Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2019, Vol. 55, No. 3, 1402–1438
https://doi.org/10.1214/18-AIHP923

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Edge of spiked beta ensembles, stochastic Airy semigroups and reflected Brownian motions

Pierre Yves Gaudreau Lamarre[1] and Mykhaylo Shkolnikov[2]

ORFE Department, Princeton University, Princeton, NJ 08544, USA. E-mail: plamarre@princeton.edu; mshkolni@gmail.com

Received 2 October 2017; revised 29 April 2018; accepted 27 July 2018

Abstract. We access the edge of Gaussian beta ensembles with one spike by analyzing high powers of the associated tridiagonal matrix models. In the classical cases $\beta = 1, 2, 4$, this corresponds to studying the fluctuations of the largest eigenvalues of additive rank one perturbations of the GOE/GUE/GSE random matrices. In the infinite-dimensional limit, we arrive at a one-parameter family of random Feynman–Kac type semigroups, which features the stochastic Airy semigroup of Gorin and Shkolnikov (*Ann. Probab.* **46** (2018) 2287–2344) as an extreme case. Our analysis also provides Feynman–Kac formulas for the spiked stochastic Airy operators, introduced by Bloemendal and Virág (*Probab. Theory Related Fields* **156** (2013) 795–825). The Feynman–Kac formulas involve functionals of a reflected Brownian motion and its local times, thus, allowing to study the limiting operators by tools of stochastic analysis. We derive a first result in this direction by obtaining a new distributional identity for a reflected Brownian bridge conditioned on its local time at zero. A key feature of our proof consists of a novel strong invariance result for certain non-negative random walks and their occupation times that is based on the Skorokhod reflection map.

Résumé. Nous accédons à l'extrémité du spectre des ensembles bêta gaussiens avec perturbation de rang un par l'entremise de grandes puissances des matrices tridiagonales qui y sont associées. Pour les valeurs traditionnelles $\beta = 1, 2, 4$, ceci correspond à l'étude des fluctuations des valeurs propres maximales des matrices aléatoires GOE/GUE/GSE assujetties à une perturbation additive de rang un. En dimensions infinies, nos résultats nous mènent vers une famille de semi-groupes de type Feynman–Kac qui, dans un cas extrême, correspond au *stochastic Airy semigroup* introduit par Gorin et Shkolnikov (*Ann. Probab.* **46** (2018) 2287–2344). De plus, nos résultats ont pour corollaire des formules de Feynman–Kac pour les *spiked stochastic Airy operators* introduits par Bloemendal et Virág (*Probab. Theory Related Fields* **156** (2013) 795–825). Ces formules sont exprimées à l'aide de certaines fonctionnelles du mouvement brownien réfléchi et de ses temps locaux. Ce faisant, les opérateurs en question peuvent être étudiés à l'aide du calcul stochastique. Nous obtenons un premier résultat dans cette lignée en démontrant une nouvelle identité décrivant la distribution du mouvement brownien réfléchi ayant été conditionné sur son temps local à zéro. La principale innovation de notre démonstration consiste en la preuve d'un nouveau résultat sur l'approximation forte du mouvement brownien réfléchi et de son temps local par une marche aléatoire non négative en utilisant la méthode de réflexion de Skorokhod.

MSC: 60B20; 60H25; 47D08; 60J55

Keywords: Beta ensembles; Feynman–Kac formulas; Local times; Low rank perturbations; Moments method; Operator limits; Path transformations; Random tridiagonal matrices; Reflected Brownian motions; Skorokhod map; Stochastic Airy semigroups; Strong invariance principles

[1]P. Y. Gaudreau Lamarre is partially supported by an NSERC Doctoral Fellowship and a Gordon Y. S. Wu Fellowship.
[2]M. Shkolnikov is partially supported by the NSF Grant DMS-1506290.

1. Introduction

A remarkable advance in the study of random matrices and related point processes has been the development of a theory of operator limits for such objects in [8,9,14,16,18,19,23,31,32,34,36–38]. This line of research originates from the publications [13,14] by Edelman and Sutton, who have realized that the random tridiagonal matrices of Dumitriu and Edelman [12] (see (1.2) below for a definition) can be viewed as finite-dimensional approximations of suitable random Schrödinger operators. Since the joint eigenvalue distributions in the Dumitriu–Edelman models for the parameter values $\beta = 1, 2, 4$ are given by the eigenvalue point processes of the Gaussian orthogonal/unitary/symplectic ensembles (GOE/GUE/GSE), respectively, the insights of [13,14] suggest that the limiting fluctuations of the largest eigenvalues of the latter can be read off from the random Schrödinger operators associated with the former. This approach has been carried out rigorously in the seminal paper [32] by Ramírez, Rider and Virág. We also refer to [23] for a corresponding universality result, to [8,9] for extensions to spiked random matrix ensembles, to [18,31,34] for operator limits describing the fluctuations of the smallest eigenvalues of large positive definite random matrices, and to [36,38] for operators arising in the study of the bulk eigenvalues of random matrices.

More recently, Gorin and Shkolnikov [16] have proposed a different operator limit approach to the study of the largest eigenvalues in the Gaussian beta ensembles. The latter are point processes on the real line, in which the joint density of the points $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$ is proportional to

$$\prod_{1\le q_1<q_2\le N} (x_{q_1} - x_{q_2})^\beta \prod_{q=1}^{N} e^{-\beta x_q^2/4}. \tag{1.1}$$

For $\beta = 1, 2, 4$, the Gaussian beta ensemble describes the eigenvalue process of a random matrix from the GOE/GUE/GSE, respectively (see e.g. [1, Section 2.5]). Gaussian beta ensembles with general values of $\beta > 0$ appear frequently in the statistical physics literature and are commonly known therein as "log-gases", see e.g. [15, Section 4.1].

The starting point of [16] is the celebrated result of Dumitriu and Edelman [12] establishing (1.1) as the joint eigenvalue distribution, for *all* values of $\beta > 0$, of the random matrix

$$H_N^\beta := \frac{1}{\sqrt{\beta}} \begin{bmatrix} \sqrt{2}G_1 & \chi_{(N-1)\beta} & & & \\ \chi_{(N-1)\beta} & \sqrt{2}G_2 & \chi_{(N-2)\beta} & & \\ & \chi_{(N-2)\beta} & \sqrt{2}G_3 & \ddots & \\ & & \ddots & \ddots & \chi_\beta \\ & & & \chi_\beta & \sqrt{2}G_N \end{bmatrix}, \tag{1.2}$$

where $G_1, G_2, \ldots, G_N$ are independent standard Gaussian random variables, $\chi_\beta, \chi_{2\beta}, \ldots, \chi_{(N-1)\beta}$ are independent chi random variables indexed by their parameters, and the chi random variables are independent of the Gaussian random variables. According to [32, Theorem 1.1], the rescaled eigenvalues of H_N^β,

$$\Lambda_{q,N} := N^{1/6}(2\sqrt{N} - \lambda_q), \quad q = 1, 2, \ldots, \tag{1.3}$$

converge in finite-dimensional distributions as $N \to \infty$ to the eigenvalues $\Lambda_1 \le \Lambda_2 \le \cdots$ of the stochastic Airy operator

$$\mathcal{H}^\beta f := \left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x + \frac{2}{\sqrt{\beta}} W_x'\right) f, \quad f \in L^2\big([0,\infty)\big), f(0) = 0, \tag{1.4}$$

where W' is the white noise on $[0, \infty)$. Thus, the simple computation

$$\left(\frac{\lambda_q}{2\sqrt{N}}\right)^{\lfloor TN^{2/3}\rfloor} = \left(1 - \frac{\Lambda_{q,N}}{2N^{2/3}}\right)^{\lfloor TN^{2/3}\rfloor} \to e^{-T\Lambda_q/2}, \quad T \ge 0 \tag{1.5}$$

suggests the convergence of the high powers

$$\left(\frac{H_N^\beta}{2\sqrt{N}}\right)^{\lfloor TN^{2/3}\rfloor}, \quad N \in \mathbb{N} \tag{1.6}$$

to $e^{-T\mathcal{H}^\beta/2}$ in a suitable operator topology. The main result of [16] (see [16, Theorem 2.8]) establishes a more general version of such an operator convergence directly, without relying on the findings of [32]. The proof is achieved by expressing the entries of (1.6) in terms of expectations of certain functionals of simple symmetric random walk bridges and their occupation times, and the asymptotics of such expressions are obtained by making use of various strong invariance principles for the considered discrete processes. In contrast to $\mathcal{H}^\beta$, the operator $e^{-T\mathcal{H}^\beta/2}$ is an integral operator, and the corresponding integral kernel can be written in terms of the Brownian motion W and an independent Brownian bridge (see [16, equation (2.4)]). This allows to study the properties of $e^{-T\mathcal{H}^\beta/2}$ by tools of stochastic analysis (see [16, Proposition 2.14], [17, Theorem 1.1] for an example).

In this paper, we continue the program initiated in [16] and consider the case of Gaussian beta ensembles with one spike. To this end, it is useful to recall that the stochastic Airy operator $\mathcal{H}^\beta$ describes the limiting behavior of the largest eigenvalues in the Laguerre beta ensemble (see [12,32]). The latter interpolates between the eigenvalue processes of sample covariance matrices $\boldsymbol{X}\boldsymbol{X}^*$, where the entries of $\boldsymbol{X}$ are independent standard Gaussian random variables. From the point of view of statistical applications, the Laguerre beta ensemble is arguably not the most interesting model, since the entries in the columns of $\boldsymbol{X}$ are uncorrelated. Instead, one often considers multiplicative perturbations of $\boldsymbol{X}\boldsymbol{X}^*$ of the form $\boldsymbol{X}\Sigma\boldsymbol{X}^*$, where $\Sigma = \widetilde{\Sigma}_r \oplus I_{N-r}$ is the direct sum of a deterministic full rank $r \times r$ matrix $\widetilde{\Sigma}_r$ and the $(N-r)\times(N-r)$ identity matrix I_{N-r}. Such models are known in the literature as the spiked covariance models, and we refer to the introduction in [8] for an excellent summary of their practical applications in statistics.

As first discovered by Baik, Ben Arous and Péché [2] in the case of complex covariance matrices, the fluctuations of the largest eigenvalues exhibit a phase transition (known as the BBP phase transition) depending on the size of the perturbation. In the subcritical regime, the perturbation Σ is so insignificant that the limiting behaviour is the same as in the unperturbed case; in the critical regime, the fluctuation exponents are the same as in the unperturbed case, but the limiting distributions are different; and in the supercritical regime, the size of the perturbation is so large that the largest eigenvalues of $\boldsymbol{X}\Sigma\boldsymbol{X}^*$ separate from the bulk of the spectrum.

The BBP phase transition was later extended to finite rank additive perturbations of the form $\boldsymbol{X}+\Sigma$, where $\boldsymbol{X}$ is a Wigner random matrix (i.e., Hermitian with independent entries) [27]. Similarly to covariance matrices, the asymptotic edge fluctuations of $\boldsymbol{X}+\Sigma$ can be characterized in terms of the size of Σ, and the same trichotomy (subcritical, critical, and supercritical phases) described above occurs. Such additive perturbation models and their generalizations find applications in physics [5,22] and signal processing problems [20].

For rank one perturbations, the critical regime of the BBP phase transition has been analyzed in detail by Bloemendal and Virág [8], and we describe their main result in the case of an additive perturbation. The corresponding tridiagonal model

$$H_N^{\beta;w} := \frac{1}{\sqrt{\beta}}\begin{bmatrix} \sqrt{2}G_1+\sqrt{\beta N}\ell_N & \chi_{(N-1)\beta} & & & \\ \chi_{(N-1)\beta} & \sqrt{2}G_2 & \chi_{(N-2)\beta} & & \\ & \chi_{(N-2)\beta} & \sqrt{2}G_3 & \ddots & \\ & & \ddots & \ddots & \chi_\beta \\ & & & \chi_\beta & \sqrt{2}G_N \end{bmatrix} \tag{1.7}$$

can be obtained for $\beta = 1,2,4$ by applying the Dumitriu–Edelman tridiagonalization procedure to the sum of a GOE/GUE/GSE matrix and a rank one matrix with non-zero eigenvalue $\sqrt{N}\ell_N$, where

$$\lim_{N\to\infty} N^{1/3}(1-\ell_N) = w \in \mathbb{R}. \tag{1.8}$$

Then, for all $\beta > 0$ and under the scaling of (1.3), the ordered eigenvalues of $H_N^{\beta;w}$ converge in finite-dimensional distribution sense, as $N \to \infty$, to the ordered eigenvalues of the spiked stochastic Airy operator

$$\mathcal{H}^{\beta;w} f := \left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x + \frac{2}{\sqrt{\beta}} W_x'\right) f, \quad f \in L^2\big([0,\infty)\big), f'(0) = w f(0). \tag{1.9}$$

The case $w = \infty$ formally corresponds to the Dirichlet boundary condition $f(0) = 0$, motivating the convention $\mathcal{H}^{\beta;\infty} := \mathcal{H}^{\beta}$.

Remark 1.1. The limit in (1.8) determines the regime in the BBP phase transition: a limit of ∞ corresponds to the subcritical regime, a finite limit to the critical regime, and a limit of $-\infty$ to the supercritical regime.

We turn to our main results. For the sake of convenience, we work with a modification of the tridiagonal model (1.7):

$$M_N^{\beta;w} := \begin{bmatrix} \sqrt{N}\ell_N & \sqrt{N}+\xi_0 & & & \\ \sqrt{N}+\xi_0 & \mathbf{a}_1 & \sqrt{N-1}+\xi_1 & & \\ & \sqrt{N-1}+\xi_1 & \mathbf{a}_2 & \ddots & \\ & & \ddots & \ddots & 1+\xi_{N-1} \\ & & & 1+\xi_{N-1} & \mathbf{a}_N \end{bmatrix}. \tag{1.10}$$

The following assumption summarizes the conditions we impose on the matrix entries throughout the paper. We emphasize that we allow the random variables $\mathbf{a}_1, \mathbf{a}_2, \dots$ and $\xi_0, \xi_1, \dots$ to vary with N, even though the dependence on N is suppressed to simplify the notation.

Assumption 1.2. The random variables $\mathbf{a}_1, \mathbf{a}_2, \dots$ and $\xi_0, \xi_1, \dots$ are mutually independent and such that:

(a) $|\mathbf{E}[\mathbf{a}_m]| = o((N-m)^{-1/3})$ and $|\mathbf{E}[\xi_m]| = o((N-m)^{-1/3})$ as $(N-m) \to \infty$,
(b) $\mathbf{E}[\mathbf{a}_m^2] = s_{\mathbf{a}}^2 + o(1)$ and $\mathbf{E}[\xi_m^2] = s_\xi^2 + o(1)$ as $(N-m) \to \infty$, where $s_{\mathbf{a}}$, s_ξ are non-negative constants satisfying $\frac{s_{\mathbf{a}}^2}{4} + s_\xi^2 = \frac{1}{\beta}$ for some $\beta > 0$,
(c) $\mathbf{E}[|\mathbf{a}_m|^p] \le C^p p^{\gamma p}$ and $\mathbf{E}[|\xi_m|^p] \le C^p p^{\gamma p}$ for all N, m and p, with some constants $C < \infty$ and $0 < \gamma < 2/3$.

Moreover, we assume that the non-random sequence ℓ_N, $N \in \mathbb{N}$ satisfies (1.8) and use the convention $\mathbf{a}_0 := 0$.

Remark 1.3. Assumption 1.2 holds, in particular, when the $\mathbf{a}_m$'s are chosen to be i.i.d. Gaussian with mean 0 and variance $2/\beta$, whereas the ξ_m's are drawn independently such that each $\sqrt{\beta}(\sqrt{N-m}+\xi_m)$ is a chi random variable with parameter $\beta(N-m)$ (see [16, Lemma 2.2]). However, we note that this is slightly different from the spiked Gaussian β-ensemble (1.7), since there is no Gaussian random variable with $\sqrt{N}\ell_N$. We show in Remark 1.8 that the results of this paper also apply to (1.7).

Motivated by the computation in (1.5), we consider the powers

$$\mathcal{M}_{T;N}^{\beta;w} := \left(\frac{M_N^{\beta;w}}{2\sqrt{N}}\right)^{\lfloor TN^{2/3}\rfloor}, \quad T \ge 0. \tag{1.11}$$

The operator limits of the latter turn out to be given by the following definition.

Definition 1.4. For every $\beta, T > 0$ consider the operator

$$\big(\mathcal{U}_T^{\beta;w} f\big)(x) := \mathbf{E}_{R^x}\left[\exp\left(-\int_0^T \frac{R_t^x}{2}\,\mathrm{d}t + \int_0^\infty \frac{L_T^a(R^x)}{\sqrt{\beta}}\,\mathrm{d}W_a - w\frac{L_T^0(R^x)}{2}\right) f\big(R_T^x\big)\right] \tag{1.12}$$

acting on the space

$$\mathcal{D} := \big\{ f \in L^1_{\mathrm{loc}}\big([0,\infty)\big) : \big|f(x)\big| \le C_1 e^{C_2 x^{1-\delta}} \text{ for some } C_1, C_2 < \infty, \delta \in (0,1)\big\}, \tag{1.13}$$

where

(a) R^x is a reflected Brownian motion started at $x \ge 0$,
(b) $\mathbf{E}_{R^x}[\cdot]$ is the expectation with respect to R^x,
(c) the local time of R^x is defined as the continuous version of

$$L^a_T\big(R^x\big) := \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^T \mathbf{1}_{[a,a+\varepsilon)}\big(R^x_t\big)\,\mathrm{d}t, \quad a \ge 0, \tag{1.14}$$

(d) W is a standard Brownian motion independent of R^x,
(e) the Itô integral with respect to W is defined pathwise, as per [21].

Remark 1.5. A trivial restatement of Definition 1.4 is that $\mathcal{U}^{\beta;w}_T$ is a random integral operator with the kernel

$$\begin{aligned} K^{\beta;w}_T(x,y) := {} & \frac{\exp(-\frac{(x-y)^2}{2T}) + \exp(-\frac{(x+y)^2}{2T})}{\sqrt{2\pi T}} \\ & \cdot \mathbf{E}_{R^x}\bigg[\exp\bigg(-\int_0^T \frac{R^x_t}{2}\,\mathrm{d}t + \int_0^\infty \frac{L^a_T(R^x)}{\sqrt{\beta}}\,\mathrm{d}W_a - w\frac{L^0_T(R^x)}{2}\bigg) \,\bigg|\, R^x_T = y\bigg]. \end{aligned} \tag{1.15}$$

Remark 1.6. For each N, the matrix $\mathcal{M}^{\beta;w}_{T;N}$ can be regarded as an integral operator acting on $L^1_{\mathrm{loc}}([0,\infty))$ by associating $\mathbb{R}^{N+1}$ with the subspace of step functions

$$L^1_N\big([0,\infty)\big) := \Bigg\{ \sum_{l=0}^N v_l \mathbf{1}_{[N^{-1/3}l, N^{-1/3}(l+1))} : v_0, v_1, \dots, v_N \in \mathbb{R} \Bigg\}, \tag{1.16}$$

and then mapping functions $f \in L^1_{\mathrm{loc}}([0,\infty))$ into $L^1_N([0,\infty))$ via

$$(\pi_N f)(x) := \sum_{l=0}^N N^{1/6} \int_{N^{-1/3}l}^{N^{-1/3}(l+1)} f(y)\,\mathrm{d}y \cdot \mathbf{1}_{[N^{-1/3}l, N^{-1/3}(l+1))}(x) \tag{1.17}$$

before acting with $\mathcal{M}^{\beta;w}_{T;N}$ on them.

Our main convergence result reads as follows.

Theorem 1.7. *For every $\beta > 0$, $w \in \mathbb{R}$, and with $\mathcal{D}$ defined in (1.13), one has*

$$\forall f, g \in \mathcal{D}, T \ge 0: \quad \lim_{N\to\infty} (\pi_N f)^\top \mathcal{M}^{\beta;w}_{T;N}(\pi_N g) = \int_0^\infty f(x)\big(\mathcal{U}^{\beta;w}_T g\big)(x)\,\mathrm{d}x, \tag{1.18}$$

where the convergence is in distribution and in the sense of moments. Moreover, these convergences hold jointly for any finite collection of T's, f's and g's, and in the case of the convergence in distribution also jointly with the convergence in distribution

$$\sqrt{\beta} \lim_{N\to\infty} N^{-1/6} \sum_{m=0}^{\lfloor N^{1/3}x \rfloor} \bigg(\frac{\mathbf{a}_m}{2} + \xi_m\bigg) = W_x, \quad x \ge 0 \tag{1.19}$$

with respect to the Skorokhod topology. Here W is the Brownian motion from (1.12).

Remark 1.8. We note that Theorem 1.7 also holds for the matrix model (1.7). Indeed, instead of ℓ_N, we may consider $\widetilde{\ell}_N := \ell_N + \mathbf{g}/\sqrt{N}$ where $\mathbf{g}$ is a fixed Gaussian random variable independent of all other entries in $M_N^{\beta;w}$. If we condition on the event $\{\mathbf{g} = y\}$ for some $y \in \mathbb{R}$, then $\widetilde{\ell}_N$ clearly satisfies (1.8). Consequently, Theorem 1.7 holds under this conditioning. Since the limit in distribution and moments thus obtained does not depend on the value of y, it immediately follows from disintegration that the convergence in distribution holds without the conditioning as well. As for the convergence of moments without the conditioning, we need only ensure that the bound (3.48) holds for $\widetilde{\ell}_N$. To this end, note that the right-hand side of (3.55) becomes $e^{-(w+o(1)-N^{-1/6}\mathbf{g})\cdot(\theta N^{-1/3}\rho_k)}$ in the present context, and that $\mathbf{E}[e^{\mathbf{g}\theta\rho_k N^{-1/2}}]$ can be controlled by using the Gaussian moment generating function and (3.51).

Remark 1.9. As in [16], the central idea in the proof of Theorem 1.7 consists of expressing the high powers (1.11) in terms of functionals of random walks and their occupation times. However, in comparison to (1.6), the presence of the *spiked* term $\sqrt{N}\ell_N$ in (1.10) induces complications in the asymptotic analysis of the random walks associated to (1.11). As a consequence, the classical strong invariance results used in the asymptotic analysis carried out in [16] cannot be applied in the present case. In this context, one of the main technical innovations of this paper is to provide a new strong invariance principle for certain non-negative random walks and their occupation times, which is based on the observation that the non-negative random walks in consideration can be viewed as images of simple random walks under the Skorokhod reflection map. We expect the same idea to apply for a wide variety of constrained discrete processes.

Next, we present some natural properties of the operators $\mathcal{U}_T^{\beta;w}$, $T \geq 0$, viewed as operators on $L^2([0,\infty))$, and, in particular, connect them to the spiked stochastic Airy operator $\mathcal{H}^{\beta;w}$ in (1.9).

Proposition 1.10. *For every $\beta > 0$ and $w \in \mathbb{R}$, the following statements hold.*

(a) *If the same Brownian motion W is used in the definitions of $\mathcal{H}^{\beta;w}$ and $\mathcal{U}_T^{\beta;w}$, then for every $T \geq 0$,*

$$\mathcal{U}_T^{\beta;w} = \exp\left(-\frac{T}{2}\mathcal{H}^{\beta;w}\right) \quad \textit{almost surely}, \tag{1.20}$$

in the sense that if $(f_q, \Lambda_q : q \in \mathbb{N})$ are the eigenfunction-eigenvalue pairs of $\mathcal{H}^{\beta;w}$, then $\mathcal{U}_T^{\beta;w}$ is the unique operator on $L^2([0,\infty))$ with eigenfunction-eigenvalue pairs $(f_q, e^{-T\Lambda_q/2} : q \in \mathbb{N})$ almost surely.

(b) *The family $(\mathcal{U}_T^{\beta;w} : T \geq 0)$ has the almost sure semigroup property in the sense that for all $T_1, T_2 \geq 0$, one has $\mathcal{U}_{T_1}^{\beta;w}\mathcal{U}_{T_2}^{\beta;w} = \mathcal{U}_{T_1+T_2}^{\beta;w}$ almost surely.*

(c) *For every $T > 0$, the operator $\mathcal{U}_T^{\beta;w}$ is symmetric, non-negative and belongs to the Hilbert–Schmidt class almost surely.*

(d) *For every $T > 0$, the operator $\mathcal{U}_T^{\beta;w}$ is almost surely trace class and obeys the trace formula*

$$\operatorname{Tr}\big(\mathcal{U}_T^{\beta;w}\big) = \int_0^\infty K_T^{\beta;w}(x,x)\,\mathrm{d}x. \tag{1.21}$$

(e) *The family $(\mathcal{U}_T^{\beta;w} : T \geq 0)$ is L^2-strongly continuous in expectation, that is, for all $p > 0$, $T \geq 0$ and $f \in L^2([0,\infty))$, one has*

$$\lim_{t \to T} \mathbf{E}\big[\big\|\mathcal{U}_T^{\beta;w} f - \mathcal{U}_t^{\beta;w} f\big\|_{L^2([0,\infty))}^p\big] = 0. \tag{1.22}$$

Remark 1.11. Proposition 1.10(a) should be viewed as a Feynman–Kac formula for the spiked stochastic Airy operator $\mathcal{H}^{\beta;w}$.

Remark 1.5 shows that one might be able to understand observables of the limiting operators $\mathcal{U}_T^{\beta;w}$, $T \geq 0$, such as moments of certain linear statistics of their spectra, by investigating the corresponding functionals of reflected

Brownian motions conditioned on their endpoints. As a first step in this direction, we consider $\mathbf{E}[K_T^{\beta;w}(0,0)]$, which, in view of the next proposition, seems to be the simplest object to study.

Proposition 1.12. *For every $\beta, T > 0$ and $w \in \mathbb{R}$,*

$$\begin{aligned}&\mathbf{E}\big[K_T^{\beta;w}(0,0)\big]\\&\quad=\sqrt{\frac{2}{\pi T}}\mathbf{E}\bigg[\exp\bigg(-\frac{T^{3/2}}{2}\bigg(\int_0^1 r_t\,\mathrm{d}t-\int_0^\infty\frac{L_1^a(r)^2}{\beta}\,\mathrm{d}a\bigg)-T^{1/2}w\frac{L_1^0(r)}{2}\bigg)\bigg],\end{aligned}\tag{1.23}$$

where r_t, $t \in [0,1]$ is a reflected Brownian bridge.

Since the density of $L_1^0(r)$ is known (see e.g. [28, equation (3)]), it suffices to find the conditional distribution of the functional

$$\int_0^1 r_t\,\mathrm{d}t-\int_0^\infty\frac{L_1^a(r)^2}{\beta}\,\mathrm{d}a\tag{1.24}$$

given $L_1^0(r)$ to compute the right-hand side in (1.12). For $\beta = 2$, this leads to the following theorem of independent interest.

Theorem 1.13. *Let r_t, $t \in [0,1]$ be a reflected Brownian bridge. Then, for every $\alpha \geq 0$, the conditional distribution of the functional*

$$\int_0^1 r_t\,\mathrm{d}t-\int_0^\infty\frac{L_1^a(r)^2}{2}\,\mathrm{d}a\tag{1.25}$$

given $L_1^0(r) = \alpha$ is Gaussian with mean $-\alpha/4$ and variance $1/12$.

Remark 1.14. Conditional on $L_1^0(r) = 0$, the process r_t, $t \in [0,1]$ is a standard Brownian excursion, so that Theorem 1.13 is a generalization of the distributional identity for the latter found in [16, Corollary 2.15] (see also [17, Theorem 1.1]).

As a consequence of Theorem 1.13, we obtain an explicit formula for $\mathbf{E}[K_T^{2;w}(0,0)]$.

Corollary 1.15. *For any $w \in \mathbb{R}$ and $T > 0$, and with*

$$C_{w;T} := \frac{\sqrt{T}(T-4w)}{4\sqrt{2}},\tag{1.26}$$

it holds

$$\begin{aligned}&\mathbf{E}\big[K_T^{2;w}(0,0)\big]\\&\quad=\sqrt{\frac{2}{\pi T}}\exp\bigg(\frac{T^3}{96}\bigg)\big(1+\sqrt{\pi}C_{w;T}\exp\big(C_{w;T}^2\big)\big(\mathrm{erf}(C_{w;T})+1\big)\big),\end{aligned}\tag{1.27}$$

where $\mathrm{erf}(z) := \frac{2}{\sqrt{\pi}}\int_0^z e^{-a^2}\,\mathrm{d}a$ denotes the error function.

For $\beta \neq 2$, we were not able to obtain an analogue of Theorem 1.13.

Open Problem 1.16. Find the conditional distribution of the functional in (1.24) given $L_1^0(r)$ for all $\beta > 0$.

Remark 1.17. A simple computation based on Theorem 1.13 shows that the *unconditional* distribution of the random variable

$$A := \sqrt{12}\left(\int_0^1 r_t \,\mathrm{d}t - \int_0^\infty \frac{L_1^a(r)^2}{2}\,\mathrm{d}a\right) \tag{1.28}$$

has a moment-generating function given by

$$\mathbf{E}\big[\exp(\kappa A)\big] = \exp\left(\frac{\kappa^2}{2}\right) - \sqrt{\frac{3}{2}}\kappa \exp\big(2\kappa^2\big)\,\mathrm{erfc}\left(\sqrt{\frac{3}{2}}\kappa\right), \tag{1.29}$$

where $\mathrm{erfc}(z) := \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-a^2}\,\mathrm{d}a$ denotes the complementary error function. Therefore, it seems natural to view the density of A as a sum of the standard Gaussian density and a function that integrates to 0, which yields a corresponding decomposition of the moments of A (see Table 1 for the first few moments). In particular, Table 1 suggests the following formula for the odd moments of A:

$$\mathbf{E}\big[A^{2n-1}\big] = -\frac{2^n(2n-1)!}{4(n-1)!}\sqrt{6\pi}, \quad n = 1, 2, \dots, \tag{1.30}$$

leading us to believe that A admits an interesting combinatorial interpretation.

The remainder of the paper is structured as follows. In Section 2, we prove the strong invariance principle for non-negative random walks that was alluded to in Remark 1.9. Section 3 is devoted to the proof of Theorem 1.7. The proof has the following general structure: first, we write the entries of the matrix $\mathcal{M}_{T;N}^{\beta;w}$ as expectations of suitable functionals of the non-negative random walks of Section 2 and the entries of the matrix $M_N^{\beta;w}$; then, we derive the limiting behavior of suitably truncated versions of such expectations using the strong invariance principle of Section 2; finally, we remove the truncation by obtaining appropriate uniform moment estimates on the functionals involved. In Section 4, we show the properties of the limiting operators $\mathcal{U}_T^{\beta;w}$, $T \geq 0$ listed in Proposition 1.10. Lastly, in Section 5 we establish Theorem 1.13, as well as Proposition 1.12 and Corollary 1.15. The proof of Theorem 1.13 combines the ideas of [17] with the analogue of Jeulin's theorem for a reflected Brownian bridge conditioned on its local time at zero from [29, Corollary 16(iii)].

Table 1
The first few moments of A

$\mathbf{E}[A]$	$-\sqrt{6\pi}/2$
$\mathbf{E}[A^2]$	$1+6$
$\mathbf{E}[A^3]$	$-6\sqrt{6\pi}$
$\mathbf{E}[A^4]$	$3+108$
$\mathbf{E}[A^5]$	$-120\sqrt{6\pi}$
$\mathbf{E}[A^6]$	$15+2646$
$\mathbf{E}[A^7]$	$-3360\sqrt{6\pi}$
$\mathbf{E}[A^8]$	$105+85{,}032$
$\mathbf{E}[A^9]$	$-120{,}960\sqrt{6\pi}$
$\mathbf{E}[A^{10}]$	$945+3404{,}430$
$\mathbf{E}[A^{11}]$	$-5322{,}240\sqrt{6\pi}$
$\mathbf{E}[A^{12}]$	$10{,}395+163{,}446{,}660$
$\mathbf{E}[A^{13}]$	$-276{,}756{,}480\sqrt{6\pi}$
$\mathbf{E}[A^{14}]$	$135{,}135+9153{,}449{,}550$

2. A strong invariance principle

This section focuses on the strong invariance principle for certain non-negative random walks and their families of occupation times, which is at the heart of the proof of Theorem 1.7. For starters, we let $Y = (Y_0, Y_1, \ldots)$ be a random walk on the non-negative integers with transition probabilities

$$\begin{aligned} &\mathbf{P}[Y_{n+1} = z+1 \mid Y_n = z] = \mathbf{P}[Y_{n+1} = z-1 \mid Y_n = z] = \frac{1}{2}, \quad z = 1, 2, \ldots, \\ &\mathbf{P}[Y_{n+1} = 1 \mid Y_n = 0] = \mathbf{P}[Y_{n+1} = 0 \mid Y_n = 0] = \frac{1}{2}. \end{aligned} \tag{2.1}$$

In other words, when Y is away from 0, it behaves like a simple symmetric random walk (SSRW), and when Y is at 0, it stays at 0 or moves to 1 with equal probability.

Given $T > 0$ and $N \in \mathbb{N}$, we let $k = k(T, N) := \lfloor T N^{2/3} \rfloor$ and $T_k := k N^{-2/3}$. Moreover, for each $x \geq 0$, we define a process $X_t^{k;x}$, $t \in [0, T_k]$ satisfying

$$\big(X_0^{k;x}, X_{N^{-2/3}}^{k;x}, \ldots, X_{T_k}^{k;x}\big) \stackrel{\mathrm{d}}{=} \big(Y_0, Y_1, \ldots, Y_k \mid Y_0 = \lfloor x N^{1/3} \rfloor\big) \tag{2.2}$$

and interpolating linearly between these time points (see Figure 1 for an illustration). We also introduce the normalized occupation times of $X^{k;x}$ for positive levels:

$$L^a\big(X^{k;x}\big) := N^{-1/3} \big|\big\{t \in [0, T_k] : X^{k;x} = a N^{1/3}\big\}\big|, \quad a > 0 \tag{2.3}$$

and use the convention

$$L^0\big(X^{k;x}\big) := \lim_{a \downarrow 0} L^a\big(X^{k;x}\big). \tag{2.4}$$

Remark 2.1. Note that the normalized occupation times as defined above have the property that for any measurable function $\varphi : \mathbb{R} \to \mathbb{R}$ that vanishes on $(-\infty, 0]$, one has

$$\int_0^{T_k} \varphi\big(X_t^{k;x}\big)\, \mathrm{d}t = \int_{\mathbb{R}} \varphi(a) L^a\big(X^{k;x}\big)\, \mathrm{d}a. \tag{2.5}$$

Finally, we let

$$H\big(X^{k;x}\big) := \big|\big\{t \in [0, T_k] \cap N^{-2/3}\mathbb{N} : X_{t-N^{-2/3}}^{k;x} = X_t^{k;x} = 0\big\}\big| \tag{2.6}$$

be the number of the horizontal steps at zero in $(X_0^{k;x}, X_{N^{-2/3}}^{k;x}, \ldots, X_{T_k}^{k;x})$. Our strong invariance principle can now be stated as follows.

Theorem 2.2. *For every $T > 0$ and $x \geq 0$, there exists a coupling of the sequence of processes $X^{k;x}$, $N \in \mathbb{N}$ and a reflected Brownian motion R^x such that*

$$\sup_{t \in [0, T_k]} \big|N^{-1/3} X_t^{k;x} - R_t^x\big| \leq C N^{-1/3} \log N, \quad N \in \mathbb{N}, \tag{2.7}$$

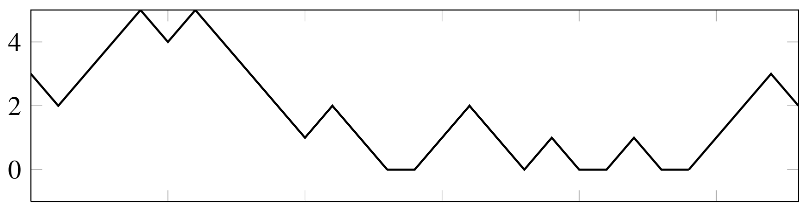

Fig. 1. A sample path of $X^{k;x}$ with $k = 28$ and $\lfloor x N^{1/3} \rfloor = 3$.

$$\sup_{a>0}\left|L^a\left(X^{k;x}\right)-L_T^a\left(R^x\right)\right|\le \mathcal{C}N^{-1/16},\quad N\in\mathbb{N},\tag{2.8}$$

$$\left|N^{-1/3}H\left(X^{k;x}\right)-L_T^0\left(R^x\right)/2\right|\le \mathcal{C}N^{-1/3}\log N,\quad N\in\mathbb{N},\tag{2.9}$$

where $\mathcal{C}$ is a suitable finite random variable.

A direct construction of the coupling in Theorem 2.2 appears to be difficult. Instead, our proof of Theorem 2.2 relies on the Komlós–Major–Tusnády coupling of a SSRW with a standard Brownian motion and an application of the Skorokhod reflection map. We briefly recall the definition and some properties of the latter.

Definition 2.3. Given a $T>0$ and a continuous process Z_t, $t\in[0,T]$, we define the Skorokhod map evaluated at Z as the continuous process

$$\Gamma(Z)_t=Z_t+\sup_{s\in[0,t]}(-Z_s)_+,\quad t\in[0,T],\tag{2.10}$$

where $(\cdot)_+:=\max(0,\cdot)$ denotes the positive part of a real number.

A reflected Brownian motion R^x can be defined by $R^x=|x+\widetilde{W}|$, where $\widetilde{W}$ is a standard Brownian motion. According to Tanaka's formula (see e.g. [33, Chapter VI, Theorem 1.2]) one has

$$R_t^x=x+\int_0^t \operatorname{sgn}(\widetilde{W}_s)\,\mathrm{d}\widetilde{W}_s+L_t^{-x}(\widetilde{W}),\quad t\in[0,T].\tag{2.11}$$

Furthermore, if we let

$$B_t^x:=x+\int_0^t \operatorname{sgn}(\widetilde{W}_s)\,\mathrm{d}\widetilde{W}_s,\quad t\in[0,T],\tag{2.12}$$

which is a Brownian motion started at x, then it follows from a classical result of Skorokhod [33, Chapter VI, Lemma 2.1 and Corollary 2.2] that

$$R_t^x=B_t^x+\sup_{s\in[0,t]}\left(-B_s^x\right)_+=\Gamma\left(B^x\right)_t,\quad t\in[0,T],\tag{2.13}$$

$$\sup_{s\in[0,t]}\left(-B_s^x\right)_+=L_t^{-x}(\widetilde{W})=L_t^0(x+\widetilde{W})=\frac{L_t^0(R^x)}{2},\quad t\in[0,T].\tag{2.14}$$

We give in the next proposition a discrete analogue of these results.

Proposition 2.4. *Let $\widetilde{Y}=(\widetilde{Y}_0,\widetilde{Y}_1,\dots)$ be a SSRW and define the process $\widetilde{X}^{k;x}$, $t\in[0,T_k]$ by using the same procedure as for $X^{k;x}$ (that is, equation (2.2) followed by a linear interpolation), but with with the SSRW $\widetilde{Y}$ instead of the random walk Y. Then, Y and $\widetilde{Y}$ can be coupled in such a way that $X_t^{k;x}=\Gamma(\widetilde{X}^{k;x})_t$, $t\in[0,T_k]$ and $H(X^{k;x})=\sup_{t\in[0,T_k]}(-\widetilde{X}_t^{k;x})_+$.*

Proof. Both $X^{k;x}$ and $\widetilde{X}^{k;x}$ can take a total of 2^k possible sample paths, and the measures that Y and $\widetilde{Y}$ induce on these paths are uniform in both cases. Therefore, we need to show that Γ is a bijection taking paths of $\widetilde{X}^{k;x}$ to paths of $X^{k;x}$ and that $H(\Gamma(\widetilde{X}^{k;x}))=\sup_{t\in[0,T_k]}(-\widetilde{X}_t^{k;x})_+$.

Whenever $\min\widetilde{X}^{k;x}\ge 0$, the Skorokhod map Γ leaves $\widetilde{X}^{k;x}$ unchanged, and trivially $H(\Gamma(\widetilde{X}^{k;x}))=0=\sup_{t\in[0,T_k]}(-\widetilde{X}_t^{k;x})_+$. On the other hand, whenever $\min\widetilde{X}^{k;x}<0$, the application of Γ can be described as follows (see Figure 2):

(a) one determines the first hitting times $\tau_{-1}<\tau_{-2}<\cdots$ of the negative integer levels by $\widetilde{X}^{k;x}$;
(b) one sets $\Gamma(\widetilde{X}^{k;x})_t:=\widetilde{X}_t^{k;x}$ for $t\in[0,\tau_{-1}-N^{-2/3}]$;

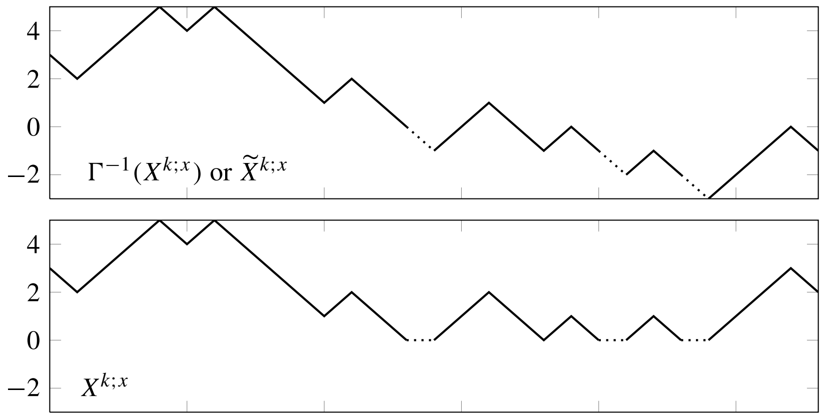

Fig. 2. Illustration of the discrete Skorokhod map.

(c) for $j = 1, 2, \dots, -\min \widetilde{X}^{k;x}$, one lets $\Gamma(\widetilde{X}^{k;x})_t := 0$ for $t \in (\tau_{-j} - N^{-2/3}, \tau_{-j}]$;
(d) for $j = 1, 2, \dots, -\min \widetilde{X}^{k;x} - 1$, one defines $\Gamma(\widetilde{X}^{k;x})_t := \widetilde{X}^{k;x}_t + j$ for $t \in (\tau_{-j}, \tau_{-j-1} - N^{-2/3}]$;
(e) one puts $\Gamma(\widetilde{X}^{k;x})_t := \widetilde{X}^{k;x}_t - \min \widetilde{X}^{k;x}$ for $t \in (\tau_{\min \widetilde{X}^{k;x}}, T_k]$.

It follows immediately from this description that the horizontal steps at zero in $(\Gamma(\widetilde{X}^{k;x})_0, \Gamma(\widetilde{X}^{k;x})_{N^{-2/3}}, \dots, \Gamma(\widetilde{X}^{k;x})_{T_k})$ occur at $\tau_{-1} - N^{-2/3}$, $\tau_{-2} - N^{-2/3}, \dots, \tau_{\min \widetilde{X}^{k;x}} - N^{-2/3}$, so that $H(\Gamma(\widetilde{X}^{k;x})) = -\min \widetilde{X}^{k;x} = \sup_{t \in [0,T_k]}(-\widetilde{X}^{k;x}_t)_+$. Moreover, for every path of $X^{k;x}$, one can uniquely reconstruct the corresponding path $\Gamma^{-1}(X^{k;x}) = \widetilde{X}^{k;x}$ by inferring the sequence $\tau_{-1} < \tau_{-2} < \cdots$ from the horizontal segments in the path of $X^{k;x}$, solving the equations in (b), (d), (e) above, and inserting the remaining $H(X^{k;x})$ downward sloping segments. □

Next, we prepare another coupling needed for the proof of Theorem 2.2.

Lemma 2.5. *For every $T > 0$ and $x \ge 0$, there exists a coupling of the sequence of processes $\widetilde{X}^{k;x}$, $N \in \mathbb{N}$ defined in Proposition 2.4 and a standard Brownian motion B such that*

$$\sup_{t \in [0,T_k]} \left| N^{-1/3} \widetilde{X}^{k;x}_t - (x + B_t) \right| \le \mathcal{C} N^{-1/3} \log N, \quad N \in \mathbb{N}, \tag{2.15}$$

where $\mathcal{C}$ is a suitable finite random variable.

Proof. Consider a probability space which supports a standard Brownian motion B and define the standard Brownian motions

$$B^{(N)}_t := N^{1/3} B_{tN^{-2/3}}, \quad t \ge 0 \tag{2.16}$$

for all $N \in \mathbb{N}$. According to a well-known procedure of Komlós, Major and Tusnády (see e.g. [24, Section 7]), one can construct random walks $(\widetilde{Y}^{(N)}_0, \widetilde{Y}^{(N)}_1, \dots)$, $N \in \mathbb{N}$ as deterministic functions of the Brownian motions $B^{(N)}$, $N \in \mathbb{N}$, respectively, such that for every $\alpha > 0$, there exists a $C < \infty$ so that

$$\mathbf{P}\Big[\max_{0 \le n \le k} \big|\widetilde{Y}^{(N)}_n - B^{(N)}_n\big| \ge C \log N\Big] \le C N^{-\alpha}, \quad N \in \mathbb{N}. \tag{2.17}$$

As a result, we can couple the sequence $\widetilde{X}^{k;x}$, $N \in \mathbb{N}$ with B ensuring

$$\mathbf{P}\Big[\max_{0 \le i \le k} \big|N^{-1/3} X^{k;x}_{N^{-2/3}i} - B_{N^{-2/3}i}\big| \ge C N^{-1/3} \log N\Big] \le C N^{-\alpha}, \quad N \in \mathbb{N}. \tag{2.18}$$

Lastly, we let $\alpha > 1$ and conclude by applying the Borel–Cantelli lemma and the Lévy modulus of continuity theorem. $\square$

Finally, we define the random walk $\overline{Y} = (\overline{Y}_0, \overline{Y}_1, \dots)$ on the integers, to be used in the proof of Theorem 2.2, by

$$\begin{aligned}
&\mathbf{P}[\overline{Y}_{n+1} = z+1 \mid \overline{Y}_n = z] = \mathbf{P}[\overline{Y}_{n+1} = z-1 \mid \overline{Y}_n = z] = \frac{1}{2}, \quad z \in \mathbb{Z}\backslash\{0\}, \\
&\mathbf{P}[\overline{Y}_{n+1} = 1 \mid \overline{Y}_n = 0] = \mathbf{P}[\overline{Y}_{n+1} = -1 \mid \overline{Y}_n = 0] = \frac{1}{4}, \qquad \mathbf{P}[\overline{Y}_{n+1} = 0 \mid \overline{Y}_n = 0] = \frac{1}{2}.
\end{aligned} \tag{2.19}$$

For every $n \in \mathbb{N}$, we let

$$H_n(\overline{Y}) = \bigl|\{1 \le i \le n : \overline{Y}_{i-1} = \overline{Y}_i = 0\}\bigr| \tag{2.20}$$

be the number of the horizontal steps at zero among the first n steps of $\overline{Y}$ and define $v(\overline{Y})_n$, $n = 0, 1, \dots$ as the process obtained from $\overline{Y}$ by removing all the horizontal steps at zero, so that

$$\overline{Y}_n = v(\overline{Y})_{n - H_n(\overline{Y})}, \quad n = 0, 1, \dots. \tag{2.21}$$

We also introduce, for $m \in \mathbb{N}$, $T > 0$ and $a \neq 0$, the normalized occupation times

$$L^a_{m;T}(\overline{Y}) := m^{-1/2}\bigl|\bigl\{0 \le t \le \lfloor mT \rfloor : \overline{Y}_t = \sqrt{m}a\bigr\}\bigr|, \tag{2.22}$$

where we define $\overline{Y}$ for non-integer times by linear interpolation. Lastly, we let $L^a_{m;T}(v(\overline{Y}))$ be given by (2.22) with $v(\overline{Y})$ in place of $\overline{Y}$.

Remark 2.6. By examining the transition probabilities of $\overline{Y}$ it becomes clear that $v(\overline{Y})_n$, $n = 0, 1, \dots$ is a SSRW, and that Y and $\overline{Y}$ can be coupled to obey

$$|\overline{Y}_n| = Y_n, \quad n = 0, 1, \dots, \tag{2.23}$$

provided the two processes have the same starting point. If we condition on the starting point $Y_0 = \overline{Y}_0 = \lfloor N^{1/3}x \rfloor$, then it holds under this coupling that

$$L^a\bigl(X^{k;x}\bigr) = L^a_{N^{2/3};T}(\overline{Y}) + L^{-a}_{N^{2/3};T}(\overline{Y}), \quad a > 0. \tag{2.24}$$

We conclude the section with the proof of Theorem 2.2.

Proof of Theorem 2.2. The maps Γ and $f \mapsto \sup_{t\in[0,T]}(-f(t))_+$ are 2-Lipschitz and 1-Lipschitz with respect to the supremum norm, respectively, so that (2.7) and (2.9) follow from (2.15) by combining Proposition 2.4 with (2.13) and (2.14), respectively.

We finish the proof of Theorem 2.2 by showing that (2.8) is a consequence of the estimate (2.7) and the regularity of the local time processes involved. To prove this, we follow the same argument as in [16, Appendix B]: Suppose that $f_1, f_2 : [0, T] \to [0, \infty)$ are measurable functions that have local times $L^a(f_i)$, in the sense that

$$\int_0^T \varphi\bigl(f_i(t)\bigr)\,\mathrm{d}t = \int_{\mathbb{R}} \varphi(a) L^a(f_i)\,\mathrm{d}a, \quad i = 1, 2 \tag{2.25}$$

for any measurable $\varphi : \mathbb{R} \to \mathbb{R}$ that vanishes on $(-\infty, 0]$. Suppose that $\kappa, \eta > 0$ satisfy

$$\sup_{a_1,a_2>0, |a_1-a_2|\le\kappa} \bigl|L^a(f_i) - L^a(f_i)\bigr| < \eta, \quad i = 1, 2. \tag{2.26}$$

Then, [4, Lemma 3.1] stipulates that[3]

$$\sup_{a>0}\big|L^a(f_1)-L^a(f_2)\big| \le \kappa^{-2} \sup_{t\in[0,T]}\big|f_1(t)-f_2(t)\big|+\eta. \tag{2.27}$$

We note that $a \mapsto L_T^a(R^x)$ inherits the regularity properties of $a \mapsto L_T^a(\widetilde{W})$ due to $R^x = |x+\widetilde{W}|$. Therefore, it follows from [35, (2.1)] that

$$\sup_{\substack{a_1,a_2>0\\ |a_1-a_2|\le N^{-2/15}}} \big|L_T^{a_1}(R^x)-L_T^{a_2}(R^x)\big| \le \mathcal{C}N^{-1/15}(\log N)^{1/2}, \quad N\in\mathbb{N} \tag{2.28}$$

almost surely for some random variable $\mathcal{C}$. At this point, if we combine (2.7) with (2.28), then the proof of (2.8) is reduced to following statement: for every $\varepsilon>0$, there exists a finite random variable $\mathcal{C}_\varepsilon$ such that

$$\sup_{\substack{a_1,a_2>0\\ |a_1-a_2|\le N^{-2/15}}} \big|L^{a_1}(X^{k;x})-L^{a_2}(X^{k;x})\big| \le \mathcal{C}_\varepsilon N^{-1/15+\varepsilon}, \quad N\in\mathbb{N}. \tag{2.29}$$

To this end and in view of (2.24), it suffices to prove that

$$\sup_{\substack{a_1,a_2\in\mathbb{R}\setminus\{0\}\\ |a_1-a_2|\le N^{-2/15}}} \big|L^{a_1}_{N^{2/3};T}(\overline{Y})-L^{a_2}_{N^{2/3};T}(\overline{Y})\big| \le \mathcal{C}_\varepsilon N^{-1/15+\varepsilon}, \quad N\in\mathbb{N}. \tag{2.30}$$

Applying [3, Proposition 3.1] to the SSRW $v(\overline{Y})$ we get

$$\sup_{t\in[0,T]}\sup_{\substack{a_1,a_2\in N^{-1/3}\mathbb{Z}\\ |a_1-a_2|\le N^{-2/15}}} \big|L^{a_1}_{N^{2/3};t}(v(\overline{Y}))-L^{a_2}_{N^{2/3};t}(v(\overline{Y}))\big| \le \mathcal{C}_\varepsilon N^{-1/15+\varepsilon}, \quad N\in\mathbb{N}, \tag{2.31}$$

and this can be extended to $a_1,a_2\in\mathbb{R}$ by the same argument used for [16, Equation (B.9)]. The desired estimate (2.30) follows then from (2.21). □

3. Proof of Theorem 1.7

This section is devoted to the proof of Theorem 1.7. Since the proof is rather long, we first present an informal overview of the arguments in Section 3.1, before rigorously carrying out the proof in Sections 3.2–3.4.

3.1. *Informal overview of the proof*

Let $\beta>0$, $w\in\mathbb{R}$, $T>0$ and $f,g\in\mathcal{D}$ be fixed. Our object of study is the scalar product

$$(\pi_N f)^\top \mathcal{M}^{\beta;w}_{T;N}(\pi_N g) = \sum_{l,l'=0}^{N} (\pi_N f)[l]\cdot \mathcal{M}^{\beta;w}_{T;N}[l,l']\cdot(\pi_N g)[l']. \tag{3.1}$$

Here and throughout the paper, we index all $(N+1)\times(N+1)$ matrices A by $l,l'\in\{0,1,\dots,N\}$ and write $A[l,l']$ for the (l,l')-entry of A. Similarly, the entries of all $(N+1)$-dimensional vectors v are denoted by $v[l]$ for $l\in$

[3] It should be noted that, as stated in [4], [4, Lemma 3.1] does not require the condition (2.26), but instead

$$\sup_{|a_1-a_2|\le\kappa}\big|L^a(f_1)-L^a(f_2)\big|<\eta$$

(see [4, Equation (3.1)]). However, this is easily remedied by implementing the following modification to the proof of [4, Lemma 3.1]: Instead of defining $\Psi(x):=[1-x\,\mathrm{sgn}(x)]\mathbf{1}_{[-1,1]}(x)$, we define $\Psi(x):=[1-(x-1)\,\mathrm{sgn}(x-1)]\mathbf{1}_{[0,2]}(x)$; the arguments in the proof of [4, Lemma 3.1] then go through by (2.25).

$\{0, 1, \dots, N\}$. By the definition of π_N in (1.17), we then see that

$$(\pi_N f)^\top \mathcal{M}_{T;N}^{\beta;w}(\pi_N g) = \int_0^{N^{2/3}} \int_0^{N^{2/3}} f(x) K_{T;N}^{\beta;w}(x, y) g(y)\,\mathrm{d}x\,\mathrm{d}y, \quad \text{where}$$

$$K_{T;N}^{\beta;w} := \sum_{l,l'=0}^{N} N^{1/3} \mathcal{M}_{T;N}^{\beta;w}\big[l, l'\big] \mathbf{1}_{[N^{-1/3}l, N^{-1/3}(l+1)) \times [N^{-1/3}l', N^{-1/3}(l'+1))}.$$

Recalling $k = k(T, N) = \lfloor T N^{2/3} \rfloor$ and the definition of $\mathcal{M}_{T;N}^{\beta;w}$ in (1.11), we find for all $l, l' \in \{0, 1, \dots, N\}$:

$$\mathcal{M}_{T;N}^{\beta;w}\big[l, l'\big] = \frac{1}{(2\sqrt{N})^k} \sum_{0 \le l_1, \dots, l_{k-1} \le N} M_N^{\beta;w}[l, l_1] M_N^{\beta;w}[l_1, l_2] \cdots M_N^{\beta;w}\big[l_{k-1}, l'\big]. \tag{3.2}$$

Since $M_N^{\beta;w}$ is tridiagonal, only $(k+1)$-tuples $(l_0, l_1, \dots, l_k)$ that satisfy $l_0 = l$, $l_k = l'$ and $|l_{s-1} - l_s| \in \{0, 1\}$ for all s contribute to $\mathcal{M}_{T;N}^{\beta;w}[l, l']$. Any such $(k+1)$-tuple can be thought of as a path from l_0 to l_k that takes steps of size $+1$ or -1 (when $|l_{s-1} - l_s| = 1$), and horizontal steps (when $l_{s-1} = l_s$). In the following, we rely on this observation to write the sum on the right-hand side of (3.2) in terms of expectations with respect to the random walks of Section 2.

For $j = 0, 1, \dots$ and $x \ge 0$, we define the random walk $X^{k-j;x}$, its normalized occupation times and its number of horizontal steps at zero by (2.2)–(2.6), with $(k-j)$ in place of k. We also let $\widehat{X}_t^{k-j;x}$, $t \in [0, T_{k-j} - N^{-2/3} H(X^{k-j;x})]$ be the path obtained from $X_t^{k-j;x}$, $t \in [0, T_{k-j}]$ by removing all horizontal segments at zero (see Figure 3). Finally, we introduce the functional

$$\begin{aligned}
&F_j\big(X^{k-j;x}, \mathbf{a}, \xi\big) \\
&\quad := \prod_{i=1}^{k-j-H(X^{k-j;x})} \frac{\sqrt{N - \widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}} \\
&\qquad \cdot \prod_{i=1}^{k-j-H(X^{k-j;x})} \left(1 + \frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N - \widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}\right) \ell_N^{H(X^{k-j;x})} \\
&\qquad \cdot \left(\frac{1}{(2\sqrt{N})^j} \sum_{0 \le i_1 \le \dots \le i_j \le k} \prod_{j'=1}^{j} \mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3} i_{j'}}}\right),
\end{aligned} \tag{3.3}$$

where the random walk $X^{k-j;x}$ is independent of all $\mathbf{a}_m$'s and ξ_m's. (We recall our convention $\mathbf{a}_0 := 0$ in Assumption 1.2, so that paths segments $(\widehat{X}^{k-j;x}_{i_1}, \dots, \widehat{X}^{k-j;x}_{i_j})$ that visit zero do not contribute to the sum on the last line of (3.3).)

If $x \in [N^{-1/3}l, N^{-1/3}(l+1))$ and $y \in [N^{-1/3}l', N^{-1/3}(l'+1))$, then by definition of $M_N^{\beta;w}$, one has

$$\mathcal{M}_{T;N}^{\beta;w}\big[l, l'\big] = \sum_{j=0}^{k} \frac{Q_{k-j}^{x,y}}{2^{k-j}} \mathbf{E}_{X^{k-j;x}}\big[F_j\big(X^{k-j;x}, \mathbf{a}, \xi\big) \mid X^{k-j;x}_{T_{k-j}} = \big\lfloor N^{1/3} y \big\rfloor\big], \tag{3.4}$$

with $Q_{k-j}^{x,y}$ being the number of paths $X^{k-j;x}$ can take such that $X^{k-j;x}_{T_{k-j}} = \lfloor N^{1/3} y \rfloor$, or equivalently,

$$Q_{k-j}^{x,y} := 2^{k-j} \mathbf{P}\big[X^{k-j;x}_{T_{k-j}} = \big\lfloor N^{1/3} y \big\rfloor\big]. \tag{3.5}$$

In the above, the parameter j represents the number of times it holds $l_{s-1} = l_s \neq 0$ within a $(k+1)$-tuple $(l_0, l_1, \dots, l_k)$. Removing the corresponding horizontal steps from the associated path leaves us with a path of $X^{k-j;x}$. At the same

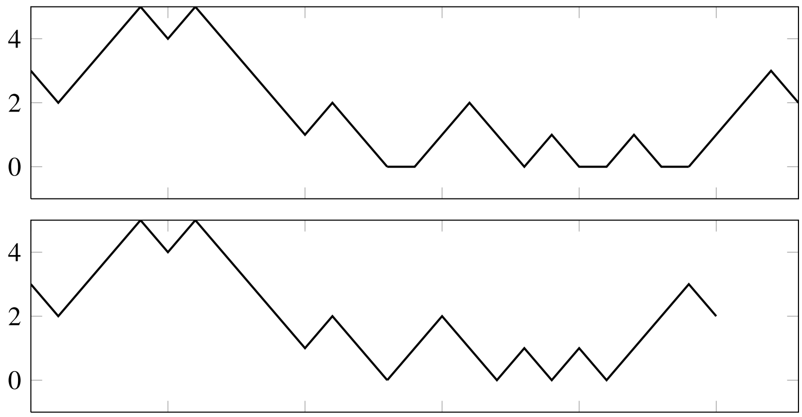

Fig. 3. Realizations of $X^{k-j;x}$ (top) and $\widehat{X}^{k-j;x}$ (bottom).

time, the term in the third line of (3.3) keeps track of all the possible ways j horizontal steps away from zero can be inserted into a given realization of $X^{k-j;x}$. Finally, the term $Q_{k-j}^{x,y}$ arises from the normalization inherent to the conditional expectation $\mathbf{E}_{X^{k-j;x}}[\cdot \mid X_{T_{k-j}}^{k-j;x} = \lfloor N^{1/3}y \rfloor]$.

Next, we denote the scalar products

$$\begin{aligned}\mathrm{Sc}_k^j(f,g) := &\int_0^{N^{2/3}}\int_0^{N^{2/3}} N^{1/3} f(x)\\ &\cdot \frac{Q_{k-j}^{x,y}}{2^{k-j}} \mathbf{E}_{X^{k-j;x}}\big[F_j\big(X^{k-j;x},\mathbf{a},\xi\big) \mid X_{T_{k-j}}^{k-j;x} = \big\lfloor N^{1/3}y\big\rfloor\big] g(y)\,\mathrm{d}x\,\mathrm{d}y, \end{aligned} \tag{3.6}$$

$j=0,1,\ldots$ and observe

$$(\pi_N f)^\top \mathcal{M}_{T;N}^{\beta;w}(\pi_N g) = \sum_{j=0}^{k} \mathrm{Sc}_k^j(f,g). \tag{3.7}$$

We also note that, by the total probability rule,

$$\mathrm{Sc}_k^j(f,g) = \int_0^{N^{2/3}} f(x)\mathbf{E}_{X^{k-j;x}}\bigg[F_j\big(X^{k-j;x},\mathbf{a},\xi\big)\cdot N^{1/3}\int_{N^{-1/3}X_{T_{k-j}}^{k-j;x}}^{N^{-1/3}(X_{T_{k-j}}^{k-j;x}+1)} g(y)\,\mathrm{d}y\bigg]\mathrm{d}x, \tag{3.8}$$

$j=0,1,\ldots$. The proof of Theorem 1.7 now hinges on justifying the following heuristic computation.

Heuristic computation 3.1. We recall the strong invariance principle of Theorem 2.2. Since $\log(1+z) = z + O(z^2)$ when $z \approx 0$, we have for $j=0,1,\ldots$ that

$$\begin{aligned} &\prod_{i=1}^{k-j-H(X^{k-j;x})} \frac{\sqrt{N - \widehat{X}_{N^{-2/3}(i-1)}^{k-j;x} \wedge \widehat{X}_{N^{-2/3}i}^{k-j;x}}}{\sqrt{N}}\\ &\quad = \exp\Bigg(\frac{1}{2}\sum_{i=1}^{k-j-H(X^{k-j;x})} \log\Bigg(1 - \frac{\widehat{X}_{N^{-2/3}(i-1)}^{k-j;x} \wedge \widehat{X}_{N^{-2/3}i}^{k-j;x}}{N}\Bigg)\Bigg)\\ &\quad \approx \exp\Bigg(-\frac{1}{2N^{2/3}}\sum_{i=1}^{k-j-H(X^{k-j;x})}\Bigg(\frac{\widehat{X}_{N^{-2/3}(i-1)}^{k-j;x} \wedge \widehat{X}_{N^{-2/3}i}^{k-j;x}}{N^{1/3}}\Bigg)\Bigg) \to \exp\bigg(-\int_0^T \frac{R_t^x}{2}\,\mathrm{d}t\bigg), \quad N\to\infty. \end{aligned} \tag{3.9}$$

At the same time, $(1-z)^{-1/2}=1+O(z)$ when $z\approx 0$ suggests for $j=0,1,\dots$ that

$$\begin{aligned}
&\prod_{i=1}^{k-j-H(X^{k-j;x})}\Bigg(1+\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N-\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}\Bigg)\\
&=\prod_{i=1}^{k-j-H(X^{k-j;x})}\Bigg(1+\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\bigg(1-\frac{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}{N}\bigg)^{-1/2}\Bigg)\\
&\approx\exp\Bigg(\sum_{i=1}^{k-j-H(X^{k-j;x})}\log\bigg(1+\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\bigg)\Bigg)\\
&\approx\exp\Bigg(\sum_{i=1}^{k-j-H(X^{k-j;x})}\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\Bigg)\\
&=\exp\bigg(\sum_{a\in N^{-1/3}(\mathbb{N}-1/2)}L^a\big(X^{k-j;x}\big)\frac{\xi_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}}\bigg)\to\exp\bigg(s_\xi\int_0^\infty L^a_T\big(R^x\big)\,\mathrm{d}W^\xi_a\bigg),
\end{aligned}\tag{3.10}$$

as $N\to\infty$, where W^ξ is the Brownian motion arising from a Donsker type invariance principle for the sequence $\xi_0,\xi_1,\dots$ (see (3.20) below). Moreover,

$$\begin{aligned}
\ell_N^{H(X^{k-j;x})}&=\bigg(1-\frac{N^{1/3}(1-\ell_N)}{N^{1/3}}\bigg)^{N^{1/3}\cdot(H(X^{k-j;x})/N^{1/3})}\\
&\to\exp\bigg(-w\frac{L^0_T(R^x)}{2}\bigg),\qquad N\to\infty,
\end{aligned}\tag{3.11}$$

for $j=0,1,\dots$.

Next, we consider $j=2$ and make the simple observation

$$\frac{1}{N}\sum_{0\le i_1\le i_2\le k}\mathbf{a}_{\widehat{X}^{k-2;x}_{N^{-2/3}i_1}}\mathbf{a}_{\widehat{X}^{k-2;x}_{N^{-2/3}i_2}}=\frac{1}{2N}\bigg(\sum_{i=0}^k\mathbf{a}_{\widehat{X}^{k-2;x}_{N^{-2/3}i}}\bigg)^2+\frac{1}{2N}\sum_{i=0}^k\mathbf{a}^2_{\widehat{X}^{k-2;x}_{N^{-2/3}i}}.\tag{3.12}$$

In addition, from $k=O(N^{2/3})$ and Assumption 1.2 we infer that the second summand on the right-hand side of (3.12) is negligible in the limit $N\to\infty$. Similar reasoning for $j=3,4,\dots$ reveals that, for all $j=1,2,\dots$, as $N\to\infty$,

$$\begin{aligned}
&\frac{1}{(2\sqrt{N})^j}\sum_{0\le i_1\le\cdots\le i_j\le k}\prod_{j'=1}^j\mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i_{j'}}}\\
&\approx\frac{1}{j!(2\sqrt{N})^j}\bigg(\sum_{i=0}^k\mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i}}\bigg)^j\\
&=\frac{1}{j!2^j}\bigg(\sum_{a\in N^{-1/3}\mathbb{N}}L^a\big(X^{k-j;x}\big)\frac{\mathbf{a}_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}}\bigg)^j\\
&\to\frac{1}{j!2^j}\bigg(s_{\mathbf{a}}\int_0^\infty L^a_T\big(R^x\big)\,\mathrm{d}W^{\mathbf{a}}_a\bigg)^j,
\end{aligned}\tag{3.13}$$

where $W^{\mathbf{a}}$ is the Brownian motion in a Donsker type invariance principle for the sequence $\mathbf{a}_1,\mathbf{a}_2,\dots$ (see (3.20) below).

Finally, the Lebesgue differentiation theorem suggests that

$$N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} g(y)\,\mathrm{d}y \to g\big(R^x_T\big), \quad N\to\infty. \tag{3.14}$$

All in all, we expect that, for each $j=0,1,\ldots$, the quantity $\mathrm{Sc}^j_k(f,g)$ converges, as $N\to\infty$, to

$$\begin{aligned}&\int_0^\infty f(x)\mathbf{E}_{R^x}\bigg[\exp\bigg(-\int_0^T \frac{R^x_t}{2}\,\mathrm{d}t\\&\quad+s_\xi\int_0^\infty L^a_T\big(R^x\big)\,\mathrm{d}W^\xi_a - w\frac{L^0_T(R^x)}{2}\bigg)\frac{1}{j!}\bigg(\frac{s_{\mathbf{a}}}{2}\int_0^\infty L^a_T\big(R^x\big)\,\mathrm{d}W^{\mathbf{a}}_a\bigg)^j g\big(R^x_T\big)\bigg]\,\mathrm{d}x.\end{aligned} \tag{3.15}$$

Summing over $j=0,1,\ldots$ and letting $W:=\sqrt{\beta}(s_\xi W^\xi+\frac{s_{\mathbf{a}}}{2}W^{\mathbf{a}})$ we end up precisely with the right-hand side of (1.18).

3.2. *Truncated convergence*

Our first step in the rigorous proof of Theorem 1.7 consists in establishing a convergence result for truncated versions of $\mathrm{Sc}^j_k(f,g)$, $j=0,1,\ldots$. To this end, we define for all $\underline{S}\in[-\infty,0]$ and $\overline{S}\in[0,\infty]$ the truncated functionals

$$\begin{aligned}&\widetilde{F}^{(\underline{S},\overline{S})}_j\big(X^{k-j;x},\mathbf{a},\xi\big)\\&=\underline{S}\vee\Bigg(\prod_{i=1}^{k-j-H(X^{k-j;x})}\frac{\sqrt{N-\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\\&\quad\cdot\prod_{i=1}^{k-j-H(X^{k-j;x})}\Bigg(1+\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N-\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge\widehat{X}^{k-j;x}_{N^{-2/3}i}}}\Bigg)\ell_N^{H(X^{k-j;x})}\\&\quad\cdot\frac{1}{j!(2\sqrt{N})^j}\Bigg(\sum_{i=0}^{k-j-H(X^{k-j;x})}\mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i}}\Bigg)^j\Bigg)\wedge\overline{S}, \quad j=0,1,\ldots\end{aligned} \tag{3.16}$$

and for all $K\in\mathbb{N}\cup\{0,\infty\}$ the truncated functions $f_K=fh_K$ and $g_K=gh_K$, where the continuous $h_K:[0,\infty)\to[0,1]$ satisfy $h_K\equiv 1$ on $[0,K)$ and $h_K\equiv 0$ on $[2K,\infty)$.

Remark 3.2. Note that, apart from the truncation at $\underline{S}$ and $\overline{S}$, the functionals F_j and $\widetilde{F}^{(\underline{S},\overline{S})}_j$ differ in the way the $\mathbf{a}_m$'s enter into them.

We now truncate the terms $\mathrm{Sc}^j_k(f,g)$, $j=0,1,\ldots$ according to

$$\begin{aligned}&\widetilde{\mathrm{Sc}}^j_k(f,g;\underline{S},\overline{S})\\&:=\int_0^{N^{2/3}} f(x)\mathbf{E}_{X^{k-j;x}}\bigg[\widetilde{F}^{(\underline{S},\overline{S})}_j\big(X^{k-j;x},\mathbf{a},\xi\big)\cdot N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} g(y)\,\mathrm{d}y\bigg]\,\mathrm{d}x.\end{aligned} \tag{3.17}$$

We also introduce the limiting operators

$$\big(\mathcal{U}^{(\underline{S},\overline{S})}_{T;j}f\big)(x):=\mathbf{E}_{R^x}\bigg[\bigg(\underline{S}\vee\exp\bigg(-\int_0^T\frac{R^x_t}{2}\,\mathrm{d}t+s_\xi\int_0^\infty L^a_T\big(R^x\big)\,\mathrm{d}W^\xi_a-w\frac{L^0_T(R^x)}{2}\bigg)$$

$$\cdot \frac{1}{j!}\biggl(\frac{s_{\mathbf{a}}}{2}\int_0^\infty L_T^a\bigl(R^x\bigr)\,\mathrm{d}W_a^{\mathbf{a}}\biggr)^j \wedge \overline{S}\biggr) f\bigl(R_T^x\bigr)\biggr] \tag{3.18}$$

for $f \in \mathcal{D}$ and $j = 0, 1, \ldots$. (The latter are truncated versions of the limits described in (3.15), whose sum over j yields $\mathcal{U}_T$.)

Proposition 3.3. *Let $\underline{S}$, $\overline{S}$ and K be finite. Then, for all functions $f \in \mathcal{D}$ and $g \in \mathcal{D} \cap C([0,\infty))$,*

$$\lim_{N\to\infty} \widetilde{\mathrm{Sc}}_k^j(f_K, g_K; \underline{S}, \overline{S}) = \int_0^\infty f_K(x)\bigl(\mathcal{U}_{T;j}^{(\underline{S},\overline{S})} g_K\bigr)(x)\,\mathrm{d}x \tag{3.19}$$

in distribution and in the sense of moments. These convergences hold jointly for any finite collection of j's, T's, f's and g's, and in the case of the convergence in distribution also jointly with the convergences in distribution

$$\lim_{N\to\infty} \sum_{m=0}^{\lfloor N^{1/3}x\rfloor} \frac{\mathbf{a}_m}{N^{1/6}} = s_{\mathbf{a}} W_x^{\mathbf{a}}, \quad x \ge 0 \quad \textit{and} \quad \lim_{N\to\infty} \sum_{m=0}^{\lfloor N^{1/3}x\rfloor} \frac{\xi_m}{N^{1/6}} = s_\xi W_x^\xi, \quad x \ge 0 \tag{3.20}$$

with respect to the Skorokhod topology.

The key ingredient in the proof of Proposition 3.3 is the next lemma. Therein and henceforth, for probability measures μ on $[0,\infty)$, we use the notations $X^{k;\mu}$ and R^μ for the random walk $X^{k;x}$ started according to the image of μ under the map $x \mapsto \lfloor xN^{1/3}\rfloor$ and the reflected Brownian motion R^x started according to μ, respectively.

Lemma 3.4. *Let $n \in \mathbb{N}$ and $\mu_1, \mu_2, \ldots, \mu_n$ be probability measures on $[0,\infty)$. Then, there exists a coupling of independent $X^{k;\mu_1}, X^{k;\mu_2}, \ldots, X^{k;\mu_n}$ with independent $R^{\mu_1}, R^{\mu_2}, \ldots, R^{\mu_n}$ such that the following limits in distribution hold jointly over $l = 1, 2, \ldots, n$, and also jointly with (3.20),*

$$\lim_{N\to\infty} \sup_{t\in[0,T_k]} \bigl|N^{-1/3}X_t^{k;\mu_l} - R_t^{\mu_l}\bigr| = 0, \tag{3.21}$$

$$\lim_{N\to\infty} \sum_{a\in N^{-1/3}(\mathbb{N}-1/2)} L^a\bigl(X^{k;\mu_l}\bigr)\frac{\xi_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}} = s_\xi \int_0^\infty L_T^a\bigl(R^{\mu_l}\bigr)\,\mathrm{d}W_a^\xi, \tag{3.22}$$

$$\lim_{N\to\infty} N^{-1/3}H\bigl(X^{k;\mu_l}\bigr) = \frac{L_T^0(R^{\mu_l})}{2}, \tag{3.23}$$

$$\lim_{N\to\infty} \sum_{a\in N^{-1/3}\mathbb{N}} L^a\bigl(X^{k;\mu_l}\bigr)\frac{\mathbf{a}_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}} = s_{\mathbf{a}} \int_0^\infty L_T^a\bigl(R^{\mu_l}\bigr)\,\mathrm{d}W_a^{\mathbf{a}}. \tag{3.24}$$

Proof. The lemma can be obtained from the coupling construction of Theorem 2.2 by the same arguments as in the derivation of [16, Proposition 4.9] from the coupling in [16, Proposition 4.1]. More specifically, one starts with the case $n = 1$ and $\mu_1 = \delta_x$ for some $x \ge 0$. Then, the joint convergences (3.21)–(3.24) in distribution are due to the convergence of the associated joint characteristic functions, which under the coupling of Theorem 2.2 is a consequence of the almost sure convergences of the *conditional* characteristic functions

$$\lim_{N\to\infty} \mathbf{E}_\xi\biggl[\exp\biggl(\mathrm{i}\theta \sum_{a\in N^{-1/3}(\mathbb{N}-1/2)} L^a\bigl(X^{k;x}\bigr)\frac{\xi_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}}\biggr)\biggr] = \mathbf{E}_{W^\xi}\biggl[\exp\biggl(\mathrm{i}\theta s_\xi \int_0^\infty L_T^a\bigl(R^x\bigr)\,\mathrm{d}W_a^\xi\biggr)\biggr], \tag{3.25}$$

$$\lim_{N\to\infty} \mathbf{E}_{\mathbf{a}}\biggl[\exp\biggl(\mathrm{i}\theta \sum_{a\in N^{-1/3}\mathbb{N}} L^a\bigl(X^{k;x}\bigr)\frac{\mathbf{a}_{\lfloor N^{1/3}a\rfloor}}{N^{1/6}}\biggr)\biggr] = \mathbf{E}_{W^{\mathbf{a}}}\biggl[\exp\biggl(\mathrm{i}\theta s_{\mathbf{a}} \int_0^\infty L_T^a\bigl(R^x\bigr)\,\mathrm{d}W_a^{\mathbf{a}}\biggr)\biggr] \tag{3.26}$$

for all $\theta \in \mathbb{R}$ (see [16, first half of p. 18] for more details). The latter follow from the central limit theorem in the form of the upper bound in [6, Theorem 8.4], the coupling of Theorem 2.2 and Assumption 1.2 (see [16, pp. 18–19] for more details).

In the case of $n=1$ and a general probability measure μ_1, the joint convergences (3.21)–(3.24) in distribution can be deduced from the previous case by integrating with respect to μ_1 and relying on the uniform boundedness of characteristic functions. Finally, in the case of $n>1$, one can repeat the same proof, but invoking the multidimensional version of the central limit theorem used before, obtaining this way also the convergences of (3.20) in the sense of convergence of finite-dimensional distributions. The latter can be improved to the desired distributional convergences of processes by applying a standard tightness result (see e.g. [7, Problem 8.4 and proof of Theorem 8.1]). □

We also prepare the following lemma needed in our proof of Proposition 3.3.

Lemma 3.5. *Let μ be a probability measure on $[0,\infty)$. Then, for each $j=0,1,\dots$, under any coupling such that $\lim_{N\to\infty} N^{-1/3}X^{k-j;\mu}_{T_{k-j}} = R^\mu_T$ almost surely, it holds*

$$\lim_{N\to\infty} N^{1/3}\int_{N^{-1/3}X^{k-j;\mu}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu}_{T_{k-j}}+1)} g(y)\,\mathrm{d}y = g\big(R^\mu_T\big) \tag{3.27}$$

with probability one, for any uniformly continuous function $g:[0,\infty)\to\mathbb{R}$.

Proof. It suffices to write

$$\begin{aligned} &N^{1/3}\int_{N^{-1/3}X^{k-j;\mu}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu}_{T_{k-j}}+1)} g(y)\,\mathrm{d}y \\ &\quad = N^{1/3}\int_{N^{-1/3}X^{k-j;\mu}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu}_{T_{k-j}}+1)} g(y) - g\big(N^{-1/3}X^{k-j;\mu}_{T_{k-j}}\big)\,\mathrm{d}y + g\big(N^{-1/3}X^{k-j;\mu}_{T_{k-j}}\big) \end{aligned} \tag{3.28}$$

and to note that the integral on the right-hand side tends to 0 with probability one, as $N\to\infty$, by the uniform continuity of g, whereas $\lim_{N\to\infty} g(N^{-1/3}X^{k-j;\mu}_{T_{k-j}}) = g(R^\mu_T)$ almost surely. □

We are now ready to prove Proposition 3.3.

Proof of Proposition 3.3. Let us first consider fixed j, T, f and g. Since the terms $\widetilde{\mathrm{Sc}}^j_k(f_K, g_K; \underline{S}, \overline{S})$ are bounded uniformly in N, it suffices to show the convergence of moments. Further, without loss of generality we may assume $f_K\ge 0$ and $\int_0^\infty f_K(x)\,\mathrm{d}x = 1$ (otherwise we write f_K as the difference of its positive and negative parts, and the latter as multiples of functions of the described kind). In particular, this allows us to define μ as the probability measure with the density f_K.

With i.i.d. copies $X^{k-j;\mu_1}, X^{k-j;\mu_2},\dots,X^{k-j;\mu_n}$ of $X^{k-j;\mu}$ and i.i.d. copies $R^{\mu_1}, R^{\mu_2},\dots,R^{\mu_n}$ of R^μ, the nth moment of $\widetilde{\mathrm{Sc}}^j_k(f_K, g_K; \underline{S}, \overline{S})$ can be expressed using Fubini's theorem as

$$\mathbf{E}\Bigg[\prod_{l=1}^n\Bigg(\widetilde{F}^{(\underline{S},\overline{S})}_j\big(X^{k-j;\mu_l},\mathbf{a},\xi\big)\cdot N^{1/3}\int_{N^{-1/3}X^{k-j;\mu_l}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu_l}_{T_{k-j}}+1)} g_K(y)\,\mathrm{d}y\Bigg)\Bigg], \tag{3.29}$$

whereas the nth moment of $\int_0^\infty f_K(x)(\mathcal{U}^{(\underline{S},\overline{S})}_{T;j} g_K)(x)\,\mathrm{d}x$ reads

$$\begin{aligned} \mathbf{E}\Bigg[\prod_{l=1}^n\Bigg(&\Bigg(\underline{S}\vee\exp\Bigg(-\int_0^T \frac{R^{\mu_l}_t}{2}\,\mathrm{d}t + s_\xi\int_0^\infty L^a_T\big(R^{\mu_l}\big)\,\mathrm{d}W^\xi_a - w\frac{L^0_T(R^{\mu_l})}{2}\Bigg) \\ &\cdot\frac{1}{j!}\Bigg(\frac{s_{\mathbf{a}}}{2}\int_0^\infty L^a_T\big(R^{\mu_l}\big)\,\mathrm{d}W^{\mathbf{a}}_a\Bigg)^j\wedge\overline{S}\Bigg) g_K\big(R^{\mu_l}_T\big)\Bigg)\Bigg]. \end{aligned} \tag{3.30}$$

To establish the convergence of the expectation in (3.29) to that in (3.2) we work under the coupling of Lemma 3.4 (that is, we upgrade the joint convergence in distribution in Lemma 3.4 to almost sure convergence by an application of the Skorokhod representation theorem [11, Theorem 3.5.1]) and view the random walks $X^{k-j;\mu_1}, X^{k-j;\mu_2}, \dots, X^{k-j;\mu_n}$ as the respective restrictions of $X^{k;\mu_1}, X^{k;\mu_2}, \dots, X^{k;\mu_n}$ to $[0, T_{k-j}]$. Then, $X^{k-j;\mu_1}, X^{k-j;\mu_2}, \dots, X^{k-j;\mu_n}$ inherit the asymptotics (3.21)–(3.24) from $X^{k;\mu_1}, X^{k;\mu_2}, \dots, X^{k;\mu_n}$, and Lemma 3.5 applies, so that

$$\lim_{N\to\infty} N^{1/3} \int_{N^{-1/3} X^{k-j;\mu_l}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu_l}_{T_{k-j}}+1)} g_K(y)\,\mathrm{d}y = g_K\big(R^{\mu_l}_T\big), \quad l = 1, 2, \dots, n \tag{3.31}$$

with probability one.

We proceed to the asymptotics of $\widetilde{F}^{(\underline{S},\overline{S})}_j(X^{k-j;\mu_l}, \mathbf{a}, \xi)$, $l = 1, 2, \dots, n$. Our first claim is that

$$\lim_{N\to\infty} \prod_{i=1}^{k-j-H(X^{k-j;\mu_l})} \frac{\sqrt{N - \widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}{\sqrt{N}} = \exp\bigg(-\int_0^T \frac{R^{\mu_l}_t}{2}\,\mathrm{d}t\bigg) \tag{3.32}$$

for $l = 1, 2, \dots, n$ almost surely. Indeed, for every such l, according to the Taylor expansion $\log(1+z) = z + O(z^2)$ about $z = 0$, an approximation as in the third line of (3.1) (with $X^{k-j;x}$, $\widehat{X}^{k-j;x}$ replaced by $X^{k-j;\mu_l}$, $\widehat{X}^{k-j;\mu_l}$) holds up to a multiplicative error of at most

$$\exp\Big(O\Big(N^{-4/3}\Big(\sup_{t\in[0,T_{k-j}]} X^{k-j;\mu_l}_t\Big)^2\Big)\Big). \tag{3.33}$$

Writing the resulting approximation in terms of $X^{k-j;\mu_l}$ we obtain (3.32) as an elementary consequence of (3.21).

Next, we prove the joint convergence in distribution

$$\begin{aligned} &\lim_{N\to\infty} \prod_{i=1}^{k-j-H(X^{k-j;\mu_l})} \Bigg(1 + \frac{\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}{\sqrt{N - \widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}\Bigg) \\ &\quad = \exp\bigg(s_\xi \int_0^\infty L^a_T\big(R^{\mu_l}\big)\,\mathrm{d}W^\xi_a\bigg) \end{aligned} \tag{3.34}$$

for $l = 1, 2, \dots, n$. To this end, we use the Taylor expansion $(1-z)^{-1/2} = 1 + O(z)$ about $z = 0$ to conclude that, for each l, an approximation as in the third line of (3.10) (with $X^{k-j;\mu_l}$, $\widehat{X}^{k-j;\mu_l}$ in place of $X^{k-j;x}$, $\widehat{X}^{k-j;x}$) applies up to a modification of each

$$\frac{\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}{\sqrt{N}} \quad \text{to} \quad \frac{\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}{\sqrt{N}} \Big(1 + O\Big(N^{-1} \sup_{t\in[0,T_{k-j}]} X^{k-j;\mu_l}_t\Big)\Big). \tag{3.35}$$

At this point, we employ the Taylor expansion $\log(1+z) = z + O(z^2)$ about $z = 0$ to obtain an expression as in the fourth line of (3.10), with the summands therein modified to

$$\begin{aligned} &\frac{\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}}{\sqrt{N}} \Big(1 + O\Big(N^{-1} \sup_{t\in[0,T_{k-j}]} X^{k-j;\mu_l}_t\Big)\Big) \\ &\quad + O\bigg(\frac{(\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}})^2}{N} \Big(1 + O\Big(N^{-1} \sup_{t\in[0,T_{k-j}]} X^{k-j;\mu_l}_t\Big)\Big)^2\bigg). \end{aligned} \tag{3.36}$$

The contribution of the first line in (3.36) can be evaluated as in the equality on the fifth line of (3.10), which leads to the limit in distribution of (3.22) after recalling (3.21). The contribution of the second line in (3.36) is asymptotically negligible due to the almost sure convergence

$$\lim_{N\to\infty} \sum_{i=1}^{k-j-H(X^{k-j;\mu_l})} \frac{(\xi_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}(i-1)}\wedge \widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}})^2}{N} = 0 \tag{3.37}$$

(simply apply the Borel–Cantelli lemma upon bounding the fourth moment of the latter sum via Assumption 1.2(c)) and (3.21). All in all, we arrive at (3.2).

Putting (3.32) and (3.2) together with the almost sure convergences

$$\ell_N^{H(X^{k-j;\mu_l})} = \left(1 - \frac{N^{1/3}(1-\ell_N)}{N^{1/3}}\right)^{N^{1/3}\cdot(H(X^{k-j;\mu_l})/N^{1/3})} \to \exp\left(-w\frac{L_T^0(R^{\mu_l})}{2}\right) \tag{3.38}$$

for $l = 1, 2, \ldots, n$ (see (3.23)), the convergences in distribution

$$\lim_{N\to\infty} \frac{1}{j!(2\sqrt{N})^j}\left(\sum_{i=0}^{k-j-H(X^{k-j;\mu_l})} \mathbf{a}_{\widehat{X}^{k-j;\mu_l}_{N^{-2/3}i}}\right)^j = \frac{1}{j!}\left(\frac{s_{\mathbf{a}}}{2}\int_0^\infty L_T^a\big(R^{\mu_l}\big)\,\mathrm{d}W_a^{\mathbf{a}}\right)^j \tag{3.39}$$

for $l = 1, 2, \ldots, n$ (see (3.24)) and (3.31) we conclude that the expectation in (3.29) converges to that in (3.2). Moreover, the joint convergence for any finitely many j's, T's, f's and g's can be shown by the same arguments, the only difference being that the formulas for moments in (3.29) and (3.2) have to be replaced by the corresponding formulas for joint moments. □

3.3. *Uniform moment bounds*

In this subsection, we establish some uniform moment estimates, which will allow us to lift the truncations and the continuity assumption on the g's of Proposition 3.3. To this end, we define the functionals

$$\begin{aligned}
&\widetilde{F}_j\big(X^{k-j;x}, \mathbf{a}, \xi\big) \\
&\quad= \prod_{i=1}^{k-j-H(X^{k-j;x})} \frac{\sqrt{N - \widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}} \\
&\qquad\cdot \prod_{i=1}^{k-j-H(X^{k-j;x})} \left(1 + \frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)}\wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N - \widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}\right)\ell_N^{H(X^{k-j;x})} \\
&\qquad\cdot \frac{1}{j!(2\sqrt{N})^j}\left(\sum_{i=0}^{k-j-H(X^{k-j;x})} \mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i}}\right)^j, \quad j = 0, 1, \ldots
\end{aligned} \tag{3.40}$$

and, with any $f, g \in \mathcal{D}$ and random variable Z_N (possibly depending on $X^{k-j;x}$, the $\mathbf{a}_m$'s and the ξ_m's), set

$$\widetilde{\mathrm{Sc}}^j_k(f, g; Z_N) := \int_0^{N^{2/3}} f(x)\mathbf{E}_{X^{k-j;x}}\left[\widetilde{F}_j\big(X^{k-j;x}, \mathbf{a}, \xi\big)\cdot N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} g(y)\,\mathrm{d}y \cdot Z_N\right]\mathrm{d}x \tag{3.41}$$

for $j = 0, 1, \ldots$. We also let

$$\overline{\mathrm{Sc}}^j_k(f, g, K; Z_N) := \widetilde{\mathrm{Sc}}^j_k(f, g; Z_N) - \widetilde{\mathrm{Sc}}^j_k(f_K, g_K; Z_N), \quad j = 0, 1, \ldots \tag{3.42}$$

for $K \in \mathbb{N}\cup\{0\}$.

Proposition 3.6. *For all $f, g \in \mathcal{D}$, one can find $N_0 \in \mathbb{N}$ and $C(K,n) < \infty$ with $\lim_{K\to\infty} C(K,n) = 0$ such that, for all $N \geq N_0$, if Z_N satisfies*

$$\mathbf{E}\big[|Z_N|^{3n}\big] \leq \Theta(3n) \tag{3.43}$$

for some function $\Theta : \mathbb{N} \to (0,\infty)$, one has

$$\mathbf{E}\big[\big|\overline{\mathrm{Sc}}_k^j(f,g,K;Z_N)\big|^n\big] \leq \frac{C(K,n)\Theta(3n)^{1/3}}{(3/2)^{jn}}, \quad K \in \mathbb{N}\cup\{0\}, j = 0,1,\dots. \tag{3.44}$$

The proof of Proposition 3.6 relies on the following lemma.

Lemma 3.7. *For all $1 \leq p < 3$ and $\theta \geq 0$, there exist constants $C = C(p,\theta) < \infty$, $c = c(\theta) > 0$ and $N_0 \in \mathbb{N}$ such that*

$$\sup_{N\geq N_0} \mathbf{E}\Bigg[\exp\Bigg(-\theta N^{-2/3}\sum_{i=1}^{k} N^{-1/3} X^{k;x}_{N^{-2/3}i}\Bigg)\Bigg] \leq Ce^{-cx}, \quad x > 0, \tag{3.45}$$

$$\sup_{N\geq N_0} \mathbf{E}\Bigg[\exp\Bigg(\theta N^{-1/3}\sum_{a\in N^{-1/3}\mathbb{N}} L^a\big(X^{k;x}\big)^p\Bigg)\Bigg] \leq C, \quad x > 0, \tag{3.46}$$

$$\sup_{N\geq N_0} \mathbf{E}\Bigg[\exp\Bigg(\theta N^{-1/3}\sum_{a\in N^{-1/3}(\mathbb{N}-1/2)} L^a\big(X^{k;x}\big)^p\Bigg)\Bigg] \leq C, \quad x > 0, \tag{3.47}$$

$$\sup_{N\geq N_0} \mathbf{E}\big[\ell_N^{\theta H(X^{x;k})}\big] \leq C, \quad x > 0. \tag{3.48}$$

Proof. Recall the random walks $\overline{Y}$ and $v(\overline{Y})$ introduced in (2.19) and the sentence following it. Throughout this proof, we condition on $\overline{Y}_0 = \lfloor N^{1/3}x\rfloor$ and assume that $\overline{Y}$ and $X^{k;x}$ are coupled as in Remark 2.6. We further write

$$\rho_k = \max_{0\leq i\leq k} v(\overline{Y})_i - \min_{0\leq i\leq k} v(\overline{Y})_i \tag{3.49}$$

for the range of the SSRW $v(\overline{Y})$ after k steps. It is clear that, for $i = 0,1,\dots,k$,

$$-X^{k;x}_{N^{-2/3}i} \leq -\min_{0\leq i\leq k} \overline{Y}_i \leq -\min_{0\leq i\leq k} v(\overline{Y})_i \leq -\big\lfloor N^{1/3}x\big\rfloor + \rho_k. \tag{3.50}$$

According to [10, inequality (6.2.3)] (the case $p = 1$ therein), one has

$$\mathbf{E}\big[\big(N^{-1/3}\rho_k\big)^n\big] \leq \sqrt{n!}\big(CT^{1/2}\big)^n, \quad n \in \mathbb{N} \tag{3.51}$$

with some uniform constant $C < \infty$. Thus, the exponential moment of $N^{-1/3}\rho_k$ can be bounded by a constant independent of N and x, yielding (3.45).

In view of (2.24),

$$L^a\big(X^{k;x}\big)^p \leq 2^{p-1}\big(L^a_{N^{2/3};T}\big(v(\overline{Y})\big)^p + L^{-a}_{N^{2/3};T}\big(v(\overline{Y})\big)^p\big). \tag{3.52}$$

Hence, it suffices to show (3.46) with θ replaced by $2^{p-1}\theta$ and $\sum_{a\in N^{-1/3}\mathbb{N}} L^a(X^{k;x})^p$ by $\sum_{a\in N^{-1/3}\mathbb{Z}\setminus\{0\}} L^a_{N^{2/3};T}(v(\overline{Y}))^p$. Repeating the proof of [16, Proposition 4.3] verbatim we find a constant $C = C(p) < \infty$ such that

$$N^{-1/3}\sum_{a\in N^{-1/3}\mathbb{Z}\setminus\{0\}} L^a_{N^{2/3};T}\big(v(\overline{Y})\big)^p \leq C\big(\big(N^{-1/3}\rho_k\big)^{p-1} + N^{-(p-1)/3}\big) \tag{3.53}$$

(note that even though [16] considers SSRW bridges, all the combinatorial identities therein apply to SSRWs as well). At this point, (3.46) follows from (3.51). Moreover, (3.47) is a consequence of

$$L^a\big(X^{k;x}\big) \le 2L^{a+N^{-1/3}/2}\big(X^{k;x}\big), \quad a \in N^{-1/3}(\mathbb{N}-1/2) \tag{3.54}$$

and (3.46).

From Proposition 2.4 we know that $H(X^{x;k}) = \sup_{t\in[0,T_k]}(-\widetilde{X}^{k;x}_t)_+$ under an appropriate coupling, where we recall that $\widetilde{X}^{x,k}$ is the rescaled SSRW introduced in Proposition 2.4. In particular, $H(X^{x;k})$ is stochastically dominated by ρ_k. Consequently, for $N\in\mathbb{N}$ large enough, the random variable inside the expectation in (3.48) is stochastically dominated by

$$\ell_N^{\theta\rho_k} = \left(1 - \frac{N^{1/3}(1-\ell_N)}{N^{1/3}}\right)^{N^{1/3}\cdot(\theta N^{-1/3}\rho_k)} = e^{-(w+o(1))\cdot(\theta N^{-1/3}\rho_k)}, \tag{3.55}$$

where $o(1)$ is non-random, so that (3.48) also follows from (3.51). □

Proof of Proposition 3.6. After observing

$$\begin{aligned}
&\overline{\mathrm{Sc}}_k^j(f,g,K;Z_N)\\
&\quad= \int_0^{N^{2/3}}\int_0^{N^{2/3}} N^{1/3} f(x)\\
&\qquad\cdot \frac{Q_{k-j}^{x,y}}{2^{k-j}}\mathbf{E}_{X^{k-j;x}}\big[\widetilde{F}_j\big(X^{k-j;x},\mathbf{a},\xi\big)\cdot Z_N \mid X^{k-j;x}_{T_{k-j}} = \lfloor N^{1/3}y\rfloor\big]g(y)\,\mathrm{d}x\,\mathrm{d}y\\
&\qquad- \int_0^{N^{2/3}}\int_0^{N^{2/3}} N^{1/3} f(x)h_K(x)\\
&\qquad\cdot \frac{Q_{k-j}^{x,y}}{2^{k-j}}\mathbf{E}_{X^{k-j;x}}\big[\widetilde{F}_j\big(X^{k-j;x},\mathbf{a},\xi\big)\cdot Z_N \mid X^{k-j;x}_{T_{k-j}} = \lfloor N^{1/3}y\rfloor\big]g(y)h_K(y)\,\mathrm{d}x\,\mathrm{d}y
\end{aligned} \tag{3.56}$$

we estimate $|\overline{\mathrm{Sc}}_k^j(f,g,K;Z_N)|$ by moving the absolute value inside the double integral and using

$$0 \le 1 - h_K(x)h_K(y) \le \mathbf{1}_{[K,\infty)}(x) + \mathbf{1}_{[K,\infty)}(y), \quad x, y \ge 0. \tag{3.57}$$

Since the roles of the variables x and y are symmetric, we only focus on the term in $\mathbf{E}[|\overline{\mathrm{Sc}}_k^j(f,g,K;Z_N)|^n]$ originating from $\mathbf{1}_{[K,\infty)}(x)$. We bound the latter by inserting an absolute value into the conditional expectation, applying Fubini's theorem and letting $\widetilde{f}_K := f\mathbf{1}_{[K,\infty)}$, thereby obtaining

$$\begin{aligned}
&\int_{[0,N^{2/3}]^n}\mathbf{E}\Bigg[\prod_{l=1}^n\big|\widetilde{f}_K(x_l)\big|\mathbf{E}_{X^{k-j;x_l}}\Bigg[\big|\widetilde{F}_j\big(X^{k-j;x_l},\mathbf{a},\xi\big)\big|\\
&\quad\cdot N^{1/3}\int_{N^{-1/3}X^{k-j;x_l}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x_l}_{T_{k-j}}+1)}\big|g(y)\big|\,\mathrm{d}y\cdot|Z_N|\Bigg]\Bigg]\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n.
\end{aligned} \tag{3.58}$$

A repeated application of Hölder's and Jensen's inequalities shows further that the quantity in (3.58) is at most

$$\begin{aligned}
&\Bigg(\int_0^{N^{2/3}}\big|\widetilde{f}_K(x)\big|\mathbf{E}\big[|Z_N|^{3n}\big]^{1/(3n)}\mathbf{E}\Bigg[\Bigg(N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)}\big|g(y)\big|\,\mathrm{d}y\Bigg)^{3n}\Bigg]^{1/(3n)}\\
&\quad\cdot\mathbf{E}\big[\big|\widetilde{F}_j\big(X^{k-j;x},\mathbf{a},\xi\big)\big|^{3n}\big]^{1/(3n)}\,\mathrm{d}x\Bigg)^n.
\end{aligned} \tag{3.59}$$

Due to $f \in \mathcal{D}$ and (3.43), we have

$$\big|\widetilde{f}_K(x)\big|\mathbf{E}\big[|Z_N|^{3n}\big]^{1/(3n)} \le C_1 e^{C_2 x^{1-\delta}} \mathbf{1}_{[K,\infty)}(x)\Theta(3n)^{1/3n}. \tag{3.60}$$

In view of $g \in \mathcal{D}$, we can choose C_1, C_2 and δ such that also

$$N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} \big|g(y)\big|\,\mathrm{d}y \le C_1 \exp\big(C_2\big(N^{-1/3}X^{k-j;x}_{T_{k-j}}\big)^{1-\delta}\big). \tag{3.61}$$

Moreover, by the argument leading to (3.50) and with the same notation as there,

$$N^{-1/3}X^{k;x}_{T_{k-j}} \le x + N^{-1/3}\rho_k. \tag{3.62}$$

It then follows from (3.51) that, with some $\widetilde{C}_1 = \widetilde{C}_1(C_1,n) < \infty$,

$$\mathbf{E}\bigg[\bigg(N^{1/3}\int_{N^{-1/3}X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} \big|g(y)\big|\,\mathrm{d}y\bigg)^{3n}\bigg]^{1/(3n)} \le \widetilde{C}_1 e^{C_2 x^{1-\delta}}. \tag{3.63}$$

It remains to control $\mathbf{E}[|\widetilde{F}_j(X^{k-j;x},\mathbf{a},\xi)|^{3n}]^{1/(3n)}$. For this purpose, we fix an $\varepsilon \in (0,\delta/3)$ and distinguish the cases $x \in (0, N^{2/3-\varepsilon}]$ and $x \in (N^{2/3-\varepsilon}, N^{2/3}]$. In the first case, we use Hölder's inequality to estimate $\mathbf{E}[|\widetilde{F}_j(X^{k-j;x},\mathbf{a},\xi)|^{3n}]^{1/(3n)}$ by the product of the four terms

$$\mathbf{E}\Bigg[\Bigg(\prod_{i=1}^{k-j-H(X^{k-j;x})} \frac{\sqrt{N-\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\Bigg)^{12n}\Bigg]^{1/(12n)}, \tag{3.64}$$

$$\mathbf{E}\Bigg[\Bigg(\prod_{i=1}^{k-j-H(X^{k-j;x})} \Bigg|1+\frac{\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N-\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}\Bigg|\Bigg)^{12n}\Bigg]^{1/(12n)}, \tag{3.65}$$

$$\mathbf{E}\big[\ell_N^{12nH(X^{k-j;x})}\big]^{1/(12n)}, \tag{3.66}$$

$$\mathbf{E}\Bigg[\Bigg(\frac{1}{j!(2\sqrt{N})^j}\Bigg|\sum_{i=0}^{k-j-H(X^{k-j;x})} \mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i}}\Bigg|^j\Bigg)^{12n}\Bigg]^{1/(12n)}. \tag{3.67}$$

Thanks to $\frac{\sqrt{N-z}}{\sqrt{N}} \le e^{-z/(2N)}$, $z \in [0,N]$ and (3.45), the quantity in (3.64) is not greater than $Ce^{-cx+O(jN^{-\varepsilon})}$. Turning to the term in (3.65), we write the expectation with respect to the ξ_m's as a product and note that, due to Assumption 1.2(c), [16, inequality (4.21)] yields for each factor a bound of the form

$$\exp\bigg(\frac{12nL^a(X^{k-j;x})|\mathbf{E}[\xi_a]|}{N^{-1/3}\sqrt{N-a}} + C'\bigg(\frac{(12nL^a(X^{k-j;x}))^2}{N^{-2/3}(N-a)} + \frac{(12nL^a(X^{k-j;x}))^{\gamma'}}{N^{-\gamma'/3}(N-a)^{\gamma'/2}}\bigg)\bigg), \tag{3.68}$$

with some $C' < \infty$ and $2 < \gamma' < 3$. For $N \in \mathbb{N}$ large enough, $N-a \ge N/2$ when $L^a(X^{k-j;x}) \ne 0$, which with Assumption 1.2(a) leads to the expectation of

$$\exp\bigg(C\sum_{a\in N^{-1/3}(\mathbb{N}-1/2)} \frac{12nL^a(X^{k-j;x})}{N^{1/2}} + \frac{(12nL^a(X^{k-j;x}))^2}{N^{1/3}} + \frac{(12nL^a(X^{k-j;x}))^{\gamma'}}{N^{\gamma'/6}}\bigg) \tag{3.69}$$

as an estimate on the expectation in (3.65). In addition, (3.48) reveals that the expression in (3.66) is at most $Ce^{O(jN^{-1/3})}$. Finally, the expectation with respect to the $\mathbf{a}_m$'s in (3.67) can be controlled via a combination of

$\frac{|z|^j}{j!} \le e^{|z|} \le e^z + e^{-z}$, $z \in \mathbb{R}$, [16, inequality (4.20)] and Assumption 1.2(a) by

$$\frac{C}{2^{12jn}} \exp\biggl(C \sum_{a \in N^{-1/3}\mathbb{N}} \frac{12nL^a(X^{k-j;x})}{N^{1/2}} + \frac{(12nL^a(X^{k-j;x}))^2}{N^{1/3}} + \frac{(12nL^a(X^{k-j;x}))^{\gamma'}}{N^{\gamma'/6}}\biggr), \tag{3.70}$$

with the same $2 < \gamma' < 3$ as before. Putting everything together, applying Hölder's inequality, and appealing to (3.47), (3.46) we arrive at

$$\mathbf{E}\bigl[\bigl|\widetilde{F}_j\bigl(X^{k-j;x}, \mathbf{a}, \xi\bigr)\bigr|^{3n}\bigr]^{1/(3n)} \le \frac{Ce^{-cx+O(jN^{-\varepsilon})}}{2^j}, \quad x \in \bigl(0, N^{2/3-\varepsilon}\bigr]. \tag{3.71}$$

In the case $x \in (N^{2/3-\varepsilon}, N^{2/3}]$, for all $N \in \mathbb{N}$ large enough, $X_t^{k-j;x} \ge N^{1-\varepsilon}/2$, $t \in [0, T_{k-j}]$, so that

$$\begin{aligned}
&\prod_{i=1}^{k-j-H(X^{k-j;x})} \left|\frac{\sqrt{N - \widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}} + \xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}}{\sqrt{N}}\right| \\
&\quad\le \prod_{i=1}^{k-j}\left(\frac{\sqrt{N - N^{1-\varepsilon}/2} + |\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}|}{\sqrt{N}}\right) \\
&\quad= \left(1 - \frac{1}{2N^{\varepsilon}}\right)^{(k-j)/2} \cdot \prod_{i=1}^{k-j}\left(1 + \frac{|\xi_{\widehat{X}^{k-j;x}_{N^{-2/3}(i-1)} \wedge \widehat{X}^{k-j;x}_{N^{-2/3}i}}|}{\sqrt{N - N^{1-\varepsilon}/2}}\right) \\
&\quad\le e^{-kN^{-\varepsilon}/4+O(jN^{-\varepsilon})} \cdot \exp\biggl(2 \sum_{a \in N^{-1/3}(\mathbb{N}-1/2)} L^a\bigl(X^{k-j;x}\bigr)\frac{|\xi_{\lfloor N^{1/3}a\rfloor}|}{N^{1/6}}\biggr).
\end{aligned} \tag{3.72}$$

Using Hölder's inequality and Lemma 3.7 as above, but changing the $\mathbf{a}_m$'s and ξ_m's to their absolute values and Assumption 1.2(a) to Assumption 1.2(c), we get

$$\mathbf{E}\bigl[\bigl|\widetilde{F}_j\bigl(X^{k-j;x}, \mathbf{a}, \xi\bigr)\bigr|^{3n}\bigr]^{1/(3n)} \le \frac{Ce^{-kN^{-\varepsilon}/4+O(jN^{-\varepsilon})+CN^{1/6}}}{2^j}, \quad x \in \bigl(N^{2/3-\varepsilon}, N^{2/3}\bigr]. \tag{3.73}$$

Lastly, we insert the right-hand sides of (3.60), (3.63), (3.71), (3.73) into (3.59):

$$\begin{aligned}
&\frac{Ce^{O(jnN^{-\varepsilon})}\Theta(3n)^{1/3}}{2^{jn}} \\
&\quad\cdot \biggl(\int_K^{N^{2/3-\varepsilon}\vee K} e^{Cx^{1-\delta}-cx}\,\mathrm{d}x + e^{-kN^{-\varepsilon}/4+CN^{1/6}} \int_{N^{2/3-\varepsilon}\vee K}^{N^{2/3}\vee K} e^{Cx^{1-\delta}}\,\mathrm{d}x\biggr)^n.
\end{aligned} \tag{3.74}$$

The estimate (3.44) readily follows upon recalling $k = \lfloor TN^{2/3}\rfloor$ and $\varepsilon \in (0, \delta/3)$. □

3.4. *Proof of Theorem* 1.7

3.4.1. *Step* 1: *Remove functional truncations*

In order to establish Theorem 1.7, we first argue that Proposition 3.3 remains true when $\underline{S} = -\infty$ and $\overline{S} = \infty$. We begin with the convergence of moments. Recall the moment formulas of (3.29) and (3.2). For any $n \in \mathbb{N}$, $\underline{S} \in [-\infty, 0]$, and $\overline{S} \in [0, \infty]$, the same arguments used in the proof of Proposition 3.3 provide a coupling under which

$$\lim_{N\to\infty} \prod_{l=1}^n \biggl(\widetilde{F}_j^{(\underline{S},\overline{S})}\bigl(X^{k-j;\mu_l}, \mathbf{a}, \xi\bigr) \cdot N^{1/3} \int_{N^{-1/3}X^{k-j;\mu_l}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;\mu_l}_{T_{k-j}}+1)} g_K(y)\,\mathrm{d}y\biggr)$$

$$= \prod_{l=1}^{n} \biggl(\underline{S} \vee \exp\biggl(- \int_0^T \frac{R_t^{\mu_l}}{2}\,\mathrm{d}t + s_\xi \int_0^\infty L_T^a\bigl(R^{\mu_l}\bigr)\,\mathrm{d}W_a^\xi - w\frac{L_T^0(R^{\mu_l})}{2} \biggr)$$

$$\cdot \frac{1}{j!}\biggl(\frac{s_{\mathbf{a}}}{2} \int_0^\infty L_T^a\bigl(R^{\mu_l}\bigr)\,\mathrm{d}W_a^{\mathbf{a}} \biggr)^j \wedge \overline{S} \biggr) g_K\bigl(R_T^{\mu_l}\bigr) \tag{3.75}$$

holds almost surely. To prove that the same limit holds in the sense of expectations, we need only show that the prelimit terms in (3.75) form a uniformly integrable sequence in $N \in \mathbb{N}$. In the cases where $\underline{S}$ and $\overline{S}$ are not both finite, this follows from (3.44) with $K = 0$ and $Z_N = 1$. The same argument gives also the joint convergence in the sense of moments.

We now prove the convergence (3.19) with $\underline{S} = -\infty$ and $\overline{S} = \infty$ in distribution. The convergence in moments of $\widetilde{\mathrm{Sc}}_k^j(f_K, g_K; \underline{S}, \infty)$ for finite $\underline{S}$ implies that the latter form a tight sequence. Thus, it is enough to prove that every weakly convergent subsequence of $\widetilde{\mathrm{Sc}}_k^j(f_K, g_K; \underline{S}, \infty)$ converge to the same limit, namely,

$$\int_0^\infty f_K(x)\bigl(\mathcal{U}_{T;j}^{(\underline{S},\infty)} g_K\bigr)(x)\,\mathrm{d}x. \tag{3.76}$$

Let $\widetilde{\mathrm{Sc}}_\infty^j(f_K, g_K; \underline{S}, \infty)$ be a limit point obtained by such a subsequence. If we assume that $f, g \ge 0$ (as we may do without loss of generality), then it is easy to see that $\widetilde{\mathrm{Sc}}_k^j(f_K, g_K; \underline{S}, \overline{S})$ is monotonically increasing in $\overline{S}$. Therefore, the limit point $\widetilde{\mathrm{Sc}}_\infty^j(f_K, g_K; \underline{S}, \infty)$ stochastically dominates the limit in distribution

$$\lim_{N\to\infty} \widetilde{\mathrm{Sc}}_\infty^j(f_K, g_K; \underline{S}, \infty) = \int_0^\infty f_K(x)\bigl(\mathcal{U}_{T;j}^{(\underline{S},\overline{S})} g_K\bigr)(x)\,\mathrm{d}x \tag{3.77}$$

for every finite $\overline{S}$. According to the monotone convergence theorem, the right-hand side of (3.77) converges as $\overline{S} \uparrow \infty$ to (3.76). We conclude that the limit point $\widetilde{\mathrm{Sc}}_\infty^j(f_K, g_K; \underline{S}, \infty)$ and (3.76) are two non-negative random variables with the same moments, and the former stochastically dominates the latter. Any two such random variables must be equal in distribution, which proves the convergence in distribution of (3.19) when $\overline{S} = \infty$.

In order to prove the convergence in distribution when $\underline{S} = -\infty$, we apply the same stochastic domination argument by exploiting the monotonicity in $\underline{S}$. As for the joint convergence for finite collections of j's, T's, f's and g's, we apply the same argument, leading to two random vectors with componentwise inequalities between them and same joint moments.

3.4.2. *Step* 2: *Continuous functions*

Next, we prove Theorem 1.7 under the additional assumption that the g's therein are continuous. Recall the definitions of Sc_k^j, $\widetilde{\mathrm{Sc}}_k^j$, and $\overline{\mathrm{Sc}}_k^j$ from (3.6), (3.3), and (3.41), respectively. Let

$$\Delta_k^j(f, g) := \mathrm{Sc}_k^j(f, g) - \widetilde{\mathrm{Sc}}_k^j(f, g; 1), \quad N \in \mathbb{N}, j = 0, 1, \ldots. \tag{3.78}$$

In view of (3.7), for all $K \in \mathbb{N}$, one has

$$(\pi_N f)^\top \mathcal{M}_{T;N}^{\beta;w}(\pi_N g) = \sum_{j=0}^{k} \widetilde{\mathrm{Sc}}_k^j(f_K, g_K; 1) + \sum_{j=0}^{k} \overline{\mathrm{Sc}}_k^j(f, g, K; 1) + \sum_{j=0}^{k} \Delta_k^j(f, g). \tag{3.79}$$

We aim to take the $N \to \infty$ limit of the right-hand side in (3.79) and start with the asymptotics of the first sum therein. For every finite set of summands $\widetilde{\mathrm{Sc}}_k^1(f_K, g_K; 1), \widetilde{\mathrm{Sc}}_k^2(f_K, g_K; 1), \ldots, \widetilde{\mathrm{Sc}}_k^J(f_K, g_K; 1)$, their joint limit in distribution and in the sense of moments is determined by the right-hand side of (3.19) with $\underline{S} = -\infty$ and $\overline{S} = \infty$. This and the moment bounds of (3.44) imply that the first sum on the right-hand side of (3.79) converges to $\int_0^\infty f_K(x)(\mathcal{U}_T^{\beta;w} g_K)(x)\,\mathrm{d}x$ in distribution and in the sense of moments. Since the moment bounds of (3.44) for

$\overline{\mathrm{Sc}}_k^j(f_K, g_K, 0; 1)$ are inherited by their $N \to \infty$ limits, we have, in addition,

$$\lim_{K\to\infty}\lim_{N\to\infty}\sum_{j=0}^{k}\widetilde{\mathrm{Sc}}_k^j(f_K, g_K; 1) = \lim_{K\to\infty}\int_0^\infty f_K(x)\big(\mathcal{U}_T^{\beta;w} g_K\big)(x)\,\mathrm{d}x = \int_0^\infty f(x)\big(\mathcal{U}_T^{\beta;w} g\big)(x)\,\mathrm{d}x \tag{3.80}$$

in distribution and in the sense of moments.

Moreover, the moment bounds of (3.44) reveal that

$$\lim_{K\to\infty}\lim_{N\to\infty}\sum_{j=0}^{k}\overline{\mathrm{Sc}}_k^j(f, g, K; 1) = 0 \tag{3.81}$$

in L^n, for all $n \in \mathbb{N}$. To analyze the third sum on the right-hand side of (3.79) we introduce, for all $j = 1, 2, \dots$, the notations

$$h_j(N) = \sum_{0\le i_1\le\cdots\le i_j\le k}\prod_{j'=1}^{j}\mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i_{j'}}}, \qquad p_j(N) = \sum_{i=0}^{k-j-H(X^{k-j;x})}\big(\mathbf{a}_{\widehat{X}^{k-j;x}_{N^{-2/3}i}}\big)^j, \tag{3.82}$$

and $[z^j]P(z)$ for the coefficient of z^j in a power series $P(z)$. Then, the Newton identities relating the complete homogeneous symmetric functions to the power sums (see e.g. [26, Chapter 1, Section 2]) yield

$$\frac{h_j(N)}{(2\sqrt{N})^j} = \frac{p_1(N)^j}{j!(2\sqrt{N})^j} + \sum_{\iota=0}^{j-1}\frac{p_1(N)^\iota}{\iota!(2\sqrt{N})^\iota}\left(\big[z^{j-\iota}\big]\exp\left(\sum_{j'=2}^{\infty}\frac{p_{j'}(N)}{j'(2\sqrt{N})^{j'}}z^{j'}\right)\right). \tag{3.83}$$

Therefore, with

$$Z_N^{j,\iota} := \big[z^{j-\iota}\big]\exp\left(\sum_{j'=2}^{\infty}\frac{p_{j'}(N)}{j'(2\sqrt{N})^{j'}}z^{j'}\right), \quad \iota = 0, 1, \dots, j-1, \tag{3.84}$$

it holds

$$\sum_{j=0}^{k}\Delta_k^j(f, g) = \sum_{j=0}^{k}\sum_{\iota=0}^{j-1}\overline{\mathrm{Sc}}_k^\iota\big(f, g; Z_N^{j,\iota}\big). \tag{3.85}$$

By [16, Lemma 4.20], one can find bounds $\mathbf{E}[|Z_N^{j,\iota}|^n] \le \Theta(3n, j-\iota, N)$ such that

$$\lim_{N\to\infty}\sum_{j=0}^{k}\sum_{\iota=0}^{j-1}\frac{\Theta(3n, j-\iota, N)^{1/(3n)}}{(3/2)^j} = 0. \tag{3.86}$$

A combination of the triangle inequality for the L^n norm, the moment bounds of (3.44) and the property (3.86) gives

$$\lim_{N\to\infty}\sum_{j=0}^{k}\Delta_k^j(f, g) = 0 \tag{3.87}$$

in L^n, for all $n \in \mathbb{N}$. This finishes the proof of Theorem 1.7 under the continuity assumption on the g's.

3.4.3. *Step* 3: *General functions*

For general $f, g \in \mathcal{D}$, the same arguments as above reveal that it suffices to identify, for all $K \in \mathbb{N}$, the limit of $\sum_{j=0}^{k} \widetilde{\mathrm{Sc}}_k^j(f_K, g_K; 1)$ in distribution and in the sense of moments as $\int_0^\infty f_K(x)(\mathcal{U}_T^{\beta;w} g_K)(x)\,\mathrm{d}x$. In fact, it is enough to establish

$$\lim_{N\to\infty} \sum_{j=0}^{k} \widetilde{\mathrm{Sc}}_k^j(f_K, g_K; 1) = \int_0^\infty f_K(x)\big(\mathcal{U}_T^{\beta;w} g_K\big)(x)\,\mathrm{d}x \tag{3.88}$$

in distribution, since the moment bounds of (3.44) for $\overline{\mathrm{Sc}}_k^j(f_K, g_K, 0; 1)$ then imply the convergence of moments. To see (3.88) we pick $g_{\eta,K}$, $\eta \in \mathbb{N}$ in $C([0,\infty))$ so that

$$\upsilon_\eta := \|g_{\eta,K} h_K - g_K\|_{L^2([0,\infty))} \underset{\eta\to\infty}{\longrightarrow} 0. \tag{3.89}$$

Recalling the symmetry of $\widetilde{\mathrm{Sc}}_k^j(\cdot,\cdot;1)$ (cf. (3.56)), applying the Cauchy–Schwarz inequality, and repeating the proof of Proposition 3.6 mutatis mutandis we get

$$\begin{aligned}
&\mathbf{E}\big[\big|\widetilde{\mathrm{Sc}}_k^j(f_K, g_K; 1) - \widetilde{\mathrm{Sc}}_k^j(f_K, g_{\eta,K} h_K; 1)\big|^2\big] \\
&\quad \le \upsilon_\eta^2 \mathbf{E}\bigg[\int_0^{N^{2/3}} \mathbf{E}_{X^{k-j;x}}\bigg[\widetilde{F}_j\big(X^{k-j;x}, \mathbf{a}, \xi\big) \cdot N^{1/3} \int_{N^{-1/3} X^{k-j;x}_{T_{k-j}}}^{N^{-1/3}(X^{k-j;x}_{T_{k-j}}+1)} f_K(y)\,\mathrm{d}y\bigg]^2 \mathrm{d}x\bigg] \\
&\quad \le \upsilon_\eta^2 \frac{C e^{O(jN^{-\varepsilon})}}{2^{2j}}.
\end{aligned} \tag{3.90}$$

To complete the proof of Theorem 1.7 we observe that

$$\forall \eta \in \mathbb{N}: \lim_{N\to\infty} \sum_{j=0}^{k} \widetilde{\mathrm{Sc}}_k^j(f_K, g_{\eta,K} h_K; 1) = \int_0^\infty f_K(x)\big(\mathcal{U}_T^{\beta;w}(g_{\eta,K} h_K)\big)(x)\,\mathrm{d}x \tag{3.91}$$

in distribution and that

$$\lim_{\eta\to\infty} \int_0^\infty f_K(x)\big(\mathcal{U}_T^{\beta;w}(g_{\eta,K} h_K)\big)(x)\,\mathrm{d}x = \int_0^\infty f_K(x)\big(\mathcal{U}_T^{\beta;w} g_K\big)(x)\,\mathrm{d}x \tag{3.92}$$

almost surely, thanks to Proposition 1.10(c).

4. Properties of the limiting operators

The goal of this section is to prove Proposition 1.10. We start by preparing some auxiliary constructions and results. Recall the spiked Gaussian β-ensemble $H_N^{\beta;w}$ defined in (1.7), and let

$$A_N^{\beta;w} := N^{1/6}\big(2\sqrt{N} - H_N^{\beta;w}\big), \tag{4.1}$$

viewed as an operator acting on $L^2([0,\infty))$ via Remark 1.6. The next proposition is a direct corollary of [8, Proposition 2.8, Remark 2.9, Lemma 2.7, Theorem 2.10 and its proof] (note that [8, Assumptions 1–3] are verified for $H_N^{\beta;w}$ in [8, Section 3]).

Proposition 4.1. *For all $\beta > 0$ and $w \in \mathbb{R}$, the operator $\mathcal{H}^{\beta;w}$ of* (1.9) *almost surely possesses a purely discrete spectrum $\Lambda_1 < \Lambda_2 < \cdots$ satisfying $\Lambda_q \to \infty$ as $q \to \infty$. The corresponding eigenspaces are one-dimensional, and* *each is therefore spanned by a normalized eigenfunction f_q. Moreover, one can couple $\mathcal{H}^{\beta;w}$ with a subsequence of $A_N^{\beta;w}$, $N\in\mathbb{N}$ along which, almost surely,*

$$\lim_{N\to\infty}\Lambda_{q,N}=\Lambda_q \quad \textit{and} \quad \lim_{N\to\infty}\|v_{q,N}-f_q\|_{L^2([0,\infty))}=0, \quad q=1,2,\ldots, \tag{4.2}$$

where $\Lambda_{1,N}<\Lambda_{2,N}<\cdots<\Lambda_{N,N}$ and $v_{1,N},v_{2,N},\ldots,v_{N,N}$ are eigenvalues and corresponding eigenfunctions of $A_N^{\beta;w}$. Along the same subsequence, the standard Brownian motion W in the definition of $\mathcal{H}^{\beta;w}$ arises in the almost sure limit (1.19).

Next, we present an alternative formula for the kernel $K_T^{\beta;w}$ of (1.15), to be used in the proof of Proposition 1.10.

Lemma 4.2. *For all $\beta>0$, $w\in\mathbb{R}$ and $T>0$, define the kernels*

$$K_T^{\beta}(x,y)=\frac{\exp(-\frac{(x-y)^2}{2T})}{\sqrt{2\pi T}} \cdot \mathbf{E}_{\widetilde{W}^x}\bigg[\mathbf{1}_{\{\min_{0\le t\le T}\widetilde{W}_t^x>0\}}\exp\bigg(-\int_0^T\frac{\widetilde{W}_t^x}{2}\,\mathrm{d}t+\int_0^\infty\frac{L_T^a(\widetilde{W}^x)}{\sqrt{\beta}}\,\mathrm{d}W_a\bigg)\,\bigg|\,\widetilde{W}_T^x=y\bigg] \tag{4.3}$$

and

$$\overline{K}_T^{\beta;w}(x,y)=\frac{2\exp(-\frac{(x-y)^2}{2T})}{\sqrt{2\pi T}}\mathbf{E}_{\widetilde{W}^x}\bigg[\mathbf{1}_{\{\min_{0\le t\le T}\widetilde{W}_t^x\le 0\}} \cdot \exp\bigg(-\int_0^T\frac{|\widetilde{W}_t^x|}{2}\,\mathrm{d}t+\int_0^\infty\frac{L_T^a(|\widetilde{W}^x|)}{\sqrt{\beta}}\,\mathrm{d}W_a-w\frac{L_T^0(|\widetilde{W}^x|)}{2}\bigg)\,\bigg|\,\widetilde{W}_T^x=y\bigg], \tag{4.4}$$

where $\widetilde{W}^x$ is a Brownian motion started at x, independent of W. Then,

$$K_T^{\beta;w}(x,y)=K_T^{\beta}(x,y)+\overline{K}_T^{\beta;w}(x,y), \quad x,y\ge 0. \tag{4.5}$$

Proof. Let $R^x:=|\widetilde{W}^x|$. For any random variable Z, we note that

$$\begin{aligned}\mathbf{E}_{\widetilde{W}^x}\big[Z\mid |\widetilde{W}_T^x|=y\big]&=\mathbf{E}_{\widetilde{W}^x}\big[Z\mid \widetilde{W}_T^x=y\big]\mathbf{P}\big[\widetilde{W}_T^x=y\mid |\widetilde{W}_T^x|=y\big]\\&\quad+\mathbf{E}_{\widetilde{W}^x}\big[Z\mid \widetilde{W}_T^x=-y\big]\mathbf{P}\big[\widetilde{W}_T^x=-y\mid |\widetilde{W}_T^x|=y\big],\end{aligned} \tag{4.6}$$

where

$$\mathbf{P}\big[\widetilde{W}_T^x=y\mid |\widetilde{W}_T^x|=y\big]=\frac{\exp(-\frac{(x-y)^2}{2T})}{\exp(-\frac{(x-y)^2}{2T})+\exp(-\frac{(x+y)^2}{2T})}. \tag{4.7}$$

Thus, the formula for $K_T^{\beta;w}$ in (1.15) can be rewritten as

$$\frac{\exp(-\frac{(x-y)^2}{2T})}{\sqrt{2\pi T}} \cdot \mathbf{E}_{\widetilde{W}^x}\bigg[\exp\bigg(-\int_0^T\frac{|\widetilde{W}_t^x|}{2}\,\mathrm{d}t+\int_0^\infty\frac{L_T^a(|\widetilde{W}^x|)}{\sqrt{\beta}}\,\mathrm{d}W_a-w\frac{L_T^0(|\widetilde{W}^x|)}{2}\bigg)\,\bigg|\,\widetilde{W}_T^x=y\bigg] + \frac{\exp(-\frac{(x+y)^2}{2T})}{\sqrt{2\pi T}}$$

$$\cdot\, \mathbf{E}_{\widetilde{W}^x}\biggl[\exp\biggl(-\int_0^T \frac{|\widetilde{W}_t^x|}{2}\,\mathrm{d}t + \int_0^\infty \frac{L_T^a(|\widetilde{W}^x|)}{\sqrt{\beta}}\,\mathrm{d}W_a - w\frac{L_T^0(|\widetilde{W}^x|)}{2}\biggr)\biggm|\widetilde{W}_T^x = -y\biggr]. \tag{4.8}$$

Next, we decompose the first expectation in (4.8) according to the events

$$\Bigl\{\min_{0\le t\le T}\widetilde{W}_t^x > 0\Bigr\} \quad \text{and} \quad \Bigl\{\min_{0\le t\le T}\widetilde{W}_t^x \le 0\Bigr\}, \tag{4.9}$$

and note that, on the former event, $|\widetilde{W}^x| = \widetilde{W}^x$ and $L_T^0(|\widetilde{W}^x|) = 0$.

In addition, by the strong Markov property and the symmetry about 0 of Brownian motion, instead of conditioning on $\widetilde{W}_T^x = -y$ in the second expectation of (4.8), we can condition on $\min_{0\le t\le T}\widetilde{W}_t^x \le 0$, $\widetilde{W}_T^x = y$. This operation, in turn, is equivalent to inserting $\mathbf{1}_{\{\min_{0\le t\le T}\widetilde{W}_t^x\le 0\}}$ into the expectation, conditioning on $\widetilde{W}_T^x = y$ and normalizing the result by $\mathbf{P}[\min_{0\le t\le T}\widetilde{W}_t^x \le 0 \mid \widetilde{W}_T^x = y]$. Computing

$$\mathbf{P}\Bigl[\min_{0\le t\le T}\widetilde{W}_t^x \le 0 \mid \widetilde{W}_T^x = y\Bigr] = e^{-2xy/T} \tag{4.10}$$

from the joint density of the running minimum and the current value of a Brownian motion (see e.g. [33, Chapter III, Exercise 3.14]) and observing

$$e^{2xy/T}\cdot\frac{\exp(-\frac{(x+y)^2}{2T})}{\sqrt{2\pi T}} = \frac{\exp(-\frac{(x-y)^2}{2T})}{\sqrt{2\pi T}} \tag{4.11}$$

we arrive at the right-hand side of (4.5). □

We are now ready to prove Proposition 1.10.

Proof of Proposition 1.10. (a) The identity (1.20) follows from Theorem 1.7 by the same arguments as were used to obtain [16, Corollary 2.12] from Theorem 2.8 therein. To summarize the argument:

1. Let $\mathcal{M}_N^{\beta;w} = (A_N^{\beta;w}/2\sqrt{N})^{\lfloor TN^{2/3}\rfloor}$, with $A_N^{\beta;w}$ as in (4.1). According to Proposition 4.1 and a computation similar to (1.5), we see that $\mathcal{M}_N^{\beta;w}$'s eigenfunctions and eigenvalues converge in the sense of Proposition 4.1 to those of $e^{-T\mathcal{H}^{\beta;w}/2}$.
2. Combining the above with Theorem 1.7 and Remark 1.8, we conclude that $\mathcal{U}_T^{\beta;w}$ and $e^{-T\mathcal{H}^{\beta;w}/2}$ must share the same eigenvalues and eigenfunctions, and we note that the Brownian motions W in $\mathcal{H}^{\beta;w}$ and $\mathcal{U}_T^{\beta;w}$ arise as the same limit of matrix entries.

We refer to the proof of [16, Corollary 2.12] for the details; for this argument to apply to the present case, one only needs to replace every reference to the main result of [32] by a reference to Proposition 4.1, the pointer to [16, Lemma 6.1] by a pointer to (1.5), and the assertion that the eigenvalues of $-\frac{1}{2}\mathcal{H}^\beta$ tend to $-\infty$ by the same statement for $-\frac{1}{2}\mathcal{H}^{\beta;w}$ (cf. Proposition 4.1).

(b), (c) We proceed to the almost sure Hilbert–Schmidt property of $\mathcal{U}_T^{\beta;w}$, for each $T > 0$. In view of Lemma 4.2, it is enough to show that

$$\mathbf{E}\biggl[\int_0^\infty\int_0^\infty \bigl(K_T^\beta(x,y) + \overline{K}_T^{\beta;w}(x,y)\bigr)^2\,\mathrm{d}x\,\mathrm{d}y\biggr] < \infty. \tag{4.12}$$

Since

$$\mathbf{E}\biggl[\int_0^\infty\int_0^\infty K_T^\beta(x,y)^2\,\mathrm{d}x\,\mathrm{d}y\biggr] < \infty \tag{4.13}$$

is established in [16, proof of Lemma 5.1], it suffices to check

$$\int_0^\infty\int_0^\infty \mathbf{E}\bigl[\overline{K}_T^{\beta;w}(x,y)^2\bigr]\,\mathrm{d}x\,\mathrm{d}y < \infty. \tag{4.14}$$

Next, we estimate $\mathbf{E}[\overline{K}_T^{\beta;w}(x,y)^2]$ by moving the square function into the expectation $\mathbf{E}_{\widetilde{W}^x}[\cdot \mid \widetilde{W}_T^x = y]$, dropping the indicator random variable, employing

$$L_T^a\big(|\widetilde{W}^x|\big)^2 = \big(L_T^a(\widetilde{W}^x) + L_T^{-a}(\widetilde{W}^x)\big)^2 \le 2L_T^a(\widetilde{W}^x)^2 + 2L_T^{-a}(\widetilde{W}^x)^2, \quad a \ge 0, \tag{4.15}$$

and evaluating the expectation with respect to W:

$$\begin{aligned}&\mathbf{E}\big[\overline{K}_T^{\beta;w}(x,y)^2\big]\\ &\quad\le \frac{2\exp(-\frac{(x-y)^2}{T})}{\pi T}\mathbf{E}\bigg[\exp\bigg(-\int_0^T |\widetilde{W}_t^x|\,\mathrm{d}t + \int_{-\infty}^{\infty}\frac{4L_T^a(\widetilde{W}^x)^2}{\beta}\,\mathrm{d}a - 2wL_T^0(\widetilde{W}^x)\bigg)\,\bigg|\,\widetilde{W}_T^x = y\bigg].\end{aligned} \tag{4.16}$$

According to Hölder's inequality, the latter expectation is at most

$$\begin{aligned}&\mathbf{E}\bigg[\exp\bigg(-\int_0^T 3|\widetilde{W}_t^x|\,\mathrm{d}t\bigg)\,\bigg|\,\widetilde{W}_T^x = y\bigg]^{1/3}\mathbf{E}\bigg[\exp\bigg(\int_{-\infty}^{\infty}\frac{12L_T^a(\widetilde{W}^x)^2}{\beta}\,\mathrm{d}a\bigg)\,\bigg|\,\widetilde{W}_T^x = y\bigg]^{1/3}\\ &\quad\cdot\mathbf{E}\big[e^{-6wL_T^0(\widetilde{W}^x)} \mid \widetilde{W}_T^x = y\big]^{1/3}.\end{aligned} \tag{4.17}$$

Thanks to $|\widetilde{W}_t^x| \ge \widetilde{W}_t^x$, the identity in distribution

$$\big(\widetilde{W}_t^x : t \in [0,T] \mid \widetilde{W}_T^x = y\big) \stackrel{\mathrm{d}}{=} \bigg(\widetilde{W}_t^0 + \bigg(1 - \frac{t}{T}\bigg)x + \frac{t}{T}y : t \in [0,T]\,\bigg|\,\widetilde{W}_T^0 = 0\bigg), \tag{4.18}$$

and $(1-\frac{t}{T})x + \frac{t}{T}y \ge x \wedge y$, the first factor in (4.17) is bounded above by

$$e^{-T(x\wedge y)}\mathbf{E}\bigg[\exp\bigg(-\int_0^T 3\widetilde{W}_t^0\,\mathrm{d}t\bigg)\,\bigg|\,\widetilde{W}_T^0 = 0\bigg]^{1/3}. \tag{4.19}$$

In addition, the coupling of [16, Proposition 4.1] reveals the random variable $\int_{-\infty}^{\infty} L_T^a(\widetilde{W}^x)^2\,\mathrm{d}a$, conditioned on $\widetilde{W}_T^x = y$, as the almost sure $N \to \infty$ limit of the left-hand side in [16, inequality (4.15)]. Thus, the second expectation in (4.17) can be controlled by the limit inferior of the corresponding exponential moment of the right-hand side in [16, inequality (4.15)]. Proceeding as therein we arrive at

$$\mathbf{E}\bigg[\exp\bigg(\int_{-\infty}^{\infty}\frac{12L_T^a(\widetilde{W}^x)^2}{\beta}\,\mathrm{d}a\bigg)\,\bigg|\,\widetilde{W}_T^x = y\bigg]^{1/3} \le Ce^{C|x-y|}, \tag{4.20}$$

with a constant $C = C(\beta,T) < \infty$. Also, we see from [28, equation (3)] that the density of the local time at 0 of a Brownian bridge from x' to y' on $[0,1]$ is

$$\big(z + x' + y'\big)\exp\bigg(\frac{1}{2}\big((x'-y')^2 - (z + x' + y')^2\big)\bigg), \quad z > 0 \tag{4.21}$$

and from (4.10) that this local time vanishes with probability $1 - e^{-2x'y'}$. Hence,

$$\begin{aligned}&\mathbf{E}\big[\exp\big(\theta L_1^0(\widetilde{W}^{x'})\big) \mid \widetilde{W}_1^{x'} = y'\big]\\ &\quad= 1 + \sqrt{\frac{\pi}{2}}\theta e^{-2x'y'}\exp\bigg(\frac{(x'+y'-\theta)^2}{2}\bigg)\bigg(2 - \operatorname{erfc}\bigg(\frac{-x'-y'+\theta}{\sqrt{2}}\bigg)\bigg)\\ &\quad\le C(\theta) < \infty\end{aligned} \tag{4.22}$$

due to standard estimates for the complementary error function. All in all, it follows that the left-hand side in (4.14) is less or equal to

$$\int_0^\infty \int_0^\infty C \exp\biggl(-\frac{(x-y)^2}{C} - \frac{x \wedge y}{C} + C|x-y|\biggr)\,\mathrm{d}x\,\mathrm{d}y < \infty, \tag{4.23}$$

where $C = C(\beta, w, T)$ is a finite positive constant.

We turn to the proof of the semigroup property in Proposition 1.10(b) and assume without loss of generality $T_1, T_2 > 0$. By the just established Hilbert–Schmidt property, it suffices to verify that, almost surely,

$$\mathcal{U}_{T_1}^{\beta;w}\mathcal{U}_{T_2}^{\beta;w} f = \mathcal{U}_{T_1+T_2}^{\beta;w} f \tag{4.24}$$

on a countable dense set of functions $f \in L^2([0,\infty))$. Fixing such a function f, writing it as the difference of its positive and negative parts, and applying Fubini's theorem we reduce the statement of Proposition 1.10(b) further to

$$\int_0^\infty K_{T_1}^{\beta;w}(x,z) K_{T_2}^{\beta;w}(z,y)\,\mathrm{d}z = K_{T_1+T_2}^{\beta;w}(x,y), \quad x, y \ge 0. \tag{4.25}$$

Let us introduce the transition kernels

$$\gamma_T(x,y) := \frac{\exp(-\frac{(x-y)^2}{2T}) + \exp(-\frac{(x+y)^2}{2T})}{\sqrt{2\pi T}}, \quad x, y \ge 0, T > 0 \tag{4.26}$$

and the additive functionals

$$F_T\bigl(R^x\bigr) := -\int_0^T \frac{R_t^x}{2}\,\mathrm{d}t + \int_0^\infty \frac{L_T^a(R^x)}{\sqrt{\beta}}\,\mathrm{d}W_a - w\frac{L_T^0(R^x)}{2}, \quad T > 0. \tag{4.27}$$

Then,

$$\begin{aligned} &\int_0^\infty K_{T_1}^{\beta;w}(x,z) K_{T_2}^{\beta;w}(z,y)\,\mathrm{d}z \\ &\quad = \gamma_{T_1+T_2}(x,y) \\ &\qquad \cdot \int_0^\infty \frac{\gamma_{T_1}(x,z)\gamma_{T_2}(z,y)}{\gamma_{T_1+T_2}(x,y)} \mathbf{E}_{R^x}\bigl[e^{F_{T_1}(R^x)} \mid R_{T_1}^x = z\bigr] \mathbf{E}_{R^z}\bigl[e^{F_{T_2}(R^z)} \mid R_{T_2}^z = y\bigr]\,\mathrm{d}z. \end{aligned} \tag{4.28}$$

To identify the right-hand side of (4) with $K_{T_1+T_2}^{\beta;w}(x,y)$ it remains to notice that the process $(R_t^x : t \in [0, T_1+T_2] \mid R_{T_1+T_2}^x = y)$ therein can be sampled by

(a) picking a random point Z according to the density $\frac{\gamma_{T_1}(x,z)\gamma_{T_2}(z,y)}{\gamma_{T_1+T_2}(x,y)}$, $z > 0$,
(b) conditional $Z = z$, sampling processes $R^{(1)}$, $R^{(2)}$ independently such that

$$\bigl(R_t^{(1)} : t \in [0, T_1]\bigr) \stackrel{\mathrm{d}}{=} \bigl(R_t^x : t \in [0, T_1] \mid R_{T_1}^x = z\bigr), \tag{4.29}$$

$$\bigl(R_t^{(2)} : t \in [0, T_2]\bigr) \stackrel{\mathrm{d}}{=} \bigl(R_t^z : t \in [0, T_2] \mid R_{T_1}^z = y\bigr), \tag{4.30}$$

(c) concatenating the paths of $R^{(1)}$ and $R^{(2)}$.

As with the semigroup property, for each $T > 0$, the symmetry property of the operator $\mathcal{U}_T$ can be reduced to an assertion about its kernel:

$$K_T^{\beta;w}(x,y) = K_T^{\beta;w}(y,x), \quad x, y \ge 0. \tag{4.31}$$

Since the transition kernels $\gamma_{T'}$, $T'>0$ of the reflected Brownian motion R are symmetric, we have

$$\big(R_t^x : t\in[0,T]\mid R_T^x=y\big)\overset{\mathrm{d}}{=}\big(R_{T-t}^y : t\in[0,T]\mid R_T^y=x\big), \tag{4.32}$$

and therefore (4.31). Finally, the non-negativity of $\mathcal{U}_T^{\beta;w}$ follows by extending

$$\begin{aligned}\int_0^\infty \big(\mathcal{U}_T^{\beta;w}f\big)(x)f(x)\,\mathrm{d}x &= \int_0^\infty \big(\mathcal{U}_{T/2}^{\beta;w}\big(\mathcal{U}_{T/2}^{\beta;w}f\big)\big)(x)f(x)\,\mathrm{d}x \\ &= \int_0^\infty \big(\big(\mathcal{U}_{T/2}^{w;\beta}f\big)(x)\big)^2\,\mathrm{d}x\ge 0,\end{aligned} \tag{4.33}$$

for a fixed $f\in L^2([0,\infty))$, to the same almost sure property for all $f\in L^2([0,\infty))$ simultaneously, by means of the almost sure Hilbert–Schmidt property of $\mathcal{U}_T^{\beta;w}$.

(d) To obtain the almost sure trace class property of $\mathcal{U}_T^{\beta;w}$ and the trace formula (1.21) we combine the spectral theorem for symmetric compact operators with the definition of the trace to find

$$\mathrm{Tr}\big(\mathcal{U}_T^{\beta;w}\big)=\sum_{q=1}^\infty e^{-T\Lambda_q/2}. \tag{4.34}$$

The latter sum is the square of the Hilbert–Schmidt norm of the symmetric Hilbert–Schmidt operator $\mathcal{U}_{T/2}^{\beta;w}$ (see e.g. [25, Section 28, Exercise 11]) and, thus, equals to

$$\int_0^\infty\int_0^\infty K_{T/2}^{\beta;w}(x,y)K_{T/2}^{\beta;w}(y,x)\,\mathrm{d}y\,\mathrm{d}x=\int_0^\infty K_T^{\beta;w}(x,x)\,\mathrm{d}x. \tag{4.35}$$

(e) For the L^2-strong continuity in expectation of (1.22), without loss of generality we fix an integer $p\ge 2$, an f with $\|f\|_{L^2([0,\infty))}=1$, and a sequence $(t_\eta)_{\eta\in\mathbb{N}}$ in $[0,T+1]$ converging to T such that $t_\eta>0$ for at least one η. Then, by applying [25, Section 28, Theorem 7] to the commuting symmetric operators $\mathcal{U}_{t_\eta}^{\beta;w}$, $\eta\in\mathbb{N}$ and $\mathcal{U}_T^{\beta;w}$, with at least one $\mathcal{U}_{t_\eta}^{\beta;w}$ being compact, we can write f as $\sum_{q=1}^\infty c_q f_q$, where f_q, $q\in\mathbb{N}$ form an orthonormal basis of common eigenfunctions for $\mathcal{U}_{t_\eta}^{\beta;w}$, $\eta\in\mathbb{N}$ and $\mathcal{U}_T^{\beta;w}$, and c_q, $q\in\mathbb{N}$ are the corresponding coefficients. By Jensen's inequality,

$$\begin{aligned}\mathbf{E}\big[\big\|\mathcal{U}_T^{\beta;w}f-\mathcal{U}_{t_\eta}^{\beta;w}f\big\|_{L^2([0,\infty))}^p\big] &= \mathbf{E}\left[\left(\sum_{q=1}^\infty c_q^2\big(e^{-T\Lambda_q/2}-e^{-t_\eta\Lambda_q/2}\big)^2\right)^{p/2}\right] \\ &\le \mathbf{E}\left[\sum_{q=1}^\infty c_q^2\big(e^{-T\Lambda_q/2}-e^{-t_\eta\Lambda_q/2}\big)^p\right].\end{aligned} \tag{4.36}$$

The random variable $(\omega,q)\mapsto(e^{-T\Lambda_q(\omega)/2}-e^{-t_\eta\Lambda_q(\omega)/2})^p$ tends to 0 in the $\eta\to\infty$ limit $\mathbf{P}\times\sum_{q=1}^\infty\delta_{c_q^2}$ almost surely. Its uniform integrability is due to

$$\mathbf{E}\left[\sum_{q=1}^\infty c_q^2\big(e^{-T\Lambda_q/2}-e^{-t_\eta\Lambda_q/2}\big)^{2p}\right]\le 2^{2p-1}\mathbf{E}\big[2\big(e^{-p(T+1)\Lambda_1}+1\big)\big] \tag{4.37}$$

and a bound on $e^{-p(T+1)\Lambda_1}$ by the squared Hilbert–Schmidt norm of $\mathcal{U}_{p(T+1)}^{\beta;w}$, whose expectation has been controlled in the proof of part (c). □

5. Functionals of the reflected Brownian bridge

In this section, we prove Theorem 1.13, Proposition 1.12 and Corollary 1.15, in the order stated. The key ingredient in the proof of Theorem 1.13 is the next lemma, which extends an argument of Hariya [17].

Lemma 5.1. *Let r_t, $t \in [0,1]$ be a reflected Brownian bridge, $\alpha > 0$ and ψ_α be the joint moment-generating function of*

$$\left(\int_0^1 r_t \,\mathrm{d}t, \int_0^\infty L_1^a(r)^2 \,\mathrm{d}a \mid L_1^0(r) = \alpha\right). \tag{5.1}$$

Then, with the three-dimensional Bessel bridge

$$\mathrm{d}b_t^{\alpha/2} = \frac{1}{b_t^{\alpha/2}}\,\mathrm{d}t - \frac{b_t^{\alpha/2}}{1-t}\,\mathrm{d}t + \mathrm{d}\widetilde{W}_t, \quad b_0^{\alpha/2} = \alpha/2 \tag{5.2}$$

and the joint moment-generating function $\widetilde{\psi}_\alpha$ of $(\int_0^1 b_t^{\alpha/2}\,\mathrm{d}t, \int_0^1 \widetilde{W}_t\,\mathrm{d}t)$, it holds

$$\psi_\alpha(\theta_1,\theta_2) = e^{-\alpha\theta_1/4}\widetilde{\psi}_\alpha(\theta_1 + 2\theta_2, -\theta_1/2), \quad \theta_1,\theta_2 \in \mathbb{R}. \tag{5.3}$$

Proof. Define the function

$$h(a) := \int_0^1 \mathbf{1}_{\{r_t \le a\}}\,\mathrm{d}t = \int_0^a L_1^{a'}(r)\,\mathrm{d}a', \tag{5.4}$$

as well as the corresponding quantile function $h^{-1}(t) := \inf\{a \ge 0 : h(a) \ge t\}$. In [29, Corollary 16(iii)] (see also [30, equation (8.20)]), Pitman shows that

$$\left(\frac{1}{2}L_1^{h^{-1}(t)}(r) : t \in [0,1] \mid L_1^0(r) = \alpha\right) \stackrel{\mathrm{d}}{=} \left(b_t^{\alpha/2} : t \in [0,1]\right), \tag{5.5}$$

which extends Jeulin's theorem beyond the $\alpha = 0$ case. Relying on (5.5) we find

$$\begin{aligned}
&\left(\frac{1}{2}\int_0^1 \frac{1-t}{b_t^{\alpha/2}}\,\mathrm{d}t, \int_0^1 b_t^{\alpha/2}\,\mathrm{d}t\right) \\
&\quad \stackrel{\mathrm{d}}{=} \left(\int_0^1 \frac{1-t}{L_1^{h^{-1}(t)}(r)}\,\mathrm{d}t, \frac{1}{2}\int_0^1 L_1^{h^{-1}(t)}(r)\,\mathrm{d}t \,\Big|\, L_1^0(r) = \alpha\right) \\
&\quad = \left(\int_0^\infty \frac{1-h(a)}{L_1^a(r)}h'(a)\,\mathrm{d}a, \frac{1}{2}\int_0^\infty L_1^a(r)h'(a)\,\mathrm{d}a \,\Big|\, L_1^0(r) = \alpha\right) \\
&\quad = \left(\int_0^\infty 1 - h(a)\,\mathrm{d}a, \frac{1}{2}\int_0^\infty L_1^a(r)^2\,\mathrm{d}a \,\Big|\, L_1^0(r) = \alpha\right) \\
&\quad = \left(\int_0^\infty \int_0^1 \mathbf{1}_{\{r_t > a\}}\,\mathrm{d}t\,\mathrm{d}a, \frac{1}{2}\int_0^\infty L_1^a(r)^2\,\mathrm{d}a \,\Big|\, L_1^0(r) = \alpha\right) \\
&\quad = \left(\int_0^1 r_t\,\mathrm{d}t, \frac{1}{2}\int_0^\infty L_1^a(r)^2\,\mathrm{d}a \,\Big|\, L_1^0(r) = \alpha\right).
\end{aligned} \tag{5.6}$$

On the other hand, (5.2) implies

$$\int_0^1 (1-t)\,\mathrm{d}b_t^{\alpha/2} = \int_0^1 \frac{1-t}{b_t^{\alpha/2}}\,\mathrm{d}t - \int_0^1 b_t^{\alpha/2}\,\mathrm{d}t + \int_0^1 (1-t)\,\mathrm{d}\widetilde{W}_t. \tag{5.7}$$

Using integration by parts for the two stochastic integrals and rearranging we get

$$\int_0^1 \frac{1-t}{b_t^{\alpha/2}}\,\mathrm{d}t = -\frac{\alpha}{2} + 2\int_0^1 b_t^{\alpha/2}\,\mathrm{d}t - \int_0^1 \widetilde{W}_t\,\mathrm{d}t. \tag{5.8}$$

Finally, a sequential application of (5.6) and (5.8) yields

$$\begin{aligned}
&\mathbf{E}\bigg[\exp\bigg(\theta_1\int_0^1 r_t\,\mathrm{d}t + \theta_2\int_0^\infty L_1^a(r)^2\,\mathrm{d}a\bigg)\,\bigg|\,L_1^0(r)=\alpha\bigg]\\
&\quad= \mathbf{E}\bigg[\exp\bigg(\frac{\theta_1}{2}\int_0^1 \frac{1-t}{b_t^{\alpha/2}}\,\mathrm{d}t + 2\theta_2\int_0^1 b_t^{\alpha/2}\,\mathrm{d}t\bigg)\bigg]\\
&\quad= e^{-\alpha\theta_1/4}\mathbf{E}\bigg[\exp\bigg((\theta_1+2\theta_2)\int_0^1 b_t^{\alpha/2}\,\mathrm{d}t - \frac{\theta_1}{2}\int_0^1 \widetilde{W}_t\,\mathrm{d}t\bigg)\bigg],
\end{aligned} \tag{5.9}$$

that is, (5.3). □

Theorem 1.13 can be now obtained from Lemma 5.1 as follows.

Proof of Theorem 1.13. The case $\alpha = 0$ is the subject of [16, Corollary 2.15], [17, Theorem 1.1], so we focus on the $\alpha > 0$ case. By Lemma 5.1, for every $\theta \in \mathbb{R}$,

$$\mathbf{E}\bigg[\exp\bigg(\theta\int_0^1 r_t\,\mathrm{d}t - \frac{\theta}{2}\int_0^\infty L_1^a(r)^2\,\mathrm{d}a\bigg)\mid L_1^0(r)=\alpha\bigg] = e^{-\alpha\theta/4}\widetilde{\psi}_\alpha(0,-\theta/2). \tag{5.10}$$

Since $\int_0^1 \widetilde{W}_t\,\mathrm{d}t$ is Gaussian with mean 0 and variance $\frac{1}{3}$, the right-hand side of (5.10) equals to $e^{-\alpha\theta/4+\theta^2/24}$, the moment-generating function of a Gaussian random variable with mean $-\alpha/4$ and variance $1/12$. □

We conclude the paper with the proofs of Proposition 1.12 and Corollary 1.15.

Proof of Proposition 1.12. Let $\widetilde{r}_t$, $t \in [0,T]$ be a reflected Brownian bridge from 0 to 0 on $[0,T]$. By the definition of $K_T^{\beta;w}$ in (1.15),

$$K_T^{\beta;w}(0,0) = \sqrt{\frac{2}{\pi T}}\mathbf{E}_{\widetilde{r}}\bigg[\exp\bigg(-\int_0^T \frac{\widetilde{r}_t}{2}\,\mathrm{d}t + \int_0^\infty \frac{L_T^a(\widetilde{r})}{\sqrt{\beta}}\,\mathrm{d}W_a - w\frac{L_T^0(\widetilde{r})}{2}\bigg)\bigg]. \tag{5.11}$$

Conditional on $\widetilde{r}$, the integral $\int_0^\infty \frac{L_T^a(\widetilde{r})}{\sqrt{\beta}}\,\mathrm{d}W_a$ is Gaussian with mean 0 and variance $\int_0^\infty \frac{L_T^a(\widetilde{r})^2}{\beta}\,\mathrm{d}a$. Hence, by taking the expectation with respect to W first, we find

$$\mathbf{E}\big[K_T^{\beta;w}(0,0)\big] = \sqrt{\frac{2}{\pi T}}\mathbf{E}\bigg[\exp\bigg(-\int_0^T \frac{\widetilde{r}_t}{2}\,\mathrm{d}t + \int_0^\infty \frac{L_T^a(\widetilde{r})^2}{2\beta}\,\mathrm{d}a - w\frac{L_T^0(\widetilde{r})}{2}\bigg)\bigg]. \tag{5.12}$$

At this point, the proposition is a consequence of

$$\begin{aligned}
&\bigg(\int_0^T \widetilde{r}_t\,\mathrm{d}t, \int_0^\infty L_T^a(\widetilde{r})^2\,\mathrm{d}a, L_T^0(\widetilde{r})\bigg)\\
&\quad\stackrel{\mathrm{d}}{=} \bigg(T^{3/2}\int_0^1 r_t\,\mathrm{d}t, T^{3/2}\int_0^\infty L_1^a(r)^2\,\mathrm{d}a, T^{1/2}L_1^0(r)\bigg),
\end{aligned} \tag{5.13}$$

which, in turn, is due to the scaling property of (reflected) Brownian bridges. □

Proof of Corollary 1.15. In view of (4.21), the local time $L_1^0(r)$ is a continuous random variable with the density $\frac{\alpha}{4}e^{-\alpha^2/8}$ on $(0,\infty)$. Using this and Theorem 1.13 for the right-hand side of (1.12) we compute

$$\begin{aligned}\mathbf{E}\big[K_T^{2;w}(0,0)\big] &= \sqrt{\frac{2}{\pi T}}\int_0^\infty \frac{\alpha}{4}e^{-\alpha^2/8}\exp\Big(-T^{1/2}w\frac{\alpha}{2}\Big)\\ &\quad\cdot \mathbf{E}\Big[\exp\Big(-\frac{T^{3/2}}{2}\Big(\int_0^T r_t\,\mathrm{d}t-\frac{1}{2}\int_0^\infty L_1^a(r)\,\mathrm{d}a\Big)\Big)\,\Big|\,L_1^0(r)=\alpha\Big]\mathrm{d}\alpha\\ &= \sqrt{\frac{2}{\pi T}}\int_0^\infty \frac{\alpha}{4}e^{-\alpha^2/8}\exp\Big(-T^{1/2}w\frac{\alpha}{2}+\frac{T^{3/2}\alpha}{8}+\frac{T^3}{96}\Big)\,\mathrm{d}\alpha\\ &= \frac{e^{\frac{T^3}{96}}(8+\sqrt{2\pi}e^{T(T-4w)^2/32}\sqrt{T}(T-4w)(\operatorname{erf}(\frac{\sqrt{T}(T-4w)}{4\sqrt{2}})+1))}{4\sqrt{2\pi T}},\end{aligned}\tag{5.14}$$

which simplifies to the right-hand side of (1.15). □

References

[1] G. Anderson, A. Guionnet and O. Zeitouni. *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge University Press, Cambridge, 2010. MR2760897

[2] J. Baik, G. Ben Arous and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** (2005) 1643–1697. MR2165575

[3] R. Bass and D. Khoshnevisan. Strong approximations to Brownian local time. *Progr. Probab.* **33** (1992) 43–65. MR1278076

[4] R. F. Bass and D. Khoshnevisan. Laws of the iterated logarithm for local times of the empirical process. *Ann. Probab.* **23** (1995) 388–399. MR1330775

[5] K. E. Bassler, P. J. Forrester and N. E. Frankel. Edge effects in some perturbations of the Gaussian unitary ensemble. *J. Math. Phys.* **51** (2010), Article ID 123305. MR2779601

[6] R. N. Bhattacharya and R. R. Rao. *Normal Approximation and Asymptotic Expansions*. John Wiley & Sons, New York, 1976. MR0436272

[7] P. Billingsley. *Convergence of Probability Measures*, 2nd edition. John Wiley & Sons, New York, 1999. MR1700749

[8] A. Bloemendal and B. Virág. Limits of spiked random matrices I. *Probab. Theory Related Fields* **156** (2013) 795–825. MR3078286

[9] A. Bloemendal and B. Virág. Limits of spiked random matrices II. *Ann. Probab.* **44** (2016) 2726–2769. MR3531679

[10] X. Chen. *Random Walk Intersections: Large Deviations and Related Topics*. *Mathematical Surveys and Monographs* **157**. American Mathematical Society, Providence, 2010. MR2584458

[11] R. Dudley. *Uniform Central Limit Theorems*. *Cambridge Studies in Advanced Mathematics* **63**. Cambridge University Press, Cambridge, 1999. MR1720712

[12] I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002) 5830–5847. MR1936554

[13] A. Edelman. Stochastic differential equations and random matrices. In *SIAM Conference on Applied Linear Algebra*. The College of William and Mary, Williamsburg, 2003. Available at http://math.mit.edu/~edelman/homepage/talks/siam2003.ppt.

[14] A. Edelman and B. Sutton. From random matrices to stochastic operators. *J. Stat. Phys.* **127** (2007) 1121–1165. MR2331033

[15] P. Forrester. *Log-Gases and Random Matrices*. *London Mathematical Society Monographs Series* **34**. Princeton University Press, Princeton, 2010. MR2641363

[16] V. Gorin and M. Shkolnikov. Stochastic Airy semigroup through tridiagonal matrices. *Ann. Probab.* **46** (2018) 2287–2344. Available at arXiv:1601.06800v1. MR3813993

[17] Y. Hariya. A pathwise interpretation of the Gorin–Shkolnikov identity. *Electron. Commun. Probab.* **21** (2016) 1–6. MR3533284

[18] D. Holcomb and G. R. M. Flores. Edge scaling of the β-Jacobi ensemble. *J. Stat. Phys.* **149** (2012) 1136–1160. MR3006323

[19] S. Jacquot and B. Valko. Bulk scaling limit of the Laguerre ensemble. *Electron. J. Probab.* **11** (2011) 314–346. MR2771139

[20] A. Javanmarda, A. Montanari and F. Ricci-Tersenghi. Phase transitions in semidefinite relaxations. *Proc. Natl. Acad. Sci. USA* **113** (2016) E2218–E2223. MR3494080

[21] R. L. Karandikar. On pathwise stochastic integration. *Stochastic Process. Appl.* **57** (1995) 11–18. MR1327950

[22] R. Kozhan. Rank one non-Hermitian perturbations of Hermitian β-ensembles of random matrices. *J. Stat. Phys.* **168** (2017) 92–108. MR3659979

[23] M. Krishnapur, B. Rider and B. Virág. Universality of the stochastic Airy operator. *Comm. Pure Appl. Math.* **69** (2016) 145–199. MR3433632

[24] G. Lawler and V. Limic. *Random Walk: A Modern Introduction*. *Cambridge Studies in Advanced Mathematics* **123**. Cambridge University Press, Cambridge, 2010. MR2677157

[25] P. Lax. *Functional Analysis*. John Wiley & Sons, New York, 2002. MR1892228

[26] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*, 2nd edition. The Clarendon Press, Oxford University Press, New York, 2015. MR3443860

[27] S. Péché. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* **134** (2006) 127–173. MR2221787

[28] J. Pitman. The distribution of local times of a Brownian bridge. In *Séminaire de Probabilités XXXIII* 388–394. *Lecture Notes in Math.* **1709**. Springer-Verlag, Berlin, 1999. MR1768012

[29] J. Pitman. The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest. *Ann. Probab.* **27** (1999) 261–283. MR1681110

[30] J. Pitman. *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer-Verlag, Berlin, 2006. MR2245368

[31] J. Ramírez and B. Rider. Diffusion at the random matrix hard edge. *Comm. Math. Phys.* **288** (2009) 887–906. MR2504858

[32] J. Ramírez, B. Rider and B. Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.* **24** (2011) 919–944. MR2813333

[33] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd edition. *Grundlehren der Mathematischen Wissenschaften* **293**. Springer-Verlag, Berlin, 1999. MR1725357

[34] B. Rider and P. Waters. Universality of the stochastic Bessel operator. Preprint, 2016. Available at arXiv:1610.01637.

[35] H. F. Trotter. A property of Brownian motion paths. *Illinois J. Math.* **2** (1958) 425–433. MR0096311

[36] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** (2009) 463–508. MR2534097

[37] B. Valkó and B. Virág. Operator limit of the circular beta ensemble, 2017. Available at arXiv:1710.06988v1.

[38] B. Valkó and B. Virág. The sine$_\beta$ operator. *Invent. Math.* **209** (2017) 275–327. MR3660310