Matrix Whittaker processes

Jonas Arista[1] · Elia Bisi[2] · Neil O'Connell[3]

Abstract

We study a discrete-time Markov process on triangular arrays of matrices of size $d \geq 1$, driven by inverse Wishart random matrices. The components of the right edge evolve as multiplicative random walks on positive definite matrices with one-sided interactions and can be viewed as a d-dimensional generalisation of log-gamma polymer partition functions. We establish intertwining relations to prove that, for suitable initial configurations of the triangular process, the bottom edge has an autonomous Markovian evolution with an explicit transition kernel. We then show that, for a special singular initial configuration, the fixed-time law of the bottom edge is a matrix Whittaker measure, which we define. To achieve this, we perform a Laplace approximation that requires solving a constrained minimisation problem for certain energy functions of matrix arguments on directed graphs.

Keywords Whittaker functions of matrix arguments · Intertwining relations · Interacting Markov dynamics · Noncommutative polymer models · Constrained minimisation · Directed graphs

Mathematics Subject Classification Primary 60K35 · 82B23 · 60B20; Secondary 33C15 · 05E05 · 22E30

✉ Elia Bisi
  elia.bisi@tuwien.ac.at

  Jonas Arista
  jarista@math.uni-bielefeld.de

  Neil O'Connell
  neil.oconnell@ucd.ie

1 Universität Bielefeld, Fakultät für Mathematik, Universitätsstrasse 25, 33615 Bielefeld, Germany

2 Technische Universität Wien, Institut für Stochastik und Wirtschaftsmathematik, E 105-07, Wiedner Hauptstraße 8-10, 1040 Wien, Austria

3 School of Mathematics and Statistics, University College Dublin, Dublin 4, Ireland

Contents

1 Introduction

In the last few decades, we have witnessed a surge of research on stochastic integrable models, often motivated by problems in mathematical physics and enriched by deep connections with algebraic combinatorics, representation theory, symmetric functions, and integrable systems [9, 10]. Some of the most intensively studied models are interacting particle systems and stochastic growth processes in the Kardar–Parisi–Zhang (KPZ) universality class [15, 45].

From a mathematical perspective, it is natural to consider noncommutative versions of these models, which have very recently received some attention. In [33] a system of interacting Brownian particles in the space of positive definite matrices was considered and shown to have an integrable structure, related to the non-Abelian Toda chain and Whittaker functions of matrix arguments (the latter introduced in that article). In the discrete-time setting, [1] proved Matsumoto-Yor and Dufresne type theorems for a random walk on positive definite matrices.

On the other hand, from the theoretical physics point of view, such matrix models may find interesting applications in quantum stochastic dynamics, as set out in [20]. In particular, [20] introduced a matrix generalisation of the classical Kesten recursion and studied a related quantum problem of interacting fermions in a Morse potential.

Quoting the authors, their initial motivation was "to explore possible matrix (non-commuting) generalizations of the famous directed polymer problem (which is related to the KPZ stochastic growth equation)".

The subject of the present article is an integrable model of random walks on positive definite matrices with local interactions. This constitutes, on the one hand, a discrete-time analogue of the matrix-valued interacting diffusions studied in [33] and, on the other hand, a matrix generalisation of the *log-gamma polymer* model.

To motivate the contributions of this article, let us first define a discrete-time *exclusion process* $\mathcal{Z}$ of $N \geq 1$ ordered particles $\mathcal{Z}^1 \leq \mathcal{Z}^2 \leq \cdots \leq \mathcal{Z}^N$ on $\mathbb{Z}$ moving to the right. Let $(\mathcal{V}^1(n), \ldots, \mathcal{V}^N(n))_{n \geq 1}$ be a collection of independent random variables supported on $\mathbb{Z}_{\geq 0}$. At each time n, the particle positions are updated sequentially from the 1-st one to the Nth one, as follows. The 1-st particle simply evolves as a random walk on $\mathbb{Z}$ with time-n increment $\mathcal{V}^1(n)$. Once the positions of the first $i-1$ particles have been updated, if the $(i-1)$th particle has overtaken the ith particle, then the latter is pushed forward to a temporary position to maintain the ordering; next, to complete its update, the ith particle takes $\mathcal{V}^i(n)$ unit jumps to the right. The particle locations then satisfy the recursive relations

$$\mathcal{Z}^1(n) = \mathcal{Z}^1(n-1) + \mathcal{V}^1(n)\,, \tag{1.1}$$

$$\mathcal{Z}^i(n) = \max\left[\mathcal{Z}^{i-1}(n), \mathcal{Z}^i(n-1)\right] + \mathcal{V}^i(n)\,, \quad 2 \leq i \leq N\,. \tag{1.2}$$

If one considers the initial state

$$\mathcal{Z}^1(0) = \mathcal{Z}^2(0) = \cdots = \mathcal{Z}^N(0) = 0\,, \tag{1.3}$$

then the following *last passage percolation* formula holds:

$$\mathcal{Z}^i(n) = \max_{\pi} \sum_{(m,k) \in \pi} \mathcal{V}^k(m)\,, \quad 1 \leq i \leq N\,,$$

where the maximum is over all directed lattice paths π in $\mathbb{Z}^2$ (i.e., at each lattice site (m,k), π is allowed to head either rightwards to $(m+1, k)$ or upwards to $(m, k+1)$) that start from $(1, 1)$ and end at (n, i). As a process of last passage percolation times, $\mathcal{Z}$ can be also associated with the *corner growth process* with step (or 'narrow wedge') initial configuration. Remarkable integrable versions of this model are those with geometrically and exponentially distributed jumps, first studied in [24].

A positive temperature version of $\mathcal{Z}$ can be obtained by formally replacing the operations $(\max, +)$ with $(+, \times)$ in the relations (1.1)–(1.2). Namely, given a collection of independent positive random variables $(V^1(n), \ldots, V^N(n))_{n \geq 1}$, we can consider the discrete-time Markov process Z defined by

$$Z^1(n) = Z^1(n-1)V^1(n)\,, \tag{1.4}$$

$$Z^i(n) = \left[Z^{i-1}(n) + Z^i(n-1)\right]V^i(n)\,, \quad 2 \leq i \leq N\,. \tag{1.5}$$

Considering the initial configuration

$$Z^1(0) = 1, \qquad Z^2(0) = \cdots = Z^N(0) = 0, \tag{1.6}$$

we have the closed-form expression

$$Z^i(n) = \sum_{\pi} \prod_{(m,k) \in \pi} V^k(m), \qquad 1 \le i \le N, \tag{1.7}$$

where the sum is over all directed lattice paths π in $\mathbb{Z}^2$ from $(1, 1)$ to (n, i). The variables (1.7) can be regarded as partition functions of the $(1 + 1)$-dimensional *directed polymer*, an intensively studied model of statistical mechanics. Of particular importance is the model with inverse gamma distributed weights $V^i(n)$, known as the *log-gamma polymer*, first considered in [39]. In [16] it was shown that the laws of log-gamma polymer partition functions are marginals of *Whittaker measures*; the latter are defined in terms of $\mathrm{GL}_d(\mathbb{R})$-Whittaker functions and were introduced in that article.

In this article, we study a noncommutative generalisation of the above Markov process of log-gamma polymer partition functions. The 'particles' of this process live in $\mathcal{P}_d$, the set of $d \times d$ positive definite real symmetric matrices. The random weights $V^i(n)$ are now independent inverse Wishart matrices (a matrix generalisation of inverse gamma random variables; see Sect. 1.2). We define Z by setting

$$Z^1(n) := Z^1(n-1)^{1/2} V^1(n) Z^1(n-1)^{1/2}, \tag{1.8}$$

$$Z^i(n) := \left[Z^{i-1}(n) + Z^i(n-1) \right]^{1/2} V^i(n) \left[Z^{i-1}(n) + Z^i(n-1) \right]^{1/2},$$
$$2 \le i \le N, \tag{1.9}$$

where, for $a \in \mathcal{P}_d$, $a^{1/2}$ denotes the unique $b \in \mathcal{P}_d$ such that $b^2 = a$. The above matrix products are symmetrised to ensure that, starting from any initial configuration $Z^i(0) \in \mathcal{P}_d$, each $Z^i(n)$ still belongs to $\mathcal{P}_d$ for all $n \ge 1$. The 1-st particle (1.8) evolves as a (GL_d-invariant) multiplicative random walk on $\mathcal{P}_d$; on the other hand, the other particles (1.9) can be viewed as analogous random walks with one-sided interactions. From this point of view, the Markov process as a whole can be also regarded as a noncommutative version of the exclusion process $\mathcal{Z}$ defined in (1.1)–(1.2). The natural generalisation of the initial configuration (1.6) is

$$Z^1(0) = I_d, \qquad Z^2(0) = \cdots = Z^N(0) = 0_d, \tag{1.10}$$

where I_d and 0_d are the $d \times d$ identity and zero matrices, respectively. Notice that, although all but the first particle are initially zero, the process Z starting from (1.10) lives in $\mathcal{P}_d^N$ at all times $n \ge 1$.

In Sect. 3, we introduce a Markov process $X = (X(n))_{n \ge 0}$, $X(n) = (X_j^i(n))_{1 \le j \le i \le N}$, on triangular arrays of positive definite matrices whose 'right edge', namely
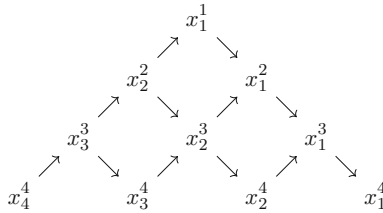
Fig. 1 Graphical representation of a 'triangular' array $x \in \mathcal{T}_d^N$ as in (2.1), for $N = 4$. Each row x^i, $1 \leq i \leq N$, consists of the matrices $(x_1^i, \dots, x_i^i)$, read from right to left. The arrows refer to the energy function $\Phi^N(x)$ in (2.3), where every summand $\mathrm{tr}[ab^{-1}]$ corresponds to an arrow pointing from a to b in the figure

$(X_1^1, \dots, X_1^N)$, equals Z. The evolution of X may be viewed as a noncommutative version of the dynamics on Gelfand-Tsetlin patterns with blocking and pushing interactions, studied in various contexts in [7, 8, 11, 31, 43, 44]. We refer to Fig. 1 for a graphical representation of such a triangular array. Moreover, as we detail in Remark 3.4, the 'left edge' of X may be regarded as a noncommutative generalisation of the *strict-weak polymer* studied in [17, 34].

The first main result of this article (Theorem 3.10) states that, for certain special (random) initial configurations $X(0)$, the 'bottom edge' $X^N = (X_1^N, \dots, X_N^N)$ of X also has an autonomous Markovian evolution. The transition kernel of X^N is explicit and has an interpretation as a Doob h-transform with h-function given by a Whittaker function of matrix arguments. To obtain this, we prove certain intertwining relations between kernels associated to the process X and use the theory of Markov functions (reviewed in Appendix B). Another consequence of these intertwinings is that Whittaker functions are eigenfunctions of certain integral operators and possess a Feynman–Kac type interpretation.

Next, in Sect. 4, we define matrix Whittaker measures on $\mathcal{P}_d^N$ after proving an integral identity of Whittaker functions of matrix arguments (Theorem 4.1), analogous to the well-known Cauchy–Littlewood identity for Schur functions. The second main result of this article (Theorem 4.8) states that, for a special initial state, the fixed-time law of the bottom edge X^N of X is a matrix Whittaker measure on $\mathcal{P}_d^N$. Such an initial state, designed to match (1.10), is *singular*, in the sense that the particles are at the 'boundary' of $\mathcal{P}_d$.

Due to the singularity of the initial configuration, the proof of Theorem 4.8 will be based on a suitable limiting procedure and a careful integral approximation via Laplace's method. This will require a digression on a constrained minimisation problem for certain energy functions of matrix arguments. We chose to include this analysis in a separate section and to present it in the more general framework of directed graphs, as it may be of independent interest; see Sect. 5. For us, the main application will be the asymptotic formula (4.8) for Whittaker functions of matrix arguments.

From our main results we deduce (see Corollary 4.10) that, under the initial configuration (1.10), the particles of the process Z defined in (1.8)–(1.9) have a fixed-time law given by the first marginal of a matrix Whittaker measure on $\mathcal{P}_d^N$. In the scalar $d = 1$ case, we recover the aforementioned result of [16] for the law of the log-

gamma polymer partition functions. In Corollary 4.10, we also obtain an analogous result concerning the fixed-time law of the 'left edge' of the triangular array X.

It is worth mentioning that the log-gamma polymer partition functions (1.7) were also studied in [16] as embedded in a dynamic on triangular arrays. However, such a dynamic was constructed via the combinatorial mechanism of the geometric Robinson–Schensted–Knuth correspondence; in particular, at each time step, the right edge is updated using N new (independent) random variables, whereas all the other components are updated via deterministic transformations of the current state and the newly updated right edge. It turns out that, for $d = 1$, the processes considered in [16] and in the present article have an identical right edge and, under the special initial configuration of Theorem 3.10, also a bottom edge process with the same Markovian evolution. However, even in the $d = 1$ case, the two processes, as a whole, differ. The dynamic introduced in this article is driven by random updates with $N(N + 1)/2$ degrees of freedom, since each particle of the triangular array is driven by an independent source of randomness (as well as by local interactions with the other particles).

1.1 Organisation of the article

In Sect. 2, we define Whittaker functions of matrix arguments. In Sect. 3, we introduce a Markov dynamic on triangular arrays of matrices and study the evolution of its bottom edge, using the theory of Markov functions; we also obtain a Feynman–Kac interpretation of Whittaker functions. In Sect. 4, we define matrix Whittaker measures (through a Whittaker integral identity) and prove that they naturally arise as fixed-time laws in the aforementioned triangular process under a singular initial configuration. To do so, we need a Laplace approximation of Whittaker functions, which can be justified by solving a constrained minimisation problem for certain energy functions of matrix arguments on directed graphs: this is the content of Sect. 5. In Appendix A, we give a proof of the Cauchy–Littlewood identity for Schur functions that resembles our proof of the Whittaker integral identity. In Appendix B, we review the theory of Markov functions for inhomogeneous discrete-time Markov processes. Finally, in Appendix C, we prove a convergence lemma related to weak convergence of probability measures.

1.2 Notation and preliminary notions

Here we introduce some notation and preliminary notions that we use throughout this work. For background and proofs, we refer to [23, 42].

1.2.1 Positive definite matrices

Let $\mathcal{P}_d$ be the set of all $d \times d$ *positive definite* matrices, i.e. $d \times d$ real symmetric matrices with positive eigenvalues. Throughout this article, for $x \in \mathcal{P}_d$, we denote by $|x|$ the determinant of x and by $\mathrm{tr}[x]$ its trace.

The following properties hold:

- $x \in \mathcal{P}_d$ if and only if $x^{-1} \in \mathcal{P}_d$;

- if $x \in \mathcal{P}_d$ and $\lambda > 0$, then $\lambda x \in \mathcal{P}_d$;
- if $x, y \in \mathcal{P}_d$, then $x + y \in \mathcal{P}_d$ (but in general $xy \notin \mathcal{P}_d$);
- $x - y \in \mathcal{P}_d$ if and only if $y^{-1} - x^{-1} \in \mathcal{P}_d$.

For $x \in \mathcal{P}_d$, there exists a unique $y \in \mathcal{P}_d$ such that $y^2 = x$; we denote such a y by $x^{1/2}$.

For any $y \in \mathcal{P}_d$, we define the (noncommutative) 'multiplication operation' by y as

$$T_y \colon \mathcal{P}_d \to \mathcal{P}_d, \qquad T_y(x) := y^{1/2} x y^{1/2}, \qquad x \in \mathcal{P}_d. \tag{1.11}$$

Such a symmetrised product will be used to construct a multiplicative random walk on $\mathcal{P}_d$ (see Definition 3.1 and Remark 3.2 below).

We also denote by I_d and 0_d the $d \times d$ identity matrix and zero matrix, respectively.

1.2.2 Measure and integration on $\mathcal{P}_d$

Let GL_d be the group of $d \times d$ invertible real matrices. Define the measure μ on $\mathcal{P}_d$ by

$$\mu(\mathrm{d}x) := |x|^{-\frac{d+1}{2}} \prod_{1 \le i \le j \le d} \mathrm{d}x_{i,j}, \tag{1.12}$$

where $\mathrm{d}x_{i,j}$ is the Lebesgue measure on $\mathbb{R}$ in the variable $x_{i,j}$. Such a measure is the GL_d-*invariant measure* on $\mathcal{P}_d$, in the sense that

$$\int_{\mathcal{P}_d} f(a^\top x a) \mu(\mathrm{d}x) = \int_{\mathcal{P}_d} f(x) \mu(\mathrm{d}x)$$

for all $a \in \mathrm{GL}_d$ and for all suitable functions f. In other words, μ is invariant under the group action of GL_d on $\mathcal{P}_d$

$$\mathrm{GL}_d \times \mathcal{P}_d \to \mathcal{P}_d, \qquad (a, x) \mapsto a^\top x a.$$

Furthermore, the measure μ is preserved under the involution $x \mapsto x^{-1}$.

1.2.3 Wishart distributions and gamma functions

For $\alpha > \frac{d-1}{2}$, we will refer to the *(d-variate) Wishart distribution* with parameter α as the probability measure

$$\frac{1}{\Gamma_d(\alpha)} |x|^\alpha \, \mathrm{e}^{-\operatorname{tr}[x]} \mu(\mathrm{d}x) \tag{1.13}$$

on $\mathcal{P}_d$, where $\Gamma_d(\alpha)$ is the *d-variate gamma function*, i.e.

$$\Gamma_d(\alpha) := \int_{\mathcal{P}_d} \mu(\mathrm{d}x)\, |x|^{\alpha}\, \mathrm{e}^{-\operatorname{tr}[x]} = \int_{\mathcal{P}_d} \mu(\mathrm{d}x)\, |x|^{-\alpha}\, \mathrm{e}^{-\operatorname{tr}[x^{-1}]}$$

$$= \pi^{\frac{d(d-1)}{4}} \prod_{k=1}^{d} \Gamma\left(\alpha - \frac{k-1}{2}\right).$$

The inverse of a Wishart matrix with parameter α has the distribution

$$\frac{1}{\Gamma_d(\alpha)}\, |x|^{-\alpha}\, \mathrm{e}^{-\operatorname{tr}[x^{-1}]}\, \mu(\mathrm{d}x) \tag{1.14}$$

on $\mathcal{P}_d$. We will refer to the latter as the *(d-variate) inverse Wishart distribution* with parameter α.

1.2.4 Kernels and integral operators

Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be two measurable spaces. Let $\mathfrak{m}\mathcal{S}$ denote the set of complex-valued measurable functions on $(S, \mathcal{S})$. For our purposes, a *kernel* from T to S will be a map $L\colon T \times \mathcal{S} \to \mathbb{C}$ such that, for each $t \in T$, $L(t; \cdot)$ is a (complex) measure on $(S, \mathcal{S})$ and, for each $A \in \mathcal{S}$, $L(\cdot; A)$ is an element of $\mathfrak{m}\mathcal{T}$. The kernel L can be also, alternatively, thought of as an *integral operator*

$$L\colon \mathfrak{m}\mathcal{S} \to \mathfrak{m}\mathcal{T}, \qquad Lf(t) := \int_S L(t; \mathrm{d}s)\, f(s) \qquad \text{for } f \in \mathfrak{m}\mathcal{S},\ t \in T, \tag{1.15}$$

whenever the integral is well defined. Clearly, the composition of kernels/operators yields another kernel/operator; such a composition is associative but, in general, not commutative. When the complex measure $L(t; \cdot)$ is a probability measure for all $t \in T$, we will talk about *Markov kernels/operators*.

Throughout this article, the measurable spaces will be usually Cartesian powers of $\mathcal{P}_d$ (which we denote by $\mathcal{P}_d^k$, $k \geq 1$), with their Borel sigma-algebras. Moreover, for a kernel L from $\mathcal{P}_d^k$ to $\mathcal{P}_d^{\ell}$, the measure $L(t; \cdot)$ will be, in most cases, absolutely continuous with respect to the reference product measure $\mu^{\otimes \ell}$ on $\mathcal{P}_d^{\ell}$, for any $t \in \mathcal{P}_d^k$; with a little abuse of notation, we will then also write $s \mapsto L(t; s)$ for the corresponding density (a measurable function on $\mathcal{P}_d^{\ell}$).

2 Whittaker functions

In this section we define Whittaker functions of matrix arguments following [33], and then extend them to a further level of generality. Notice also that the kernels (2.8) and (2.15) defined below are matrix versions of certain kernels defined in [16, Sect. 3.1] and [35, Sect. 2] (see also references therein).

2.1 Whittaker functions of matrix arguments

We define Whittaker functions of matrix arguments as integrals over 'triangular arrays' of $d \times d$ positive definite matrices. For $N \geq 1$, denote by $\mathcal{T}_d^N := \mathcal{P}_d \times \mathcal{P}_d^2 \times \cdots \times \mathcal{P}_d^N$ the set of height-N triangular arrays

$$x = (x^1, \ldots, x^N) = (x_j^i)_{1 \leq j \leq i \leq N}, \tag{2.1}$$

where $x^i = (x_1^i, \ldots, x_i^i) \in \mathcal{P}_d^i$ will be referred to as the ith row of x, for $1 \leq i \leq N$. For $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$ and $x \in \mathcal{T}_d^N$, let

$$\Delta_\lambda^N(x) := |x_1^1|^{-\lambda_1} \prod_{i=2}^N \left(\frac{|x_1^i \cdots x_i^i|}{|x_1^{i-1} \cdots x_{i-1}^{i-1}|} \right)^{-\lambda_i}, \tag{2.2}$$

$$\Phi^N(x) := \sum_{i=1}^{N-1} \sum_{j=1}^i \left(\operatorname{tr} \left[x_{j+1}^{i+1} (x_j^i)^{-1} \right] + \operatorname{tr} \left[x_j^i (x_j^{i+1})^{-1} \right] \right). \tag{2.3}$$

For a graphical representation of the array (2.1) and of the 'energy function' Φ^N, see Fig. 1. For $z = (z_1, \ldots, z_N) \in \mathcal{P}_d^N$, let $\mathcal{T}_d^N(z) \subset \mathcal{T}_d^N$ be the set of all height-N triangular arrays x with Nth row $x^N = z$. We define the *Whittaker function* $\psi_\lambda^N(z)$ with argument $z \in \mathcal{P}_d^N$ and parameter $\lambda \in \mathbb{C}^N$ as

$$\psi_\lambda^N(z) := \int_{\mathcal{T}_d^N(z)} \left(\prod_{i=1}^{N-1} \prod_{j=1}^i \mu(\mathrm{d}x_j^i) \right) \Delta_\lambda^N(x) \, \mathrm{e}^{-\Phi^N(x)}. \tag{2.4}$$

Notice that, for $N = 1$, the expression above reduces to $\psi_\lambda^1(z) = |z|^{-\lambda}$. As proved in [33], the integral (2.4) is absolutely convergent for all $\lambda \in \mathbb{C}^N$, so that Whittaker functions are well defined.

For our purposes, it is convenient to rewrite Whittaker functions in terms of certain kernels that we now introduce. For $N \geq 1$, $\lambda \in \mathbb{C}^N$ and $x \in \mathcal{T}_d^N$, define the kernel

$$\Sigma_\lambda^N(x^N; \mathrm{d}x^{1:(N-1)}) := \Delta_\lambda^N(x) \, \mathrm{e}^{-\Phi^N(x)} \prod_{i=1}^{N-1} \prod_{j=1}^i \mu(\mathrm{d}x_j^i), \tag{2.5}$$

where, as always from now on, $i : j$ denotes the tuple $(i, i + 1, \ldots, j - 1, j)$ for $i \leq j$, so that $x^{1:(N-1)} \in \mathcal{T}_d^{N-1}$ is the triangular array consisting of the first $N - 1$ rows of x. Notice that, for $N = 1$, (2.5) reduces to $\Sigma_\lambda^1(z; \varnothing) = |z|^{-\lambda} = \psi_\lambda^1(z)$. For $z \in \mathcal{P}_d^N$, let us also define the kernel

$$\tilde{\Sigma}_\lambda^N(z; \mathrm{d}x) := \delta(z; \mathrm{d}x^N) \, \Sigma_\lambda^N(x^N; \mathrm{d}x^{1:(N-1)}), \tag{2.6}$$

where δ is the Dirac delta kernel on $\mathcal{P}_d^N$. Then, the Whittaker function (2.4) can be written as

$$\psi_\lambda^N(z) = \int_{\mathcal{T}_d^{N-1}} \Sigma_\lambda^N(z; \mathrm{d}x) = \int_{\mathcal{T}_d^N} \tilde{\Sigma}_\lambda^N(z; \mathrm{d}x). \qquad (2.7)$$

Moreover, for $N \geq 2$, $b \in \mathbb{C}$, $z = (z_1, \ldots, z_N) \in \mathcal{P}_d^N$, and $y = (y_1, \ldots, y_{N-1}) \in \mathcal{P}_d^{N-1}$, let

$$K_b^N(z; y) = \left(\prod_{i=1}^N |z_i|^{-b}\right) \prod_{j=1}^{N-1} |y_j|^b \, \mathrm{e}^{-\mathrm{tr}\left[z_{j+1}y_j^{-1} + y_j z_j^{-1}\right]}. \qquad (2.8)$$

We will usually regard (2.8) as a kernel by setting $K_b^N(z; \mathrm{d}y) := K_b^N(z; y)\mu^{\otimes(N-1)}$ $(\mathrm{d}y)$. We then have, for $\lambda \in \mathbb{C}^N$, $z \in \mathcal{P}_d^N$, and $x \in \mathcal{T}_d^{N-1}$,

$$\begin{aligned}
\Sigma_\lambda^N(z; \mathrm{d}x) &= K_{\lambda_N}^N(z; \mathrm{d}x^{N-1}) K_{\lambda_{N-1}}^{N-1}(x^{N-1}; \mathrm{d}x^{N-2}) \cdots K_{\lambda_2}^2(x^2; \mathrm{d}x^1)\psi_\lambda^1(x^1) \\
&= K_{\lambda_N}^N(z; \mathrm{d}x^{N-1}) \Sigma_{(\lambda_1, \ldots, \lambda_{N-1})}^{N-1}(x^{N-1}; \mathrm{d}x^{1:(N-2)}). \qquad (2.9)
\end{aligned}$$

This yields a recursive definition of Whittaker functions:

$$\psi_\lambda^N(z) = \begin{cases} |z|^{-\lambda} & N = 1, \\ K_{\lambda_N}^N \psi_{(\lambda_1, \ldots, \lambda_{N-1})}^{N-1}(z) = K_{\lambda_N}^N K_{\lambda_{N-1}}^{N-1} \cdots K_{\lambda_2}^2 \psi_{\lambda_1}^1(z) & N \geq 2. \end{cases} \qquad (2.10)$$

2.2 A generalisation of Whittaker functions

We now introduce a generalisation of Whittaker functions of matrix arguments, which will naturally emerge in Sect. 4.3 and, in the scalar case $d = 1$, corresponds to the one considered in [35]. These generalised Whittaker functions are integrals over *trapezoidal* arrays of positive definite matrices, similarly to how the Whittaker functions of Sect. 2.1 are defined as integrals over triangular arrays.
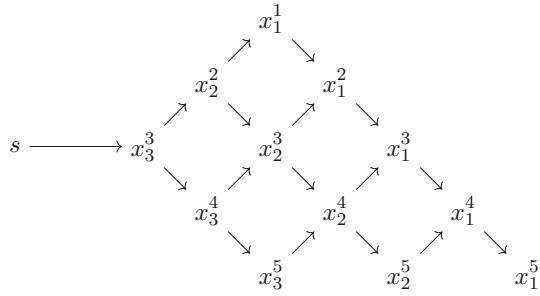
Let $n \geq N \geq 1$ and denote by

$$\mathcal{T}_d^{N,n} := \mathcal{P}_d \times \mathcal{P}_d^2 \times \cdots \times \mathcal{P}_d^N \times \underbrace{\mathcal{P}_d^N \times \cdots \times \mathcal{P}_d^N}_{n-N \text{ times}}$$

the set of trapezoidal arrays

$$x = (x^1, \ldots, x^n) = (x_j^i : 1 \leq i \leq n, \, 1 \leq j \leq i \wedge N), \qquad (2.11)$$

Fig. 2 Graphical representation of a trapezoidal array $x \in \mathcal{T}_d^{N,n}$ as in (2.11), for $N = 3$ and $n = 5$. The additional variable s appears in the definition (2.13) of the energy function $\Phi_s^{N,n}(x)$, in which every summand $\mathrm{tr}[ab^{-1}]$ corresponds to an arrow pointing from a to b in the figure

with ith row $x^i = (x_1^i, \ldots, x_{i \wedge N}^i) \in \mathcal{P}_d^{i \wedge N}$, for $1 \leq i \leq n$ (here $i \wedge N$ denotes the minimum between i and N). For $\lambda \in \mathbb{C}^n$, $x \in \mathcal{T}_d^{N,n}$ and $s \in \mathcal{P}_d$, let

$$
\Delta_\lambda^{N,n}(x) := |x_1^1|^{-\lambda_1} \prod_{i=2}^{N} \left(\frac{|x_1^i \cdots x_i^i|}{|x_1^{i-1} \cdots x_{i-1}^{i-1}|} \right)^{-\lambda_i} \prod_{i=N+1}^{n} \left(\frac{|x_1^i \cdots x_N^i|}{|x_1^{i-1} \cdots x_N^{i-1}|} \right)^{-\lambda_i},
$$
(2.12)

$$
\Phi_s^{N,n}(x) := \mathrm{tr}\left[s(x_N^N)^{-1} \right] + \sum_{i=1}^{n-1} \left(\sum_{j=1}^{i \wedge (N-1)} \mathrm{tr}\left[x_{j+1}^{i+1}(x_j^i)^{-1} \right] + \sum_{j=1}^{i \wedge N} \mathrm{tr}\left[x_j^i(x_j^{i+1})^{-1} \right] \right).
$$
(2.13)

See Fig. 2 for a graphical representation of the array (2.11) and of the energy function $\Phi_s^{N,n}$. For $z \in \mathcal{P}_d^N$, let $\mathcal{T}_d^{N,n}(z) \subset \mathcal{T}_d^{N,n}$ be the set of all trapezoidal arrays x with nth row $x^n = z$. For $n \geq N$, $\lambda \in \mathbb{C}^n$, $s \in \mathcal{P}_d$ and $z \in \mathcal{P}_d^N$, we define

$$
\psi_{\lambda;s}^{N,n}(z) := \int_{\mathcal{T}_d^{N,n}(z)} \left(\prod_{i=1}^{n-1} \prod_{j=1}^{i \wedge N} \mu(\mathrm{d}x_j^i) \right) \Delta_\lambda^{N,n}(x) \, \mathrm{e}^{-\Phi_s^{N,n}(x)}.
$$
(2.14)

Notice that, if $s = 0_d$ and $n = N$, $\psi_{\lambda;0}^{N,N} = \psi_\lambda^N$ corresponds to the Whittaker function defined in (2.4). The absolute convergence of the integral in (2.14), for all $\lambda \in \mathbb{C}^n$, can be shown by adapting the proof of [33, Prop. 6-(i)].

Let us now give an equivalent representation of these generalised Whittaker functions. The following kernel will play a central role in this work. For $a \in \mathbb{C}$ and $z, \tilde{z} \in \mathcal{P}_d^N$, set

$$
P_a^N(z; \tilde{z}) := \left(\prod_{i=1}^{N-1} \mathrm{e}^{-\mathrm{tr}[\tilde{z}_{i+1} z_i^{-1}]} \right) \prod_{j=1}^{N} \left| z_j \tilde{z}_j^{-1} \right|^a \mathrm{e}^{-\mathrm{tr}\left[z_j \tilde{z}_j^{-1} \right]}.
$$
(2.15)

We will see $P_a^N(z; \tilde{z})$ as a measure in either of the two arguments, defining

$$P_a^N(z; \mathrm{d}\tilde{z}) := P_a^N(z; \tilde{z})\mu^{\otimes N}(\mathrm{d}\tilde{z}) \quad \text{and} \quad \bar{P}_a^N(z; \mathrm{d}\tilde{z}) := P_a^N(\tilde{z}; z)\mu^{\otimes N}(\mathrm{d}\tilde{z}).$$
$$(2.16)$$

We then have

$$\psi_{\lambda;s}^{N,n}(z) := \begin{cases} \mathrm{e}^{-\operatorname{tr}[sz_N^{-1}]}\psi_\lambda^N(z) & n = N, \\ \bar{P}_{\lambda_n}^N \psi_{(\lambda_1,\dots,\lambda_{n-1});s}^{N,n-1}(z) = \bar{P}_{\lambda_n}^N \bar{P}_{\lambda_{n-1}}^N \cdots \bar{P}_{\lambda_{N+1}}^N \psi_{(\lambda_1,\dots,\lambda_N);s}^{N,N}(z) & n > N. \end{cases}$$
$$(2.17)$$

We also record here two relations between the kernels (2.8) and (2.15), which follow directly from the definitions:

$$K_a^N(z; y) = |s|^{-a} \mathrm{e}^{\operatorname{tr}[sz_N^{-1}]} P_a^N(y_1, \dots, y_{N-1}, s; z),$$
$$(2.18)$$

$$K_a^N(z; y) = |z_N|^{-a} \mathrm{e}^{-\operatorname{tr}[z_N y_{N-1}^{-1}]} P_a^{N-1}(y; z_1, \dots, z_{N-1}),$$
$$(2.19)$$

for $y = (y_1, \dots, y_{N-1}) \in \mathcal{P}_d^{N-1}$, $s \in \mathcal{P}_d$, and $z = (z_1, \dots, z_N) \in \mathcal{P}_d^N$. Taking $a = \lambda_N$ in (2.19), multiplying both sides by $\psi_{(\lambda_1,\dots,\lambda_{N-1})}^{N-1}(y)$, integrating over $\mathcal{P}_d^{N-1}$ with respect to $\mu^{\otimes(N-1)}(\mathrm{d}y)$, and using (2.10) and (2.17), we obtain the identity

$$\psi_\lambda^N(z) = |z_N|^{-\lambda_N} \psi_{\lambda;z_N}^{N-1,N}(z_1, \dots, z_{N-1}).$$
$$(2.20)$$

Remark 2.1 Let us mention that we anticipate the function $\psi_{\lambda;s}^{N,n}$ to be symmetric in the parameters $\lambda_1, \dots, \lambda_n$. This is not obvious from the definition, but it is suggested by an integral identity of Whittaker functions of matrix arguments that will be proven later on (see (4.1)). As argued in [33, § 7.1], this symmetry is true at least in the case $N = n = 2$. Moreover, it is known for $d = 1$ and arbitrary n, N; see, for example, [27], [21] and [35, pp. 369–370].

3 Markov dynamics

In this section, we define a Markov process X on triangular arrays, which can be viewed as a system of interacting random walks on $\mathcal{P}_d$. Next, we prove intertwining relations between certain transition kernels related to this process. This implies, via the theory of Markov functions, that, under certain random initial configurations, the bottom edge of the triangular process X has an autonomous stochastic evolution. A consequence of these results is that Whittaker functions of matrix arguments are eigenfunctions of certain integral operators and, thereupon, admit a Feynman–Kac interpretation.

3.1 Interacting Markov dynamics on triangular arrays

Let $\mathcal{O}_d$ be the real orthogonal group in dimension d. Recall that a random matrix Y in $\mathcal{P}_d$ is said to be $\mathcal{O}_d$-invariant (or orthogonally invariant) if $k^\top Y k$ has the same distribution of Y, for every $k \in \mathcal{O}_d$.

Definition 3.1 Let $(W(n))_{n \geq 1}$ be a family of independent and $\mathcal{O}_d$-invariant random matrices in $\mathcal{P}_d$. The GL_d-*invariant random walk on* $\mathcal{P}_d$ with initial state $r \in \mathcal{P}_d$ and increments $(W(n))_{n \geq 1}$ is the $\mathcal{P}_d$-valued process $R = (R(n))_{n \geq 0}$ such that $R(0) = r$ and

$$R(n) := T_{R(n-1)}(W(n)) = R(n-1)^{1/2} W(n) R(n-1)^{1/2}, \qquad n \geq 1. \quad (3.1)$$

Remark 3.2 The random walk R of Definition 3.1 is indeed GL_d-*invariant*, in the sense that the conjugated walk $(g^\top R(n)g)_{n \geq 0}$ has the same transition kernels for any choice of $g \in \mathrm{GL}_d$ (cf. [1, § 3]). Instead of (3.1), one could consider a different process through the alternative symmetrisation

$$R'(n) := T_{W(n)}(R'(n-1)) = W(n)^{1/2} R'(n-1) W(n)^{1/2}.$$

One can check that the resulting random walk R' is $\mathcal{O}_d$-invariant, but in general not GL_d-invariant. In principle, one could proceed to obtain analogous results to those presented in the present article using this alternative symmetrisation (for a similar approach in the continuous Brownian setting, see [33, Prop. 3.5]). However, from our point of view, the choice (3.1) is the most natural and leads to more explicit transition kernels throughout.

It is well known that the Wishart distribution (1.13) and the inverse Wishart distribution (1.14) are $\mathcal{O}_d$-invariant. In this article, we will focus on GL_d-invariant random walks with inverse Wishart increments.

Recall from definition (2.15) that $P_a^1(z; \mathrm{d}\tilde{z}) = \left| z\tilde{z}^{-1} \right|^a \mathrm{e}^{-\operatorname{tr}[z\tilde{z}^{-1}]} \mu(\mathrm{d}\tilde{z})$ for $a \in \mathbb{C}$. Using a straightforward change of variables, we see that, if $\Re(a) > \frac{d-1}{2}$,

$$\int_{\mathcal{P}_d} P_a^1(z; \mathrm{d}\tilde{z}) = \Gamma_d(a) \qquad \text{for any } z \in \mathcal{P}_d. \quad (3.2)$$

Define then the renormalised kernel

$$\overline{P}_a^1(z; \mathrm{d}\tilde{z}) := \frac{1}{\Gamma_d(a)} P_a^1(z; \mathrm{d}\tilde{z}). \quad (3.3)$$

It is immediate to see that the (time-homogeneous) GL_d-invariant random walk on $\mathcal{P}_d$ with inverse Wishart increments of parameter $a > \frac{d-1}{2}$ has transition kernel $\overline{P}_a^1$.

We now define a discrete-time Markov process $X = (X(n))_{n \geq 0}$ on the set $\mathcal{T}_d^N$ of height-N triangular arrays whose components are elements of $\mathcal{P}_d$.

Definition 3.3 Fix a sequence of real parameters $\alpha = (\alpha(n))_{n \geq 1}$, an integer $N \geq 1$, and a real N-tuple $\beta = (\beta^1, \dots, \beta^N)$ such that $\alpha(n) + \beta^i > (d-1)/2$ for all n, i. Denote by $\alpha(n) + \beta$ the N-tuple $(\alpha(n) + \beta^1, \dots, \alpha(n) + \beta^N)$. For $n \geq 1$ and $1 \leq j \leq i \leq N$, let $W_j^i(n)$ be an inverse Wishart random matrix with parameter $\alpha(n) + \beta^i$ (the same parameter across j); assume further that all these random matrices are independent of each other. We define the process $X = (X(n))_{n \geq 0}$, where $X(n) = (X_j^i(n))_{1 \leq j \leq i \leq N}$ is a random element of $\mathcal{T}_d^N$, as follows: given an initial state $X(0)$ in $\mathcal{T}_d^N$, for $n \geq 1$ we set recursively

$$
X_j^i(n) := \begin{cases}
T_{X_1^1(n-1)}(W_1^1(n)) & 1 = j = i \\
T_{X_1^{i-1}(n) + X_1^i(n-1)}(W_1^i(n)) & 1 = j < i \leq N \\
\left[X_{i-1}^{i-1}(n-1)^{-1} + T_{X_i^i(n-1)}(W_i^i(n))^{-1} \right]^{-1} & 1 < j = i \leq N \\
\left[X_{j-1}^{i-1}(n-1)^{-1} + T_{X_j^{i-1}(n) + X_j^i(n-1)}(W_j^i(n))^{-1} \right]^{-1} & 1 < j < i \leq N
\end{cases}
\tag{3.4}
$$

The i-tuple $X^i := (X_1^i, \dots, X_i^i)$ will be referred to as the ith *row* of X.

The fact that each $X_j^i(n)$ takes values in $\mathcal{P}_d$ follows by standard properties of positive definite matrices (cf. Sect. 1.2). Notice that, adopting the convention $X_0^i(n)^{-1} = X_{i+1}^i(n) = 0_d$ for all $i \geq 0$ and $n \geq 0$, then the last formula in (3.4) can be taken as the definition of $X_j^i(n)$ for all $1 \leq j \leq i \leq N$.

The dynamic on $\mathcal{T}_d^N$ defined by (3.4) implies that the 'top particle' X_1^1 evolves as a GL_d-invariant random walk in $\mathcal{P}_d$ with inverse Wishart increments $(W_1^1(n))_{n \geq 1}$.

Furthermore, the 'right edge' process $(X_1^1, X_1^2, \dots, X_1^N)$ equals the system $(Z^1, \dots, Z^N)$ of random particles in $\mathcal{P}_d$ with one-sided interactions defined in (1.8)–(1.9), where the random weight $V^i(n)$ equals $W_1^i(n)$.

The 'left edge' process $(X_1^1, X_2^2, \dots, X_N^N)$ also evolves as a system of particles in $\mathcal{P}_d$ with one-sided interactions, as we now explain. Set $L^i(n) := X_i^i(n)^{-1}$ and $U^i(n) := W_i^i(n)^{-1}$ for all $1 \leq i \leq N$ and $n \geq 0$. Then, $U^i(n)$ has the Wishart distribution with parameter $\alpha(n) + \beta^i$, and the process $L = (L^1, \dots, L^N)$ satisfies the recursions

$$
L^1(n) = T_{L^1(n-1)}(U^1(n)), \tag{3.5}
$$
$$
L^i(n) = L^{i-1}(n-1) + T_{L^i(n-1)}(U^i(n)), \quad 2 \leq i \leq N. \tag{3.6}
$$

Under the (singular) initial configuration

$$
L^1(0) = I_d, \qquad L^2(0) = \cdots = L^N(0) = 0_d, \tag{3.7}
$$

one can see by induction that $L^i(n) = 0_d$ for all $n < i - 1$ and $L^i(i-1) = I_d$, while $L^i(i)$ reduces to a sum of independent Wishart matrices:

$$
L^i(i) = U^1(1) + U^2(2) + \cdots + U^i(i), \quad 1 \leq i \leq N.
$$

In particular, $L^i(i)$ has the Wishart distribution with parameter $\sum_{j=1}^{i}(\alpha(j) + \beta^j)$.

Remark 3.4 We make a few remarks about various specialisations of the process X and related Markov dynamics:

(i) The interacting diffusion model on positive definite matrices studied in [33] (see also [32, § 9] for the $d = 1$ case) can be regarded as a continuous-time analogue of the process X defined in (3.4).

(ii) It seems that even the $d = 1$ case of the dynamic (3.4) has not been explicitly considered elsewhere. It is related, even though not identical, to the process constructed in [16] via the geometric Robinson–Schensted–Knuth correspondence; see the discussion in the introduction for further details.

(iii) For $d = 1$, under the 'step' initial configuration, the right edge can be regarded as a process of log-gamma polymer partition functions; see (1.6)–(1.7) and the discussion therein.

(iv) For $d = 1$, under the 'step' initial configuration (3.7), the left edge can be regarded as a process of strict-weak polymer partition functions in a gamma environment, studied in [17, 34]. A strict-weak path is a lattice path π that, at each lattice site (m, k), is allowed to head either horizontally to the right to $(m + 1, k)$ or diagonally up-right to $(m + 1, k + 1)$. It is easily seen that the process L defined in (3.5)–(3.6), in the $d = 1$ case, takes the closed form expression

$$L^i(n) = \sum_{\pi} \prod_{e \in \pi} d_e \,, \tag{3.8}$$

where the sum is over all strict-weak paths π from $(0, 1)$ to (n, i), the product is over all edges e in the path π, and d_e is a weight attached to the edge e and defined as follows: $d_e := 1$ if e is a diagonal edge from (m, k) to $(m+1, k+1)$; $d_e := U^k(m + 1)$ (gamma distributed with parameter $\alpha(m + 1) + \beta^k$) if e is a horizontal edge from (m, k) to $(m+1, k)$. Formula (3.8) defines the *strict-weak polymer partition function*.

(v) The $d = 1$ case of (3.4) is a 'positive temperature' analogue (equivalently, a $(+, \times)$ version) of the process defined by

$$\mathcal{X}_j^i(n) := \min\left(\mathcal{X}_{j-1}^{i-1}(n - 1), \max\left(\mathcal{X}_j^{i-1}(n), \mathcal{X}_j^i(n - 1)\right) + \mathcal{W}_j^i(n)\right) \,,$$

where $\mathcal{W}_j^i(n)$ are non-negative random variables representing jumps to the right (see e.g. [44]). Roughly speaking, particle $\mathcal{X}_j^i$ performs a random walk subject to certain interactions with other particles: it is pushed by $\mathcal{X}_j^{i-1}$ and blocked by $\mathcal{X}_{j-1}^{i-1}$.

(vi) Besides [44], other works [7, 8, 11, 31, 43] studied, in various discrete and continuous settings, similar push-and-block dynamics on Gelfand–Tsetlin patterns driven by random updates with $N(N + 1)/2$ degrees of freedom. In particular, again in the case $d = 1$, the process X should correspond to a certain $q \to 1$ scaling limit of the q-Whittaker processes studied in [7, 11].

Motivated to obtain the explicit Markovian evolution of X, we now introduce the following kernels. For $a \in \mathbb{C}$, $y = (y_1, \ldots, y_{N-1}) \in \mathcal{P}_d^{N-1}$, $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_{N-1}) \in \mathcal{P}_d^{N-1}$, $z = (z_1, \ldots, z_N) \in \mathcal{P}_d^N$, and $\tilde{z} = (\tilde{z}_1, \ldots, \tilde{z}_N) \in \mathcal{P}_d^N$, we set

$$Q_a^N(y, \tilde{y}, z; d\tilde{z}) := \prod_{j=1}^{N} \left| (\tilde{y}_j + z_j) \left(\tilde{z}_j^{-1} - y_{j-1}^{-1} \right) \right|^a e^{-\operatorname{tr}\left[(\tilde{y}_j + z_j)\left(\tilde{z}_j^{-1} - y_{j-1}^{-1} \right) \right]}$$

$$\left| I_d - \tilde{z}_j y_{j-1}^{-1} \right|^{-\frac{d+1}{2}} \mathbb{1}_{\mathcal{P}_d}\left(\tilde{z}_j^{-1} - y_{j-1}^{-1} \right) \mu(d\tilde{z}_j), \qquad (3.9)$$

with the convention $y_0^{-1} = \tilde{y}_N = 0$. Moreover, for $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$, we set

$$\Pi_\lambda^N(x; d\tilde{x})$$
$$:= \begin{cases} P_\lambda^1(x; d\tilde{x}) & \text{if } N = 1, \\ \Pi_{(\lambda_1, \ldots, \lambda_{N-1})}^{N-1}(x^{1:(N-1)}; d\tilde{x}^{1:(N-1)}) \, Q_{\lambda_N}^N(x^{N-1}, \tilde{x}^{N-1}, x^N; d\tilde{x}^N) & \text{if } N \geq 2, \end{cases}$$
$$(3.10)$$

where $x \in \mathcal{T}_d^N$ (resp., $\tilde{x} \in \mathcal{T}_d^N$) is a height-$N$ triangular array of $d \times d$ positive definite matrices with ith row $x^i \in \mathcal{P}_d^i$ (resp., $\tilde{x}^i \in \mathcal{P}_d^i$), according to the notation of Sect. 2.1. One can show (an analogous computation is made in the proof of Prop. 3.6) that, if $\Re(a) > (d-1)/2$,

$$\int_{\mathcal{P}_d^N} Q_a^N(y, \tilde{y}, z; d\tilde{z}) = \Gamma_d(a)^N \qquad \text{for any } y, \tilde{y} \in \mathcal{P}_d^{N-1} \text{ and } z \in \mathcal{P}_d^N .$$
$$(3.11)$$

Using (3.2) and (3.11), we see that, if $\Re(\lambda_i) > (d-1)/2$ for all i, then

$$\int_{\mathcal{T}_d^N} \Pi_\lambda^N(x; d\tilde{x}) = \prod_{i=1}^{N} \Gamma_d(\lambda_i)^i \qquad \text{for any } x \in \mathcal{T}_d^N . \qquad (3.12)$$

Therefore, under the above conditions on the parameters, one can renormalise these kernels, so that they integrate to 1:

$$\overline{Q}_a^N(y, \tilde{y}, z; d\tilde{z}) := \frac{1}{\Gamma_d(a)^N} Q_a^N(y, \tilde{y}, z; d\tilde{z}), \qquad (3.13)$$

$$\overline{\Pi}_\lambda^N(x; d\tilde{x}) := \frac{1}{\prod_{i=1}^{N} \Gamma_d(\lambda_i)^i} \Pi_\lambda^N(x; d\tilde{x}). \qquad (3.14)$$

The following result can be easily verified using the construction of X in Definition 3.3.

Proposition 3.5 *Let X as in Definition 3.3. Then, the conditional distribution of $X^N(n)$ given $X^{N-1}(n-1) = y$, $X^{N-1}(n) = \tilde{y}$ and $X^N(n-1) = z$, is $\overline{Q}_{\alpha(n)+\beta^N}^N(y, \tilde{y}, z; \cdot)$.*

Consequently, the process $X = (X(n))_{n \geq 0}$ is a time-inhomogeneous Markov process with state space T_d^N and time-n transition kernel $\overline{\Pi}_{\alpha(n)+\beta}^N$.

3.2 Intertwining relations

We will now show that the Markov dynamic on X (see Definition 3.3), when started from an appropriate random initial state, induces an autonomous Markov dynamic on the Nth row, or 'bottom edge', of X. This will be a consequence of an intertwining relation between kernels through the theory of Markov functions, which is reviewed in Appendix B for the reader's convenience.

Let $N \geq 2$ and $a, b \in \mathbb{C}$. Recalling the definitions (2.8) and (2.15) of the kernels K_b^N and P_a^N, respectively, and denoting by δ the Dirac delta kernel on $\mathcal{P}_d^N$, let us set

$$\tilde{K}_b^N(z; \mathrm{d}y \, \mathrm{d}\tilde{z}) := \delta(z; \mathrm{d}\tilde{z}) \, K_b^N(\tilde{z}; \mathrm{d}y), \tag{3.15}$$

$$\Lambda_{a,b}^N(y, z; \mathrm{d}\tilde{y} \, \mathrm{d}\tilde{z}) := P_a^{N-1}(y; \mathrm{d}\tilde{y}) \, Q_{a+b}^N(y, \tilde{y}, z; \mathrm{d}\tilde{z}), \tag{3.16}$$

for $z, \tilde{z} \in \mathcal{P}_d^N$ and $y, \tilde{y} \in \mathcal{P}_d^{N-1}$. We then have the following intertwining relation.

Proposition 3.6 *Let $N \geq 2$ and $a, b \in \mathbb{C}$ such that $\Re(a + b) > (d - 1)/2$. Then,*

$$\tilde{K}_b^N \Lambda_{a,b}^N = \Gamma_d(a + b)^{N-1} P_a^N \tilde{K}_b^N \tag{3.17}$$

holds as an equality between kernels from $\mathcal{P}_d^N$ to $\mathcal{P}_d^{N-1} \times \mathcal{P}_d^N$.

Proof We have to prove that $\tilde{K}_b^N \Lambda_{a,b}^N f(z) = \Gamma_d(a + b)^{N-1} P_a^N \tilde{K}_b^N f(z)$, for any suitable test function $f \colon \mathcal{P}_d^{N-1} \times \mathcal{P}_d^N \to \mathbb{R}$ and any $z \in \mathcal{P}_d^N$. Using (3.15), we see that this is equivalent to the identity

$$\int_{\mathcal{P}_d^{N-1}} K_b^N(z; \mathrm{d}y) \int_{\mathcal{P}_d^{N-1} \times \mathcal{P}_d^N} \Lambda_{a,b}^N(y, z; \mathrm{d}\tilde{y} \, \mathrm{d}\tilde{z}) f(\tilde{y}, \tilde{z})$$

$$= \Gamma_d(a + b)^{N-1} \int_{\mathcal{P}_d^N} P_a^N(z; \mathrm{d}\tilde{z}) \int_{\mathcal{P}_d^{N-1}} K_b^N(\tilde{z}; \mathrm{d}\tilde{y}) f(\tilde{y}, \tilde{z}). \tag{3.18}$$

Using the definitions of K_b^N and $\Lambda_{a,b}^N$, we obtain, after some rearrangements and cancellations, that the left-hand side of (3.18) equals

$$\int_{\mathcal{P}_d^{N-1}} \mu^{\otimes(N-1)}(\mathrm{d}y) \int_{\mathcal{P}_d^{N-1}} \mu^{\otimes(N-1)}(\mathrm{d}\tilde{y}) \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}\tilde{z}) f(\tilde{y}, \tilde{z})$$

$$\prod_{i=1}^{N-1} \left(|\tilde{y}_i|^{-a} \, |y_i - \tilde{z}_{i+1}|^{a+b} \, \mathrm{e}^{-\mathrm{tr}\left[y_i(z_i^{-1} + \tilde{y}_i^{-1}) \right]} \left| (y_i - \tilde{z}_{i+1}) \, y_i^{-1} \right|^{-\frac{d+1}{2}} \mathbb{1}_{\mathcal{P}_d} \left(\tilde{z}_{i+1}^{-1} - y_i^{-1} \right) \right)$$

$$\prod_{j=1}^{N} \left(|z_j|^{-b} \left| (\tilde{y}_j + z_j) \, \tilde{z}_j^{-1} \right|^{a+b} \mathrm{e}^{-\mathrm{tr}\left[(\tilde{y}_j + z_j) \tilde{z}_j^{-1} \right]} \right),$$

with the usual convention $\tilde{y}_N = 0$. By interchanging the order of integration, we see that the latter display equals

$$
\int_{\mathcal{P}_d^{N-1}} \mu^{\otimes(N-1)}(d\tilde{y}) \int_{\mathcal{P}_d^N} \mu^{\otimes N}(d\tilde{z}) f(\tilde{y}, \tilde{z}) \prod_{i=1}^{N-1} \left(|\tilde{y}_i|^{-a} \, \mathfrak{J}(z_i, \tilde{y}_i, \tilde{z}_{i+1}) \right)
$$

$$
\prod_{j=1}^{N} \left(|z_j|^{-b} \left| (\tilde{y}_j + z_j) \, \tilde{z}_j^{-1} \right|^{a+b} e^{-\operatorname{tr}\left[(\tilde{y}_j + z_j) \tilde{z}_j^{-1} \right]} \right),
$$

where $\mathfrak{J} \colon \mathcal{P}_d^3 \to \mathbb{C}$ is defined by

$$
\mathfrak{J}(u, v, w)
$$
$$
:= \int_{\mathcal{P}_d} \mu(ds) \, |s - w|^{a+b} \, e^{-\operatorname{tr}\left[s(u^{-1} + v^{-1}) \right]} \left| (s - w) \, s^{-1} \right|^{-\frac{d+1}{2}} \mathbb{1}_{\mathcal{P}_d}\left(w^{-1} - s^{-1} \right).
$$

By the properties of positive definite matrices (see Sect. 1.2), we have that $w^{-1} - s^{-1} \in \mathcal{P}_d$ if and only if $s - w \in \mathcal{P}_d$; moreover, for $w \in \mathcal{P}_d$, the latter condition is stronger than $s \in \mathcal{P}_d$. We then make the change of variables $s' := s - w$, which preserves the Lebesgue measure on the 'independent' entries of the symmetric matrix s, so that

$$
|s|^{\frac{d+1}{2}} \mu(ds) = \left| s' \right|^{\frac{d+1}{2}} \mu(ds').
$$

Therefore, we have

$$
\mathfrak{J}(u, v, w) = \int_{\mathcal{P}_d} \mu(ds') \left| s' \right|^{a+b} e^{-\operatorname{tr}\left[(s' + w)(u^{-1} + v^{-1}) \right]}.
$$

After the further, this time μ-preserving, change of variables $s'' := T_{u^{-1} + v^{-1}}(s')$, we obtain

$$
\mathfrak{J}(u, v, w) = \left| uv \, (u + v)^{-1} \right|^{a+b} e^{-\operatorname{tr}\left[w(u^{-1} + v^{-1}) \right]} \int_{\mathcal{P}_d} \mu(ds'') \left| s'' \right|^{a+b} e^{-\operatorname{tr}[s'']}
$$
$$
= \left| uv \, (u + v)^{-1} \right|^{a+b} e^{-\operatorname{tr}\left[w(u^{-1} + v^{-1}) \right]} \Gamma_d(a + b),
$$

where the gamma function is well defined since by hypothesis $\Re(a + b) > (d - 1)/2$. After a few cancellations, we then see that the left-hand side of (3.18) equals

$$
\Gamma_d(a + b)^{N-1} \int_{\mathcal{P}_d^{N-1}} \mu^{\otimes(N-1)}(d\tilde{y}) \int_{\mathcal{P}_d^N} \mu^{\otimes N}(d\tilde{z}) f(\tilde{y}, \tilde{z})
$$

$$
\times \prod_{i=1}^{N-1} \left(|\tilde{y}_i|^b \, e^{-\operatorname{tr}\left[\tilde{z}_{i+1} z_i^{-1} + \tilde{z}_{i+1} \tilde{y}_i^{-1} + \tilde{y}_i \tilde{z}_i^{-1} \right]} \right) \prod_{j=1}^{N} \left(|z_j|^a \, |\tilde{z}_j|^{-a-b} \, e^{-\operatorname{tr}\left[z_j \tilde{z}_j^{-1} \right]} \right).
$$

It now follows from the definitions that this equals the right-hand side of (3.18), thus concluding the proof. □

A simple inductive argument shows that the intertwining (3.17) can be extended to an intertwining that involves the Π-kernel (3.10) and the $\tilde{\Sigma}$-kernel (2.6). From now on, we fix $N \geq 1$, $a \in \mathbb{C}$ and $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$ such that $\Re(a + \lambda_i) > (d-1)/2$ for all i. As usual, we also use the notation $a + \lambda := (a + \lambda_1, \ldots, a + \lambda_N)$.

Corollary 3.7 *The intertwining relation*

$$\tilde{\Sigma}_\lambda^N \Pi_{a+\lambda}^N = \left(\prod_{i=1}^N \Gamma_d(a + \lambda_i)^{i-1} \right) P_a^N \tilde{\Sigma}_\lambda^N \tag{3.19}$$

holds as an equality between kernels from $\mathcal{P}_d^N$ to $\mathcal{T}_d^N$.

Proof Taking into account (2.6), it is immediate to see that (3.19) is equivalent to

$$\int_{\mathcal{T}_d^{N-1}} \Sigma_\lambda^N(z; \mathrm{d}x) \int_{\mathcal{T}_d^{N-1} \times \mathcal{P}_d^N} \Pi_{a+\lambda}^N(x, z; \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{z}) f(\tilde{x}, \tilde{z})$$

$$= \kappa_{a+\lambda} \int_{\mathcal{P}_d^N} P_a^N(z; \mathrm{d}\tilde{z}) \int_{\mathcal{T}_d^{N-1}} \Sigma_\lambda^N(\tilde{z}; \mathrm{d}\tilde{x}) f(\tilde{x}, \tilde{z}) \tag{3.20}$$

for all $z \in \mathcal{P}_d^N$ and test function $f : \mathcal{T}_d^{N-1} \times \mathcal{P}_d^N \to \mathbb{R}$, where we set

$$\kappa_{(\xi_1, \ldots, \xi_N)} := \prod_{i=1}^N \Gamma_d(\xi_i)^{i-1} \qquad \text{if} \quad \Re(\xi_i) > \frac{d-1}{2} \quad \text{for all } i \,.$$

To prove (3.20), we proceed by induction. For $N = 1$, (3.20) amounts to the identity

$$\psi_\lambda^1(z) \int_{\mathcal{P}_d} P_{a+\lambda}^1(z; \mathrm{d}\tilde{z}) f(\tilde{z}) = \int_{\mathcal{P}_d} P_a^1(z; \mathrm{d}\tilde{z}) \psi_\lambda^1(\tilde{z}) f(\tilde{z})$$

for $z \in \mathcal{P}_d$ and $f : \mathcal{P}_d \to \mathbb{R}$. Using (2.10) and (2.15), one can easily verify that the latter is true, as both sides equal $|z|^a \int_{\mathcal{P}_d} \mu(\mathrm{d}\tilde{z}) |\tilde{z}|^{-a-\lambda} e^{-\mathrm{tr}[z\tilde{z}^{-1}]} f(\tilde{z})$.

Let now $N \geq 2$ and $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{N-1})$. Assume by induction that

$$\int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(y; \mathrm{d}x) \int_{\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1}} \Pi_{a+\tilde{\lambda}}^{N-1}(x, y; \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}) g(\tilde{x}, \tilde{y})$$

$$= \kappa_{a+\tilde{\lambda}} \int_{\mathcal{P}_d^{N-1}} P_a^{N-1}(y; \mathrm{d}\tilde{y}) \int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(\tilde{y}; \mathrm{d}\tilde{x}) g(\tilde{x}, \tilde{y}) \tag{3.21}$$

for any $y \in \mathcal{P}_d^{N-1}$ and any test function $g \colon \mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1} \to \mathbb{R}$. Fix $z \in \mathcal{P}_d^N$ and $f \colon \mathcal{T}_d^{N-1} \times \mathcal{P}_d^N \to \mathbb{R}$ (which we view as $f \colon \mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1} \times \mathcal{P}_d^N \to \mathbb{R}$). Choosing

$$g(\tilde{x}, \tilde{y}) := \int_{\mathcal{P}_d^N} Q_{a+\lambda_N}^N(y, \tilde{y}, z; \mathrm{d}\tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z})$$

in (3.21) and integrating both sides with respect to the measure $K_{\lambda_N}^N(z; \cdot)$, we obtain

$$\int_{\mathcal{P}_d^{N-1}} K_{\lambda_N}^N(z; \mathrm{d}y) \int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(y; \mathrm{d}x) \int_{\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1}} \Pi_{a+\tilde{\lambda}}^{N-1}(x, y; \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y})$$

$$\int_{\mathcal{P}_d^N} Q_{a+\lambda_N}^N(y, \tilde{y}, z; \mathrm{d}\tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z})$$

$$= \kappa_{a+\tilde{\lambda}} \int_{\mathcal{P}_d^{N-1}} K_{\lambda_N}^N(z; \mathrm{d}y) \int_{\mathcal{P}_d^{N-1}} P_a^{N-1}(y; \mathrm{d}\tilde{y}) \int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(\tilde{y}; \mathrm{d}\tilde{x})$$

$$\int_{\mathcal{P}_d^N} Q_{a+\lambda_N}^N(y, \tilde{y}, z; \mathrm{d}\tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z}).$$

Using (2.9) and (3.10) for the left-hand side and (3.16) for the right-hand side, and interchanging the integration order, we then have

$$\int_{\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1}} \Sigma_{\lambda}^N(z; \mathrm{d}x\,\mathrm{d}y) \int_{\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1} \times \mathcal{P}_d^N} \Pi_{a+\lambda}^N(x, y, z; \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}\,\mathrm{d}\tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z})$$

$$= \kappa_{a+\tilde{\lambda}} \int_{\mathcal{P}_d^{N-1}} K_{\lambda_N}^N(z; \mathrm{d}y) \int_{\mathcal{P}_d^{N-1} \times \mathcal{P}_d^N} \Lambda_{a,\lambda_N}^N(y, z; \mathrm{d}\tilde{y}\,\mathrm{d}\tilde{z})$$

$$\left(\int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(\tilde{y}; \mathrm{d}\tilde{x}) f(\tilde{x}, \tilde{y}, \tilde{z}) \right)$$

$$= \kappa_{a+\tilde{\lambda}} \, \Gamma_d(a + \lambda_N)^{N-1} \int_{\mathcal{P}_d^N} P_a^N(z; \mathrm{d}\tilde{z}) \int_{\mathcal{P}_d^{N-1}} K_{\lambda_N}^N(\tilde{z}; \mathrm{d}\tilde{y})$$

$$\left(\int_{\mathcal{T}_d^{N-2}} \Sigma_{\tilde{\lambda}}^{N-1}(\tilde{y}; \mathrm{d}\tilde{x}) f(\tilde{x}, \tilde{y}, \tilde{z}) \right)$$

$$= \kappa_{a+\lambda} \int_{\mathcal{P}_d^N} P_a^N(z; \mathrm{d}\tilde{z}) \int_{\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1}} \Sigma_{\lambda}^N(\tilde{z}; \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}) f(\tilde{x}, \tilde{y}, \tilde{z}),$$

where the latter two equalities follow from (3.18) and (2.9), respectively. The identification $\mathcal{T}_d^{N-2} \times \mathcal{P}_d^{N-1} = \mathcal{T}_d^{N-1}$ concludes the proof of (3.20). $\qquad\square$

Recall now that the $\tilde{\Sigma}$-kernels generate Whittaker functions of matrix arguments, in the sense of (2.7). By integrating the intertwining relation (3.19) and using (3.12),

we immediately deduce that Whittaker functions are eigenfunctions of the *integral P*-operators:

Corollary 3.8 *We have*

$$P_a^N \psi_\lambda^N = \left(\prod_{i=1}^N \Gamma_d(a + \lambda_i) \right) \psi_\lambda^N. \tag{3.22}$$

We note that this complements the interpretation of the Whittaker functions ψ_λ^N, given in [33], as eigenfunctions of a *differential* operator, namely the Hamiltonian of a quantisation in $\mathcal{P}_d^N$ of the N-particle non-Abelian Toda chain.

For $x \in \mathcal{T}_d^N$ and $z, \tilde{z} \in \mathcal{P}_d^N$, we now define

$$\overline{\Sigma}_\lambda^N(z; dx) := \frac{1}{\psi_\lambda^N(z)} \tilde{\Sigma}_\lambda^N(z; dx), \tag{3.23}$$

$$\boldsymbol{P}_{a,\lambda}^N(z; d\tilde{z}) := \frac{1}{\prod_{i=1}^N \Gamma_d(a + \lambda_i)} \frac{\psi_\lambda^N(\tilde{z})}{\psi_\lambda^N(z)} P_a^N(z; d\tilde{z}). \tag{3.24}$$

It follows from (2.7) and (3.22) that the above kernels are normalised; therefore, they are Markov kernels when the parameters $a, \lambda_1, \ldots, \lambda_N$ are real. Notice that (3.24) may be seen as a Doob *h*-transform of the *P*-kernel (2.15). It is now immediate to deduce a renormalised version of (3.19):

Corollary 3.9 *The intertwining relation*

$$\overline{\Sigma}_\lambda^N \overline{\Pi}_{a+\lambda}^N = \boldsymbol{P}_{a,\lambda}^N \overline{\Sigma}_\lambda^N \tag{3.25}$$

holds as an equality between kernels from $\mathcal{P}_d^N$ to $\mathcal{T}_d^N$.

From a probabilistic point of view, (3.25) states that, for any fixed $z \in \mathcal{P}_d^N$, the two following update rules are equivalent: (i) starting the process X from a (random) initial configuration dictated by the intertwining kernel $\overline{\Sigma}(z; \cdot)$ and letting it evolve according to the dynamic $\overline{\Pi}$; and (ii) running the dynamic $\boldsymbol{P}$ on the bottom edge (started at z) and then updating the whole triangular array according to the intertwining kernel $\overline{\Sigma}$. The main result of this section is a precise account of this interpretation.

Theorem 3.10 *Let $X = (X(n))_{n \geq 0}$ be the Markov process on $\mathcal{T}_d^N$ as in Definition 3.3. Assume that, for an arbitrary $z \in \mathcal{P}_d^N$, the initial state $X(0)$ of X is distributed according to the measure $\overline{\Sigma}_\beta^N(z; \cdot)$. Then, the Nth row $X^N = (X^N(n))_{n \geq 0}$ is a time-inhomogeneous Markov process (in its own filtration) on the state space $\mathcal{P}_d^N$, with initial state z and time-n transition kernel $\boldsymbol{P}_{\alpha(n),\beta}^N$. Moreover, for any bounded measurable function $f : \mathcal{T}_d^N \to \mathbb{R}$ and $n \geq 0$, we have*

$$\mathbb{E}\left[f(X(n)) \,\Big|\, X^N(0), \ldots, X^N(n-1), X^N(n)\right] = \overline{\Sigma}_\beta^N f\left(X^N(n)\right) \qquad a.s.$$
(3.26)

Proof The statement is an application of Theorem B.1, where the state spaces are $S = \mathcal{T}_d^N$ and $T = \mathcal{P}_d^N$, and the function $\varphi \colon \mathcal{T}_d^N \to \mathcal{P}_d^N$ is the projection $\varphi(x) := x^N$ onto the Nth row of x, so that $X^N(n) = \varphi(X(n))$. Hypothesis (i) of Theorem B.1, i.e. the fact that $\overline{\Sigma}_\beta(z; \varphi^{-1}\{z\}) = 1$ for any $z \in \mathcal{P}_d^N$, holds because, by definition, the measure $\overline{\Sigma}_\beta(z; \cdot)$ is supported on the set $\mathcal{T}_d^N(z)$ of height-N triangular arrays with Nth row equal to z. On the other hand, by Prop. 3.5, the time-n transition kernel of X is $\overline{\Pi}_{\alpha(n)+\beta}^N$. Therefore, in this case, hypothesis (ii) of Theorem B.1 reads as the set of intertwining relations

$$\overline{\Sigma}_\beta^N \overline{\Pi}_{\alpha(n)+\beta}^N = \boldsymbol{P}_{\alpha(n),\beta}^N \overline{\Sigma}_\beta^N \qquad \text{for all } n \geq 1 .$$

These follow from Corollary 3.9.   □

Remark 3.11 By letting N vary, it is immediate to deduce from Theorem 3.10 that *every* row of X evolves as a Markov process in its own filtration, under an appropriate (random) initial configuration on the previous rows. Therefore, the focus on the Nth row should only be seen as a convenient choice.

3.3 Feynman–Kac interpretation

Here we provide a Feynman–Kac type interpretation of Whittaker functions based on the eigenfunction equation (3.22). Our result should be compared to the one obtained in [33, Prop. 9] in the continuous setting of Brownian particles.

Definition 3.12 Let $\lambda \in \mathbb{R}^N$ with

$$\min(\lambda_1, \lambda_2 - \lambda_1, \ldots, \lambda_N - \lambda_{N-1}) > \frac{d-1}{2} .$$

Let $y \in \mathcal{P}_d^N$. We define $Y = (Y(n))_{n \geq 0} = (Y_1(n), \ldots, Y_N(n))_{n \geq 0}$ to be a process in $\mathcal{P}_d^N$ with independent components, such that each component $Y_i = (Y_i(n))_{n \geq 0}$ is a GL_d-invariant random walk on $\mathcal{P}_d$ with initial state $Y_i(0) = y_i$ and inverse Wishart increments with parameter λ_i.

Recalling (3.3), Y is then a time-homogeneous Markov process starting at y with transition kernel

$$\Theta_\lambda^N(z; \mathrm{d}\tilde{z}) := \overline{P}_{\lambda_1}^1(z_1; \mathrm{d}\tilde{z}_1) \cdots \overline{P}_{\lambda_N}^1(z_N; \mathrm{d}\tilde{z}_N) \qquad \text{for } z, \tilde{z} \in \mathcal{P}_d^N .$$

For $z, \tilde{z} \in \mathcal{P}_d^N$, define the sub-Markov kernel

$$\hat{\Theta}_\lambda^N(z; \tilde{z}) := \mathrm{e}^{-V(z;\tilde{z})} \Theta_\lambda^N(z; \tilde{z}) ,$$
(3.27)

where V is the 'killing potential'

$$V(z; \tilde{z}) := \sum_{i=1}^{N-1} \text{tr} \left[\tilde{z}_{i+1} z_i^{-1} \right]. \tag{3.28}$$

Denote by $\mathbb{P}_y$ and $\mathbb{E}_y$ the probability and expectation, respectively, with respect to the law of Y with initial state y.

Theorem 3.13 *For all $y \in \mathcal{P}_d^N$, we have*

$$\psi_\lambda^N(y) = \prod_{1 \le i < j \le N} \Gamma_d(\lambda_j - \lambda_i) \left(\prod_{i=1}^{N} |y_i|^{-\lambda_i} \right) \mathbb{E}_y \left[e^{-\sum_{n=0}^{\infty} V(Y(n); Y(n+1))} \right]. \tag{3.29}$$

The main purpose of this subsection is to prove (3.29). In a nutshell, using a fairly standard martingale argument, we will show that the expectation in (3.29) is the unique solution to an eigenproblem; the latter is also, essentially, solved by Whittaker functions.

Lemma 3.14 *Fix an integer $\ell \ge 0$. For any $y \in \mathcal{P}_d^N$, we have*

$$\limsup_{n \to \infty} \frac{1}{n} \log V(Y(n); Y(n + \ell)) < 0 \qquad \mathbb{P}_y\text{-a.s.}$$

Remark 3.15 In particular Lemma 3.14 with $\ell = 1$ implies that the infinite series inside the expectation in (3.29) converges $\mathbb{P}_y$-a.s.

Proof of Lemma 3.14 Since

$$V(Y(n); Y(n + \ell)) = \sum_{i=1}^{N-1} \text{tr} \left[Y_{i+1}(n + \ell) Y_i(n)^{-1} \right],$$

it suffices to show that, for each $1 \le i \le N - 1$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \text{tr} \left[Y_{i+1}(n + \ell) Y_i(n)^{-1} \right] < 0 \qquad \mathbb{P}_y\text{-a.s.}$$

Let us record the following properties, which hold for any $a, b \in \mathcal{P}_d$:

- $\text{tr}[ab] \le \text{tr}[a] \, \text{tr}[b]$ (submultiplicativity of the trace);
- $\text{tr}[a] \le d \, \lambda_{\max}(a)$;
- $\lambda_{\max}(a^{-1}) = \lambda_{\min}(a)^{-1}$.

Here, $\lambda_{\max}$ and $\lambda_{\min}$ denote the maximum and minimum eigenvalue, respectively. Using these facts, we have, for $1 \le i \le N - 1$:

$$\text{tr} \left[Y_{i+1}(n + \ell) Y_i(n)^{-1} \right] \le \text{tr} \left[Y_{i+1}(n + \ell) \right] \text{tr} \left[Y_i(n)^{-1} \right] \le d^2 \frac{\lambda_{\max}(Y_{i+1}(n + \ell))}{\lambda_{\min}(Y_i(n))}.$$

Now, using for example [1, Corollary B.4], we have

$$\lim_{n\to\infty} \frac{1}{n} \log \lambda_{\max}(Y_{i+1}(n)) = -\psi\left(\lambda_{i+1} - \frac{d-1}{2}\right),$$

$$\lim_{n\to\infty} \frac{1}{n} \log \lambda_{\min}(Y_i(n)) = -\psi(\lambda_i),$$

$\mathbb{P}_y$-a.s., where ψ is the digamma function. These are the maximum (respectively, minimum) Lyapunov exponent of a GL_d-invariant random walk with inverse Wishart increments of parameter λ_{i+1} (respectively, λ_i). We then obtain

$$\limsup_{n\to\infty} \frac{1}{n} \log \operatorname{tr}\left[Y_{i+1}(n+\ell)Y_i(n)^{-1}\right] \le \psi(\lambda_i) - \psi\left(\lambda_{i+1} - \frac{d-1}{2}\right) < 0,$$

since the digamma function is strictly increasing and, by Definition 3.12, $\lambda_{i+1} - \lambda_i > (d-1)/2$. □

Lemma 3.16 *Let* $u\colon \mathcal{P}_d^N \to \mathbb{R}$ *be a measurable function such that*

(i) $\hat{\Theta}_\lambda^N u = u$ *(eigenfunction equation);*
(ii) *u is bounded (boundedness property);*
(iii) $u(y) \to 1$ *as* $V(y;y) \to 0$ *(boundary condition).*

Then, for all $y \in \mathcal{P}_d^N$,

$$u(y) = \mathbb{E}_y\left[e^{-\sum_{n=0}^{\infty} V(Y(n);Y(n+1))}\right].$$

Proof Consider the process Y as in Definition 3.12, with initial state $y \in \mathcal{P}_d^N$ and transition kernel Θ_λ^N. Denote by $(\mathcal{F}(k))_{k\ge 0}$ its natural filtration. It follows from the eigenfunction equation that

$$\mathbb{E}_y\left[u(Y(k+1))e^{-V(Y(k);Y(k+1))} \,\middle|\, \mathcal{F}(k)\right] = \int_{\mathcal{P}_d^N} \Theta_\lambda^N(Y(k);d\tilde{z})\,e^{-V(Y(k);\tilde{z})}\,u(\tilde{z})$$

$$= \hat{\Theta}_\lambda^N u(Y(k)) = u(Y(k)).$$

Therefore, the process $M = (M(k))_{k\ge 0}$ defined by

$$M(k) := \begin{cases} u(Y(0)) = u(y) & k = 0, \\ u(Y(k))\,e^{-\sum_{n=0}^{k-1} V(Y(n);Y(n+1))} & k \ge 1 \end{cases} \tag{3.30}$$

is an $(\mathcal{F}(k))_{k\ge 0}$-martingale. By the boundedness property, M is uniformly bounded and, thus, a uniformly integrable martingale. Therefore, M converges $\mathbb{P}_y$-a.s. and in 1-norm to a certain limit $M(\infty)$ and, for all $k \ge 0$, we have $M(k) = \mathbb{E}_y\left[M(\infty) \,\middle|\, \mathcal{F}(k)\right]$.

By Lemma 3.14 (with $\ell = 0$), we have $\lim_{k \to \infty} V(Y(k); Y(k)) = 0$, $\mathbb{P}_y$-a.s. The boundary condition then implies $\lim_{k \to \infty} u(Y(k)) = 1$, $\mathbb{P}_y$-a.s., whence

$$M(\infty) = \mathrm{e}^{-\sum_{n=0}^{\infty} V(Y(n); Y(n+1))} \; .$$

We conclude that, for any $y \in \mathcal{P}_d^N$,

$$u(y) = M(0) = \mathbb{E}_y[M(\infty)] = \mathbb{E}_y\left[\mathrm{e}^{-\sum_{n=0}^{\infty} V(Y(n); Y(n+1))} \right].$$

$\square$

Proof of Theorem 3.13 It was proven in [33, proof of Prop. 9] that the function

$$v(y) := \psi_\lambda^N(y) \prod_{i=1}^N |y_i|^{\lambda_i} \;, \qquad y \in \mathcal{P}_d^N \;,$$

is bounded and satisfies

$$\lim_{V(y;y) \to 0} v(y) = \prod_{1 \le i < j \le N} \Gamma_d(\lambda_j - \lambda_i) \,.$$

By Lemma 3.16, it then remains to prove that $\hat{\Theta}_\lambda^N v = v$. It follows from the definition (2.15) of the kernel P_a^N that

$$\hat{\Theta}_\lambda^N v(z) = \left(\prod_{i=1}^N \frac{|z_i|^{\lambda_i}}{\Gamma_d(\lambda_i)} \right) P_0^N \psi_\lambda^N(z)$$

for $z \in \mathcal{P}_d^N$. Using the eigenfunction equation (3.22), we see that the right-hand side above equals $v(z)$, as desired. $\square$

Corollary 3.17 *Under $\mathbb{P}_y$, we have the distributional equality*

$$\sum_{n=0}^{\infty} \mathrm{tr} \left[Y_2(n+1) Y_1(n)^{-1} \right] \overset{d}{=} \mathrm{tr} \left[a Z \right] \,, \tag{3.31}$$

where $a := y_1^{-1} y_2 y_1^{-1/2}$ and Z has the inverse Wishart distribution of parameter $\lambda_2 - \lambda_1$.

Proof We may assume that $N = 2$, so that $Y = (Y_1, Y_2)$ starts at $y = (y_1, y_2)$. Using Theorem 3.13 and the definition of Whittaker functions, we compute the Laplace transform of the left-hand side of (3.31) as

$$\mathbb{E}_{(y_1, y_2)} \left[e^{-s \sum_{n=0}^{\infty} \operatorname{tr}[Y_2(n+1)Y_1(n)^{-1}]} \right] = \mathbb{E}_{(y_1, sy_2)} \left[e^{-\sum_{n=0}^{\infty} \operatorname{tr}[Y_2(n+1)Y_1(n)^{-1}]} \right]$$

$$= \frac{|y_1|^{\lambda_1} |sy_2|^{\lambda_2}}{\Gamma_d(\lambda_2 - \lambda_1)} \int_{\mathcal{P}_d} \mu(\mathrm{d}x) \, |x|^{-\lambda_1} \left(\frac{|sy_1 y_2|}{|x|} \right)^{-\lambda_2} e^{-\operatorname{tr}[sy_2 x^{-1} + x y_1^{-1}]}$$

$$= \int_{\mathcal{P}_d} \mu(\mathrm{d}z) \, e^{-s \operatorname{tr}[y_1^{-1/2} y_2 y_1^{-1/2} z]} \frac{|z|^{-(\lambda_2 - \lambda_1)} e^{-\operatorname{tr}[z^{-1}]}}{\Gamma_d(\lambda_2 - \lambda_1)}$$

for $s \in \mathbb{R}$, where we used the change of variables $z = y_1^{1/2} x^{-1} y_1^{1/2}$. The last integral equals $\mathbb{E} \, e^{-s \operatorname{tr}[aZ]}$, where Z is inverse Wishart of parameter $\lambda_2 - \lambda_1$. We conclude that the two sides of (3.31) have the same Laplace transform and, hence, the same law. $\square$

Remark 3.18 Up to some technical details, identity (3.31) may be also deduced from the Dufresne type identity for a random walk on $\mathcal{P}_d$ proved in [1]. Let $(R(n))_{n \geq 0}$ be a GL_d-invariant random walk on $\mathcal{P}_d$ whose initial state $R(0)$ is an inverse Wishart matrix with parameter λ_2 and whose increments are Beta type II matrices with parameters λ_1 and λ_2 (see [1] for more details). It is then natural to expect that the eigenvalue processes of the two processes $(Y_1(n)^{-1/2} Y_2(n+1) Y_1(n)^{-1/2})_{n \geq 0}$ and $(a^{1/2} R(n) a^{1/2})_{n \geq 0}$, where $a = y_1^{-1} y_2 y_1^{-1/2}$, have the same law; this is certainly true at least in the case $d = 1$. By summing the traces of these two processes over all $n \geq 0$, [1, Theorem 4.10] would then immediately provide a proof of (3.31) that does *not* rely upon the Feynman–Kac formula (3.29). See [33, Lemma 8] for an analogous argument in the Brownian setting.

4 Fixed-time laws and matrix Whittaker measures

In this section, we first prove a Whittaker integral identity that allows us to introduce matrix Whittaker measures. We then obtain an asymptotic formula for a Whittaker function whose arguments go to zero or infinity in norm. Using the latter result, we next show that, for a certain singular initial state, matrix Whittaker measures appear naturally as the fixed-time laws of the bottom edge of the triangular process X introduced in Sect. 3.1. Finally, under the same singular initial state, we study the fixed-time law of the right edge and of the left edge of X.

4.1 Matrix Whittaker measures

Whittaker functions of matrix arguments satisfy a remarkable integral identity:

Theorem 4.1 Let $n \geq N \geq 1$. Let $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ and $\varrho = (\varrho_1, \ldots, \varrho_N) \in \mathbb{C}^N$ such that $\Re(\lambda_\ell + \varrho_i) > \frac{d-1}{2}$ for all $1 \leq \ell \leq n$, $1 \leq i \leq N$. Let $s \in \mathcal{P}_d$. Then,

$$\int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \psi_{\lambda; s}^{N, n}(z) \psi_\varrho^N(z) = |s|^{-\sum_{j=1}^N (\lambda_j + \varrho_j)} \prod_{\ell=1}^n \prod_{i=1}^N \Gamma_d(\lambda_\ell + \varrho_i). \quad (4.1)$$

The case $N = n$ of (4.1) was noted in [33, Prop. 10]; however, the identity did not play a key role in that article, and the details of the proof were not provided therein. Below we provide a proof of the general case $n \geq N$ that involves the generalised Whittaker functions introduced in Sect. 2.2.

In the scalar $d = 1$ setting, (4.1) goes back to [35, Corollary 3.5]. For $d = 1$ and $N = n$, it is equivalent to an identity that was originally found in the number theoretic literature [12, 40].

Theorem 4.1 can be also seen as an analogue, in the context of Whittaker functions, of the celebrated Cauchy–Littlewood identity for Schur functions. In the literature on symmetric functions, the latter is usually proved using either the determinantal structure of Schur functions (see [29, I-(4.3)]) or the Robinson–Schensted–Knuth correspondence, a combinatorial bijection (see [41, Theorem 7.12.1]). None of these tools is available, so far, in our context. To prove (4.1), we will rather proceed inductively, using the recursive definition of Whittaker functions and the eigenfunction equation (3.22). For the reader's convenience, we also include in Appendix A a proof of the classical Cauchy–Littlewood identity that similarly relies on an eigenfunction equation for Schur functions (which can be seen as a version of the so-called Pieri rule).

Proof of Theorem 4.1 We will prove (4.1) by induction on n. For a fixed integer $n \geq 1$, let S(n) be the statement that (4.1) holds for all N such that $n \geq N \geq 1$ and for any choice of λ and ϱ satisfying the assumptions of the theorem.

For $n = N = 1$ we have

$$
\int_{\mathcal{P}_d} \mu(\mathrm{d}z) \psi_{\lambda;s}^{1,1}(z) \psi_{\varrho}^1(z) = \int_{\mathcal{P}_d} \mu(\mathrm{d}z) \, \mathrm{e}^{-\operatorname{tr}[sz^{-1}]} |z|^{-\lambda_1} |z|^{-\varrho_1}
$$

$$
= |s|^{-\lambda_1-\varrho_1} \int_{\mathcal{P}_d} \mu(\mathrm{d}\tilde{z}) \, |\tilde{z}|^{-(\lambda_1+\varrho_1)} \, \mathrm{e}^{-\operatorname{tr}[\tilde{z}^{-1}]} = |s|^{-\lambda_1-\varrho_1} \, \Gamma_d(\lambda_1 + \varrho_1) \, ,
$$

where we have used the definitions of Whittaker functions and gamma functions and the μ-preserving change of variables $\tilde{z} = T_{s^{-1}}(z)$. This proves the base case S(1).

Suppose now by induction that S($n - 1$) holds for some fixed $n \geq 2$. To prove that S(n) holds, let us first prove that (4.1) is valid for all N such that $n > N \geq 1$. It follows from (2.17), Fubini's theorem, (2.16) and (3.22), that

$$
\int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \psi_{\lambda;s}^{N,n}(z) \psi_{\varrho}^N(z) = \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \big(\bar{P}_{\lambda_n}^N \psi_{(\lambda_1,\ldots,\lambda_{n-1});s}^{N,n-1} \big)(z) \psi_{\varrho}^N(z)
$$

$$
= \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}\tilde{z}) \psi_{(\lambda_1,\ldots,\lambda_{n-1});s}^{N,n-1}(\tilde{z}) \big(P_{\lambda_n}^N \psi_{\varrho}^N \big)(\tilde{z})
$$

$$
= \prod_{i=1}^N \Gamma_d(\lambda_n + \varrho_i) \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}\tilde{z}) \psi_{(\lambda_1,\ldots,\lambda_{n-1});s}^{N,n-1}(\tilde{z}) \psi_{\varrho}^N(\tilde{z}) \, .
$$

Since $n - 1 \geq N$, using the assumption S($n - 1$) in the latter integral we obtain (4.1).

To conclude S(n), we are left to prove the case $N = n$. Using (2.17), (2.10), Fubini's theorem, (2.18), (3.22), and (2.20), we have

$$
\int_{\mathcal{P}_d^n} \mu^{\otimes n}(\mathrm{d}z) \psi_{\lambda;s}^{n,n}(z) \psi_\varrho^n(z)
$$

$$
= \int_{\mathcal{P}_d^n} \mu^{\otimes n}(\mathrm{d}z) \, \mathrm{e}^{-\operatorname{tr}[sz_n^{-1}]} \left(K_{\lambda_n}^n \psi_{(\lambda_1,\dots,\lambda_{n-1})}^{n-1} \right)(z) \psi_\varrho^n(z)
$$

$$
= |s|^{-\lambda_n} \int_{\mathcal{P}_d^{n-1}} \mu^{\otimes(n-1)}(\mathrm{d}y) \psi_{(\lambda_1,\dots,\lambda_{n-1})}^{n-1}(y) \left(P_{\lambda_n}^n \psi_\varrho^n \right)(y_1,\dots,y_{n-1},s)
$$

$$
= |s|^{-\lambda_n} \left(\prod_{i=1}^n \Gamma_d(\lambda_n + \varrho_i) \right) \int_{\mathcal{P}_d^{n-1}} \mu^{\otimes(n-1)}(\mathrm{d}y) \psi_{(\lambda_1,\dots,\lambda_{n-1})}^{n-1}(y) \psi_\varrho^n(y_1,\dots,y_{n-1},s)
$$

$$
= |s|^{-\lambda_n - \varrho_n} \left(\prod_{i=1}^n \Gamma_d(\lambda_n + \varrho_i) \right) \int_{\mathcal{P}_d^{n-1}} \mu^{\otimes(n-1)}(\mathrm{d}y) \psi_{(\lambda_1,\dots,\lambda_{n-1})}^{n-1}(y) \psi_{\varrho;s}^{n-1,n}(y) .
$$

Recall that we have already proved (4.1) for all N such that $n > N \geq 1$. Applying this, for $N = n - 1$, to the latter integral, we conclude that (4.1) holds also for $N = n$. $\square$

Definition 4.2 For $n \geq N \geq 1$. Let $\lambda = (\lambda_1,\dots,\lambda_n) \in \mathbb{R}^n$ and $\varrho = (\varrho_1,\dots,\varrho_N) \in \mathbb{R}^N$ such that $\lambda_\ell + \varrho_i > \frac{d-1}{2}$ for all $1 \leq \ell \leq n$, $1 \leq i \leq N$. We call *matrix Whittaker measure with parameters λ and ϱ* the measure on $\mathcal{P}_d^N$ that is absolutely continuous with respect to $\mu^{\otimes N}(\mathrm{d}z)$ with density

$$
W_{\lambda,\varrho}^{N,n}(z) := \left(\prod_{\ell=1}^n \prod_{i=1}^N \frac{1}{\Gamma_d(\lambda_\ell + \varrho_i)} \right) \psi_{\lambda;I_d}^{N,n}(z) \psi_\varrho^N(z) \qquad \text{for } z \in \mathcal{P}_d^N , \qquad (4.2)
$$

where I_d is the $d \times d$ identity matrix. According to the usual convention, we also denote by $W_{\lambda,\varrho}^{N,n}(\mathrm{d}z)$ the measure itself.

By Theorem 4.1, (4.2) defines a probability distribution on $\mathcal{P}_d^N$. This extends the definition of matrix Whittaker measures given in [33, § 7.4], which corresponds to the case $n = N$:

$$
W_{\lambda,\varrho}^{N,N}(z) = \left(\prod_{\ell,i=1}^N \frac{1}{\Gamma_d(\lambda_\ell + \varrho_i)} \right) \mathrm{e}^{-\operatorname{tr}[z_N^{-1}]} \psi_\lambda^N(z) \psi_\varrho^N(z) . \qquad (4.3)
$$

4.2 Asymptotics of Whittaker functions

For any real $k > 0$, let

$$
r_j^i(k) := k^{2j-i-1} I_d \qquad \text{for } 1 \leq j \leq i \qquad (4.4)
$$

and let $r^i(k) := (r^i_1(k), \ldots, r^i_i(k))$. Our ultimate goal is to obtain the $k \to \infty$ leading order approximation of the Whittaker function $\psi^N_\lambda(r^N(k))$.

We rely on some results (Theorem 5.13 and Prop. 5.14) that we will prove, in a more general setting, in Sect. 5. With this purpose in mind, we use the graphical representations of the set of height-N triangular arrays $\mathcal{T}^N_d$ and of the energy function Φ^N, both involved in the definition of the Whittaker function (2.4) (see Fig. 1). Given $N \geq 2$, we set

$$V := \{(i, j) \in \mathbb{Z}^2 \colon 1 \leq j \leq i \leq N\}$$

and consider the finite graph $G = (V, E)$, where E consists of all (directed) edges $(i, j) \to (i + 1, j)$ and $(i + 1, j + 1) \to (i, j)$, for $1 \leq j \leq i \leq N - 1$. Then, $\mathcal{T}^N_d$ may be identified as the set $\mathcal{P}^V_d$ of arrays $x = (x_v)_{v \in V}$, where each $x_v \in \mathcal{P}_d$. Let also

$$\Gamma := \{(N, j) \colon 1 \leq j \leq N\}.$$

We may thus identify $z \in \mathcal{P}^N_d$ with $z \in \mathcal{P}^\Gamma_d$, so that the set $\mathcal{T}^N_d(z)$ of all height-N triangular arrays whose Nth row equals z coincides with the set $\mathcal{P}^V_d(z)$, according to the notation (5.4). Furthermore, the energy function (2.3) can be equivalently rewritten as

$$\Phi^N(x) = \sum_{\substack{(i,j),(k,\ell) \in V \colon \\ (i,j) \to (k,\ell)}} \mathrm{tr}\left[x^i_j(x^k_\ell)^{-1}\right] = \sum_{\substack{v,w \in V \colon \\ v \to w}} \mathrm{tr}[x_v x_w^{-1}] \qquad \text{for all } x \in \mathcal{P}^V_d.$$

All the results of Sect. 5 hold for the above 'triangular graph' structure, since:

- $G = (V, E)$ is an *acyclic finite directed graph*;
- Γ is a proper subset of V containing the only *source* (N, N) and *sink* $(N, 1)$ of G;
- the energy function Φ^N is of the form (5.9).

We first prove a property of the critical points of Φ^N that, in the scalar $d = 1$ setting, was observed in [32].

Lemma 4.3 *Let $z \in \mathcal{P}^N_d$. Let x be any critical point of Φ^N on $\mathcal{T}^N_d(z)$. For all $1 \leq i \leq N$, let $p_i := \left|x^i_1 \cdots x^i_i\right|$ be the determinant of the product of the ith row of x. Then,*

$$p_1 = \sqrt[2]{p_2} = \cdots = \sqrt[N-1]{p_{N-1}} = \sqrt[N]{p_N} = \sqrt[N]{|z_1 \cdots z_N|}. \qquad (4.5)$$

Proof The critical point equations of the energy function Φ^N are

$$(x^i_j)^{-1}(x^{i+1}_{j+1} + x^{i-1}_j)(x^i_j)^{-1} = (x^{i-1}_{j-1})^{-1} + (x^{i+1}_j)^{-1} \qquad \text{for all } 1 \leq j \leq i < N,$$

with the convention $x_i^{i-1} = (x_0^{i-1})^{-1} = 0$ for all $1 \leq i < N$ (these correspond to (5.15) in the case of the triangular graph G). Taking determinants of both sides, we obtain

$$|x_j^i|^2 = \frac{\left| x_{j+1}^{i+1} + x_j^{i-1} \right|}{\left| (x_{j-1}^{i-1})^{-1} + (x_j^{i+1})^{-1} \right|} \qquad \text{for all } 1 \leq j \leq i < N .$$

Taking the product over j in the latter, many terms cancel out, yielding

$$\prod_{j=1}^{i} |x_j^i|^2 = \prod_{j=1}^{i-1} |x_j^{i-1}| \prod_{j=1}^{i+1} |x_j^{i+1}| .$$

By definition of $p_1, \ldots, p_N$, the latter can be written as

$$p_i^2 = p_{i-1} p_{i+1} \qquad \text{for all } 1 \leq j \leq i < N , \tag{4.6}$$

with the convention $p_0 := 1$. Finally, it is straightforward to see that equations (4.6) are equivalent to (4.5). $\qquad \square$

Let now

$$I_d^N := \underbrace{(I_d, \ldots, I_d)}_{N \text{ times}} = r^N(1) \in \mathcal{P}_d^N .$$

As the components of I_d^N are scalar matrices, Theorem 5.13 implies:

Corollary 4.4 *The function Φ^N on $\mathcal{T}_d^N(I_d^N)$ has a unique global minimiser, at which the Hessian is positive definite. Moreover, each component m_j^i of the minimiser $m = (m_j^i)_{1 \leq j \leq i \leq N}$ is a positive scalar matrix.*

Throughout this subsection, m will always denote the above minimiser.

Corollary 4.5 *We have $m_1^1 = I_d$ and*

$$\left| m_1^i \cdots m_i^i \right| = 1 \qquad \text{for all } i = 1, \ldots, N . \tag{4.7}$$

Proof Since $m \in \mathcal{T}_d^N(I_d^N)$, we have $m_j^N = I_d$ for all $j = 1, \ldots, N$, hence $\left| m_1^N \cdots m_N^N \right| = 1$. On the other hand, as a minimiser, m is a critical point of Φ^N on $\mathcal{T}_d^N(I_d^N)$, hence (4.7) follows from Lemma 4.3. Furthermore, since m_1^1 is a multiple of I_d with determinant 1, we have $m_1^1 = I_d$. $\qquad \square$

Theorem 4.6 *For any $\lambda \in \mathbb{C}^N$, we have*

$$\psi_\lambda^N(r^N(k)) \overset{k \to \infty}{\sim} \frac{1}{\sqrt{|\mathcal{H}(m)|}} \left(\frac{2\pi}{k} \right)^{\frac{N(N-1)d(d+1)}{8}} e^{-k\Phi^N(m)} , \tag{4.8}$$

where $|\mathcal{H}(m)| > 0$ is the Hessian determinant of Φ^N at m.

The case $d = 1$, $N = 2$ of this asymptotic result is classical; the case $d = 1$ and general N can be found in [32, eq. 20]. Finally, the case $d > 1$, $N = 2$ may be inferred from the Laplace approximation of Bessel functions of matrix arguments studied in [14] (see also [22, Appendix B] and [33, Section 2.6]).

An important feature of (4.8) is that the leading order asymptotics does *not* depend on the parameter λ. This was already remarked in [13] in the special case $d = 1$ and $N = 3$, for which the full asymptotic expansion was obtained.

Proof of Theorem 4.6 By (2.4), we have

$$\psi_\lambda^N(r^N(k)) = \int_{\mathcal{T}_d^N(r^N(k))} \left(\prod_{i=1}^{N-1} \prod_{j=1}^{i} \mu(\mathrm{d}x_j^i) \right) \Delta_\lambda^N(x) \, \mathrm{e}^{-\Phi^N(x)} \, .$$

Recalling (4.4), let us change variables by setting

$$\tilde{x}_j^i = r_j^i(k)^{-1} x_j^i = k^{i-2j+1} x_j^i \qquad \text{for } 1 \le j \le i \le N \, . \tag{4.9}$$

One can then easily verify, using also the invariance property of the measure μ, that

$$\psi_\lambda^N(r^N(k)) = \int_{\mathcal{T}_d^N(I_d^N)} \left(\prod_{i=1}^{N-1} \prod_{j=1}^{i} \mu(\mathrm{d}\tilde{x}_j^i) \right) \Delta_\lambda^N(\tilde{x}) \, \mathrm{e}^{-k\Phi^N(\tilde{x})} \, . \tag{4.10}$$

Applying Prop. 5.14 with $g := \Delta_\lambda^N$, we obtain

$$\psi_\lambda^N(r^N(k)) \overset{k\to\infty}{\sim} \frac{\Delta_\lambda^N(m)}{\sqrt{|\mathcal{H}(m)|}} \left(\prod_{i=1}^{N-1} |m_1^i \cdots m_i^i| \right) \left(\frac{2\pi}{k} \right)^{\frac{N(N-1)d(d+1)}{8}} \mathrm{e}^{-k\Phi^N(m)} \, ,$$

since the number of vertices of G that do *not* belong to Γ is $N(N-1)/2$. The claim then follows from Corollary 4.5 (which, in particular, implies that $\Delta_\lambda^N(m) = 1$). $\qquad\square$

Recall now the definition (2.6) of the $\tilde{\Sigma}$-kernel.

Corollary 4.7 *Let $f : \mathcal{T}_d^N \to \mathbb{R}$ be a bounded and continuous function and let*

$$f_k(x) := f\big((r_j^i(k)x_j^i)_{1\le j\le i\le N}\big) \quad \text{for } k > 0 \text{ and } x \in \mathcal{T}_d^N \, . \tag{4.11}$$

Assume that $f_k \overset{k\to\infty}{\longrightarrow} f_\infty$ uniformly on any compact subsets of $\mathcal{T}_d^N(I_d^N)$. Then, for any $\lambda, \varrho \in \mathbb{R}^N$,

$$\lim_{k\to\infty} \frac{\tilde{\Sigma}_\lambda f(r^N(k))}{\psi_\varrho^N(r^N(k))} = f_\infty(m) \, . \tag{4.12}$$

Proof As the leading order asymptotics of the Whittaker function $\psi_\lambda^N(r^N(k))$ does not depend on λ by Theorem 4.6, we have

$$\lim_{k\to\infty} \frac{\psi_\lambda^N(r^N(k))}{\psi_\varrho^N(r^N(k))} = 1 \,.$$

Therefore, it suffices to prove (4.12) for $\varrho = \lambda$.

Note that, using (4.10) and the fact that $\lambda \in \mathbb{R}^N$, the measure μ_k^N defined by

$$\mu_k^N(\mathrm{d}\tilde{x}) := \frac{1}{\psi_\lambda^N(r^N(k))} \Delta_\lambda^N(\tilde{x}) \, \mathrm{e}^{-k\Phi^N(\tilde{x})} \left(\prod_{i=1}^{N-1} \prod_{j=1}^{i} \mu(\mathrm{d}\tilde{x}_j^i) \right)$$

is a probability measure on $\mathcal{T}_d^N(I_d^N)$. By definition of $\tilde{\Sigma}_\lambda$, we then have

$$\begin{aligned}
\frac{\tilde{\Sigma}_\lambda f(r^N(k))}{\psi_\lambda^N(r^N(k))} &= \int_{\mathcal{T}_d^N(r^N(k))} \left(\prod_{i=1}^{N-1} \prod_{j=1}^{i} \mu(\mathrm{d}x_j^i) \right) \frac{\Delta_\lambda^N(x) \, \mathrm{e}^{-\Phi^N(x)}}{\psi_\lambda^N(r^N(k))} f(x) \\
&= \int_{\mathcal{T}_d^N(I_d^N)} \mu_k^N(\mathrm{d}\tilde{x}) f_k(\tilde{x}) \,,
\end{aligned}$$

where in the integral we performed the change of variables (4.9). Since f is bounded and continuous, the functions $\{f_k\}_{k>0}$ are uniformly bounded and continuous; moreover, by assumption, they converge as $k \to \infty$ to f_∞ uniformly on any compact subsets of $\mathcal{T}_d^N(I_d^N)$. Therefore, by Lemma C.1, it is now enough to show that μ_k^N converges weakly as $k \to \infty$ to the Dirac measure δ_m, i.e. that

$$\lim_{k\to\infty} \int_{\mathcal{T}_d^N(I_d^N)} \mu_k^N(\mathrm{d}x) g(x) = g(m)$$

for every bounded and continuous function $g: \mathcal{T}_d^N(I_d^N) \to \mathbb{R}$. This claim, in turn, follows readily from Prop. 5.14, since, without loss of generality, one can assume $g(m) \neq 0$. □

4.3 Fixed-time law of the 'bottom edge' process

Let us now go back to the Markov process X on $\mathcal{T}_d^N$ from Definition 3.3. Recall that, under the hypotheses of Theorem 3.10, the Nth row X^N of the process X has an autonomous Markov evolution with time-n transition kernel $P_{\alpha(n),\beta}^N$ (cf. (3.24)). The transition kernel of X^N from time 0 to time n is then given by the composition

$$U_{\alpha,\beta}^{N,n} := P_{\alpha(1),\beta}^N P_{\alpha(2),\beta}^N \cdots P_{\alpha(n),\beta}^N \,. \tag{4.13}$$

Thus, if the initial state of X^N is $X^N(0) = z$, then the law of $X^N(n)$ is $U_{\alpha,\beta}^{N,n}(z; \cdot)$.

Let now $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$ such that $\alpha(\ell) + \Re(\lambda_i) > \frac{d-1}{2}$ for all $1 \le \ell \le n$ and $1 \le i \le N$. Iterating the eigenfunction equation (3.22) n times, we obtain the following eigenfunction equation for $U_{\alpha,\beta}^{N,n}$:

$$U_{\alpha,\beta}^{N,n} \frac{\psi_\lambda^N}{\psi_\beta^N} = \left(\prod_{\ell=1}^{n} \prod_{i=1}^{N} \frac{\Gamma_d(\alpha(\ell) + \lambda_i)}{\Gamma_d(\alpha(\ell) + \beta^i)} \right) \frac{\psi_\lambda^N}{\psi_\beta^N}. \tag{4.14}$$

Consider now the initial state $X^N(0) = r^N(k)$ (cf. (4.4)), which becomes singular in the limit as $k \to \infty$. We will show that the measure $U_{\alpha,\beta}^{N,n}(r^N(k); \cdot)$ converges, as $k \to \infty$, to the matrix Whittaker measure with parameters $(\alpha(1), \ldots, \alpha(n))$ and β. An intuition about this fact is provided by (4.14). It follows from Theorem 4.6 that the ratio of Whittaker functions on the right-hand side of (4.14), evaluated at $r^N(k)$, converges to 1 as $k \to \infty$. It is then easy to see that, if the convergence to matrix Whittaker measures holds as claimed above, then (4.14) reduces to the Whittaker integral identity proved in Sect. 4.1.

Theorem 4.8 *Let $n \ge N$. As $k \to \infty$, the distribution $U_{\alpha,\beta}^{N,n}(r^N(k); \cdot)$ converges in total variation distance (and, hence, weakly) to the matrix Whittaker measure with parameters $(\alpha(1), \ldots, \alpha(n))$ and β (which we denote by $W_{\alpha,\beta}^{N,n}$ for simplicity). Namely, we have*

$$\lim_{k \to \infty} \sup_A \left| U_{\alpha,\beta}^{N,n}(r^N(k); A) - W_{\alpha,\beta}^{N,n}(A) \right| = 0, \tag{4.15}$$

where the supremum is taken over all measurable sets $A \subseteq \mathcal{P}_d^N$.

Proof We will prove that

$$\lim_{k \to \infty} \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \left| U_{\alpha,\beta}^{N,n}(r^N(k); z) - W_{\alpha,\beta}^{N,n}(z) \right| = 0. \tag{4.16}$$

This statement is stronger than (4.15), as the supremum in (4.15) is clearly bounded from above by the integral in (4.16).

Let us fix N and prove (4.16) by induction on $n \ge N$. Before proving the base case, let us verify the (simpler) induction step. Assume that (4.16) holds for a certain $n \ge N$. Using (4.2), (3.24), (2.16) and (2.17), we obtain

$$(W_{\alpha,\beta}^{N,n} \boldsymbol{P}_{\alpha(n+1),\beta}^N)(z) = \left(\prod_{\ell=1}^{n+1} \prod_{i=1}^{N} \frac{1}{\Gamma_d(\alpha(\ell) + \beta^i)} \right) (\bar{P}_{\alpha(n+1)}^N \psi_{\alpha(1:n); I_d}^{N,n})(z) \psi_\beta^N(z)$$

$$= \left(\prod_{\ell=1}^{n+1} \prod_{i=1}^{N} \frac{1}{\Gamma_d(\alpha(\ell) + \beta^i)} \right) \psi_{\alpha(1:n+1); I_d}^{N,n+1}(z) \psi_\beta^N(z) = W_{\alpha,\beta}^{N,n+1}(z)$$

for $z \in \mathcal{P}_d^N$. On the other hand, by (4.13) we have $U_{\alpha,\beta}^{N,n+1} = U_{\alpha,\beta}^{N,n} \boldsymbol{P}_{\alpha(n+1),\beta}^N$. Applying Fubini's theorem and recalling that $\boldsymbol{P}_{\alpha(n+1),\beta}^N$ is a Markov kernel, we then obtain

$$
\int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \left| U_{\alpha,\beta}^{N,n+1}(r^N(k); z) - W_{\alpha,\beta}^{N,n+1}(z) \right|
$$

$$
= \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \left| \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}\tilde{z}) \left(U_{\alpha,\beta}^{N,n}(r^N(k); \tilde{z}) - W_{\alpha,\beta}^{N,n}(\tilde{z}) \right) \boldsymbol{P}_{\alpha(n+1),\beta}^N(\tilde{z}; z) \right|
$$

$$
\leq \int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}\tilde{z}) \left| U_{\alpha,\beta}^{N,n}(r^N(k); \tilde{z}) - W_{\alpha,\beta}^{N,n}(\tilde{z}) \right| \underbrace{\int_{\mathcal{P}_d^N} \mu^{\otimes N}(\mathrm{d}z) \boldsymbol{P}_{\alpha(n+1),\beta}^N(\tilde{z}; z)}_{=1 \quad \text{for all } \tilde{z}} .
$$

The latter expression vanishes as $k \to \infty$ by the induction hypothesis, thus proving the induction step.

It remains to prove the base case, i.e. (4.16) for $n = N$. Recall that the measures $U_{\alpha,\beta}^{N,N}(r^N(k); \cdot)$, for any $k > 0$, and $W_{\alpha,\beta}^{N,N}$ have the same finite total mass, since they are all probability distributions, and are absolutely continuous with respect to $\mu^{\otimes N}$. By Scheffé's theorem (see e.g. [4, Theorem 16.12]), it then suffices to show the convergence of the densities:

$$
\lim_{k \to \infty} U_{\alpha,\beta}^{N,N}(r^N(k); z) = W_{\alpha,\beta}^{N,N}(z) \qquad \text{for } \mu^{\otimes N}\text{-almost every } z \in \mathcal{P}_d^N .
$$

(4.17)
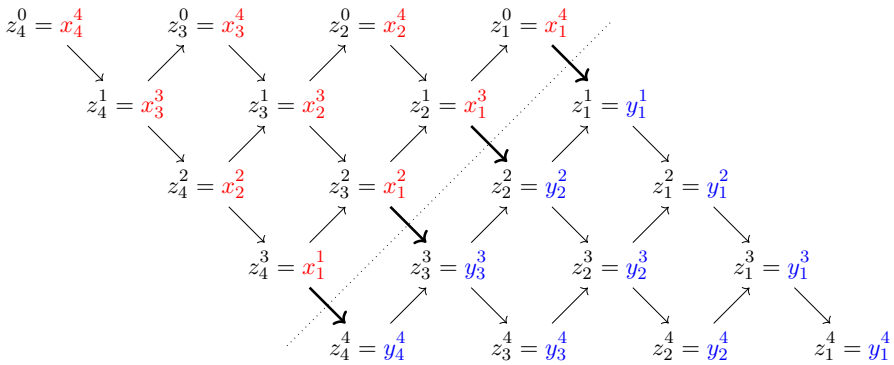
Fix $z \in \mathcal{P}_d^N$ once for all. Using (4.13), we write

$$
U_{\alpha,\beta}^{N,N}(r^N(k); z) = \int_{\mathcal{P}_d^N} \boldsymbol{P}_{\alpha(1),\beta}^N(r^N(k); \mathrm{d}z^1) \int_{\mathcal{P}_d^N} \boldsymbol{P}_{\alpha(2),\beta}^N(z^1; \mathrm{d}z^2) \cdots
$$

$$
\cdots \int_{\mathcal{P}_d^N} \boldsymbol{P}_{\alpha(N-1),\beta}^N(z^{N-2}; \mathrm{d}z^{N-1}) \boldsymbol{P}_{\alpha(N),\beta}^N(z^{N-1}; z)
$$

Define now

$$
J_\alpha^N(z^0; z^N) := \int_{\mathcal{P}_d^{N(N-1)}} \left(\prod_{\ell=1}^{N-1} \mu^{\otimes N}(\mathrm{d}z^\ell) \right) \prod_{i,j=1}^N \frac{|z_j^{i-1}|^{\alpha(i)}}{|z_j^i|^{\alpha(i)}} \mathrm{e}^{-\mathrm{tr}\left[z_{j+1}^i (z_j^{i-1})^{-1} + z_j^{i-1}(z_j^i)^{-1} \right]}
$$

(4.18)

for $z^0, z^N \in \mathcal{P}_d^N$, with the usual conventions $z_{N+1}^i := 0$ for all $i = 0, \ldots, N$. Using the definition (3.24) of the $\boldsymbol{P}$-kernels, we then have

$$z_4^0 = x_4^4 \qquad z_3^0 = x_3^4 \qquad z_2^0 = x_2^4 \qquad z_1^0 = x_1^4$$

$$z_4^1 = x_3^3 \qquad z_3^1 = x_2^3 \qquad z_2^1 = x_1^3 \qquad z_1^1 = y_1^1$$

$$z_4^2 = x_2^2 \qquad z_3^2 = x_1^2 \qquad z_2^2 = y_2^2 \qquad z_1^2 = y_1^2$$

$$z_4^3 = x_1^1 \qquad z_3^3 = y_3^3 \qquad z_2^3 = y_2^3 \qquad z_1^3 = y_1^3$$

$$z_4^4 = y_4^4 \qquad z_3^4 = y_3^4 \qquad z_2^4 = y_2^4 \qquad z_1^4 = y_1^4$$

Fig. 3 Graphical representation of the set of variables z_j^i, $0 \le i \le N$ and $1 \le j \le N$ (here $N = 4$) appearing in (4.18). Each arrow $a \to b$ corresponds to the term $\mathrm{e}^{-\operatorname{tr}[ab^{-1}]}$ in the integral. Relabelling the z_j^i as in (4.21)–(4.22) yields two triangular arrays x (coloured in red) and y (coloured in blue) (colour figure online)

$$U_{\alpha,\beta}^{N,N}(r^N(k); z) = \left(\prod_{\ell,i=1}^N \frac{1}{\Gamma_d(\alpha(\ell) + \beta^i)} \right) \frac{\psi_\beta^N(z)}{\psi_\beta^N(r^N(k))} J_\alpha^N(r^N(k); z) \,.$$

$$(4.19)$$

Comparing (4.19) with (4.3), we are reduced to show that

$$\lim_{k \to \infty} \frac{J_\alpha^N(r^N(k); z)}{\psi_\beta^N(r^N(k))} = \mathrm{e}^{-\operatorname{tr}\left[z_N^{-1}\right]} \psi_{\alpha(1:N)}^N(z) \qquad \text{for } \mu^{\otimes N}\text{-almost every } z \in \mathcal{P}_d^N.$$

$$(4.20)$$

Let us relabel the variables in the integral (4.18) by setting

$$z_j^i = x_{j-i}^{N-i} \qquad \text{for } 0 \le i \le N-1 \text{ and } i+1 \le j \le N \,, \qquad (4.21)$$

$$z_j^i = y_j^i \qquad \text{for } 1 \le i \le N \text{ and } 1 \le j \le i \,. \qquad (4.22)$$

This relabelling yields two triangular arrays $x, y \in \mathcal{T}_d^N$. See Fig. 3 for a graphical representation of the variables z_j^i and the corresponding arrays x and y. Recalling the definition (2.6) of the $\tilde{\Sigma}$-kernel, we have

$$J_\alpha^N(r^N(k); z) = \tilde{\Sigma}_{\hat{\alpha}(1:N)}^N f(r^N(k)) \,,$$

where $\hat{\alpha}(1:N) := (-\alpha(N), \ldots, -\alpha(1))$ and the function $f \colon \mathcal{T}_d^N \to \mathbb{R}$ is defined by

$$f(x) := \int_{\mathcal{T}_d^N(z)} \left(\prod_{i=1}^{N-1} \prod_{j=1}^i \mu(\mathrm{d}y_j^i) \right) \left(\prod_{i=1}^N \mathrm{e}^{-\operatorname{tr}\left[x_1^{N-i+1}(y_i^i)^{-1}\right]} \right) \Delta_{\alpha(1:N)}^N(y) \, \mathrm{e}^{-\Phi^N(y)} \,.$$

Here, each term $e^{-\operatorname{tr}\left[x_1^{N-i+1}(y_i^i)^{-1}\right]}$ corresponds, graphically, to a bold arrow in Fig. 3.

We now wish to apply Corollary 4.7. Notice first that f is a continuous function of x; moreover, it is bounded below by 0 and above by $\psi_{\alpha(1:N)}^N(z)$ (cf. (2.4)). The associated functions f_k defined in (4.11) are

$$
f_k(x) = \int_{\mathcal{T}_d^N(z)} \left(\prod_{i=1}^{N-1} \prod_{j=1}^{i} \mu(dy_j^i) \right) \left(\prod_{i=1}^{N} e^{-k^{-(N-i)}\operatorname{tr}\left[x_1^{N-i+1}(y_i^i)^{-1}\right]} \right) \Delta_{\alpha(1:N)}^N(y) \, e^{-\Phi^N(y)} .
$$

By dominated convergence and by the definition (2.4) of Whittaker function, we have the pointwise convergence

$$
\lim_{k \to \infty} f_k(x) = e^{-\operatorname{tr}\left[x_1^1 z_N^{-1}\right]} \psi_{\alpha(1:N)}^N(z) =: f_\infty(x) .
$$

Notice that $\{f_k\}_{k>0}$ is a collection of continuous functions, increasing with k, that converges pointwise to a continuous limit; hence, by Dini's theorem (see e.g. [38, Theorem 7.13]), the convergence is uniform on compacts. Then, the assumptions of Corollary 4.7 are satisfied and we have

$$
\lim_{k \to \infty} \frac{J_\alpha^N(r^N(k); z)}{\psi_\beta^N(r^N(k))} = \lim_{k \to \infty} \frac{\tilde{\Sigma}_{\tilde{\alpha}(1:N)}^N f(r^N(k))}{\psi_\beta^N(r^N(k))} = f_\infty(m) ,
$$

where m is the unique global minimiser of Φ^N on $\mathcal{T}_d^N(I_d^N)$ (cf. Corollary 4.4). Since $m_1^1 = I_d$ by Corollary 4.5, we have

$$
f_\infty(m) = e^{-\operatorname{tr}\left[z_N^{-1}\right]} \psi_{\alpha(1:N)}^N(z) .
$$

This yields the desired limit (4.20).   □

4.4 Fixed-time laws of the 'right edge' and 'left edge' processes

Throughout this subsection, it will be convenient to work with the space of $d \times d$ *positive semidefinite* matrices, i.e. $d \times d$ real symmetric matrices with nonnegative eigenvalues; such a space is the closure of $\mathcal{P}_d$ under the standard Euclidean topology, and we thus denote it by $\overline{\mathcal{P}}_d$.

It is clear from the definition given in Sect. 3.1 that the 'right edge' $X_1 = (X_1^1, \ldots, X_1^N)$ of X is a Markov process in its own filtration. Furthermore, as mentioned before, X_1 equals the system $Z = (Z^1, \ldots, Z^N)$ of random particles in $\mathcal{P}_d$ with one-sided interactions defined in (1.8)–(1.9), where the random weight $V^i(n)$ equals $W_1^i(n)$, an inverse Wishart random matrix with parameter $\alpha(n) + \beta^i$. If the initial state $Z(0)$ of this process is in $\mathcal{P}_d^N$ (respectively, $\overline{\mathcal{P}}_d^N$), then clearly Z evolves as a process in $\mathcal{P}_d^N$ (respectively, $\overline{\mathcal{P}}_d^N$).

Analogously, the 'left edge' of X is a Markov process in its own filtration. Its 'inverse' $L = (L^1, \ldots, L^N) := ((X_1^1)^{-1}, \ldots, (X_N^N)^{-1})$ is given by (3.5)–(3.6), where $U^i(n) := (W_i^i(n))^{-1}$ is a Wishart random matrix with parameter $\alpha(n) + \beta^i$. If the initial state $L(0)$ of this process is in $\mathcal{P}_d^N$ (respectively, $\overline{\mathcal{P}}_d^N$), then clearly L evolves as a process in $\mathcal{P}_d^N$ (respectively, $\overline{\mathcal{P}}_d^N$).

As the next lemma shows, the singular initial state of the bottom edge of X considered in Sect. 4.3 induces (through Theorem 3.10) the initial state (1.10) on the right edge X_1, which resembles the step or 'narrow wedge' initial configuration in systems of interacting particles/random walks. A similar statement holds for the left edge.

Lemma 4.9 *Let $X(0)$ be distributed according to $\overline{\Sigma}_\beta(r^N(k); \cdot)$. Then, on the space $\overline{\mathcal{P}}_d^N$, both $(X_1^1(0), \ldots, X_1^N(0))$ and $((X_1^1(0))^{-1}, \ldots, (X_N^N(0))^{-1})$ converge in law, as $k \to \infty$, to $(I_d, 0_d, \ldots, 0_d)$.*

Proof We prove the claim about $(X_1^1(0), \ldots, X_1^N(0))$, as the proof of the claim about $((X_1^1(0))^{-1}, \ldots, (X_N^N(0))^{-1})$ is completely analogous.

Let $g \colon \overline{\mathcal{P}}_d^N \to \mathbb{R}$ be a bounded and continuous test function. We need to prove that

$$\lim_{k \to \infty} \mathbb{E}[g(X_1(0))] = g(I_d, 0_d, \ldots, 0_d). \tag{4.23}$$

Let $f \colon \mathcal{T}_d^N \to \mathbb{R}$, $f(x) := g(x_1) = g(x_1^1, \ldots, x_1^N)$ for all $x \in \mathcal{T}_d^N$. By definition (3.23) of $\overline{\Sigma}_\beta$, we then have

$$\mathbb{E}[g(X_1(0))] = \mathbb{E}[f(X(0))] = \overline{\Sigma}_\beta f(r^N(k)) = \frac{\tilde{\Sigma}_\beta f(r^N(k))}{\psi_\beta^N(r^N(k))}.$$

We now wish to apply Corollary 4.7. Since g is bounded and continuous, f also is. The associated functions f_k defined in (4.11) are

$$f_k(x) := g\left(x_1^1, \frac{x_1^2}{k}, \ldots, \frac{x_1^N}{k^{N-1}}\right), \qquad x \in \mathcal{T}_d^N.$$

These functions converge as $k \to \infty$ to $f_\infty(x) := g(x_1^1, 0_d, \ldots, 0_d)$ uniformly on compacts, since g is continuous on $\overline{\mathcal{P}}_d^N$. Therefore, by Corollary 4.7, $\mathbb{E}[g(X_1(0))]$ converges as $k \to \infty$ to $g(m_1^1, 0_d, \ldots, 0_d)$, where m is the minimiser of Φ^N on $\mathcal{T}_d^N(I_d^N)$. By Corollary 4.5 we have $m_1^1 = I_d$, and the claim (4.23) follows. □

As a consequence of Theorem 4.8 and Lemma 4.9, we obtain:

Corollary 4.10 *As above, let $Z = (Z^1, \ldots, Z^N)$ and $L = (L^1, \ldots, L^N)$ be the right edge process and the (inverse) left edge process, respectively, with initial states $Z(0) = L(0) = (I_d, 0_d, \ldots, 0_d) \in \overline{\mathcal{P}}_d^N$. Then, for $n \geq N$, $Z^N(n)$ and $L^N(n)$ are distributed as the first marginal and the Nth marginal, respectively, of the matrix Whittaker measure with parameters $(\alpha(1), \ldots, \alpha(n))$ and β.*

Proof Again, we only prove the claim about the right edge, as the proof of the claim about the left edge is completely analogous.

Let the process X be as in Definition 3.3, with initial state $X(0)$ distributed according to $\overline{\Sigma}_\beta(r^N(k); \cdot)$. It is clear from the definition that $X_1^N(n)$ can be written as a continuous, deterministic function of the right edge initial state $X_1(0)$ and of the collection of random matrices $(W_1^1(\ell), \ldots, W_1^N(\ell))_{1 \le \ell \le n}$. Therefore, by Lemma 4.9 and the continuous mapping theorem [5, Theorem 2.7], $X_1^N(n)$ converges in law as $k \to \infty$ to $Z^N(n)$.

On the other hand, by Theorem 4.8, for $n \ge N$, $X_1^N(n)$ converges in law as $k \to \infty$ to the first marginal of a matrix Whittaker measure with parameters $(\alpha(1), \ldots, \alpha(n))$ and β. □

Remark 4.11 The following generalisation of Corollary 4.10 is immediate: under the same hypotheses, for every $1 \le i \le N$ and $n \ge i$, $Z^i(n)$ is distributed as the first marginal of the matrix Whittaker measure with parameters $(\alpha(1), \ldots, \alpha(n))$ and $(\beta^1, \ldots, \beta^i)$. This is due to the fact that, by definition, for any fixed $i \ge 1$, the process $(Z_1, \ldots, Z_i)$ has both an initial configuration $(I_d, 0_d, \ldots, 0_d)$ and a Markov evolution that do not depend on the choice of $N \ge i$. Analogously, for every $1 \le i \le N$ and $n \ge i$, $L^i(n)$ is distributed as the ith marginal of the same matrix Whittaker measure.

5 Minimisation of energy functions and Laplace approximations

In this section, we study minimisation problems for certain energy functions of matrix arguments on directed graphs. As a consequence, we obtain Laplace approximations for integrals of exponentials of these energy functions. For our purposes, the most important application of such results consists in certain asymptotics of Whittaker functions of matrix arguments; see Sect. 4.2. However, the results of this section may be of independent interest. For instance, the general framework we work with may be applied to obtain analogous asymptotics for *orthogonal* Whittaker functions, which also appeared in the study of stochastic systems—see [2, 6].

5.1 Energy functions on directed graphs

Let us recall some terminology of graph theory that will be useful throughout this section. A *finite directed graph $G = (V, E)$* is a pair consisting of a nonempty finite set V of *vertices* and a set $E \subset \{(v, w) \in V^2 : v \ne w\}$ of *edges*. Note that edges connecting a vertex to itself are not allowed, nor are multiple edges. The direction of an edge (v, w) connecting v to w is given by the ordering of the pair. For the sake of notational convenience, we also write $v \to w$ when $(v, w) \in E$, and $v \nrightarrow w$ when $(v, w) \notin E$. A vertex v is called a *sink* if it has no outcoming edges (i.e. if $v \nrightarrow w$ for all $w \in V$) and a *source* if it has no incoming edges (i.e. if $w \nrightarrow v$ for all $w \in V$). For any $v, w \in V$ and $0 \le l < \infty$, we call *path* of length l in G from v to w any sequence $(v_0, v_1, \ldots, v_l)$ such that $v_0 = v$, $v_l = w$, and $v_{i-1} \to v_i$ for all $1 \le i \le l$. A *cycle* is any path $(v_0, v_1, \ldots, v_l)$ such that $v_0 = v_l$ and any other two vertices are distinct. We

say that G is *acyclic* if it has no cycles. From now on, throughout the whole section, $G = (V, E)$ will always be an acyclic finite directed graph.

Lemma 5.1 *For all $v \in V$, there exists a path in G from v to a sink; moreover, there exists a path in G from a source to v.*

Proof We will prove the existence of the first path only, as the existence of the second path follows from a similar argument. We construct the path algorithmically. Set $v_0 := v$. For all $i = 0, 1, 2, \ldots$, we proceed as follows: if v_i is a sink, then we stop the algorithm; otherwise, we pick v_{i+1} to be any vertex such that $v_i \to v_{i+1}$. If the algorithm never terminates, then there exist two distinct indices i, j with $v_i = v_j$, since G is finite; this implies that G has a cycle, against the hypotheses. Therefore, the procedure must stop in a finite number l of steps, thus yielding a path $(v_0, v_1, \ldots, v_l)$ from $v_0 = v$ to a sink v_l.   □

For any integer $d \geq 1$, let Sym_d, Diag_d, and Scal_d be the sets of $d \times d$ real symmetric matrices, real diagonal matrices, and real scalar matrices (i.e. multiples of the $d \times d$ identity matrix I_d), respectively. We will write Sym_d^V for the set of arrays $x = (x_v)_{v \in V}$, where each $x_v \in \mathrm{Sym}_d$. We will use the notations Diag_d^V and Scal_d^V in a similar way.

Let us define the 'energy functions'

$$\varphi_d : \mathrm{Sym}_d^V \to \mathbb{R}, \qquad \varphi_d(x) := \sum_{\substack{v,w \in V: \\ v \to w}} \mathrm{tr}[e^{x_v} e^{-x_w}], \qquad (5.1)$$

$$\chi_d : \mathrm{Sym}_d^V \to \mathbb{R}, \qquad \chi_d(x) := \sum_{\substack{v,w \in V: \\ v \to w}} \mathrm{tr}[e^{x_v - x_w}], \qquad (5.2)$$

where e^a denotes the usual exponential of the matrix a. The Golden-Thompson inequality (see e.g. [3]) states that $\mathrm{tr}[e^a e^b] \geq \mathrm{tr}[e^{a+b}]$ if a and b are symmetric matrices. It follows that

$$\varphi_d(x) \geq \chi_d(x) \qquad \text{for all } x \in \mathrm{Sym}_d^V . \qquad (5.3)$$

However, the two energy functions are identical only for $d = 1$.

Notice that, by Lemma 5.1, G has at least one sink and one source, possibly coinciding. Throughout, we also assume that there exists at least one vertex of G that is neither a source nor a sink. We can thus fix a subset $\Gamma \subset V$ that contains all the sinks and sources and such that Γ^c, the complement of Γ in V, is nonempty. For any set S and any fixed array $z = (z_v)_{v \in \Gamma} \in S^\Gamma$, let

$$S^V(z) := \{x = (x_v)_{v \in V} \in S^V : x_v = z_v \text{ for all } v \in \Gamma\}. \qquad (5.4)$$

Our first result concerns the asymptotic behaviour of the energy functions on $\mathrm{Sym}_d^V(z)$. Let $\|\cdot\|$ denote any norm on Sym_d^V.

Proposition 5.2 *Let $z \in \mathrm{Sym}_d^{\Gamma}$. For $x \in \mathrm{Sym}_d^V(z)$, we have $\varphi_d(x) \to \infty$ and $\chi_d(x) \to \infty$ as $\|x\| \to \infty$.*

Proof By inequality (5.3), it suffices to prove the claim for χ_d. As all norms on a finite-dimensional space are equivalent, we may arbitrarily take

$$\|x\| := \sum_{v \in V} \varrho(x_v) \qquad \text{for } x \in \mathrm{Sym}_d^V \,, \tag{5.5}$$

where $\varrho(a)$ denotes the spectral radius of a symmetric matrix a (i.e. the largest absolute value of its eigenvalues). As the spectral radius is a norm on Sym_d, it can be easily verified that (5.5) defines a norm on Sym_d^V. We will show that, for any sequence $(x^{(n)})_{n \geq 1} \subseteq \mathrm{Sym}_d^V(z)$ such that $\|x^{(n)}\| \to \infty$ as $n \to \infty$, we have $\chi_d(x^{(n)}) \to \infty$ as $n \to \infty$. For the sake of notational simplicity, we will drop the superscript of $x^{(n)}$ and leave the dependence on n implicit.

By contradiction, assume that there exists a positive constant C such that, along a subsequence, $\chi_d(x) \leq C$. Since $\|x\| \to \infty$, there exists $w \in \Gamma^{\mathsf{c}}$ such that, along a further subsequence, $\varrho(x_w) \to \infty$. This implies that, passing to a final subsequence, either $\lambda_{\max}(x_w) \to \infty$ or $\lambda_{\max}(-x_w) \to \infty$, where $\lambda_{\max}(a)$ denotes the maximum eigenvalue of a symmetric matrix a. As $w \in \Gamma^{\mathsf{c}}$, it is neither a source nor a sink. By Lemma 5.1, there exists a path $(v_0, v_1, \ldots, v_l)$ of length $l \geq 1$ in G from $v_0 = w$ to a sink $v_l \in \Gamma$. Since G has no cycles, we have $v_i \neq v_j$ for all $i \neq j$; therefore, all directed edges $v_{i-1} \to v_i$ $(1 \leq i \leq d)$ are distinct. We thus have

$$C \geq \chi_d(x) \geq \sum_{i=1}^{l} \mathrm{tr}[e^{x_{v_{i-1}} - x_{v_i}}] \geq \sum_{i=1}^{l} e^{\lambda_{\max}(x_{v_{i-1}} - x_{v_i})} \geq \sum_{i=1}^{l} \lambda_{\max}(x_{v_{i-1}} - x_{v_i}) \,, \tag{5.6}$$

where we used the bounds $\mathrm{tr}[e^y] \geq \lambda_{\max}(e^y) = e^{\lambda_{\max}(y)}$ for $y \in \mathrm{Sym}_d$ and $e^{\alpha} \geq \alpha$ for $\alpha \in \mathbb{R}$. Recall now that, for any $a, b \in \mathrm{Sym}_d$,

$$\lambda_{\max}(a + b) \leq \lambda_{\max}(a) + \lambda_{\max}(b) \,. \tag{5.7}$$

By iterating (5.7) several times and using (5.6), we obtain

$$\lambda_{\max}(x_w) \leq \sum_{i=1}^{l} \lambda_{\max}(x_{v_{i-1}} - x_{v_i}) + \lambda_{\max}(x_{v_l}) \leq C + \lambda_{\max}(x_{v_l}) \,.$$

By considering a path $(u_0, u_1, \ldots, u_m)$ of length $m \geq 1$ from a source $u_0 \in \Gamma$ to $u_m = w$ (which again exists by Lemma 5.1) and using similar bounds, we also have

$$\lambda_{\max}(-x_w) \leq \lambda_{\max}(-x_{u_0}) + C \,.$$

Since either $\lambda_{\max}(x_w) \to \infty$ or $\lambda_{\max}(-x_w) \to \infty$, it follows that either $\lambda_{\max}(x_{v_l}) \to \infty$ or $\lambda_{\max}(-x_{u_0}) \to \infty$. This contradicts the fact that $x_{v_l} = z_{v_l}$ and $x_{u_0} = z_{u_0}$ are both fixed for all $x \in \mathrm{Sym}_d^V(z)$, since $v_l, u_0 \in \Gamma$. $\qquad\square$

Remark 5.3 Above we have assumed that $G = (V, E)$ is acyclic and that Γ is a subset of V containing all the sinks and sources of G. We stress that both hypotheses are necessary for Prop. 5.2 to hold. As a counterexample, let G be the cycle graph with n vertices and let $\Gamma = \varnothing$. If $a \in \mathrm{Sym}_d$ and $x = (x_v)_{v \in V}$ is the array with $x_v = a$ for all v, then

$$\varphi_d(x) = \chi_d(x) = \underbrace{\mathrm{tr}[I_d] + \cdots + \mathrm{tr}[I_d]}_{n \text{ times}} = dn$$

is constant in a; however, for the norm $\|\cdot\|$ defined in (5.5), if $\varrho(a) \to \infty$, then $\|x\| \to \infty$.

5.2 Minima of energy functions

We now study the minima of the functions (5.1)–(5.2) on the set $\mathrm{Sym}_d^V(z)$, where $z \in \mathrm{Sym}_d^\Gamma$. In words, we wish to minimise the energy functions subject to the constraint that some of the entries of the input array (precisely, those indexed by the vertices of the subset Γ) are fixed.

We start with the simplest case $d = 1$, in which $\mathrm{Sym}_1 = \mathrm{Diag}_1 = \mathrm{Scal}_1 = \mathbb{R}$ and the two energy functions coincide:

$$\varphi_1 = \chi_1 \colon \mathbb{R}^V \to \mathbb{R}, \qquad \varphi_1(x) = \chi_1(x) = \sum_{v \to w} e^{x_v - x_w}.$$

We denote by ∂_v the partial derivative of a function on $\mathbb{R}^V$ with respect to the variable x_v.

Lemma 5.4 Let $z \in \mathbb{R}^\Gamma$. The Hessian matrix of φ_1 on $\mathbb{R}^V(z)$ is positive definite everywhere. In particular, φ_1 is strictly convex on $\mathbb{R}^V(z)$.

Proof On $\mathbb{R}^V(z)$ the variables indexed by Γ are fixed to the assigned values z, hence we can consider φ_1 and its Hessian as functions of $(x_v)_{v \in \Gamma^c}$. For $v, w \in \Gamma^c$, we have

$$\partial_v \partial_w \varphi_1 = \begin{cases} \sum_{u \in V} \left(e^{x_v - x_u} \mathbb{1}_{v \to u} + e^{x_u - x_v} \mathbb{1}_{u \to v} \right) & \text{if } v = w, \\ -e^{x_v - x_w} \mathbb{1}_{v \to w} - e^{x_w - x_v} \mathbb{1}_{w \to v} & \text{if } v \neq w. \end{cases}$$

Thus, the quadratic form of the Hessian of φ_1 on $\mathbb{R}^V(z)$, as a function of $\alpha = (\alpha_v)_{v \in \Gamma^c}$, is

$$\sum_{v, w \in \Gamma^c} \alpha_v \alpha_w \partial_v \partial_w \varphi_1 = \sum_{v \in \Gamma^c} \alpha_v^2 \sum_{u \in V} \left(e^{x_v - x_u} \mathbb{1}_{v \to u} + e^{x_u - x_v} \mathbb{1}_{u \to v} \right)$$
$$+ 2 \sum_{v, w \in \Gamma^c} \alpha_v \alpha_w \left(-e^{x_v - x_w} \mathbb{1}_{v \to w} - e^{x_w - x_v} \mathbb{1}_{w \to v} \right).$$

Setting $\alpha_v := 0$ for all $v \in \Gamma$, it is easy to see that the latter expression equals

$$\sum_{\substack{v,w \in V: \\ v \to w}} e^{x_v - x_w} \left(\alpha_v^2 + \alpha_w^2 - 2\alpha_v \alpha_w \right) = \sum_{\substack{v,w \in V: \\ v \to w}} e^{x_v - x_w} (\alpha_v - \alpha_w)^2 \geq 0 \,.$$

Therefore, the Hessian is positive semidefinite everywhere. To prove that it is in fact positive definite, we will show that, if the quadratic form of the Hessian vanishes at α, then $\alpha = 0$. If the above expression vanishes, then $\alpha_v = \alpha_w$ for all $v, w \in V$ such that $v \to w$. Let $v \in \Gamma^c$. By Lemma 5.1, there exists a path from v to a sink $s \in \Gamma$. The value α_w is then the same for all the vertices w along such a path. We then have $\alpha_v = \alpha_s = 0$, since $s \in \Gamma$. As $v \in \Gamma^c$ was arbitrary, it follows that $\alpha = (\alpha_v)_{v \in \Gamma^c} = 0$. $\qquad\square$

Proposition 5.5 *The function $\varphi_1 = \chi_1$ has a unique (global) minimiser on $\mathbb{R}^V(z)$.*

Proof By Lemma 5.4, φ_1 is a strictly convex function over the convex set $\mathbb{R}^V(z)$; therefore, it has at most one minimiser. It remains to show the existence of a minimiser. Since φ_1 is a continuous function, it admits at least one minimiser on every closed ball $B_r := \{x \in \mathbb{R}^V(z) \colon \|x\| \leq r\}$. By Prop. 5.2, for r large enough, the minimiser on B_r is also a (global) minimiser on $\mathbb{R}^V(z)$. $\qquad\square$

The case $d > 1$ is much more challenging, and we are able to deal with it only under rather strong assumptions on the fixed array z. Nonetheless, this is sufficient for our ultimate purposes.

We will be using the fact that the relation between the eigenvalues and the diagonal entries of a symmetric matrix is completely characterised by the *majorisation* relation. Let us briefly explain this statement, referring to [23, § 4.3] for proofs and details. For any $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, let us denote by $\alpha^\downarrow = (\alpha_1^\downarrow, \ldots, \alpha_d^\downarrow)$ its nonincreasing rearrangement, i.e. the permutation of the coordinates of α such that $\alpha_1^\downarrow \geq \alpha_2^\downarrow \geq \cdots \geq \alpha_d^\downarrow$. Given $\alpha, \beta \in \mathbb{R}^d$, we say that α *majorises* β, and write $\alpha \succ \beta$, if

$$\sum_{i=1}^k \alpha_i^\downarrow \geq \sum_{i=1}^k \beta_i^\downarrow \quad \text{for } 1 \leq k \leq d-1 \quad \text{and} \quad \sum_{i=1}^d \alpha_i = \sum_{i=1}^d \beta_i \,. \tag{5.8}$$

Theorem 5.6 ([23, Theorem 4.3.45]) *Let $x \in \mathrm{Sym}_d$. Let $\lambda = (\lambda_1, \ldots \lambda_d)$ be the vector of the (real) eigenvalues of x, taken in any order. Let $\delta_i := x(i, i)$ for $1 \leq i \leq d$, so that $\delta = (\delta_1, \ldots, \delta_d)$ is the vector of the diagonal entries of x. Then we have $\lambda \succ \delta$, and the equality $\lambda^\downarrow = \delta^\downarrow$ holds if and only if x is a diagonal matrix.*

We now briefly introduce the concept of Schur convexity and state the criterion that is useful for our purposes, referring e.g. to [30, Ch. I.3] for more details. A function $H \colon \mathbb{R}^d \to \mathbb{R}$ is called *Schur-convex* if $H(\alpha) \geq H(\beta)$ for all $\alpha, \beta \in \mathbb{R}^d$ such that $\alpha \succ \beta$. In particular, for all α, β such that $\alpha^\downarrow = \beta^\downarrow$, we have $\alpha \succ \beta \succ \alpha$, hence $H(\alpha) = H(\beta)$; in other words, every Schur-convex function is a symmetric function. Additionally, H is called *strictly Schur-convex* if $H(\alpha) > H(\beta)$ for all $\alpha, \beta \in \mathbb{R}^d$ such that $\alpha \succ \beta$ and $\alpha^\downarrow \neq \beta^\downarrow$.

Theorem 5.7 ([30], Ch. I.3, § C]) *Let* $h: \mathbb{R} \to \mathbb{R}$ *and*

$$H: \mathbb{R}^d \to \mathbb{R}, \qquad H(\alpha_1, \ldots, \alpha_d) = \sum_{i=1}^{d} h(\alpha_i).$$

If h is convex, then H is Schur-convex. If h is strictly convex, then H is strictly Schur-convex.

As a consequence of the results just stated, we obtain:

Proposition 5.8 *Suppose that* $x \in \mathrm{Sym}_d$ *and* $y \in \mathrm{Diag}_d$ *have the same diagonal entries. Then* $\mathrm{tr}[e^x] \geq \mathrm{tr}[e^y]$, *and the equality holds if and only if* $x = y$.

Proof Let $\lambda = (\lambda_1, \ldots, \lambda_d)$ be the vector of the eigenvalues of x, taken in any order. Let $\delta = (\delta_1, \ldots, \delta_d)$ be the vector of (common) diagonal entries of x and y, i.e. $\delta_i = x(i, i) = y(i, i)$ for all $1 \leq i \leq d$. Since y is diagonal, notice that the δ_i's are also its eigenvalues. Therefore, the claimed inequality $\mathrm{tr}[e^x] \geq \mathrm{tr}[e^y]$ reads as $H(\lambda) \geq H(\delta)$, where

$$H: \mathbb{R}^d \to \mathbb{R}, \qquad H(\alpha) = H(\alpha_1, \ldots, \alpha_d) := \sum_{i=1}^{d} e^{\alpha_i}.$$

The function H is strictly Schur-convex by Theorem 5.7, since the exponential function is strictly convex. Since $\lambda \succ \delta$ by Theorem 5.6, we then have $H(\lambda) \geq H(\delta)$, as required. Moreover, assume that $H(\lambda) = H(\delta)$. Then, by *strict* Schur-convexity of H, we have $\lambda^{\downarrow} = \delta^{\downarrow}$. Again by Theorem 5.6, we conclude that x is diagonal, which in turn implies $x = y$. $\qquad\square$

From the latter proposition we deduce the existence and uniqueness of a minimiser of χ_d on $\mathrm{Sym}_d^V(z)$, under the assumption that all the 'fixed' entries z are diagonal matrices.

Theorem 5.9 *Let* $z = (z_v)_{v \in \Gamma} \in \mathrm{Diag}_d^{\Gamma}$ *and set* $z(i, i) := (z_v(i, i))_{v \in \Gamma} \in \mathbb{R}^{\Gamma}$ *for all* $1 \leq i \leq d$. *Then, the function* χ_d *admits a unique minimiser on* $\mathrm{Sym}_d^V(z)$. *Such a minimiser is of the form* $m = (m_v)_{v \in V} \in \mathrm{Diag}_d^V(z)$, *where* $m(i, i) := (m_v(i, i))_{v \in V} \in \mathbb{R}^V$ *denotes the unique minimiser of* χ_1 *on* $\mathbb{R}^V(z(i, i))$ *for all* i.

Proof The claim will immediately follow from the two following facts:

(i) for any $x \in \mathrm{Sym}_d^V(z)$, there exists $y \in \mathrm{Diag}_d^V(z)$ such that $\chi_d(x) \geq \chi_d(y)$, with equality if and only if $x = y$;

(ii) there exists $m \in \mathrm{Diag}_d^V(z)$ (as in the statement of the theorem) such that $\chi_d(x) \geq \chi_d(m)$ for any $x \in \mathrm{Diag}_d^V(z)$, with equality if and only if $x = m$.

Proof of (i). Fix any $x \in \mathrm{Sym}_d^V(z)$. Define $y = (y_v)_{v \in V}$ so that, for all $v \in V$, y_v is the diagonal matrix with the same diagonal entries as x, i.e. $y_v(i, i) = x_v(i, i)$ for $1 \leq i \leq d$. Since each z_v (for $v \in \Gamma$) is diagonal by hypothesis, we have

$y \in \mathrm{Diag}_d^V(z)$. For any $v, w \in V$, the matrices $x_v - x_w \in \mathrm{Sym}_d$ and $y_v - y_w \in \mathrm{Diag}_d$ have the same diagonal entries, hence $\mathrm{tr}[e^{x_v - x_w}] \geq \mathrm{tr}[e^{y_v - y_w}]$ by Prop. 5.8; summing over $v \to w$, we obtain that $\chi_d(x) \geq \chi_d(y)$. Assume now that $\chi_d(x) = \chi_d(y)$. Then, $\mathrm{tr}[e^{x_v - x_w}] = \mathrm{tr}[e^{y_v - y_w}]$ whenever $v \to w$. Again by Prop. 5.8, we then have $x_v - y_v = x_w - y_w$ for all $v \to w$. For any $v \in V$, by Lemma 5.1 there exists a path $(v_0, v_1, \ldots, v_l)$ in G from $v_0 = v$ to a sink v_l. Since all sinks are in Γ by assumption (see Sect. 5.1) and both x and y are in $\mathrm{Sym}_d^V(z)$, we have $x_{v_l} = z_{v_l} = y_{v_l}$. Therefore, $x_v - y_v = x_{v_1} - y_{v_1} = \cdots = x_{v_l} - y_{v_l} = 0$; in particular, $x_v = y_v$. As $v \in V$ is arbitrary, we conclude that $x = y$.

   Proof of (ii). For $x \in \mathrm{Diag}_d^V(z)$, set $x(i, i) := (x_v(i, i))_{v \in V} \in \mathbb{R}^V$. As each x_v is diagonal, we have

$$\chi_d(x) = \sum_{v \to w} \mathrm{tr}[e^{x_v - x_w}] = \sum_{v \to w} \sum_{i=1}^{d} e^{x_v(i,i) - x_w(i,i)} = \sum_{i=1}^{d} \chi_1(x(i, i)).$$

By Prop. 5.5, for all i, χ_1 has a unique minimiser $m(i, i)$ on $\mathbb{R}^V(z(i, i))$. Therefore, we have

$$\chi_d(x) = \sum_{i=1}^{d} \chi_1(x(i, i)) \geq \sum_{i=1}^{d} \chi_1(m(i, i)) = \chi_d(m),$$

and the inequality is strict whenever $x \neq m$.   □

   In the case where the 'fixed' entries z are scalar matrices, the inequality (5.3) immediately implies the existence and uniqueness of a minimiser of φ_d.

Corollary 5.10 *Let $z = (z_v)_{v \in \Gamma} \in \mathrm{Scal}_d^\Gamma$, so that $z_v = \zeta_v I_d$ for all $v \in \Gamma$ and for a certain $\zeta = (\zeta_v)_{v \in V} \in \mathbb{R}^\Gamma$. Then, the function φ_d admits a unique minimiser on $\mathrm{Sym}_d^V(z)$. Such a minimiser is of the form $m = (m_v)_{v \in V} = (\mu_v I_d)_{v \in V} \in \mathrm{Scal}_d^V(z)$, where $\mu = (\mu_v)_{v \in V} \in \mathbb{R}^V$ is the unique minimiser of φ_1 on $\mathbb{R}^V(\zeta)$.*

Proof Since $\mathrm{Scal}_d^\Gamma \subseteq \mathrm{Diag}_d^\Gamma$, it follows from Theorem 5.9 that χ_d has a unique minimiser m on $\mathrm{Sym}_d^V(z)$, which is of the form specified above. Since the scalar matrices m_v and m_w commute for any $v, w \in V$, we have $e^{m_v} e^{-m_w} = e^{m_v - m_w}$, hence $\varphi_d(m) = \chi_d(m)$. By (5.3), we then have

$$\varphi_d(x) \geq \chi_d(x) \geq \chi_d(m) = \varphi_d(m) \qquad \text{for all } x \in \mathrm{Sym}_d^V(z),$$

where the second inequality is strict if $x \neq m$. It follows that m is also the unique minimiser of φ_d on $\mathrm{Sym}_d^V(z)$.   □

5.3 Energy functions in logarithmic variables

It is a well-known fact that the functions

$$\mathrm{Sym}_d \to \mathcal{P}_d, \quad a \mapsto e^a \qquad \text{and} \qquad \mathcal{P}_d \to \mathrm{Sym}_d, \quad a \mapsto \log a,$$

namely the matrix exponential and the matrix logarithm, are both bijections on the stated domains and inverse to each other. From now on, for any set S, we will use the following compact notations: $\log x := (\log x_v)_{v \in S} \in \mathrm{Sym}_d^S$ for $x = (x_v)_{v \in S} \in \mathcal{P}_d^S$, and $\mathrm{e}^x := (\mathrm{e}^{x_v})_{v \in S} \in \mathcal{P}_d^S$ for $x = (x_v)_{v \in S} \in \mathrm{Sym}_d^S$.

Let us consider the analogue of φ_d 'in logarithmic variables', that is the energy function $\Phi_d(x) := \varphi_d(\log x)$ for $x \in \mathcal{P}_d^V$. More explicitly, recalling (5.1), we define

$$\Phi_d : \mathcal{P}_d^V \to \mathbb{R}, \qquad \Phi_d(x) := \sum_{\substack{v, w \in V: \\ v \to w}} \mathrm{tr}[x_v x_w^{-1}] \qquad \text{for all } x = (x_v)_{v \in V} \in \mathcal{P}_d^V.$$

$$(5.9)$$

Take now $z = (z_v)_{v \in \Gamma}$ such that each z_v is a positive multiple of I_d, or equivalently $\log z \in \mathrm{Scal}_d^\Gamma$. By Corollary 5.10, Φ_d has a unique minimiser m on $\mathcal{P}_d^V(z)$, where $\log m$ is the unique minimiser of φ_d on $\mathrm{Sym}_d^V(\log z)$. This implies that, on $\mathcal{P}_d^V(z)$, the Hessian of Φ_d at m is positive semidefinite. We now aim to prove the stronger statement that the Hessian of Φ_d at m is *positive definite*.

As in the previous subsections, we first work with $d = 1$. Recall that, by Lemma 5.4, φ_1 is strictly convex on $\mathbb{R}^V(\log z)$, for $z \in \mathcal{P}_1^\Gamma$. The analogous statement does not hold for Φ_1 on $\mathcal{P}_1^V(z)$; however, the following is still true:

Lemma 5.11 *For $z \in \mathcal{P}_1^\Gamma$, the Hessian of Φ_1 on $\mathcal{P}_1^V(z)$ is positive definite at any critical point.*

Proof We prove the claim by simply expressing the derivatives of Φ_1 in terms of the derivatives of φ_1. For $v \in \Gamma^{\mathsf{c}}$, the first partial derivative of Φ_1 w.r.t. x_v is

$$\partial_v \Phi_1(x) = \frac{1}{x_v} \partial_v \varphi_1(\log x) \qquad \text{for any } x \in \mathcal{P}_1^V(z). \qquad (5.10)$$

Therefore, for $v, w \in \Gamma^{\mathsf{c}}$,

$$\partial_v \partial_w \Phi_1(x) = \begin{cases} \frac{1}{x_v^2} \left[\partial_v^2 \varphi_1 - \partial_v \varphi_1 \right] (\log x) & \text{if } v = w, \\ \frac{1}{x_v x_w} \partial_v \partial_w \varphi_1(\log x) & \text{if } v \neq w. \end{cases}$$

Assume now that x is a critical point of Φ_1 on $\mathcal{P}_1^V(z)$, i.e. that both sides of (5.10) vanish for all $v \in \Gamma^{\mathsf{c}}$. Then, we have

$$\partial_v \partial_w \Phi_1(x) = \frac{1}{x_v x_w} \partial_v \partial_w \varphi_1(\log x) \qquad \text{for all } v, w \in \Gamma^{\mathsf{c}}.$$

As the Hessian of φ_1 on $\mathbb{R}^V(\log z)$ is positive definite everywhere by Lemma 5.4, it follows that the Hessian of Φ_1 on $\mathcal{P}_1^V(z)$ is positive definite at x. $\qquad \square$

To compute the Hessian in the general case $d \geq 1$, we will use the following basic formulas (see e.g. [36]) that hold for any $a, b \in \mathcal{P}_d$, $1 \leq i \leq j \leq d$, and $1 \leq k \leq \ell \leq d$:

$$\frac{\partial (a^{-1})(k, \ell)}{\partial a(i, j)} = -\frac{1}{1 + \delta(i, j)} \left[a^{-1}(k, i)a^{-1}(\ell, j) + a^{-1}(k, j)a^{-1}(\ell, i) \right], \quad (5.11)$$

$$\frac{\partial (a^{-1}ba^{-1})(k, \ell)}{\partial a(i, j)}$$
$$= -\frac{a^{-1}(k, i)(a^{-1}ba^{-1})(\ell, j) + (a^{-1}ba^{-1})(k, i)a^{-1}(\ell, j) + \boxed{i \leftrightarrow j}}{1 + \delta(i, j)} \quad (5.12)$$

$$\frac{\partial}{\partial a(i, j)} \operatorname{tr}[ab^{-1}] = \frac{2}{1 + \delta(i, j)} b^{-1}(i, j), \quad (5.13)$$

$$\frac{\partial}{\partial b(i, j)} \operatorname{tr}[ab^{-1}] = -\frac{2}{1 + \delta(i, j)} \left[b^{-1}ab^{-1} \right](i, j), \quad (5.14)$$

where $\delta(i, j)$ is 1 if $i = j$ and 0 otherwise. In (5.12), $\boxed{i \leftrightarrow j}$ denotes the preceding expression with the indices i and j swapped. Notice that (5.12) and (5.14) can be deduced from (5.11).

Let Scal_d^+ be the set of positive definite scalar matrices, i.e. positive multiples of I_d.

Lemma 5.12 *Let $z \in (\operatorname{Scal}_d^+)^\Gamma$. Then, the Hessian of Φ_d on $\mathcal{P}_d^V(z)$ is positive definite at any critical point x such that $x \in (\operatorname{Scal}_d^+)^V(z)$.*

Proof We will prove that, under the stated assumptions, the Hessian of Φ_d (for $d \geq 1$) can be expressed in terms of the Hessian of Φ_1; the claim will then follow from Lemma 5.11.

For ease of notation, given any $v \in V$ and $1 \leq i \leq j \leq d$, we will denote by $\partial_{v;i,j}$ the partial derivative of a function of $x \in \mathcal{P}_d^V$ with respect to the real variable $x_v(i, j)$.

It follows from the definition (5.9) and from the formulas (5.13)–(5.14) that

$$\partial_{v;i,j} \Phi_d(x) = \frac{2}{1 + \delta(i, j)} \left(\sum_{\substack{w \in V: \\ v \to w}} x_w^{-1}(i, j) - \sum_{\substack{u \in V: \\ u \to v}} \left[x_v^{-1} x_u x_v^{-1} \right](i, j) \right)$$

for $v \in \Gamma^c$ and $1 \leq i \leq j \leq d$. The critical point equations of Φ_d on $\mathcal{P}_d^V(z)$ are then

$$x_v^{-1} \left(\sum_{\substack{u \in V: \\ u \to v}} x_u \right) x_v^{-1} = \sum_{\substack{w \in V: \\ v \to w}} x_w^{-1} \qquad \text{for all } v \in \Gamma^c. \quad (5.15)$$

We will now compute the second derivatives at any critical point $x = (x_v)_{v \in V} \in \mathcal{P}_d^V(z)$. Using (5.12) and (5.15), we have

$$
\begin{aligned}
\partial_{v;k,\ell}\partial_{v;i,j}\Phi_d(x) &= \frac{2}{1 + \delta(i,j)}\frac{1}{1 + \delta(k,\ell)}\left(x_v^{-1}(i,k)\left[x_v^{-1}\left(\sum_{\substack{u \in V: \\ u \to v}}x_u\right)x_v^{-1}\right](j,\ell)\right. \\
&\quad + \left[x_v^{-1}\left(\sum_{\substack{u \in V: \\ u \to v}}x_u\right)x_v^{-1}\right](i,k)x_v^{-1}(j,\ell) + \boxed{k \leftrightarrow \ell} \\
&= \frac{2}{1 + \delta(i,j)}\frac{1}{1 + \delta(k,\ell)}\left(x_v^{-1}(i,k)\sum_{\substack{u \in V: \\ v \to u}}x_u^{-1}(j,\ell)\right. \\
&\quad + \sum_{\substack{u \in V: \\ v \to u}}x_u^{-1}(i,k)x_v^{-1}(j,\ell) + \boxed{k \leftrightarrow \ell}\bigg)
\end{aligned}
$$

for $v \in \Gamma^c$, $1 \le i \le j \le d$, and $1 \le k \le \ell \le d$. Recall now that the acyclic structure of the underlying graph guarantees that, if $v \to w$, then $w \not\to v$. Therefore, for $v, w \in \Gamma^c$ such that $v \to w$, $1 \le i \le j \le d$, and $1 \le k \le \ell \le d$, we have

$$
\partial_{w;k,\ell}\partial_{v;i,j}\Phi_d(x) = -\frac{2}{1 + \delta(i,j)}\frac{1}{1 + \delta(k,\ell)}\left[x_w^{-1}(i,k)x_w^{-1}(j,\ell) + x_w^{-1}(i,\ell)x_w^{-1}(j,k)\right].
$$

On the other hand, the second derivative w.r.t. $x_v(i,j)$ and $x_w(k,\ell)$ vanishes for all $v, w \in \Gamma^c$ such that $v \not\to w$ and $w \not\to v$.

According to the hypotheses of the theorem, we further assume from now on that there exists $\zeta = (\zeta_v)_{v \in V} \in \mathcal{P}_1^\Gamma$ such that $z_v = \zeta_v I_d$ for all $v \in \Gamma$, and there exists $\xi = (\xi_v)_{v \in V} \in \mathcal{P}_1^V$ such that $x_v = \xi_v I_d$ for all $v \in V$. Using the identity

$$
\frac{\delta(i,k)\delta(j,\ell) + \delta(i,\ell)\delta(j,k)}{1 + \delta(k,\ell)} = \delta((i,j),(k,\ell)) \qquad \text{for } 1 \le i \le j \le d, \ 1 \le k \le \ell \le d,
$$

we see that the second derivatives at x factorise as

$$
\partial_{w;k,\ell}\partial_{v;i,j}\Phi_d(x) = f((i,j),(k,\ell))\,g_\xi(v,w), \qquad \text{with}
$$
$$
f((i,j),(k,\ell)) = \frac{2\delta((i,j),(k,\ell))}{1 + \delta(i,j)}.
$$

Here, $g_\xi(v,w)$ is an explicit function of ξ, v and w; we stress that it is the same function for all $d \ge 1$. It follows from (5.15) that, since $x = (\xi_v I_d)_{v \in V}$ is a critical point of Φ_d on $\mathcal{P}_d^V(z)$, ξ is a critical point of Φ_1 on $\mathcal{P}_1^V(\zeta)$. Therefore, the matrix

$$
g_\xi(v,w) = \partial_v\partial_w\Phi_1(\xi)
$$

(with 'row index' v and 'column index' w), which is the Hessian matrix of Φ_1 on $\mathcal{P}_1^V(\zeta)$ at ξ, is positive definite by Lemma 5.11. On the other hand, the matrix $f((i, j), (k, \ell))$ (with 'row index' (i, j) and 'column index' (k, ℓ)) is clearly positive definite as a diagonal matrix with positive diagonal entries. Therefore, the Hessian of Φ_d on $\mathcal{P}_d^V(z)$ at x is positive definite, as it can be written as a Kronecker product of two positive definite matrices. □

As any minimiser is a critical point, the main result of this section follows immediately from Corollary 5.10 and Lemma 5.12.

Theorem 5.13 *Let* $z \in (\mathrm{Scal}_d^+)^\Gamma$. *Then, the function* Φ_d *on* $\mathcal{P}_d^V(z)$ *has a unique (global) minimiser* m, *at which the Hessian is positive definite. Moreover, we have* $m \in (\mathrm{Scal}_d^+)^V(z)$.

5.4 Laplace approximation

We will now use Theorem 5.13 to study the asymptotic behaviour of integrals of exponentials of Φ_d, via Laplace's approximation method. Recall the definition (1.12) of the measure μ on $\mathcal{P}_d$.

Proposition 5.14 *Let* $z \in (\mathrm{Scal}_d^+)^\Gamma$ *and let* m *be the unique global minimiser of* Φ_d *on* $\mathcal{P}_d^V(z)$ *(see Theorem 5.13). Let* $g : \mathcal{P}_d^V(z) \to \mathbb{C}$ *be a continuous function in a neighbourhood of* m, *with* $g(m) \neq 0$, *and such that*

$$\int_{\mathcal{P}_d^V(z)} \left(\prod_{v \in \Gamma^c} \mu(\mathrm{d}x_v) \right) |g(x)| \, e^{-k\Phi_d(x)} < \infty \qquad \text{for some } k > 0 \, .$$

Then

$$\int_{\mathcal{P}_d^V(z)} \left(\prod_{v \in \Gamma^c} \mu(\mathrm{d}x_v) \right) g(x) \, e^{-k\Phi_d(x)}$$
$$\xrightarrow{k \to \infty} \frac{g(m)}{\sqrt{|\mathcal{H}(m)|}} \left(\prod_{v \in \Gamma^c} |m_v|^{-\frac{d(d+1)}{2}} \right) \left(\frac{2\pi}{k} \right)^{|\Gamma^c| \frac{d(d+1)}{4}} e^{-k\Phi_d(m)}, \quad (5.16)$$

where $|\mathcal{H}(m)| > 0$ *is the Hessian determinant of* Φ_d *at* m *and* $|\Gamma^c|$ *is the number of vertices in* Γ^c.

We start by stating the Laplace approximation integral formula in the multivariate context, which can be found e.g. in [19].

Theorem 5.15 ([19, Theorem 4.14]) *Let* A *be an open subset of the p-dimensional space* $\mathbb{R}^p$. *Let* $h : A \to \mathbb{C}$ *and* $\varrho : A \to \mathbb{R}$ *be functions such that*

  (i) $\int_A |h(x)| \, e^{-k\varrho(x)} \, \mathrm{d}x < \infty$ *for some* $k > 0$.
  (ii) ϱ *has a global minimiser* $x_0 \in A$ *such that, for every* $\varepsilon > 0$,

$$\inf\{\varrho(x) - \varrho(x_0) : x \in A, \, |x - x_0| \geq \varepsilon\} > 0 \, . \qquad (5.17)$$

(iii) h is continuous in a neighbourhood of x_0 and $h(x_0) \neq 0$.

(iv) ϱ is twice continuously differentiable on A and its Hessian matrix $\mathcal{H}(x_0)$ at x_0 is positive definite (in particular, its determinant $|\mathcal{H}(x_0)|$ is positive).

Then,

$$\int_A h(x)\, \mathrm{e}^{-k\varrho(x)}\, \mathrm{d}x \xrightarrow{k \to \infty} \frac{h(x_0)}{\sqrt{|\mathcal{H}(x_0)|}} \left(\frac{2\pi}{k} \right)^{\frac{p}{2}} \mathrm{e}^{-k\varrho(x_0)}\,. \qquad (5.18)$$

Proof of Prop. 5.14 We will apply Theorem 5.15 with

$$A = \mathcal{P}_d^V(z)\,, \qquad h(x) = g(x) \prod_{v \in \Gamma^{\mathrm{c}}} |x_v|^{-\frac{d(d+1)}{2}}\,, \qquad \varrho = \Phi_d\,, \qquad x_0 = m\,.$$

The set $\mathcal{P}_d^V(z)$ can be clearly viewed as an open subset of $\mathbb{R}^p$, where $p = \left|\Gamma^{\mathrm{c}}\right| d(d + 1)/2$ is the number of 'free' real variables in A and d is the dimension of each matrix in the array. The extra product in the definition of h is the density of the measure $\prod_{v \in \Gamma^{\mathrm{c}}} \mu(\mathrm{d}x_v)$ with respect to the Lebesgue measure on $\mathcal{P}_d^V(z)$.

Hypothesis (ii) of Theorem 5.15 is satisfied due to Theorem 5.13 and Prop. 5.2. Hypotheses (i) and (iii) are matched by the assumptions of Prop. 5.14. Finally, hypothesis (iv) also holds because of Theorem 5.13. The asymptotic formula (5.16) then follows from (5.18). ∎

Appendix A: A proof of the Cauchy–Littlewood identity

In this appendix we include a proof of the classical Cauchy–Littlewood identity for Schur functions that is based on a version of the Pieri rule. The proof of the Whittaker integral identity (4.1) is based, *mutatis mutandis*, on the same line of reasoning.

For any two integer partitions $\mu = (\mu_1 \geq \mu_2 \geq \cdots)$ and $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots)$, we write $\mu \prec \lambda$ if $\lambda_i \geq \mu_i \geq \lambda_{i+1}$ for all $i \geq 1$. The *size* of λ is $|\lambda| := \lambda_1 + \lambda_2 + \ldots$, while its *length* is the smallest $i \geq 0$ such that $\lambda_{i+1} = 0$. We will adopt the following recursive definition of Schur polynomials: for a partition λ of length $\leq n$, we set

$$s_\lambda(x_1, \ldots, x_n) := \begin{cases} x_1^{|\lambda|} & \text{if } n = 1, \\ \sum_{\mu \prec \lambda} x_n^{|\lambda|-|\mu|} s_\mu(x_1, \ldots, x_{n-1}) & \text{if } n > 1. \end{cases} \tag{A.1}$$

For the sake of convenience, we also set $s_\lambda(x_1, \ldots, x_n) := 0$ if the length of λ exceeds n. This definition is easily seen to be equivalent to the classical combinatorial definition of Schur polynomials as generating functions of semistandard Young tableaux.

We will use the following version of the Pieri rule:

$$\sum_{\lambda: \, \mu \prec \lambda} s_\lambda(x_1, \ldots, x_n) = \left(\prod_{i=1}^{n} \frac{1}{1 - x_i} \right) s_\mu(x_1, \ldots, x_n). \tag{A.2}$$

The latter can be deduced from the usual Pieri rule (see e.g. [29, I-(5.16)])

$$\sum_{\substack{\lambda: \, \mu \prec \lambda, \\ |\lambda|-|\mu|=r}} s_\lambda(x_1, \ldots, x_n) = \left(\sum_{\substack{k_1,\ldots,k_n \geq 0: \\ k_1+\cdots+k_n=r}} x_1^{k_1} \cdots x_n^{k_n} \right) s_\mu(x_1, \ldots, x_n)$$

by summing over all $r \geq 0$. Notice that (A.2) can be read as an eigenfunction equation for the operator defined through the kernel $I(\mu; \lambda) := \mathbb{1}_{\mu \prec \lambda}$, with the Schur function s_μ (viewed as a function of the partition μ) as an eigenfunction.

Theorem A.1 (Cauchy–Littlewood identity) *For any* $n, N \geq 1$, *we have*

$$\sum_\lambda s_\lambda(x_1, \ldots, x_n) s_\lambda(y_1, \ldots, y_N) = \prod_{\ell=1}^{n} \prod_{i=1}^{N} \frac{1}{1 - x_\ell y_i}, \tag{A.3}$$

where the sum is over all integer partitions λ.

Proof Note first that we can restrict the sum on the left-hand side of (A.3) to the partitions λ with length $\leq \min(n, N)$. When $N = n = 1$, the identity reduces to a geometric sum:

$$\sum_{\lambda_1 \geq 0} x_1^{\lambda_1} y_1^{\lambda_1} = \frac{1}{1 - x_1 y_1}. \tag{A.4}$$

We can then proceed by induction on $n + N$.

Let $n + N > 2$ and assume, without loss of generality, that $n > 1$. Using the definition (A.1), the fact that s_λ is a homogeneous polynomial of degree $|\lambda|$, and

identity (A.2), we obtain

$$
\sum_\lambda s_\lambda(x_1, \ldots, x_n) s_\lambda(y_1, \ldots, y_N)
$$

$$
= \sum_\lambda \left(\sum_{\mu \prec \lambda} x_n^{|\lambda| - |\mu|} s_\mu(x_1, \ldots, x_{n-1}) \right) s_\lambda(y_1, \ldots, y_N)
$$

$$
= \sum_\mu \left(\sum_{\lambda \,:\, \mu \prec \lambda} s_\lambda(x_n y_1, \ldots, x_n y_N) \right) x_n^{-|\mu|} s_\mu(x_1, \ldots, x_{n-1})
$$

$$
= \sum_\mu \left(\prod_{i=1}^N \frac{1}{1 - x_n y_i} \right) s_\mu(x_n y_1, \ldots, x_n y_N) x_n^{-|\mu|} s_\mu(x_1, \ldots, x_{n-1})
$$

$$
= \left(\prod_{i=1}^N \frac{1}{1 - x_n y_i} \right) \sum_\mu s_\mu(x_1, \ldots, x_{n-1}) s_\mu(y_1, \ldots, y_N) .
$$

The claim then follows from the induction hypothesis applied to the latter sum. □

Appendix B: Markov functions and intertwinings

In this appendix we review the theory of Markov functions, in the case of inhomogeneous discrete-time Markov processes, which we are concerned with in the present article.

Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be measurable spaces and $\varphi \colon S \to T$ be a measurable function. Let $X = (X(n))_{n \geq 0}$ be a time-inhomogeneous Markov process with state space S, time-n transition kernel Π_n and any initial distribution on $X(0)$. Defining $Z(n) := \varphi(X(n))$ for all $n \geq 0$, we will give conditions under which the transformed process $Z = (Z(n))_{n \geq 0}$ with state space T is still Markov in its own filtration. The well-known Dynkin criterion [18] ensures that Z satisfies the Markov property for any possible initial distribution on X. On the other hand, the theory of Markov functions (developed at various levels of generality in [25, 26, 28, 37]) provides a more subtle criterion, in which the Markov property of Z is guaranteed only under certain specific initial states of X.

Let $\mathfrak{b}\mathcal{S}$ be the space of bounded measurable functions from $(S, \mathcal{S})$ to $\mathbb{R}$.

Theorem B.1 *Let $X = (X(n))_{n \geq 0}$ be a time-inhomogeneous Markov process on S with time-n transition kernel Π_n. Let $\varphi \colon S \to T$ be a measurable function. Let $Z = (Z(n))_{n \geq 0}$, where $Z(n) = \varphi(X(n))$ for all $n \geq 0$. Assume that $\mathcal{T}$ contains all the singleton sets $\{z\}$. Let Σ be a Markov kernel from T to S and, for all $n \geq 1$, let P_n be a Markov kernel from T to T. Suppose:*

(i) $\Sigma \left(z; \varphi^{-1}\{z\} \right) = 1$ for every $z \in T$;
(ii) $\Sigma \Pi_n = P_n \Sigma$ for all $n \geq 1$.

Assume that, for an arbitrary $z \in T$, the initial state $X(0)$ of X is distributed according to the measure $\Sigma(z; \cdot)$. Then, Z is a time-inhomogeneous Markov process (in its own filtration), with initial state z and time-n transition kernel P_n. Moreover, for all $f \in \mathfrak{b}\mathcal{S}$ and $n \geq 0$, we have

$$\mathbb{E}[f(X(n)) \mid Z(0), \ldots, Z(n-1), Z(n)] = \Sigma f(Z(n)) \quad a.s. \tag{B.1}$$

Proof This proof is an inhomogeneous discrete-time version of the argument given for continuous-time Markov processes in [37]. Note that (i) implies

$$\int_S \Sigma(z; \mathrm{d}x) g(\varphi(x)) f(x) = g(z) \int_S \Sigma(z; \mathrm{d}x) f(x)$$

for all $g \in \mathfrak{b}\mathcal{T}$, $f \in \mathfrak{b}\mathcal{S}$, and $z \in T$. Letting $\Phi \colon \mathfrak{b}\mathcal{T} \to \mathfrak{b}\mathcal{S}$ be the Markov operator defined by $\Phi g := g \circ \varphi$ for $g \in \mathfrak{b}\mathcal{T}$, we may rewrite the above identity as

$$\Sigma(\Phi g) f = g \Sigma f . \tag{B.2}$$

Here, as in the following, the operations should be read from right to left, prioritising the brackets (for example, on the left-hand side of (B.2), one first multiplies the two functions f and Φg and then applies the operator Σ to the resulting function). Applying P_i to both sides of (B.2) and using hypothesis (ii), we have

$$\Sigma \Pi_i (\Phi g) f = P_i g \Sigma f \qquad \text{for all } i \geq 1 . \tag{B.3}$$

Consider now test functions $g_0, \ldots, g_n$ in $\mathfrak{b}\mathcal{T}$ and $f \in \mathfrak{b}\mathcal{S}$. Using (B.2) and (B.3) several times, we obtain

$$\begin{aligned}
&\Sigma(\Phi g_0) \Pi_1 (\Phi g_1) \Pi_2 (\Phi g_2) \cdots \Pi_n (\Phi g_n) f \\
&= g_0 \Sigma \Pi_1 (\Phi g_1) \Pi_2 (\Phi g_2) \cdots \Pi_n (\Phi g_n) f \\
&= g_0 P_1 g_1 \Sigma \Pi_2 (\Phi g_2) \cdots \Pi_n (\Phi g_n) f = \cdots = g_0 P_1 g_1 P_2 g_2 \cdots P_n g_n \Sigma f .
\end{aligned} \tag{B.4}$$

Fix now an arbitrary $z \in T$ and assume that $X(0)$ is distributed according to $\Sigma(z; \cdot)$. Then, (B.4) yields

$$\mathbb{E}\left[g_0(Z(0)) \, g_1(Z(1)) \cdots g_n(Z(n)) \, f(X(n))\right] = g_0 P_1 g_1 \cdots P_n g_n \Sigma f(z) .$$

Taking $f \equiv 1$, we deduce that Z is a Markov process started at z with time-n transition kernel P_n. For general f, the right-hand side of the equation above agrees with

$$\mathbb{E}\left[g_0(Z(0)) \, g_1(Z(1)) \cdots g_n(Z(n)) \, \Sigma f(Z(n))\right] .$$

This, by definition of conditional expectation, proves (B.1). $\qquad\square$

Remark B.2 Taking $f \equiv 1$ in (B.2), we see that $\Sigma \Phi$ is the identity on $\mathfrak{b}\mathcal{T}$. Combining this with hypothesis (ii) of Theorem B.1, it is immediate to deduce that every kernel P_n is uniquely determined by the relation $P_n = \Sigma \Pi_n \Phi$.

Appendix C: A convergence lemma

Here we state a useful convergence lemma. For completeness we also include its proof, which follows from the properties of weak convergence and standard estimates.

Lemma C.1 *Let S be a locally compact metric space equipped with its Borel σ-algebra. Let $(\mu_k)_{k>0}$ be a collection of probability measures on S that converges weakly, as $k \to \infty$, to a Dirac measure δ_s for some $s \in S$. Let $(f_k)_{k>0}$ be a uniformly bounded collection of continuous functions $S \to \mathbb{R}$ such that $f_k \xrightarrow{k\to\infty} f_\infty$ uniformly on any compact subset of S. Then*

$$\lim_{k\to\infty} \int_S \mu_k(\mathrm{d}x) f_k(x) = f_\infty(s). \tag{C.1}$$

Proof Fix $\varepsilon > 0$. For any Borel set $U \subset S$, we may write

$$\left| \int_S \mu_k(\mathrm{d}x) f_k(x) - f_\infty(s) \right| \leq \left| \int_U \mu_k(\mathrm{d}x)[f_k(x) - f_\infty(s)] \right|$$
$$+ \left| \int_{U^c} \mu_k(\mathrm{d}x)[f_k(x) - f_\infty(s)] \right|$$
$$\leq \int_{\overline{U}} \mu_k(\mathrm{d}x)\, |f_k(x) - f_\infty(x)| + \int_U \mu_k(\mathrm{d}x)\, |f_\infty(x) - f_\infty(s)|$$
$$+ \left(\sup_{x\in S} |f_k(x)| + |f_\infty(s)| \right) \mu_k(U^c),$$

where $\overline{U}$ and U^c are the closure and the complement of U, respectively. Since f_∞ is continuous (as a uniform limit of continuous functions) and S is a locally compact metric space, we can choose U to be a precompact open neighbourhood of s such that $|f_\infty(x) - f_\infty(s)| \leq \varepsilon$ for all $x \in U$. Moreover, as $\overline{U}$ is compact, for k large enough we have $|f_k(x) - f_\infty(x)| \leq \varepsilon$ for all $x \in \overline{U}$. Finally, the Portmanteau theorem (see [5, § 2]) yields

$$\limsup_{k\to\infty} \mu_k(U^c) = \delta_s(U^c) = 0,$$

since μ_k converges weakly to δ_s, U^c is closed, and $s \notin U^c$; therefore, for large enough k we also have $\mu_k(U^c) \leq \varepsilon$. By the hypothesis of uniform boundedness, there exists $M > 0$ such that $|f_k(x)| \leq M$ for all $x \in S$ and $k > 0$. Hence, we have

$$\left| \int_S \mu_k(\mathrm{d}x) f_k(x) - f_\infty(s) \right| \leq \mu_k(\overline{U})\varepsilon + \mu_k(U)\varepsilon + [M + |f_\infty(s)|]\,\varepsilon$$
$$\leq [2 + M + |f_\infty(s)|]\,\varepsilon$$

for k large enough. As $\varepsilon > 0$ is arbitrary, the claim follows.   □

References

1. Arista, J., Bisi, E., O'Connell, N.: Matsumoto-Yor and Dufresne type theorems for a random walk on positive definite matrices. Ann. Inst. H. Poincaré (B) Probab. Statist. (2023+). arXiv: 2112.12558
2. Barraquand, G., Wang, S.: An identity in distribution between full-space and half-space log-gamma polymers. Int. Math. Res. Not. (2023+). arXiv:2108.08737
3. Bhatia, R.: Matrix Analysis Graduate. Texts in Mathematics. Springer, New York (1997)
4. Billingsley, P.: Probability and Measure. Wiley Series in Probability and Mathematical Statistics, 3rd ed. Wiley, Hoboken (1995)
5. Billingsley, P.: Convergence of Probability. Measures Wiley Series in Probability and Statistics, 2nd ed. Wiley, Hoboken (1999)
6. Bisi, E., Zygouras, N.: Point-to-line polymers and orthogonal Whittaker functions. Trans. Am. Math. Soc. **371**(12), 8339–8379 (2019)
7. Borodin, A., Corwin, I.: Macdonald processes. Probab. Theory Relat. Fields **158**(1–2), 225–400 (2014)
8. Borodin, A., Ferrari, P.L.: Anisotropic growth of random surfaces in 2+1 dimensions. Commun. Math. Phys. **325**, 603–684 (2014)
9. Borodin, A., Gorin, V.: Lectures on integrable probability. In: Sidoravicius, V., Smirnov, S. (eds.) Probability and Statistical Physics in St. Petersburg, Vol. 91. Proceedings of Symposia in Pure Mathematics. AMS, pp. 155– 214 (2016)
10. Borodin, A., Petrov, L.: Integrable probability: from representation theory to Macdonald processes. Probab. Surveys **11**, 1–58 (2014)
11. Borodin, A., Petrov, L.: Nearest neighbor Markov dynamics on Macdonald processes. Adv. Math. **300**, 71–155 (2016)
12. Bump, D.: The Rankin-Selberg method: a survey. In: Aubert, K.E., Bombieri, E., Goldfeld, D. (eds.) Number Theory, Trace Formulas and Discrete Groups, pp. 49–109. Academic Press, London (1989)
13. Bump, D., Huntley, J.: Unramified Whittaker functions for $GL(3, \mathbb{R})$. J. Anal. Math. **65**, 19–44 (1995)
14. Butler, R.W., Wood, A.T.A.: Laplace approximation for Bessel functions of matrix argument. J. Comput. Appl. Math. **155**(2), 359–382 (2003)
15. Corwin, I.: Kardar–Parisi–Zhang Universality. Not. Am. Math. Soc. **63**(3), 230–239 (2016)
16. Corwin, I., O'Connell, N., Seppäläinen, T., Zygouras, N.: Tropical combinatorics and Whittaker functions. Duke Math. J. **163**(3), 513–563 (2014)
17. Corwin, I., Seppäläinen, T., Shen, H.: The strict-weak lattice polymer. J. Stat. Phys. **160**, 1027–1053 (2015)
18. Dynkin, E.B.: Non-negative eigenfunctions of the Laplace-Beltrami operator and Brownian motion in certain symmetric spaces. Dokl. Akad. Nauk SSSR **141**(2), 288–291 (1961)
19. Evans, M., Swartz, T.: Approximating Integrals via Monte Carlo and Deterministic Methods. Oxford Statistical Science Series, Oxford Univ Press, Oxford (2000)
20. Gautié, T., Bouchaud, J.-P., Le Doussal, P.: Matrix Kesten recursion, inverse-Wishart ensemble and fermions in a Morse potential. J. Phys. A: Math. Theor. **54**(25), 255201 (2021)
21. Gerasimov, A., Lebedev, D., Oblezin, S.: Baxter operator and Archimedean Hecke algebra. Commun. Math. Phys. **284**(3), 867–896 (2008)
22. Grabsch, A.: Théorie des matrices aléatoires en physique statistique: théorie quantique de la diffusion et systèmes désordonnés. PhD thesis. Université Paris- Saclay (2018)
23. Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd ed Cambridge University Press, Cambridge (2013)
24. Johansson, K.: Shape fluctuations and random matrices. Commun. Math. Phys. **209**(2), 437–476 (2000)
25. Kelly, F.P.: Markovian Functions of a Markov Chain. Sankhyā: Indian J. Stat. Ser. A **44**(3), 372–379 (1982)
26. Kemeny, J.G., Snell, J.L.: Finite Markov Chains. Undergraduate Texts in Mathematics, Springer-Verlag, New York (1976)
27. Kharchev, S., Lebedev, D.: Integral representations for the eigenfunctions of quantum open and periodic Toda chains from the QISM formalism. J. Phys. A **34**(11), 2247–2258 (2001)
28. Kurtz, T.: Martingale problems for conditional distributions of Markov processes. Electron. J. Probab. **3**, 1–29 (1998)
29. Macdonald, I.G.: Symmetric Functions and Hall Polynomials. Oxford Univ. Press, London (1979)
30. Marshall, A.W., Olkin, I., Arnold, B.C.: Inequalities: Theory of Majorization and its Applications. Springer Series in Statistics, 2nd ed. Springer, New York (2011)
31. Nordenstam, E.: On the shuffling algorithm for domino tilings. Electron. J. Probab. **15**, 75–95 (2010)

32. O'Connell, N.: Directed polymers and the quantum Toda lattice. Ann. Probab. **40**(2), 437–458 (2012)
33. O'Connell, N.: Interacting diffusions on positive definite matrices. Probab. Theory Relat. Fields **180**, 679–726 (2021)
34. O'Connell, N., Ortmann, J.: Tracy-Widom asymptotics for a random polymer model with gamma-distributed weights. Electron. J. Probab. **20**(25), 1–18 (2015)
35. O'Connell, N., Seppäläinen, T., Zygouras, N.: Geometric RSK correspondence, Whittaker functions and symmetrized random polymers. Invent. Math. **197**(2), 361–416 (2014)
36. Petersen, K.B., Pedersen, M.S.: The Matrix Cookbook (2012). http://www2.imm.dtu.dk/pubdb/pubs/3274-full.html
37. Rogers, L.C.G., Pitman, J.W.: Markov functions. Ann. Probab. **9**(4), 573–582 (1981)
38. Rudin, W.: Principles of Mathematical Analysis, 3rd edn. McGraw-Hill, New York (1976)
39. Seppäläinen, T.: Scaling for a one-dimensional directed polymer with boundary conditions. Ann. Probab. **40**(1), 19–73 (2012)
40. Stade, E.: Archimedean L-factors on $GL(n) \times GL(n)$ and generalized Barnes integrals. Israel J. Math. **127**(1), 201–219 (2002)
41. Stanley, R. P.: Enumerative Combinatorics: volume 2. Vol. 62. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1999)
42. Terras, A.: Harmonic Analysis on Symmetric Spaces-Higher Rank Spaces, Positive Definite Matrix Space and Generalizations, 2nd edn. Springer, New York (2016)
43. Warren, J.: Dyson's Brownian motions, intertwining and interlacing. Electron. J. Probab. **12**, 573–590 (2007)
44. Warren, J., Windridge, P.: Some examples of dynamics for Gelfand-Tsetlin patterns. Electron. J. Probab. **14**, 1745–1769 (2009)
45. Zygouras, N.: Some algebraic structures in the KPZ universality. Probab. Surv. **19**, 590–700 (2022)