemma 22 below)—is not immediately transparent. More simply, it is natural to ask what matrix laws beyond the Wishart can be constructed from of a “Dufresne procedure” (back in the vein of Theorem 1).

2 The Matrix Dufresne Identity

We prove the basic matrix Dufresne identity in two different ways. Stated above as Theorems 1 and 2, they fall below under the headings “diffusion” and “Feynman–Kac” proof. We also provide a sketch of a “diffusion” proof of Corollary 10.

First, we summarize some of the properties of $M_t = M_t^{(\mu)}$. Using the Taylor expansion of the determinant near I and Itô’s formula, one finds that

$$d \det M_t = \det M_t (\operatorname{tr} dB + r(\frac{1}{2} + \mu) dt), \tag{30}$$

for $t \geq 0$. Hence, $\det M_t = \exp(\operatorname{tr} B_t + \mu r t)$, and M_t is almost surely invertible for all time. Then, by the linearity of the sde (4) it follows that, for any $s > 0$, the process $t \rightarrow M_{s,t} = M_s^{-1} M_t$, $t \geq s$ satisfies the same equation subject to $M_{t,t} = I$. But that means that for fixed $s \geq 0$:

$$\{M_{s,s+t}, t \geq 0\} \stackrel{(\text{law})}{=} \{M_t, t \geq 0\}, \quad \{M_{s,s+t}, t \geq 0\} \text{ is independent of } \{M_r, r \leq s\}. \tag{31}$$

These properties allow a natural extension of M_t to all $t \in \mathbb{R}$. First, extend the matrix Brownian motion B_t for all $t \in \mathbb{R}$, and consider the strong solution \tilde{M}_t of the sde:

$$d\tilde{M}_t = \tilde{M}_t dB_t + \tilde{M}_t (\frac{1}{2} + \mu) dt, \quad \tilde{M}_{-1} = I, \quad t \in [-1, 0].$$

Setting $M_t = \tilde{M}_0^{-1} \tilde{M}_t$ for $t \in [-1, 0)$, the extended M_t , $t \geq -1$ process satisfies (31) for any $s \geq -1$. Repeating this procedure for earlier starting points defines a version of M_t over the whole line. Importantly, the resulting process has properties (31) for each $s \in \mathbb{R}$ and $t \rightarrow M_{s,t}$ satisfies (4) for $t \geq s$ with $M_{t,t} = I$.

Last, we state the Lemma 11 on the norm growth of M_t which we will use repeatedly. The proof is deferred to the very end of the section.

Lemma 11. Let $m_1 \leq m_2 \leq \dots \leq m_r$ be the singular values of $M_t = M_t^{(\mu)}$. It holds that $\lim_{t \rightarrow \infty} \frac{1}{t} \log m_i(t) = \mu + \frac{i-1}{2}$ with probability one for each $i = 1, \dots, r$. □

2.1 Diffusion proof

We also actually prove Theorem 1 in two different ways. For completeness, we first indicate how everything works directly through the matrix coordinates. The proof is somewhat more transparent in eigenvalue/eigenvector coordinates, and we carry out that approach afterwards.

2.1.1 Via matrix coordinates

Recall the definition of $M_t = M_t^{(-\mu)}$ with the convergent choice of sign for the drift:

$$dM_t = M_t dB_t + \left(\frac{1}{2} - \mu\right) M_t dt, \quad M_0 = I, \quad t \geq 0 \tag{32}$$

and of course $2\mu > r - 1$.

Consider the version of this process extended to the whole line (as described just above), and then introduce the time reversed process $N_t = M_{-t}$. We claim that N_t is also a Brownian motion on GL_r , but with drift μ instead of $-\mu$. In particular, for $t \geq 0$, it solves the SDE

$$dN_t = N_t d\tilde{B}_t + \left(\frac{1}{2} + \mu\right) N_t dt, \quad N_0 = I, \quad t \geq 0, \tag{33}$$

where $d\tilde{B}_t = -dB_{-t}$. A quick (but formal) explanation for this statement would follow from $(I + dB_t + (\frac{1}{2} - \mu)Idt)^{-1} - I \approx -dB_t - (\frac{1}{2} - \mu)Idt + dB_t dB_t$ and $dB_t dB_t = Idt$. For the precise proof, one needs to first verify that N_t also satisfies the stationary and independent increment property as M_t , and then to show that if M_t solves (32) on say $t \in [0, 1]$, the process $\tilde{N}_t = M_1^{-1} M_{1-t}$, $t \in [0, 1]$ will solve the same sde with $+\mu$ instead of $-\mu$ and with $-dB_{1-t}$ playing the role of dB_t .

From this point (with a small abuse of notation), we will drop the tilde from $d\tilde{B}_t$ and just use dB_t for the noise in N_t .

Next, consider $N_{s,t} = N_t^{-1} N_{s+t}$ for any fixed t . Then, $N_{s,t} \stackrel{\text{(law)}}{=} N_s$, again as processes in $s \in \mathbb{R}$, and if we further define

$$Q_t := \int_{-\infty}^0 N_{s,t} N_{s,t}^T ds = N_t^{-1} \left(\int_{-\infty}^t N_s N_s^T ds \right) N_t^{-T}, \tag{34}$$

we have at last a process that is stationary with marginal law given by that of

$$Q_0 = \int_{-\infty}^0 N_s N_s^T ds \stackrel{\text{(law)}}{=} \int_0^{\infty} M_t M_t^T dt.$$

The proof of Theorem 1 then comes down to the following.

Proposition 12. Let $2\mu > r - 1$. The process $t \mapsto Q_t$ is a diffusion corresponding to the matrix sde

$$dQ_t = Idt + (1 - 2\mu)Q_t dt + \text{tr}Q_t Idt - dB_t Q_t - Q_t dB_t^T. \tag{35}$$

If $Q_0 \in \mathcal{P}$, then Q_t remains in \mathcal{P} for all $t > 0$ with $\gamma_{2\mu}^{-1}$ as its unique invariant measure. \square

Proof. Applying Itô's formula in (34), we find that

$$\begin{aligned} dQ_t &= Idt + dN_t^{-1} \left(\int_{-\infty}^t N_s N_s^T ds \right) N_t^{-T} + N_t^{-1} \left(\int_{-\infty}^t N_s N_s^T ds \right) dN_t^{-T} \\ &\quad + dN_t^{-1} \left(\int_{-\infty}^t N_s N_s^T ds \right) dN_t^{-T}. \end{aligned}$$

We also have that

$$\begin{aligned} dN_t^{-1} &= -N_t^{-1} dN_t N_t^{-1} + N_t^{-1} dN_t N_t^{-1} dN_t N_t^{-1} \\ &= -dB_t N_t^{-1} + \left(\frac{1}{2} - \mu\right) N_t^{-1} dt, \end{aligned}$$

after substituting in (33) and using $dB_t dB_t = Idt$ in the second term of line one. Combined, this produces

$$dQ_t = Idt - (dB + (\mu - \frac{1}{2})Idt)Q_t - Q_t(dB + (\mu - \frac{1}{2})Idt)^T + dBQ_t dB^T,$$

which simplifies to (35) on account of the rule $dB_t C dB_t^T = (\text{tr}C)Idt$ for any matrix C .

To see that Q_t , defined by (35) and a fixed starting point $Q_0 \in \mathcal{P}$, remains in \mathcal{P} for all time, apply Itô's formula yet again, now to $t \mapsto \det(Q_t)$:

$$\begin{aligned} d \det(Q_t) &= \det(Q_t)(\text{tr}[Q_t^{-1} dQ_t] + \frac{1}{2}(\text{tr}[Q_t^{-1} dQ_t])^2 - \frac{1}{2}\text{tr}[Q_t^{-1} dQ_t Q_t^{-1} dQ_t]) \\ &= \det(Q_t)(-2\text{tr}dB_t + \text{tr}Q_t^{-1} dt + (2r - 2\mu)dt). \end{aligned} \tag{36}$$

For the second line, we used

$$\text{tr}[Q_t^{-1} dQ_t Q_t^{-1} dQ_t] = \text{tr}[Q_t^{-1} (dB)^2 Q_t] + \text{tr}(dB)^2 + 2\text{tr}[Q_t^{-1} dB Q_t dB^T].$$

Introducing $b_t = -\text{tr}B_t/\sqrt{r}$ and

$$dz_t = 2\sqrt{r}z_t db_t + (2r - 2\mu)z_t dt, \quad z_0 = \det(Q_0) > 0,$$

we see that $t \mapsto \det(Q_t)$ is bounded below by the geometric Brownian motion z_t up to the first passage of $\det(Q_t)$ to zero. But, z_t never vanishes, and so that passage time must be infinite.

The next ingredients are to show that Q_t admits a smooth positive transition density on \mathcal{P} , and then to identify the inverse Wishart law $\gamma_{2\mu}^{-1}$ as an invariant measure. Since the latter also has a positive density with respect to Lebesgue measure on \mathcal{P} , it will follow that this is the unique invariant measure for Q_t .

The generator of Q_t can be succinctly expressed in matrix coordinates as in

$$G_Q = 2\text{tr}\left(Q^2\left(\frac{\partial}{\partial Q}\right)^2\right) + (1 - 2\mu)\text{tr}\left(Q\frac{\partial}{\partial Q}\right) + (1 + \text{tr}Q)\text{tr}\left(\frac{\partial}{\partial Q}\right), \tag{37}$$

where again $\frac{\partial}{\partial Q}$ is the matrix-valued operator $[\frac{\partial}{\partial Q}]_{ij} = (\frac{1}{2} + \frac{1}{2}\delta_{ij})\frac{\partial}{\partial Q_{ij}}$. The second-order part of (37) is verified by writing out

$$(dBQ + QdB^T)_{ij}(dBQ + QdB^T)_{kl} = [Q^2]_{j\ell}\delta_{ik} + [Q^2]_{i\ell}\delta_{jk} + [Q^2]_{jk}\delta_{i\ell} + [Q^2]_{ik}\delta_{j\ell},$$

and summing over $i \leq j$ and $k \leq \ell$. One then checks that the adjoint takes the form

$$G_Q^* = 2\text{tr}\left(Q^2\left(\frac{\partial}{\partial Q}\right)^2\right) + (2\mu + 2r + 3)\text{tr}\left(Q\frac{\partial}{\partial Q}\right) + (\text{tr}Q - 1)\text{tr}\left(\frac{\partial}{\partial Q}\right) + \mu r(r + 1) + \frac{1}{2}r(r + 1)^2, \tag{38}$$

when restricted to act on functions of a symmetric matrix variable.

To invoke the necessary regularity estimates, we temporarily consider the “vectorized” Q_t , or $\text{vect}(Q_t) = (Q_{11}(t), Q_{22}(t), \dots) \in \mathbb{R}^{\frac{r(r+1)}{2}}$. In particular, we show that the diffusion matrix written in these coordinates is positive definite on the open set $\mathcal{P} \subset \mathbb{R}^{\frac{r(r+1)}{2}}$. This comes down to proving: if Q is a positive definite matrix and Z is an $r \times r$ matrix normal with iid entries then the covariance matrix of $\text{vect}(ZQ + QZ^T)$ is positive definite. To see this, take any $r \times r$ orthogonal matrix O and let K_O be the linear map on $\mathbb{R}^{\frac{r(r+1)}{2}}$ defined by $\text{vect}(O^T Q O) = K_O \text{vect}(Q)$. It is easy to verify that $|\det K_O| = 1$. Next write the spectral decomposition of $Q = O^T \Lambda O$, and note that the desired covariance matrix satisfies

$$\begin{aligned} & \text{E} \text{vect}(ZQ + QZ^T) \text{vect}(ZQ + QZ^T)^T \\ &= \text{E} K_O \text{vect}(Z\Lambda + \Lambda Z^T) \text{vect}(Z\Lambda + \Lambda Z^T)^T K_O^T. \end{aligned}$$

Thus, by the first remark, we may assume that Q is diagonal with entries $Q_{ii} > 0$. But in that case the covariance matrix is diagonal with entries $4[Q^2]_{ii}, 1 \leq i \leq r$ and $[Q^2]_{ii} + [Q^2]_{jj}, 1 \leq i < j \leq r$ which is clearly positive definite.

Finally, [33, Theorem 3.4.1] (though see also Remark 3.4.2 there) implies that $\partial_t - G_Q^*$ is hypoelliptic on \mathcal{P} . At the same time, a straightforward but tedious calculation will show that $G_Q^* f(Q) = 0$ for $f(Q) = (\det Q)^{-\mu - \frac{r+1}{2}} e^{-\frac{1}{2}\text{tr}Q^{-1}}$, the $\gamma_{2\mu}^{-1}$ density. For smooth

test functions h of compact support on \mathcal{P} and T_t the semigroup of Q_t , we have

$$\frac{\partial}{\partial t} \int_{\mathcal{P}} (T_t h)(Q) f(Q) dQ = \int_{\mathcal{P}} G_Q (T_t h)(Q) f(Q) dQ.$$

An integration by parts would continue the equality as $\int_{\mathcal{P}} (T_t h)(Q) (G_Q^* f)(Q) dQ = 0$ and complete the proof. To justify this, that is, that there are no boundary terms, requires two facts. The first is that $T_t h(Q)$ and its normal derivative are bounded along the boundary $\det(Q) = 0$. This can be established by writing G_Q in local coordinates in the vicinity of $\det(Q) = 0$, and working by comparison with 1D process (36) whose semigroup is readily seen to have the desired property at the origin. The second is to check by a simple computation that $\frac{\partial}{\partial Q_{ij}} f(Q)|_{\det(Q)=0} = 0$ for all i, j , noting that $f(Q)|_{\det(Q)=0} = 0$ is obvious. ■

2.1.2 Via the eigenvalue law

One could alternately argue that, since the law of Q_t defined in (35) is invariant under rotations by the orthogonal group, it is enough to consider the motion of the eigenvalues. More convenient still, is to identify the Wishart law $\gamma_{2\mu}$ itself by considering instead the eigenvalues of $P_t = Q_t^{-1}$.

We have that

$$dP_t = -P_t^2 dt + (1 + 2\mu) P_t dt + \text{tr} P_t I dt - dB_t^T P_t - P_t dB_t, \tag{39}$$

the solution of which, by similar reasoning as above, also remains in \mathcal{P} for all time after starting from any point in the interior. Further, $p_r \geq p_{r-1} \geq \dots \geq p_1 \geq 0$, the ordered eigenvalues perform the joint diffusion,

$$dp_i = -p_i^2 dt + (2\mu + 2) p_i dt + p_i \sum_{j \neq i} \frac{p_i + p_j}{p_i - p_j} dt + 2 p_i db_i, \tag{40}$$

with $\{b_i\}_{i=1, \dots, r}$ independent standard Brownian motions. The above can be derived from (39) by computing the Itô differential of the corresponding spectral representation. We will do a sample of such a calculation below in a slightly more complicated context. It is by now standard that system (40) possesses a strong solution, that the paths $p_i = p_i(t)$ do not intersect for $t > 0$, and any initial condition P_0 with some $p_i(0) = p_{i+1}(0)$ is an entrance point, see for example [1, §4].

The action of the corresponding generator G_p can be expressed in the form

$$\begin{aligned}
 G_p f &= \sum_{i=1}^r 2p_i^2 \partial_i^2 f + \sum_i \left(-p_i^2 + (2\mu + 2)p_i + p_i \sum_{j \neq i} \frac{p_i + p_j}{p_i - p_j} \right) \partial_i f \\
 &= \sum_{i=1}^r \partial_i (\phi(p_i) \partial_i f) - (\phi(p_i) \partial_i V) \partial_i f,
 \end{aligned} \tag{41}$$

where

$$\phi(p_i) = 2p_i^2, \quad V(p) = - \sum_{i>j} \log(p_i - p_j) + \frac{1}{2} \sum_{i=1}^r p_i - \left(\mu - \frac{r+1}{2} \right) \sum_{i=1}^r \log p_i, \tag{42}$$

restricted to the Weyl chamber $\Sigma_r = \{p_r \geq \dots \geq p_1 \geq p_0 = 0\} \subset \mathbb{R}^r$. Now, $e^{-V(p)} \mathbf{1}_{\Sigma_r}(p)$ is recognized as (after a suitable normalization) the joint density of eigenvalues for the real Wishart ensemble $\gamma_{2\mu}$, and the form of G_p in line two of (41) is particularly suited for verifying that $G_p^*(e^{-V(p)}) = 0$.

To identify $e^{-V(p)}$ as the invariant measure, we have only to deal with the same integration by parts issue that came up when working with matrix coordinates. In this case, Appendix A of [15] explains why for smooth h of compact support in Σ_r we have that $x \mapsto E_x[h(p_1(t), \dots, p_r(t))]$ along with its normal derivative are bounded at the seams $p_{i+1} = p_i$ (or boundary of Σ_r). After that, a quick calculation yields $\lim_{p_{i+1} \rightarrow p_i} \frac{\partial}{\partial(p_{i+1} + p_i)} e^{-V(p)} = 0$.

Note that using the gradient form of G_p in line two of (41) one can also check that the process p is reversible with respect to the found invariant measure. Using the rotational invariance of our matrix-valued processes one gets the reversibility of the process Q_t with respect to the invariant measure $\gamma_{2\mu}^{-1}$.

2.1.3 Proof of Corollary 10

For the complex and quaternion cases, it is a bit more constructive to go through eigenvalue/eigenvector coordinates. The starting point remains the same: a matrix diffusion $t \mapsto Q_t$ is constructed which has the desired distribution as its invariant measure (assuming the latter exists and is unique).

Corollary 10 considers two setting. The analogs of (35) are

$$dQ_t = Idt + \left(\frac{2}{\beta} - 1 - 2\mu \right) Q_t dt + \text{tr} Q_t dt + dB_t Q_t + Q_t dB_t^\dagger, \tag{43}$$

for the $U_r(\mathbb{F})$ -invariant noise $t \mapsto B_t$, and

$$dQ_t = Idt + \left(\frac{1}{\beta} - 2\mu\right) Q_t dt + \left(\frac{1}{\beta} - 1\right) \text{diag}(Q_t) dt + \text{tr} Q_t dt + dW_t Q_t + Q_t dW_t^\dagger, \quad (44)$$

for the case of the particular structured $\beta = 1, 2, 4$ noise $t \mapsto W_t$ appearing in the original spiked random matrix problem. Here, $\text{diag}(Q_t)$ is the diagonal matrix with the same diagonal as Q_t . This corresponding term in the sde shows up because of $dW_t Q_t dW_t^\dagger = (\frac{1}{\beta} - 1)\text{diag}(Q_t) dt + \text{tr} Q_t dt$.

The key observation is the following.

Proposition 13. In either setting (43) or (44), the eigenvalues of $P_t = Q_t^{-1}$ are Markovian with common sde

$$dp_i = -p_i^2 dt + \left(2\mu + \frac{2}{\beta}\right) p_i dt + p_i \sum_{j \neq i} \frac{p_i + p_j}{p_i - p_j} dt + \frac{2}{\sqrt{\beta}} p_i db_i, \quad (45)$$

for $\beta = 1, 2$, or 4 . □

Compare (40), noting the overlap at $\beta = 1$. That the eigenvalues of (43) are Markov is self-evident. For (44), we see it only by going through the calculation, which we defer to the end of the proof. Note that one can use (45) after that fact to see that P_t (and so Q_t) remains in \mathcal{P} for all time.

The upshot is that, for the eigenvalue motion(s), the argument is now precisely the same as in the $\beta = 1$ case. The corresponding $\beta = 2$ or 4 generator $G_{\beta,p}$ has the same form as (41), with (42) replaced by

$$\phi(p_i) = \frac{2}{\beta} p_i^2, \quad V(p) = -\beta \sum_{i>j} \log(p_i - p_j) + \frac{\beta}{2} \sum_{i=1}^r p_i - \beta \left(\mu - \frac{r + 2/\beta - 1}{2}\right) \sum_{i=1}^r \log p_i.$$

It follows that $G_{\beta,p}^*(e^{-V(p)}) = 0$. And, as it has to be, $e^{-V(p)} \mathbf{1}_{\Sigma_r}(p)$ is proportional to the complex/quaternion Wishart eigenvalue density. For the isotropic setting, one then has the full \mathbb{F} -Wishart law, as the eigenvector process of (43) clearly has the Haar measure on $U_r(\mathbb{F})$ as its unique invariant measure.

Proof of Proposition 13. Itô’s formula shows that if Q solves (43) or (44) then P solves the sde analog to (39), with \dagger in place of the transpose and an extra $(\frac{1}{\beta} - 1)\text{diag}(P) dt$ term in the second case.

At this point, we make two simplifications. For clarity, we carry out the computation for $\beta = 2$ only (that the $\beta = 4$ case will go through in the same way will be clear).

Also, we consider the simplified matrix sde:

$$dP_t = F(P_t)dt + dW_t P_t + P_t dW_t^\dagger, \quad F(P) = -\frac{1}{2} \text{diag}(P), \tag{46}$$

where again W_t has independent real Brownian motions with variance $\frac{1}{2}t$ on the diagonal and independent unit complex Brownian motions elsewhere. The point is that (46) retains everything “nonisotropic” in (44). That the corresponding isotropic case (with B_t replacing W_t and no $\text{diag}(\cdot)$ term in the drift) produces the same answer will also become clear in the course of the proof.

Either way, the strategy is standard. Write $P_t = U_t^\dagger \Lambda_t U_t$ for Λ_t the diagonal matrix of eigenvalues $(\lambda_{1,t}, \dots, \lambda_{r,t})$ and a unitary matrix U_t . Also, introduce the notation

$$W_t = U_t W_t U_t^\dagger, \quad dU_t U_t^\dagger = d\Gamma_t + dG_t. \tag{47}$$

Note that the former is that it is not simply a copy of W_t . The latter is the Doob–Meyer decomposition, with Γ_t a local martingale and G_t of finite variation. Since $d(U_t U_t^\dagger) = 0$, one finds that

$$d\Gamma_t^\dagger = -d\Gamma_t, \quad \text{and} \quad dG_t + dG_t^\dagger = -d\Gamma_t d\Gamma_t^\dagger, \tag{48}$$

in particular, $d\Gamma_{ii} = 0$. With this in hand, an application of Itô’s formula produces

$$\begin{aligned} d\Lambda &= dW\Lambda + \Lambda dW^\dagger + UF(U^\dagger \Lambda U)U^\dagger dt + (d\Gamma\Lambda + \Lambda d\Gamma^\dagger) + (dG\Lambda + \Lambda dG^\dagger) \\ &\quad + d\Gamma\Lambda d\Gamma^\dagger + d\Gamma dW\Lambda + d\Gamma\Lambda dW^\dagger + dW\Lambda d\Gamma^\dagger + \Lambda dW^\dagger d\Gamma^\dagger. \end{aligned} \tag{49}$$

As the martingale part of the right-hand side must vanish off the diagonal, we infer that

$$d\Gamma_{ij} = \frac{\lambda_j dW_{ij} + \lambda_i d\bar{W}_{ji}}{\lambda_i - \lambda_j}. \tag{50}$$

for $i \neq j$. Next, we write out (49) on the diagonal:

$$\begin{aligned} d\lambda_i &= \lambda_i (dW_{ii} + d\bar{W}_{ii}) + \sum_j u_{ij} F_{jj}(U^\dagger \Lambda U) \bar{u}_{ij} dt - \sum_{j \neq i} (\lambda_i - \lambda_j) d\Gamma_{ij} d\bar{\Gamma}_{ij} \\ &\quad + \sum_{j \neq i} (\lambda_i dW_{ji} + \lambda_j d\bar{W}_{ij}) d\Gamma_{ij} + \sum_{j \neq i} (\lambda_i d\bar{W}_{ji} + \lambda_j dW_{ij}) d\bar{\Gamma}_{ij}, \end{aligned} \tag{51}$$

having used (48). We also record that

$$\begin{aligned} \sum_j \mathbf{u}_{ij} F_{jj}(U^\dagger \Lambda U) \bar{\mathbf{u}}_{ij} &= -\frac{1}{2} \sum_{j=1}^r \sum_{\ell=1}^r \mathbf{u}_{ij} \bar{\mathbf{u}}_{\ell j} \lambda_\ell \mathbf{u}_{\ell j} \bar{\mathbf{u}}_{ij} = -\frac{1}{2} \sum_{j=1}^r \sum_{\ell=1}^r \lambda_\ell |\mathbf{u}_{ij}|^2 |\mathbf{u}_{\ell j}|^2, \\ &= -\frac{1}{2} \lambda_i + \frac{1}{2} \sum_{j \neq i} (\lambda_i - \lambda_j) \sum_{\ell=1}^r |\mathbf{u}_{i\ell}|^2 |\mathbf{u}_{j\ell}|^2, \end{aligned} \tag{52}$$

where u_{ab} are the entires of U .

To finish, first note that $(W_{ii} + \bar{W}_{ii}, i = 1, \dots, r) \stackrel{\text{(law)}}{=} (\sqrt{2}b_i, i = 1, \dots, r)$ for independent standard real Brownian motions b_i . Next, one may check that for $i \neq j$

$$dW_{ij} d\bar{W}_{ij} = dt - \frac{1}{2} \sum_{\ell=1}^r |\mathbf{u}_{i\ell}|^2 |\mathbf{u}_{j\ell}|^2 dt, \quad dW_{ij} dW_{ji} = \frac{1}{2} \sum_{\ell=1}^r |\mathbf{u}_{i\ell}|^2 |\mathbf{u}_{j\ell}|^2 dt. \tag{53}$$

These rules may then be used along with (50) in (51) to find that the contribution of the final two lines there is equal to

$$\begin{aligned} &\sum_{j \neq i} (\lambda_i - \lambda_j)^{-1} (\lambda_i^2 |dW_{ji}|^2 + \lambda_j^2 |dW_{ij}|^2 + \lambda_i \lambda_j (dW_{ij} dW_{ji} + d\bar{W}_{ij} d\bar{W}_{ji})) \\ &= \sum_{j \neq i} \frac{\lambda_i^2 + \lambda_j^2}{\lambda_i - \lambda_j} - \frac{1}{2} \sum_{j \neq i} (\lambda_i - \lambda_j) \sum_{\ell=1}^r |\mathbf{u}_{i\ell}|^2 |\mathbf{u}_{j\ell}|^2. \end{aligned} \tag{54}$$

The last term here now cancels with the last term in (52). This produces the system $d\lambda_i = \sqrt{2}\lambda_i db_i - \frac{1}{2}\lambda_i dt + \sum_{j \neq i} \frac{\lambda_i^2 + \lambda_j^2}{\lambda_i - \lambda_j} dt$. Now putting back the additional drift terms $(\text{tr}P - P^2 + (\frac{1}{2} - 2\mu)P)$ will produce (45) and complete the proof.

Note also that in the case of the isotropic noise the cross variations in (53) simplify as dt and 0 respectively, and thus in (54), only the term $\sum_{j \neq i} \frac{\lambda_i^2 + \lambda_j^2}{\lambda_i - \lambda_j}$ remains (as it should). ■

2.2 Feynman–Kac proof

We prove Theorem 2.

Taking M_t with the convergent choice of drift, as in (32) and again with $2\mu > r - 1$, the task is to show that: with $Y_t = M_t M_t^T$,

$$\mathbb{E}[e^{-\frac{1}{2} \int_0^\infty Y_t dt} | Y_0 = Y] = \frac{K_r(-\mu | Y, I)}{2^{\mu-1} \Gamma_r(\mu)}, \tag{55}$$

via the partial differential equation that characterizes the left-hand side.

This relies on the Markov property of $t \mapsto Y_t \in \mathcal{P}$. Applying Itô's formula gives

$$dY_t = M_t(dB_t + dB_t^T)M_t^T + (r + 1 - 2\mu)Y_t dt, \tag{56}$$

which does not appear to close. But, a check of the matrix entry covariances produces the generator

$$G_Y = \sum_{i \leq j} \sum_{k \leq \ell} (Y_{ik}Y_{j\ell} + Y_{i\ell}Y_{jk}) \frac{\partial^2}{\partial Y_{ij} \partial Y_{k\ell}} + (r + 1 - 2\mu) \sum_{i \leq j} Y_{ij} \frac{\partial}{\partial Y_{ij}},$$

which can then be put into the abbreviated form (10): $G_Y = 2\text{tr}(Y \frac{\partial}{\partial Y})^2 - 2\mu \text{tr}(Y \frac{\partial}{\partial Y})$. (One can also argue that by the rotation invariance of dB_t each appearance of M_t in (56) can be replaced by a positive square root $\sqrt{Y_t}$.)

The standard martingale argument used to derive the Feynman–Kac formula shows that the left-hand side $:= U(Y)$ of (55) solves $(G_Y - \frac{1}{2}\text{tr}Y)U(Y) = 0$. We have $U(0) = 1$ since $Y_t = 0$ for all time if $Y_0 = 0$. As for uniqueness, we claim that if $\tilde{U}(Y)$ is bounded and satisfies $(G_Y - \frac{1}{2}\text{tr}Y)\tilde{U}(Y) = 0$ and $\tilde{U}(0) = 0$, then $\tilde{U}(Y)$ is identically zero. Consider the bounded martingale $S_t^Y = \tilde{U}(Y_t)e^{-\frac{1}{2}\int_0^t \text{tr}(Y_s) ds}$. By Lemma 11, the matrix norm of $Y_t \rightarrow 0$ almost surely as $t \rightarrow \infty$. Hence, S_t^Y also converges to zero as $t \rightarrow \infty$. But, then $S_t^Y \equiv 0$ as S_t^Y is a regular martingale, and so $\tilde{U}(Y) \equiv 0$.

We are left to check that the K -Bessel function on the right-hand side of (55) satisfies the PDE. First, note

$$e^{\frac{1}{2}\text{tr}(xy)}(2G_Y - \text{tr}(y))e^{-\frac{1}{2}\text{tr}(xy)} = \text{tr}(xyxy) + (2\mu - r - 1)\text{tr}(xy) - \text{tr}(y),$$

where from here on we revert to lower case for matrix variables. Differentiating under the integral defining $K_r(-\mu|y, I)$, we are left to show that

$$0 = \int_{\mathcal{P}} (\text{tr}(xyxy) + (2\mu - r - 1)\text{tr}(xy) - \text{tr}(y)) \det(x)^{-\mu} e^{-\frac{1}{2}\text{tr}(xy+x^{-1})} d\mu_r(x). \tag{57}$$

Here, we have introduced the invariant measure on \mathcal{P} :

$$d\mu_r(x) = (\det x)^{-\frac{r+1}{2}} dx, \tag{58}$$

so called because

$$z = t^T x t \rightsquigarrow d\mu_r(z) = d\mu_r(x), \quad z = x^{-1} \rightsquigarrow d\mu_r(z) = d\mu_r(x), \tag{59}$$

with any invertible t in the first relation, see [34]. Unfortunately, integrating back up inside the integral (57) appears cumbersome. Instead, we rely on the uniqueness of the associated Mellin transform.

Proposition 14. [See Theorem 1 of [34, §4.3.1]] Introduce the power function

$$p_s(y) = \prod_{k=1}^r \det(y_k)^{s_k}, \quad s = (s_1, \dots, s_r) \in \mathbb{C}^r, \tag{60}$$

where $y_k \in \mathcal{P}_k$ (the $k \times k$ positive definite matrices) is the k th minor of $y \in \mathcal{P} = \mathcal{P}_r$. Then, the Mellin transform,

$$\hat{h}(s) := \int_{\mathcal{P}} p_s(y) h(y) \mu_r(dy),$$

defines an invertible map on the subspace of rotation invariant functions—those h for which $h(x) = h(O^T x O)$ for orthogonal O – of $L^2(\mathcal{P}, \mu_r)$. \square

The necessary isometry is actually stated in the cited theorem for the more general Helgason–Fourier transform on \mathcal{P} , but this reduces to the above Mellin transform on rotation invariant functions. The right-hand side of (57) is clearly rotation invariant in y since it is a function of the eigenvalues of y alone.

Setting

$$V(y) = \int_{\mathcal{P}} \det(x)^{-\mu} e^{-\frac{1}{2} \text{tr}(xy+x^{-1})} d\mu_r(x) = \int_{\mathcal{P}} \det(x)^{\mu} e^{-\frac{1}{2} \text{tr}(x^{-1}y+x)} d\mu_r(x), \tag{61}$$

that is, $V(y) = 2K_r(-\mu|y, I) = 2K_r(\mu|I, y)$, we compute first $\hat{V}(s)$ and then track the multipliers to this result produced after multiplying by, or integrating against, the additional factors $\text{tr}(y)$, $\text{tr}(\cdot y)$, and $\text{tr}(\cdot y \cdot y)$.

Step 1: this formula can be found in [34], but it guides the later computations, so we record it here. The trick is to introduce the change of variables $x = t^T t$ for t upper triangular with $t_{i,j} \in \mathbb{R}$, $t_{ii} \in \mathbb{R}_+$. This results in the rules

$$d\mu_r(x) = 2^r \prod_{i=1}^r t_{ii}^{-i} \prod_{i \leq j} dt_{ij}, \quad p_s(x) = \prod_{j=1}^r t_{jj}^{r_j}, \quad \text{for } r_j = 2(s_j + \dots + s_n). \tag{62}$$

Using two such changes of variables $y = t^T t$ and $x = q^T q$ produces

$$\begin{aligned} 2^{-2r} \hat{V}(s) &= \iint p_s(t^T t) \det(q^T q)^{\mu} e^{-\frac{1}{2} \text{tr} q^T q - \frac{1}{2} \text{tr} q^{-T} t^T t q^{-1}} \prod_j t_{jj}^{-j} q_{jj}^{-j} dt dq \\ &= \iint p_s(t^T t) p_s(q^T q) \det(q^T q)^{\mu} e^{-\frac{1}{2} \text{tr} q^T q - \frac{1}{2} \text{tr} t^T t} \prod_j t_{jj}^{-j} q_{jj}^{-j} dt dq \\ &= \iint \prod_{j=1}^r t_{jj}^{r_j-j} q_{jj}^{r_j+2\mu-j} e^{-\frac{1}{2} \sum_{i \leq j} (q_{ii}^2 + t_{ii}^2)} dt dq. \end{aligned} \tag{63}$$

In going from line one to line two, we replaced $y = t^T t$ with $q^T y q$, used the first invariance in (59), and then also the fact $p_s(q^T y q) = p_s(y) p_s(q^T q)$, see Proposition 1 of [34, §4.2.1]. The final line can be computed explicitly (each diagonal component producing a gamma function, the Gaussian integral over each off diagonals producing a factor of 2π), but for what we do here it is better to leave the answer in this form.

Step 2: Denote $V_1(y) = \text{tr}(y)V(y)$. By comparison to the last step, we easily find that

$$2^{-2r} \hat{V}_1(s) = \int \int \text{tr}(t^T t q q^T) \prod_{j=1}^r t_{jj}^{r_j-j} q_{jj}^{r_j+2\mu-j} e^{-\frac{1}{2} \sum_{i \leq j} (q_{ij}^2 + t_{ij}^2)} dq dt. \tag{64}$$

Now, expanding out, $\text{tr}(t^T t q q^T) = \sum_{k \leq i, j \leq \ell} t_{ki} t_{kj} q_{i\ell} q_{j\ell}$, we see that any term with $i < j$ will vanish by producing a factor of the form $\int u e^{-\frac{1}{2} u^2} du = 0$. Hence, we can replace the trace inside the integral (64) with $\sum_{k \leq i \leq \ell} t_{ki}^2 q_{i\ell}^2$. Each integral in the resulting sum (over $k \leq i \leq \ell$) can be reduced to (63) after an integration by parts (or two), yielding a different multiplicative factor of \hat{V} for different pairings of indices ($k < i < \ell$ versus $k = i = \ell$, for example). Those multipliers are as follows.

Multiplier:	1,	for	$k < i < \ell,$
	$r_i + 2\mu - i + 1,$		$k < i = \ell,$
	$r - i + 1,$		$k = i < \ell,$
	$(r_i - i + 1)(r_i + 2\mu - i + 1),$		$k = i = \ell.$

Summing up, we find a total multiplier of

$$c_1 = \sum_{i=1}^r r_i^2 + (2\mu + r + 1) \sum_{i=1}^r r_i - 2 \sum_{i=1}^r i r_i, \tag{65}$$

that is, $\hat{V}_1 = c_1 \hat{V}$.

Step 3: Now, let $\hat{V}_2(s) = \int p_s(y) \int (\det x)^{-\mu} \text{tr}(xy) e^{-\frac{1}{2} \text{tr}(yx+x^{-1})} d\mu(x) d\mu(y)$, for which we have that

$$2^{-2r} \hat{V}_2(s) = \int \int \text{tr}(t^T t) \prod_{j=1}^r t_{jj}^{r_j-j} q_{jj}^{r_j+2\mu-j} e^{-\frac{1}{2} \sum_{i \leq j} (q_{ij}^2 + t_{ij}^2)} dq dt. \tag{66}$$

With $\text{tr}(t^T t) = \sum_{k \leq i, k \leq j} t_{ki} t_{kj}$, the considerations are even simpler than above. Comparing (66) to (63), there are just two different cases.

Multiplier:	1,	for	$i < j,$
	$r - i + 1,$		$i = j.$

Summing over all possible i, j , we find that $\hat{V}_2 = c_2 \hat{V}$ with

$$c_2 = \sum_{i=1}^r r_i. \tag{67}$$

Step 4: Finally, set $\hat{V}_3(s) = \int p_s(y) \int (\det x)^{-\mu} \text{tr}(xyxY) e^{-\frac{1}{2} \text{tr}(yx+x^{-1})} d\mu(x) d\mu(y)$, and write

$$2^{-2r} \hat{V}_3(s) = \iint \text{tr}(t^T t t^T t) \prod_{j=1}^r t_{jj}^{r_j-j} q_{jj}^{r_j+2\mu-j} e^{-\frac{1}{2} \sum_{i \leq j} (q_{ij}^2 + t_{ij}^2)} dq dt. \tag{68}$$

Now, the expansion is

$$\text{tr}(t^T t t^T t) = \sum_{k, \ell \leq i, j} t_{ki} t_{kj} t_{\ell i} t_{\ell j},$$

and similar to the \hat{V}_1 calculation, all terms corresponding to $k \neq \ell < i \neq j$ terms will vanish. The six remaining choices yield

Multiplier:	$2 \times 1,$	for	$k = \ell < i < j$ or $k < \ell < i = j,$
	$2 \times (r - i + 1),$		$k = \ell = i < j$ or $k < \ell = i = j,$
	$3,$		$k = \ell < i = j,$
	$(r_i - i + 3)(r_i - i + 1),$		$k = \ell = i = j.$

The additional factor of two in lines one and two count the ordering of (i, j) or (k, ℓ) . A little algebra shows that then $\hat{V}_3 = c_3 \hat{V}$ with

$$c_3 = \sum_{i=1}^r r_i^2 + 2(r + 1) \sum_{i=1}^r r_i - 2 \sum_{i=1}^r i r_i. \tag{69}$$

The proof of Theorem 2 is finished by checking that

$$c_3 + (2\mu - r - 1)c_2 - c_1 = 0,$$

recall (65) and (67).

To finish the proof of Theorem 2, we now return to the

Proof of Lemma 11. This calculation can basically be found in [26]. It would be nice to have a way to control at least the matrix norm as sharply without going to eigenvalue coordinates.

Staying in the setting just considered, we let $y_1 \leq y_2 \leq \dots \leq y_r$ be the eigenvalues of Y_t and show that $\lim_{t \rightarrow \infty} \frac{1}{t} \log y_i(t) = -2\mu + i - 1$ with probability one for each $i = 1, \dots, r$.

With $\gamma_i = \log y_i$, we find from (56) and considerations similar to those behind Proposition 13 that,

$$d\gamma_i = 2db_i - 2\mu dt + \sum_{j \neq i} \frac{e^{\gamma_i} + e^{\gamma_j}}{e^{\gamma_i} - e^{\gamma_j}} dt.$$

One checks that $\sum_{j \neq r} \frac{e^{\gamma_r} + e^{\gamma_j}}{e^{\gamma_r} - e^{\gamma_j}} \geq r - 1$ and $\sum_{j \neq 1} \frac{e^{\gamma_1} + e^{\gamma_j}}{e^{\gamma_1} - e^{\gamma_j}} \leq 1 - r$. Moreover, if we change i to $i + 1$, then the interaction term will change by at most $2 \frac{e^{\gamma_{i+1}} + e^{\gamma_i}}{e^{\gamma_{i+1}} - e^{\gamma_i}}$. Thus, $\gamma_{i+1} - \gamma_i$ is bounded above by the solution to

$$dz_i = 2d(b_{i+1} - b_i) + 2 \left(\frac{1 + e^{-z_i}}{1 - e^{-z_i}} \right) dt.$$

The proof is finished by remarking that $P(\lim_{t \rightarrow \infty} \frac{z_i}{t} = 2, i = 1, \dots, r - 1) = 1$. ■

3 Process-Level Identities

We prove Theorem 4 and the Burke property statements of Theorems 8 and 9.

3.1 $A_t^{(\mu)}$ and $A_t^{(-\mu)}$

Here, we again start by taking $M_t = M_t^{(-\mu)}$ defined by (32) for $t \geq 0$ with driving matrix Brownian motion $t \mapsto B_t$. Throughout this section, we drop the superscript on the corresponding additive functional, $A_t = A_t^{(-\mu)}$, and it is always assumed that $2\mu > r - 1$.

We need the following two facts.

Proposition 15. Denote by $\mathcal{B}_t = \sigma(B_s, s \leq t)$ and by $\hat{\mathcal{B}}_t$ the initial enlargement $\mathcal{B}_t \vee \sigma(A_\infty)$. Then,

$$\hat{B}_t := B_t - \int_0^t (2\mu I - M_s^T (A_\infty - A_s)^{-1} M_s) ds \tag{70}$$

is a standard matrix Brownian motion with respect to $\hat{\mathcal{B}}_t$ and is independent of A_∞ . □

Proposition 16. Almost surely,

$$(A_t^{-1} - A_\infty^{-1})^{-1} = \int_0^t N_s N_s^T ds, \quad \text{for } N_t := A_\infty (A_\infty - A_t)^{-1} M_t. \tag{71}$$

Furthermore, conditioned on the value of A_∞ , the process N_t satisfies

$$dN_t = N_t d\hat{B}_t + \left(\frac{1}{2} + \mu\right) N_t dt, \quad N_0 = I, \tag{72}$$

where \hat{B}_t is as defined in (70). □

Granted these propositions the result is immediate.

Proof of Theorem 4. By (72), $\int_0^t N_s N_s^T ds$ has the same distribution as $A_t^{(\mu)}$ as a process. Thus, conditioned on A_∞ , we have that $\{A_t^{-1} - A_\infty^{-1}, t \geq 0\}$ is equal in law to $\{(A_t^{(\mu)})^{-1}, t \geq 0\}$. But, then this is also true unconditionally, which is precisely the desired statement. Afterward, one can read (71) as an almost sure version of the identity (12). ■

Proof of Proposition 15. We start with the representation

$$A_\infty = A_t + M_t \left(\int_0^\infty M_t^{-1} M_{s+t} M_{s+t}^T M_t^{-T} ds \right) M_t^T,$$

noting that by the matrix Dufresne identity (Theorem 1) the integral on the right-hand side has the $\gamma_{2\mu}^{-1}$ distribution. Then, with f denoting the corresponding density function,

$$P(A_\infty \in dA | \mathcal{B}_t) = f(M_t^{-1}(A - A_t)M_t^{-T})(\det M_t)^{-(r+1)} dA. \tag{73}$$

Here, we have used that the Jacobian of the map $A \mapsto MAM^T$ on symmetric matrices is given by $(\det M)^{r+1}$, see for example [17, Lemma 2.2].

Next, by the tower property of conditional expectations, the Radon–Nikodym derivative

$$\begin{aligned} R_{t,A} &= \frac{dP(A_\infty \in dA | \mathcal{B}_t)}{dP(A_\infty \in dA)} \\ &= c_A \det(A - A_t)^{-\mu - \frac{r+1}{2}} \det(M_t)^{2\mu} \exp \left[-\frac{1}{2} \text{tr}((A - A_t)^{-1} M_t M_t^T) \right] \end{aligned} \tag{74}$$

is a \mathcal{B}_t -martingale. The formula in line two follows from writing out the appearance of $f(M_t^{-1}(A - A_t)M_t^{-T})$ in (73). The factor c_A is then the density $f(A)$ without the normalizing constant. As noted in (30), $t \mapsto \det M_t^{(-\mu)}$ is the geometric Brownian motion $\det M_t = e^{\text{tr} B_t - \mu r t}$. Substituting this in the preceding formula, we record

$$dR_{t,A} = R_{t,A} (2\mu \text{tr} dB_t - \text{tr}[dB_t M_t^T (A - A_t)^{-1} M_t]) \tag{75}$$

for later use.

Next, we will show that for any bounded continuous test function h and event $\Lambda_s \in \mathcal{B}_s, s < t$:

$$E \left[\mathbf{1}_{\Lambda_s} h(A_\infty) (B_t - B_s) \right] = E \left[\mathbf{1}_{\Lambda_s} h(A_\infty) \int_s^t (2\mu - M_u^T (A_\infty - A_u)^{-1} A_u) du \right]. \tag{76}$$

Granted (76) the monotone class theorem will imply that

$$E[B_t - B_s | \hat{\mathcal{B}}_s] = \int_s^t (2\mu - M_u^T(A_\infty - A_u)^{-1} M_u) du,$$

or in other words that \hat{B}_t defined in (70) is a local martingale (with respect to the filtration $\hat{\mathcal{B}}_t$). To see that it is actually a matrix Brownian motion, and so complete the proof, one checks the quadratic covariation of the entries and invokes Lévy’s theorem.

Returning now to (76), we introduce $\lambda_t(h) := E[h(A_\infty) | \mathcal{B}_t]$ and write the left-hand side of that equality as in

$$E[\mathbf{1}_{A_s} h(A_\infty)(B_t - B_s)] = E[\mathbf{1}_{A_s} (\lambda_t(h) B_t - \lambda_s(h) B_s)],$$

as follows by conditioning separately with respect to both \mathcal{B}_t and then \mathcal{B}_s . By (74), we have that,

$$\begin{aligned} \lambda_t(h) &= \int_{\mathcal{P}} h(A) f(M_t^{-1}(A - A_t)[M_t]^T) (\det M_t)^{-2} dA \\ &= \int_{\mathcal{P}} h(A) R_{t,A} f(A) dA, \end{aligned}$$

where again f is the $\gamma_{r,2\mu}^{-1}$ density function. Then, (75) implies that

$$\lambda_t(h) - \lambda_s(h) = \int_{\mathcal{P}} \int_s^t h(A) R_{A,u} (2\mu \text{tr} dB_u - \text{tr}[dB_u M_u^T (A - A_u)^{-1} M_u]) f(A) dA.$$

To continue, we compute the cross variation of $\lambda_t(h)$ and $[B_t]_{i,j}$, for which we just need to check the coefficient of $dB_{i,j}$ in the previous integral with the result that

$$\langle \lambda_t(h), [B_t]_{ij} \rangle = \int_{\mathcal{P}} h(A) R_{A,u} (2\mu \mathbf{1}_{i=j} - [M_u^T (A - A_u)^{-1} M_u]_{i,j}) f(A) dA.$$

Then, writing $B_t - B_s = \int_s^T dB_u$ leads to

$$\begin{aligned} E[(\lambda_t(h) B_t - \lambda_s(h) B_s) | \mathcal{B}_s] &= \int_{\mathcal{P}} \int_s^t h(A) R_{A,u} (2\mu I - M_u^T (A - A_u)^{-1} M_u) du f(A) dA \\ &= \int_s^t E[h(A_\infty) (2\mu I - M_u^T (A_\infty - A_u)^{-1} M_u) | \mathcal{B}_u] du. \end{aligned}$$

From here, we see that $E[\mathbf{1}_{A_s} (\lambda_t(h) B_t - \lambda_s(h) B_s)] = E[\mathbf{1}_{A_s} E[(\lambda_t(h) B_t - \lambda_s(h) B_s) | \mathcal{B}_s]]$ is equal to the right-hand side of (76), as required. ■

Proof of Proposition 16. Since $t \mapsto A_t^{-1}$ is almost surely once differentiable and $\|A_t - tI\| = o(t)$ as $t \rightarrow 0$ (as $M_0 = I$ and M_t is continuous), we have that

$$(A_t^{-1} - A_\infty^{-1})^{-1} = \int_0^t d[(A_s^{-1} - A_\infty^{-1})^{-1}], \tag{77}$$

also almost surely. On the other hand,

$$\begin{aligned} d[(A_t^{-1} - A_\infty^{-1})^{-1}] &= dA_t(A_\infty - A_t)^{-1}A_\infty + A_t(A_\infty - A_t)^{-1}dA_t(A_\infty - A_t)^{-1}A_\infty \\ &= M_tM_t^T(A_\infty - A_t)^{-1}A_\infty dt + A_t(A_\infty - A_t)^{-1}M_tM_t^T(A_\infty - A_t)^{-1}A_\infty dt \\ &= A_\infty(A_\infty - A_t)^{-1}M_tM_t^T(A_\infty - A_t)^{-1}A_\infty dt. \end{aligned}$$

Here, we have used the matrix identity $(C^{-1} - D^{-1})^{-1} = C(D - C)^{-1}D$ in line one, and $I + C(D - C)^{-1} = D(D - C)^{-1}$ to go from line two to line three. Substituting back into (77) yields (71), identifying $N_t = A_\infty(A_\infty - A_t)^{-1}M_t$ at the same time.

To see (72), note that $N_0 = A_\infty A_\infty^{-1}M_0 = I$ and compute,

$$\begin{aligned} d[A_\infty(A_\infty - A_t)^{-1}M_t] &= A_\infty(A_\infty - A_t)^{-1}M_tM_t^T(A_\infty - A_t)^{-1}M_t dt + A_\infty(A_\infty - A_t)^{-1}M_t(dB_t + (\tfrac{1}{2} - \mu)Idt) \\ &= N_t[M_t^T(A_\infty - A_t)^{-1}M_t dt + dB_t + (\tfrac{1}{2} - \mu)Idt]. \end{aligned}$$

By Proposition 15, the quantity inside the final bracket equals $d\hat{B}_t + (\tfrac{1}{2} + \mu)Idt$, as desired. ■

3.2 Burke properties

Both results are consequences of Propositions 15 and 16. Let N_t be defined as in the latter statement and introduce the temporary shorthand $\hat{A}_t = \int_0^t N_s N_s^T ds$, reserving A_t for $\int_0^t M_s M_s^T ds$. Importantly, M_t is also chosen as throughout the proofs of those same propositions. In particular, the underlying matrix Brownian motions \hat{B}_t and B_t stand in the same relationship.

Proof of Theorem 8. We start by rewriting the almost sure identity (71) as in

$$\hat{A}_t^{-1} = A_t^{-1} - A_\infty^{-1}.$$

Repeated use of this along with the resolvent identity then produces

$$\begin{aligned}
 A_\infty (A_\infty - A_t)^{-1} &= \left(I - \left(\hat{A}_t^{-1} + A_\infty^{-1} \right)^{-1} A_\infty^{-1} \right)^{-1} \\
 &= \left(I - \left(I + A_\infty \hat{A}_t^{-1} \right)^{-1} \right)^{-1} \\
 &= \hat{A}_t A_\infty^{-1} \left(I + A_\infty \hat{A}_t^{-1} \right) \\
 &= \left(A_\infty + \hat{A}_t \right) A_\infty^{-1}.
 \end{aligned} \tag{78}$$

The important observations are: the left-hand side of (78) times M_t (from the right) is the definition of N_t , and, on the right-hand side, we have that A_∞ is independent of \hat{A}_t (on account of being independent of \hat{B}_t and so N_t).

To exploit the second point, we extend N_t to $t \in (-\infty, 0)$ as in Section 2. By the same reasoning used in the proof of Theorem 1, $\int_{-\infty}^0 N_s N_s^T ds$ is an independent copy of A_∞ .

Now, modifying the introduced notation to let $\hat{A}_{(-\infty, t)}$ denote $\int_{-\infty}^t N_s N_s^T ds$, we conclude from the above comments and (78) that

$$\hat{A}_{(-\infty, 0)} \left(\hat{A}_{(-\infty, t)} \right)^{-1} N_t \stackrel{\text{(law)}}{=} M_t,$$

as processes for $t \geq 0$. In order to recognize this as equivalent to the identity (26) announced in Theorem 8, take inverse-transposes throughout the above to find that

$$\left(\hat{A}_{(-\infty, 0)} \right)^{-1} N_t \left(N_t^{-1} \hat{A}_{(-\infty, t)} N_t^{-T} \right) \stackrel{\text{(law)}}{=} M_t^{-T}. \tag{79}$$

An application of Itô’s formula will then show that M_t^{-1} is a copy of N_t . In particular, it satisfies $dM_t^{-1} = M_t^{-1} d(-B_t^T) + (\frac{1}{2} + \mu) M_t^{-1} dt$, for $t \geq 0$.

Finally, the independence statement follows from the same trick used in [28] for the 1D case (see the proof of Theorem 6 there). Bringing in yet more notation, let L_t denote the left-hand side of (79). We will show that

$$N_t^{-1} \hat{A}_{(-\infty, t)} N_t^{-T} = L_t^T \left(\int_t^\infty L_s^{-T} L_s^{-1} ds \right) L_t, \tag{80}$$

with probability one. The independence of $\{L_s, s \leq t\}$ and $\{N_s^{-1} \hat{A}_{(-\infty, s)} N_s^{-T}, s > t\}$ being made clear by writing the right-hand side of (80) as $\int_t^\infty (L_t^{-1} L_s)^{-T} (L_t^{-1} L_s)^{-1} ds$. On account of (79) the process L_t inherits the independence of multiplicative increments from M_t .

To verify (80) note that

$$\begin{aligned} L_t^T \left(\int_t^\infty L_s^{-T} L_s^{-1} ds \right) L_t &= N_t^{-1} \hat{A}_{(-\infty,t)} \left(\int_t^\infty \hat{A}_{(-\infty,s)}^{-1} N_s N_s^T \hat{A}_{(-\infty,s)}^{-1} ds \right) \hat{A}_{(-\infty,t)} N_t^{-T} \\ &= N_t^{-1} \hat{A}_{(-\infty,t)} \left(- \int_t^\infty d(\hat{A}_{(-\infty,s)}^{-1}) \right) \hat{A}_{(-\infty,t)} N_t^{-T}. \end{aligned}$$

The proof is finished upon integrating and using the fact that $\|\hat{A}_{(-\infty,t)}^{-1}\| \rightarrow 0$ as $t \rightarrow \infty$ with probability one (which follows by the computation behind Lemma 11). ■

As for Theorem 9, the identity (78) together with the definition of N_t from (71) gives that

$$M_s^T (A_\infty - A_s)^{-1} M_s = N_s^T (A_\infty + \hat{A}_s)^{-1} N_s = N_s^T \hat{A}_{(-\infty,s)}^{-1} N_s, \tag{81}$$

where we continue using the notation introduced in the previous proof. One immediate consequence is another proof for the reversibility of the process X_t with respect to its stationary distribution. The process $(N_s^T \hat{A}_{(-\infty,s)}^{-1} N_s, s \geq 0)$ is the stationary version of X_s (with parameter μ), whereas the process $(M_s^T (A_\infty - A_s)^{-1} M_s, s \geq 0)$ is distributed as the time-reversed version of the same process.

Another consequence of (81) is that Proposition 15 can be rewritten as in

$$B_t = \hat{B}_t + \int_0^t (2\mu I - N_s^T \hat{A}_{(-\infty,s)}^{-1} N_s) ds.$$

Since the right-hand side only depends on \hat{B}_t , this identity provides a nonlinear transformation producing one matrix Brownian motion from another. Reversing the roles of the Brownian motions and reverting to our original notation yields

Corollary 17. Let $2\mu > r - 1$ and now take $M_t = M_t^{(\mu)}$, extended to $t \in \mathbb{R}$ as described in Section 2. Denote the (two-sided) driving matrix Brownian motion for M_t by B_t . Then,

$$\hat{B}_t = B_t + 2\mu It - \int_0^t M_s^T \left(\int_{-\infty}^s M_u M_u^T du \right)^{-1} M_s ds \stackrel{\text{(law)}}{=} B_t$$

as processes for $t \geq 0$. □

From this, the proof of Theorem 9 is straightforward.

Proof of Theorem 9. The introduced process $H_{t/2}$ can be equated with M_t , though driven by the standard matrix Brownian motion $B_{t/2} + C_{t/2}$. Then, by the corollary, we

have that

$$\begin{aligned} \hat{B}_t &= B_{t/2} + C_{t/2} + 2\mu It - \int_0^t M_s^T \left(\int_{-\infty}^s M_u M_u^T du \right)^{-1} M_s ds \\ &= B_{t/2} + C_{t/2} + 2\mu It - \int_0^{t/2} H_s^T \left(\int_{-\infty}^s H_u H_u^T du \right)^{-1} H_s ds \end{aligned}$$

is another standard matrix Brownian motion. Checking the definitions, we see that $F_t = \frac{\hat{B}_{2t}}{2} + \frac{B_t - C_t}{2}$ and $G_t = \frac{\hat{B}_{2t}}{2} - \frac{B_t - C_t}{2}$. Since \hat{B}_t is constructed from $B_t + C_t$, it is independent (as a process) from the process $B_t - C_t$. But this means that $\{F_t, t \geq 0\}$ and $\{G_t, t \geq 0\}$ are independent of each other and they are both standard matrix Brownian motions. ■

4 The Matrix X_t and Z_t Processes

The process X_t is actually equivalent to the Q_t encountered in the proof of Theorem 1, and hence its SDE has already been recorded in (35). The fact that $\gamma_{2\mu}^{-1}$ is the stationary measure was proved in Proposition 12 in Subsection 2.1. The reversibility was shown in the same subsection at the end of the discussion via the eigenvalue law, in this section we will give another (more direct) proof.

As for Z_t , we will again rely in part on the technology developed in the last section. Recall M_t and N_t from Propositions 15 and 16 and define two versions of the Z_t process:

$$Z_t = M_t^{-1} \int_0^t M_s M_s^T ds, \quad \hat{Z}_t = N_t^{-1} \int_0^t N_s N_s^T ds. \tag{82}$$

That is, Z_t corresponds to $-\mu$ and is driven by B_t , \hat{Z}_t to $+\mu$ and \hat{B}_t , and B_t and \hat{B}_t are related by Proposition 15. (Note that in the discussion of Section 1.2, Z_t was defined through M_t with the $+\mu$ drift term. The choice to flip things around here is natural given the course of the previous arguments.) Itô's formula yields

$$dZ_t = \left(\frac{1}{2} + \mu\right)Z_t dt + M_t^T dt - dB_t Z_t, \quad d\hat{Z}_t = \left(\frac{1}{2} - \mu\right)\hat{Z}_t dt + N_t^T dt - d\hat{B}_t \hat{Z}_t. \tag{83}$$

We first show how to close these equations using properties of the matrix GIG distribution (proving Theorem 5 and then Corollary 6). After that, we consider the asymptotics of the underlying eigenvalue processes (proving Theorem 7).

4.1 The role of the generalized inverse Gaussian

The following rather surprising fact already implies the invariance in law of Z_t under the map $\mu \mapsto -\mu$.

Proposition 18. $Z_t = \hat{Z}_t$ almost surely for $t \geq 0$. □

Proof. From Proposition 16, we have that $\hat{Z}_t = N_t^{-1} \int_0^t N_s N_s^T ds = N_t^{-1} (A_t^{-1} - A_\infty^{-1})^{-1}$ as well as $N_t = A_\infty (A_\infty - A_t)^{-1} M_t$. Hence,

$$\begin{aligned} \hat{Z}_t &= (A_\infty (A_\infty - A_t)^{-1} M_t)^{-1} (A_t^{-1} - A_\infty^{-1})^{-1} \\ &= M_t^{-1} (A_\infty - A_t) A_\infty^{-1} (A_t^{-1} - A_\infty^{-1})^{-1} A_t^{-1} A_t \\ &= M_t^{-1} A_t, \end{aligned}$$

which is the definition of Z_t . ■

For the Markov property, we need the following.

Proposition 19. The conditional distribution of $M_t^T A_t^{-1} M_t$ given $\{Z_s, s \leq t, Z_t = Z\}$ is the matrix GIG law $\eta_{-\mu, I, (ZZ^T)^{-1}}$. The conditional distribution of $Z_t (A_t^{-1} - A_\infty^{-1}) Z_t^T$ given the same σ -field is $\eta_{-\mu, (ZZ^T)^{-1}, I}$. □

We go ahead with the consequences of Proposition 19, and return to its proof at the end of the section.

Proof of Theorem 5. We follow the strategy of Matsumoto and Yor [21], employing [18, Theorem 7.12] to close the equation for Z_t in (83) as follows. Taking conditional expectation throughout with respect to $\mathcal{Z}_t = \sigma(Z_s, s \leq t)$, the ideas there show that

$$dZ_t = \left(\frac{1}{2} + \mu\right) Z_t dt + E[M_t^T | Z_s, s \leq t] dt - dW_t Z_t, \tag{84}$$

where W_t is matrix Brownian motion adapted to $\mathcal{Z}_t \subsetneq \mathcal{B}_t = \sigma(B_s, s \leq t)$. Next, since $M_t^T A_t^{-1} M_t = M_t^T Z_t^{-1} \sim \eta_{-\mu, I, (Z_t Z_t^T)^{-1}}$ by Proposition 19, there is the formula $E[M_t^T | Z_s, s \leq t] = \kappa_{-\mu}(I, (Z_t Z_t^T)^{-1}) Z_t$. We can then rewrite the above as in

$$dZ_t = \left(\frac{1}{2} + \mu\right) Z_t dt + \kappa_{-\mu}(I, (Z_t Z_t^T)^{-1}) Z_t dt - dW_t Z_t, \tag{85}$$

and believing that one can let $\mu \mapsto -\mu$, Z_t is the stated diffusion (in its own filtration).

Note that above construction entails that

$$W_t = B_t + \int_0^t [M_s^T Z_s^{-1} - \left(\frac{1}{2} + \mu\right) - \kappa_{-\mu}(I, (Z_s Z_s^T)^{-1})] ds.$$

Similar to the case considered in [18], one checks that this is a matrix Brownian motion by using of Itô's formula to find an ordinary differential equation for $t \mapsto E[e^{\text{itr}(C W_t)}]$, $C \in \text{GL}_r$.

To verify that the formulas themselves hold up under the change of sign, consider \hat{Z}_t for which the above procedure leads to the following analog of (84):

$$d\hat{Z}_t = \left(\frac{1}{2} - \mu\right) \hat{Z}_t dt + E[N_t^T | \hat{Z}_s, s \leq t] dt - d\hat{W}_t \hat{Z}_t dt, \tag{86}$$

with a similar considerations for the new matrix Brownian motion \hat{W}_t . Proposition 19 again applies after noting that

$$N_t^T = M_t^T (A_\infty - A_t)^{-1} A_\infty = M_t^T A_t^{-1} A_t (A_\infty - A_t)^{-1} A_\infty = Z_t^{-T} (A_t^{-1} - A_\infty^{-1})^{-1}.$$

That is, $N_t^T Z_t^{-1} = (Z_t (A_t^{-1} - A_\infty^{-1}) Z_t^T)^{-1}$ which has law $\eta_{\mu, I, (Z_t Z_t^T)^{-1}}$ conditional on $\{Z_s, s \leq t\}$. Then, by Proposition 18, we also have that $N_t^T \hat{Z}_t^{-1}$ has this same law conditional on $\{\hat{Z}_s, s \leq t\}$. Substituting into (86) gives the desired sde, that is, (85) with a sign flip on the parameter μ . ■

Proof of Corollary 6. Proposition 2.1 of [10] provides a soft criteria for two processes X_t and Y_t defined on the same probability space to intertwine. With X_t taking values in E and Y_t taking values in F (possibly separate measure spaces), it is assumed that

- (i) X_t is Markovian with respect to a filtration \mathcal{F}_t , and Y_t is Markovian with respect to a filtration \mathcal{G}_t such that $\mathcal{G}_t \subset \mathcal{F}_t$,
- (ii) There exists a Markov kernel $\Lambda : E \mapsto F$ for which $E[h(X_t) | \mathcal{G}_t] = (\Lambda h)(Y_t)$ for all $t > 0$ and integrable $h : E \mapsto \mathbb{R}_+$.

Given this the outcome is that for any bounded continuous function $f : E \rightarrow \mathbb{R}$ and $s > 0$ we have $T_t^Y \Lambda f(Y_t) = \Lambda T_t^X f(Y_t)$ almost surely. From this, $T_t^Y \Lambda = \Lambda T_t^X$ follows under additional “mild continuity assumptions”.

From what we have just shown, the above applies directly to $X_t = N_t$ and $Y_t = \hat{Z}_t$. To go from the almost sure statement to the operator identity, it is enough to show that \hat{Z}_t possesses a smooth positive density function of full support on GL_r^+ . Note that $\hat{Z}_t = (N_t^{-1} \int_0^t N_s N_s^T ds N_t^{-T}) N_t^T$ and the joint distribution of $N_t^{-1} \int_0^t N_s N_s^T ds N_t^{-T}$ and N_t^T is the same as that of X_t and M_t . But (X_t, M_t) satisfies a linear sde system, and using the same hypoellipticity argument as the one at the end of the proof of Proposition 12 it follows that (X_t, M_t) has a positive joint density of full support on $\mathcal{P}_r \times GL_r^+$. Hence, $\hat{Z}_t \stackrel{(law)}{=} X_t M_t^{-T}$ has a positive density function of full support on GL_r^+ . (Note since the original statement takes the positive μ drift, it is consistent to consider here (N_t, \hat{Z}_t) rather than (M_t, Z_t) for which there is a corresponding result.)

For (i), N_t is Markovian with respect to $\hat{B}_t = \sigma(\hat{B}_s, s \leq t)$ and \hat{Z}_t is Markovian with respect to its own filtration \mathcal{Z}_t which the proof of Theorem 5 shows is contained strictly inside \hat{B}_t . And, (ii) is the second point of Proposition (19): $\hat{Z}_t^{-T} N_t \sim \eta_{\mu, I, (ZZ^T)^{-1}}$ and so,

$$E[h(N_t) | \hat{\mathcal{Z}}_t, \hat{Z}_t = Z] = \int_{GL_r} h(Z^T X) \eta(dX) := \Lambda h(Z),$$

where $\eta = \eta_{\mu, I, (ZZ^T)^{-1}}$,

For the second part, take $X_t = \hat{Z}_t N_t^{-T}$. Then, the representation of the intertwining kernel by $\Lambda h(Z) = \int_{\mathcal{P}} h(X^{-1}) \eta(dX)$, now viewed as from \mathcal{P} into GL_r , follows from $(N_t^T \hat{Z}_t^{-1})^{-1}$ having conditional law $\eta_{-\mu, I, (ZZ^T)^{-1}, I} = \eta^{-1}$. ■

We now return to the proof of Proposition 19. The first step, Lemma 20, is designed to implement Bernadac’s characterization of the matrix GIG [4, Theorem 5.1] to this end. The notation $\gamma_{p,A}$ introduced in the statement refers to the noncentral Wishart law, which has density proportional $(\det X)^{\frac{p-r-1}{2}} e^{-\frac{1}{2} \text{tr}(A^{-1}X)} \mathbf{1}_{\mathcal{P}}(X)$ for $A \in \mathcal{P}$. It is worth pointing out that the posited independence structure in the statement (of $(X + Y)^{-1}$ and $X^{-1} - (X + Y)^{-1}$ given that of X and Y) is now commonly referred to as the Matsumoto–Yor property. The proof of Proposition 19 is then completed by specifying our particular choice of X and Y in Lemma 21.

Lemma 20. Suppose that the \mathcal{P} -valued random variables X and Y are independent, and that $(X + Y)^{-1}$ and $X^{-1} - (X + Y)^{-1}$ are also independent. Suppose further that $Y \sim \gamma_{2p, A^{-1}}$ and $X^{-1} - (X + Y)^{-1} \sim \gamma_{2p, B^{-1}}$ for $A, B \in \mathcal{P}$. Then, $X \sim \eta_{-p, A, B}$ and $(X + Y)^{-1} \sim \eta_{-p, B, A}$. □

Proof. Set $U = (X + Y)^{-1}$ and $V = X^{-1} - (X + Y)^{-1}$. Let Y' be a copy of V (that is, $Y' \sim \gamma_{2p, B^{-1}}$) independent of both X and Y . Then,

$$X = (U + V)^{-1} \stackrel{(\text{law})}{=} (Y' + U)^{-1} = (Y' + (Y + X)^{-1})^{-1}.$$

Bernadac’s result is that the above distributional identity holds with X, Y , and Y' all independent, and Y and Y' having the corresponding Wishart distributions if and only if $X \sim \eta_{-p, A, B}$.

Alternatively, if we let V' be a copy of Y (so $V' \sim \gamma_{2p, A^{-1}}$) independent of U and V we will have that

$$U = (Y + (V + U)^{-1})^{-1} \stackrel{(\text{law})}{=} (V' + (V + U)^{-1})^{-1},$$

and the result is that $U = (X + Y)^{-1} \sim \eta_{-p, B, A}$. ■

Lemma 21. Set

$$X = M_t^T Z_t^{-1} = M_t^T A_t^{-1} M_t, \quad Y = M_t^T (A_\infty - A_t)^{-1} M_t.$$

Then,

$$(X + Y)^{-1} = Z_t (A_t^{-1} - A_\infty^{-1}) Z_t^T, \quad X^{-1} - (X + Y)^{-1} = Z_t A_\infty^{-1} Z_t^T.$$

Conditioned on the σ -field generated by $\{Z_s, s \leq t\}$ the random matrices X and Y are independent, as are $(X + Y)^{-1}$ and $X^{-1} - (X + Y)^{-1}$. Further, the conditional distribution of Y is $\gamma_{2\mu, A^{-1}}$, where $A = I$ and that of $X^{-1} - (X + Y)^{-1}$ is $\gamma_{2\mu, B^{-1}}$, where $B = (Z_t Z_t^T)^{-1}$. \square

Proof. The formulas for $(X + Y)^{-1}$ and $X^{-1} - (X + Y)^{-1}$ follow from some simple algebra.

Clearly, Y is independent of $\sigma(B_s, s \leq t)$, and thus is independent of both X and $\{Z_s, s \leq t\}$. For the independence of $(X + Y)^{-1}$ and $X^{-1} - (X + Y)^{-1}$, we need to prove the conditional independence of $A_t^{-1} - A_\infty^{-1}$ and A_∞ . By Proposition 16, we know that $A_t^{-1} - A_\infty^{-1} = \int_0^t N_s N_s^T ds$ is measurable $\sigma(\hat{B}_s, s \leq t)$ and so is independent of A_∞ . But, Z_t is in this σ -field as well ($Z_t = \hat{Z}_t$) implying there is also conditional independence. \blacksquare

4.2 Asymptotics

We start by identifying the underlying eigenvalue processes:

Lemma 22. Denote the (ordered, nonintersecting) eigenvalues of X_t by $0 \leq x_r \leq x_{r-1} \leq \dots \leq x_1$. These perform the joint diffusion:

$$dx_i = 2x_i db_i + \left(1 + (2 - 2\mu)x_i + x_i \sum_{j \neq i} \frac{x_i + x_j}{x_i - x_j} \right) dt. \tag{87}$$

For Z_t , consider instead the similarly ordered singular values $\{z_i\}$ of Z_t . This family is also Markov, and is governed by

$$dz_i = z_i db_i + \left(\left(\frac{r}{2} - \mu \right) z_i + \frac{1}{z_i} [k_\mu(\Lambda_z, I)]_{ii} + z_i \sum_{j \neq i} \frac{z_j^2}{z_i^2 - z_j^2} \right) dt, \tag{88}$$

where Λ_z denotes the diagonal matrix $[\Lambda_z]_{ii} = z_i^{-2}$. \square

Proof. As X_t is equivalent in law to Q_t used throughout Section 2, Proposition 13 describes the process for the inverse eigenvalues of X_t . That is to say that (87) follows from (45) upon making the substitution $x_i = p_i^{-1}$ in the $\beta = 1$ instance of that equation.

The argument for Z_t requires only a couple of additional observations. Write

$$d(Z_t^T Z_t) = Z_t^T (dB_t + dB_t^T) Z_t + (1 - 2\mu + r) Z_t^T Z_t dt + 2Z_t^T \kappa_\mu(I, (Z_t Z_t^T)^{-1}) Z_t dt, \quad (89)$$

which is now rotation invariant. In particular, setting $Z_t = V_t \Lambda_t^{1/2} U_t^T$ for orthogonal U, V and Λ_t the diagonal of square-singular values of Z_t , the right-hand side of (89) equals

$$U_t (\Lambda_t^{1/2} dB_t \Lambda_t^{1/2} + (1 - 2\mu + r) \Lambda_t dt + 2\Lambda_t^{1/2} \kappa_\mu(I, \Lambda_t^{-1}) \Lambda_t^{1/2} dt) U_t^T.$$

Here, $B_t = V_t^T (B_t + B_t^T) V_t$ is equal in law to twice a symmetric (or ‘‘Dyson’’) Brownian motion, and we have used that $\kappa_\mu(I, UAU^T) = U\kappa_\mu(I, A)U^T$, for any symmetric A . From here, the standard method used before will yield the system,

$$d\lambda_i = 2\lambda_i db_i + \left[(1 - 2\mu + r)\lambda_i + 2[\kappa_\mu(\Lambda^{-1}, I)]_{ii} + 2 \sum_{j \neq i} \frac{\lambda_i \lambda_j}{\lambda_i - \lambda_j} \right] dt, \quad (90)$$

having employed the identity $\Lambda^{1/2} \kappa_\mu(I, \Lambda^{-1}) \Lambda^{1/2} = \kappa_\mu(\Lambda^{-1}, I)$ along the way. Setting $z_i = \sqrt{\lambda_i}$ completes the proof. ■

Before continuing with the proof of Theorem 7, we take a quick detour to show that the generator of the process $(Z_t^T Z_t)^{-1}$ can be expressed as the Doob transform of the generator of the process $Y_t = M_t M_t^T$ with an additional killing term.

Proposition 23. Let $2\mu > r - 1$ and consider $M_t = M_t^{(-\mu)}$ and Z_t defined via this M_t according to (13). Then, the process $(Z_t^T Z_t)^{-1}$ is Markovian with generator given by the Doob transform $U(Y)^{-1}(G_Y - \frac{1}{2}\text{tr}Y)U(Y)$. Here, G_Y is the generator of the process $Y_t = M_t M_t^T$, the operator defined in (10), and $U(Y)$ is the normalized K -Bessel function $U(Y) = \frac{K_r(-\mu|Y, I)}{2^{\mu-1} \Gamma_r(\mu)}$. □

Proof. From (89) (noting that we have $-\mu$ instead of μ now), we get

$$d(Z_t^T Z_t)^{-1} = -Z_t^{-1} (dB_t + dB_t^T) Z_t^{-T} + (-2\mu + r + 1) (Z_t^T Z_t)^{-1} dt - 2Z_t^{-1} \kappa_{-\mu}(I, (Z_t Z_t^T)^{-1}) Z_t^{-T} dt.$$

Setting $Z_t = V_t \Lambda_t^{1/2} U_t^T$ for orthogonal U, V the last drift term can be rewritten as

$$\begin{aligned} -2Z_t^{-1} \kappa_{-\mu}(I, (Z_t Z_t^T)^{-1}) Z_t^{-T} &= -2U_t \Lambda_t^{-1/2} \kappa_{-\mu}(I, \Lambda_t^{-1}) \Lambda_t^{-1/2} U_t^T \\ &= -2U_t \Lambda_t^{-1} \kappa_{-\mu}(\Lambda_t^{-1}, I) \Lambda_t^{-1} U_t^T \\ &= -2(Z_t^T Z_t)^{-1} \kappa_{-\mu}((Z_t^T Z_t)^{-1}, I) (Z_t^T Z_t)^{-1}. \end{aligned}$$

Mimicking the first steps of the Proof of Theorem 2 now, it is straightforward to show that $(Z_t^T Z_t)^{-1}$ is Markovian with a generator given by

$$2\text{tr}\left(Y\frac{\partial}{\partial Y}\right)^2 - 2\mu\text{tr}\left(Y\frac{\partial}{\partial Y}\right) - 2\text{tr}\left(Y\kappa_{-\mu}(Y, I)Y\frac{\partial}{\partial Y}\right) = G_Y - 2\text{tr}\left(Y\kappa_{-\mu}(Y, I)Y\frac{\partial}{\partial Y}\right).$$

To show that this operator is the same as $U(Y)^{-1}(G_Y - \frac{1}{2}\text{tr}Y)U(Y)$, one just has to apply the chain rule, the identity (11) and the fact that $2U(Y)^{-1}\frac{\partial}{\partial Y}U(Y) = -\kappa_{-\mu}(Y, I)$. ■

We now return to the proof of Theorem 7 which is split into two parts.

Proposition 24. Set $\mu = \frac{r-1}{2} + \gamma$ in (87) and denote the maximal eigenvalue by x_t^γ . Then, as processes, $\frac{1}{2c} \log x_{c^2t}^{\gamma/c}$ converges as $c \rightarrow \infty$ to the Brownian motion with drift $-\gamma$ reflected at the origin. (The lower eigenvalues converge to the zero process in this scaling). □

Proof. Changing to logarithmic coordinates, $y_i = \log x_i$, and introducing the scaling as in $\gamma \mapsto \gamma/c$ (after putting $\mu = \frac{r-1}{2} + \gamma$) and $y_i(t) \mapsto y_i(c^2t)/(2c)$, we can work with the system

$$dy_i = 2db_i + \left(-\gamma + \frac{c}{2}e^{-cy_i} + c\sum_{j \neq i} \frac{e^{cy_j}}{e^{cy_i} - e^{cy_j}}\right) dt. \tag{91}$$

Here, the drift was simplified ahead of time by using $-(r-1) + \sum_{j \neq i} \frac{x_i + x_j}{x_i - x_j} = \sum_{j \neq i} \frac{2x_j}{x_i - x_j}$.

On the other hand, back in the original coordinates we have that

$$P\left(\lim_{t \rightarrow \infty} (\log x_i(t) - \log x_j(t)) = +\infty\right) = 1,$$

for any pair $i > j$. This follows again by the proof of Lemma 11. From here, we see that the top point in (91) shares whatever $c \rightarrow \infty$ process limit it may have with that for $y_c(t)$ defined by

$$y_c(t) = b_t^{-\gamma} + L_c(t), \quad L_c(t) = \frac{c}{2} \int_0^t e^{-cy_c(s)} ds, \tag{92}$$

and we want to show that $L_c(t)$ produces a local time contribution in the limit.

For $\epsilon > 0$, decompose the path $t \mapsto y_c(t)$ at the time $s \leq t$ at which it was last beneath level ϵ to find that

$$y_c(t) \leq \epsilon + \max_{s \leq t} (b_t^{-\gamma} - b_s^{-\gamma}) + \frac{c}{2} e^{-\alpha \epsilon} t.$$

Thus,

$$L_c(t) \leq \epsilon + \max_{s \leq t} (b_t^{-\gamma} - b_s^{-\gamma}) + a_\epsilon t - b_t^{-\gamma},$$

and for each fixed t , it holds that $\sup_{c>0} L_c(t) < \infty$ with probability one. Decomposing instead at the last time that the path exceeds $-\epsilon$ similar reasoning shows that $\liminf_{c \rightarrow \infty} \inf_{0 \leq s \leq t} Y_c(s) \geq 0$. With both sequences bounded above and below, by passing to a subsequence if needed there exist $(\gamma(t), L(t))$ with $Y_c(t) \rightarrow \gamma(t)$ and $L_c(t) \rightarrow L(t)$ at all but countably many t for which $\gamma(t) = b_t^{-\gamma} + L(t)$. Since $\int_0^t \mathbf{1}_{[\epsilon, \infty)}(Y_c(s)) dL_c(s) \rightarrow 0$, any limiting $L(t)$ is nondecreasing and increases only when $\gamma(t) = 0$. As $\gamma(t) \geq 0$, we see that any such pair $(\gamma(t), L(t))$ is the (unique) solution to the Skorohod problem for $b_t^{-\gamma}$. This precisely what we wanted to show. ■

Proposition 25. Now, set $\mu = \frac{r-1}{2} + \gamma$ in (88) and denote by z_t^γ the minimal singular value there. Then, again in the Skorohod topology, $\lim_{c \rightarrow \infty} \frac{1}{c} \log z_{c^2 t}^{\gamma/c} \Rightarrow r_t$, where $t \mapsto r_t$ is the diffusion on the positive half-line with generator $\frac{1}{2} \frac{d^2}{dr^2} + \gamma \coth(\gamma r) \frac{d}{dr}$. □

Proof. In order to get a workable formula for the matrix GIG mean, we bring in the more general K -Bessel functions. For $\mathbf{s} = (s_1, \dots, s_r) \in \mathbb{C}^r$ recall the power function $p_{\mathbf{s}}(X)$ from (60) and the invariant measure μ_r on \mathcal{P} from (58), and set

$$K_r(\mathbf{s}|A, B) = \frac{1}{2} \int_{\mathcal{P}} p_{\mathbf{s}}(X) e^{-\frac{1}{2} \text{tr}AX - \frac{1}{2} \text{tr}BX^{-1}} d\mu_r(X).$$

This is actually how Terras introduces the K -Bessel functions from the start (see [34, Section 4.2.2], though keep in mind our inclusion of various factors of $\frac{1}{2}$ not used there), and reduces to our earlier defined $K_r(\mathbf{s}|A, B)$ when $\mathbf{s} = (0, \dots, 0, s)$.

In terms of the above we have: for any (positive) diagonal matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$,

$$[\kappa_\mu(\Lambda, I)]_{11} = \frac{K_r(\mathbf{s}|\Lambda, I)}{K_r(\mathbf{s}'|\Lambda, I)}, \tag{93}$$

where now $\mathbf{s} = (1, 0, \dots, 0, \mu)$ and $\mathbf{s}' = (0, 0, \dots, 0, \mu)$. Likewise,

$$[\kappa_\mu(\Lambda, I)]_{ii} = \frac{K_r(\mathbf{s}|\Lambda_{\sigma_i}, I)}{K_r(\mathbf{s}'|\Lambda, I)}, \tag{94}$$

in which Λ_{σ_i} is the matrix arrived at from Λ by swapping λ_1 and λ_i . This uses (again) that $\kappa_\mu(I, \Lambda) = U^T \kappa_\mu(I, U \Lambda U^T) U$ for orthogonal U , here with the choice of U being the corresponding permutation matrix. Note that taking U to be a diagonal orthogonal matrix with ± 1 entries so that $U_{ii} U_{jj} = -1$ for given $i \neq j$, yields $[\kappa_\mu(I, \Lambda)]_{i,j} = -[\kappa_\mu(I, \Lambda)]_{i,j}$, so we get that $\kappa_\mu(I, \Lambda)$ is actually diagonal.

The ratios in (93) and (94) can then be expanded with the help of Terras' induction formula (see Exercise 20 of [34, §4.2.2] though note the typo, $\frac{m-n}{2}$ should be $\frac{n-m}{2}$). In the

present setting, this implies that

$$\begin{aligned}
 K_r(\mathbf{s}|\Lambda, I) &= \int_{\mathbb{R}^{r-1}} K_1\left(\mu - \frac{r-3}{2} \middle| \lambda_1 + \sum_{i=2}^r \lambda_i x_i^2, 1\right) \\
 &\quad \times K_{r-1}(\mathbf{s}''|\Lambda^{(1)}, I + \mathbf{x}\mathbf{x}^T) d\mathbf{x}_2 \cdots d\mathbf{x}_r,
 \end{aligned}
 \tag{95}$$

with $\Lambda^{(1)} = \text{diag}(\lambda_2, \dots, \lambda_r)$ and $\mathbf{s}'' = (0, \dots, 0, \mu - \frac{1}{2}) \in \mathbb{R}^{(r-1)}$. Applied to $K_r(\mathbf{s}'|\Lambda, I)$ the result is of course similar, the only difference that the K_1 factor on the right-hand side of (95) is replaced by $K_1(\mu - \frac{r-1}{2}|\lambda_1 + \sum_{i=2}^r \lambda_i x_i^2, 1)$.

Writing both occurrences of $K_1(\cdot|\cdot, \cdot)$ in terms of the standard Macdonald function we have that $K_1(\mu - \frac{r-3}{2}|\psi^2, 1) = \psi^{-\mu + \frac{r-3}{2}} K_{\mu - \frac{r-3}{2}}(\psi)$ and $K_1(\mu - \frac{r-1}{2}|\psi^2, 1) = \psi^{-\mu + \frac{r-1}{2}} K_{\mu - \frac{r-1}{2}}(\psi)$. So, with the shorthand,

$$\psi = \psi(\mathbf{x}, \Lambda) = \sqrt{\lambda_1 + \sum_{i=2}^r \lambda_i x_i^2}, \quad \mathcal{K}(\mathbf{x}, \Lambda^{(1)}) = K_{r-1}(\mathbf{s}''|\Lambda^{(1)}, I + \mathbf{x}\mathbf{x}^T)$$

we record the new expression for (93): making the substitution $\mu = \frac{r-1}{2} + \gamma$,

$$\begin{aligned}
 [K_{\frac{r-1}{2} + \gamma}(\Lambda, I)]_{11} &= \frac{\int_{\mathbb{R}^{r-1}} \psi^{-1-\gamma} K_{1+\gamma}(\psi) \mathcal{K}(\mathbf{x}, \Lambda^{(1)}) d\mathbf{x}}{\int_{\mathbb{R}^{r-1}} \psi^{-\gamma} K_{\gamma}(\psi) \mathcal{K}(\mathbf{x}, \Lambda^{(1)}) d\mathbf{x}}, \\
 &:= \frac{1}{\sqrt{\lambda_1}} \frac{\int_{\mathbb{R}^{r-1}} \psi_0^{-1} K_{1+\gamma}(\sqrt{\lambda_1} \psi_0) \rho_{\Lambda}^{(1)}(d\mathbf{x})}{\int_{\mathbb{R}^{r-1}} K_{\gamma}(\sqrt{\lambda_1} \psi_0) \rho_{\Lambda}^{(1)}(d\mathbf{x})}.
 \end{aligned}
 \tag{96}$$

In line two, we have made the change of variables $T_{\Lambda} : \mathbf{x}_i \mapsto \sqrt{\frac{\lambda_1}{\lambda_i}} \mathbf{x}_i$, and have introduced

$$\psi_0(\mathbf{x}) = \sqrt{1 + \sum_{i=2}^r x_i^2}, \quad \rho_{\Lambda}^{(1)}(d\mathbf{x}) = \psi_0(\mathbf{x})^{-\gamma} \mathcal{K}(T_{\Lambda} \mathbf{x}, \Lambda^{(1)}) d\mathbf{x}.
 \tag{97}$$

By way of (94), there are allied expressions for the other diagonal components of the mean.

Finally, returning to (88) and setting $y_i = \log z_i$ that equation becomes

$$\begin{aligned}
 dy_i &= db_i + \left(-\gamma + e^{-y_i} \frac{\int_{\mathbb{R}^{r-1}} \psi_0^{-1} K_{1+\gamma}(e^{-y_i} \psi_0) \rho_{\Lambda}^{(i)}(d\mathbf{x})}{\int_{\mathbb{R}^{r-1}} K_{\gamma}(e^{-y_i} \psi_0) \rho_{\Lambda}^{(i)}(d\mathbf{x})} \right) dt \\
 &\quad + \left(\frac{r-1}{2} + \frac{1}{2} \sum_{j \neq i} \frac{e^{2y_i} + e^{2y_j}}{e^{2y_i} - e^{2y_j}} \right) dt.
 \end{aligned}
 \tag{98}$$

Here, Λ is now the diagonal matrix $\Lambda_{ii} = e^{-2y_i}$, and we have employed (96) while being a little fluid with notation: $\rho_\Lambda^{(i)}$ stands for the comparable object to $\rho_\Lambda^{(1)}$ defined in the same way as in (97) but for the i th coordinate.

The strategy from this point is

- (i) Show yet again a separation of scales. That is, for long time, it holds that $y_i \ll y_j$ for all $i < j$ with probability tending to one. Without the presence of the Macdonald function term in the drift, the same calculation used in Lemma 11 would (yet again) show that the solution of (98) satisfies $\frac{1}{t} \log y_i \rightarrow (r - i) - \gamma$ for all i with probability one. The claim is that the added drift doesn't affect this appraisal too much
- (ii) Show that

$$\frac{\int_{\mathbb{R}^{r-1}} \psi_0^{-1} K_{1+\gamma}(e^{-y_1} \psi_0) \rho_\Lambda^{(1)}(dx)}{\int_{\mathbb{R}^{r-1}} K_\gamma(e^{-y_1} \psi_0) \rho_\Lambda^{(1)}(dx)} = \frac{K_{1+\gamma}(e^{-y_1})}{K_\gamma(e^{-y_1})} (1 + o(1)),$$

uniformly in y_1 as $y_2, \dots, y_r \rightarrow \infty$.

The estimate for (i) follows from known bounds for the K -Bessel function at ∞ . For (ii), using the explicit formulas, it is easy to see that the measure ρ_Λ^1 has Gaussian concentration at the point $x = 0$ (where one notes that $\psi_0 = 1$).

Put together, and after the required scaling, the drift in the equation for $y = \frac{1}{c} Y_1(c^2 t, \gamma/c)$ equals $-\gamma + e^{-cY} \frac{K_{1+\gamma/c}(e^{-cY})}{K_{\gamma/c}(e^{-cY})}$ up to vanishing errors as $c \rightarrow \infty$. That is, we recover the same calculation needed by Matsumoto-Yor in the 1D case, and so also the same limit. ■

Funding

B.R. was supported in part by NSF grants DMS-1340489 and DMS-1406107, as well as grant 229249 from the Simons Foundation. B.V. was supported in part by the NSF CAREER award DMS-1053280.

Acknowledgement

We thank F. Baudoin, G. Letác, and N. O'Connell for their interest and many helpful discussions. Thanks as well to D.W. Stroock for pointers to the PDE literature, and T. Kurtz for assistance with the proof of Proposition 24. We also thank the anonymous referees for their useful comments and suggestions.

References

- [1] Anderson, G. W., A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics 118. Cambridge University Press, 2010.

- [2] Baik, J., G. Ben Arous, and S. Péché. "Phase transition of the largest eigenvalue for non-null complex sample covariance matrices." *The Annals of Probability* 33, no. 5 (2005): 1643–97.
- [3] Baudoin, F. and N. O’Connell. "Exponential functionals of Brownian motion and class-one Whittaker functions." *Annales de l’Institut Henri Poincaré* 47, no. 4 (2011): 1096–120.
- [4] Bernadac, E. "Random continued fractions and inverse Gaussian distribution on a symmetric cone." *Journal of Theoretical Probability* 8, no. 2 (1995): 221–59.
- [5] Bloemendal, A. and B. Virág. "Limits of spiked random matrices I." *Probability Theory and Related Fields* 156, no. 3–4 (2013): 795–825.
- [6] Bloemendal, A. and B. Virág. "Limits of spiked random matrices II." *The Annals of Probability*, arXiv:1109.3704.
- [7] Borodin, A., I. Corwin, and P. Ferrari. "Free energy fluctuations for directed polymers in random media in $1 + 1$ dimension." *Communications on Pure and Applied Mathematics* 67, no. 7 (2014): 1129–1214.
- [8] Borodin, A., I. Corwin, and D. Remenik. "Log-gamma polymer free energy fluctuations via a Fredholm determinant identity." *Communications in Mathematical Physics* 324 (2013): 215–232.
- [9] Bougerol, Ph. "The Matsumoto and Yor process and infinite dimensional hyperbolic space." (2014): preprint arXiv:1408.2108.
- [10] Carmona, P., F. Petit, and M. Yor. "Beta-gamma random variables and intertwining relations between certain Markov processes." *Revista Matemática Iberoamericana* 14, no. 2 (1998): 311–67.
- [11] Chhaibi, R. "Littlemann path model for geometric crystals: Whittaker functions on Lie groups and Brownian motion." PhD thesis, Université Paris VI – Pierre et Marie Curie, (2012).
- [12] Desrosiers, P. and P. Forrester. "Asymptotic correlations for Gaussian and Wishart matrices with external source." *International Mathematics Research Notices* 2006 (2006), Art. ID 27395.
- [13] Dufresne, D. "The distribution of a perpetuity, with application to risk theory and pension funding." *Scandinavian Actuarial Journal* 1990, no. 1 (1990): 39–79.
- [14] Dufresne, D. "An affine property of the reciprocal Asian option process." *Osaka Journal of Mathematics* 38, no. 2 (2001): 379–81.
- [15] Erdős, L., B. Schlein, and H.-T. Yau. "Universality of random matrices and local relaxation flow." *Inventiones Mathematicae* 185, no. 1 (2011): 75–119.
- [16] Herz, C. S. "Bessel functions of matrix argument." *Annals of Mathematics* 61, no. 3 (1955): 474–523.
- [17] Letac, G. and J. Wesolowski. "An independence property for the product of GIG and gamma laws." *The Annals of Probability* 28, no. 3 (2000): 1371–83.
- [18] Lipster, R. S. and A. N. Shiryaev. *Statistics of Random Processes*. Berlin, Springer, 2000.
- [19] Matsumoto, H. and M. Yor. "A version of Pitman’s $2M - X$ theorem for geometric Brownian motions." *Comptes Rendus de l’Académie des Sciences Paris Series I Math.* 328 (1999): 1067–74.

- [20] Matsumoto, H. and M. Yor. "An analogue of Pitman's $2M - X$ theorem for exponential wiener functionals, Part I: A time-inversion approach." *Nagoya Mathematical Journal* 159 (2000): 125–66.
- [21] Matsumoto, H. and M. Yor. "An analogue of Pitman's $2M - X$ theorem for exponential wiener functionals, Part 2: the role of the generalized inverse Gaussian laws." *Nagoya Mathematical Journal* 162 (2001): 65–86.
- [22] Matsumoto, H. and M. Yor. "A relationship between Brownian motions of opposite drifts via certain enlargements of the Brownian filtration." *Osaka Journal of Mathematics* 38, no. 2 (2001): 383–98.
- [23] Matsumoto, H. and M. Yor. "Exponential functionals of Brownian motion, I: probability laws at fixed time." *Probability Surveys* 2 (2005): 312–47.
- [24] Mo, M. Y. "The rank 1 real spiked model." *Communications on Pure and Applied Mathematics* 65, no. 11 (2011): 1528–638.
- [25] Muirhead, R. J. *Aspects of Multivariate Statistical Theory*. London: Wiley, 2005.
- [26] Norris, J. R., L. C. G. Rogers, and D. Williams. "Brownian motions of ellipsoids." *Transactions of the American Mathematical Society* 294, no. 2 (1986): 757–65.
- [27] O'Connell, N. "Directed polymers and the quantum Toda lattice." *The Annals of Probability* 40, no. 2 (2012): 437–58.
- [28] O'Connell, N. and M. Yor. "Brownian analogues of Burke's theorem." *Stochastic Processes and their Applications* 96 (2001): 285–304.
- [29] Ramírez, J. and B. Rider. "Diffusion at the random matrix hard edge." *Communications in Mathematical Physics* 288, no. 3 (2009): 887–906.
- [30] Ramírez, J. and B. Rider. "Spiking the random matrix hard edge." (2015). In preparation.
- [31] Ramírez, J., B. Rider, and B. Virág. "Beta ensembles, stochastic Airy spectrum, and a diffusion." *Journal of the American Mathematical Society* 24, no. 4 (2011): 919–44.
- [32] Seppäläinen, T. "Scaling limit for a one-dimensional directed polymer with boundary conditions." *The Annals of Probability* 40, no. 1 (2012): 19–73.
- [33] Stroock, D. W. *Partial Differential Equations for Probabilists*. Cambridge, UK: Cambridge University Press, 2008.
- [34] Terras, A. *Harmonic Analysis on Symmetric Spaces and Applications II*. Berlin: Springer, 1988.
- [35] Wang, D. "The largest sample eigenvalue distribution in the rank one quaternionic spiked model of Wishart ensemble." *The Annals of Probability* 37, no. 4 (2009): 1273–328.