

# OPERATOR LIMIT OF THE CIRCULAR BETA ENSEMBLE

BY BENEDEK VALKÓ<sup>1</sup> AND BÁLINT VIRÁG<sup>2</sup>

<sup>1</sup>*Department of Mathematics, University of Wisconsin—Madison, [valko@math.wisc.edu](mailto:valko@math.wisc.edu)*

<sup>2</sup>*Department of Mathematics, University of Toronto, [balint@math.toronto.edu](mailto:balint@math.toronto.edu)*

We provide a precise coupling of the finite circular beta ensembles and their limit process via their operator representations. We prove explicit bounds on the distance of the operators and the corresponding point processes. We also prove an estimate on the beta-dependence of the  $\text{Sine}_\beta$  process.

**1. Introduction.** Valkó and Virág (2017) introduced a family of differential operators parametrized by a path  $\gamma : [0, 1) \rightarrow \mathbb{H}$  in the upper half plane and two points on the boundary.

When the path is a certain hyperbolic random walk in the Poincaré half-plane model, the operator  $\text{Circ}_{\beta,n}$  has eigenvalues given by the points of the circular beta ensemble scaled and lifted periodically to the real line. With the path  $\gamma(t) = \mathcal{B}(-\frac{4}{\beta} \log(1-t))$  where  $\mathcal{B}$  is standard hyperbolic Brownian motion, the operator  $\text{Sine}_\beta$  has eigenvalues given by the  $\text{Sine}_\beta$  process, the limit of the circular beta ensembles. (See Theorems 7 and 8.)

The inverses of these operators in a compatible basis are integral operators denoted by  $\text{rCirc}_{\beta,n}$  and  $\text{rSine}_\beta$ , respectively. Our main result is a coupling which gives  $\text{Circ}_{\beta,n} \rightarrow \text{Sine}_\beta$  with an explicit rate of convergence. (See Section 2 for additional details, and Figure 1 for an illustration of the coupling.)

**THEOREM 1.** *There is a probability space with a standard hyperbolic Brownian motion  $\mathcal{B}$  and an array of stopping times  $0 = \tau_{n,n} < \tau_{n,n-1} < \dots < \tau_{n,0} = \infty$  so that  $\mathcal{B}(\tau_{n, \lceil(1-t)n\rceil})$ ,  $t \in [0, 1)$  has the law of the random walk on  $\mathbb{H}$  used to generate  $\text{Circ}_{\beta,n}$ . This provides a coupling of  $\text{Sine}_\beta$  and the sequence of operators  $\text{Circ}_{\beta,n}$ .*

*There exists an a.s. finite positive random variable  $N$  so that in this coupling*

$$(1) \quad \|\text{rSine}_\beta - \text{rCirc}_{\beta,n}\|_{\text{HS}}^2 \leq \frac{\log^6 n}{n}$$

*a.s. in the Hilbert–Schmidt norm for all  $n \geq N$ .*

As a corollary, we get new results about the rate of convergence of the eigenvalue processes. Let  $\lambda_k, k \in \mathbb{Z}$  be the ordered sequence of eigenvalues of  $\text{Sine}_\beta$  with  $\lambda_0 < 0 \leq \lambda_1$ , the sequence  $\lambda_{k,n}, k \in \mathbb{Z}$  is defined analogously for  $\text{Circ}_{\beta,n}$ .

**COROLLARY 2.** *In the coupling of Theorem 1, we have a.s.,*

$$(2) \quad \sum_{k \in \mathbb{Z}} (\lambda_k^{-1} - \lambda_{n,k}^{-1})^2 \leq \frac{\log^6 n}{n}$$

*for all  $n \geq N$ . (Here,  $N$  is the finite random variable from Theorem 1.) Moreover, as  $n \rightarrow \infty$  we have a.s.*

$$(3) \quad \max_{|k| \leq \frac{n^{1/4}}{\log^2 n}} |\lambda_k - \lambda_{k,n}| \rightarrow 0.$$

Received February 2018; revised July 2019.

*MSC2010 subject classifications.* Primary 60B20, 47B80; secondary 47E05.

*Key words and phrases.* Random matrices, random operators, coupling.

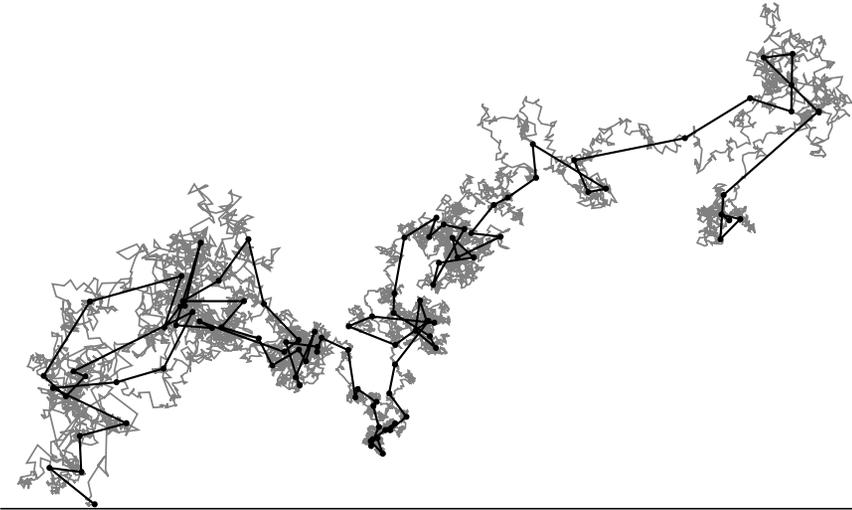


FIG. 1. Simulation of hyperbolic Brownian motion and a coupled random walk.

For all  $\varepsilon > 0$ , there is a random  $N_\varepsilon$  so that for  $n \geq N_\varepsilon$  and  $|k| \leq n^{1/2-\varepsilon}$  we have a.s.

$$(4) \quad |\lambda_k - \lambda_{k,n}| \leq \frac{1 + k^2}{n^{1/2-\varepsilon}}.$$

This provides the best known coupling of the circular  $\beta$ -ensemble to the  $\text{Sine}_\beta$  process, even for  $\beta = 2$ , when both processes are determinantal with explicitly given kernels. For  $\beta = 2$ , the bound (4) improves on a coupling given in [Maples, Najnudel and Nikeghbali \(2019\)](#) in which the inequality holds with the exponent  $1/3$  instead of  $1/2$ , for  $|k| \leq n^{1/4}$ .

Using the techniques introduced in our proof, we also give an estimate on the dependence on  $\beta$  for the  $\text{Sine}_\beta$  process.

**THEOREM 3.** *Construct the  $\text{Sine}_\beta$  operators for all  $\beta > 0$  with the same hyperbolic Brownian motion. Denote the eigenvalues corresponding to  $\beta$  by  $\{\lambda_{k,\beta}, k \in \mathbb{Z}\}$  with  $\lambda_{0,\beta} < 0 \leq \lambda_{1,\beta}$ . Then for  $0 < \theta$ , there is an a.s. finite  $C = C_\theta$  depending only on  $\theta$  and  $\mathcal{B}$  so that if  $\theta < \beta < \beta' \leq \infty$  and  $\delta = \frac{4}{\beta} - \frac{4}{\beta'} \leq 1/3$  then*

$$(5) \quad \sum_k \left( \frac{1}{\lambda_{k,\beta}} - \frac{1}{\lambda_{k,\beta'}} \right)^2 \leq \|\text{rSine}_\beta - \text{rSine}_{\beta'}\|_{\text{HS}}^2 \leq C\delta \log(\delta^{-1}).$$

The theorem allows the choice  $\beta' = \infty$ . In this case, the driving path  $\gamma(t)$  is just a constant, and the corresponding point process  $\text{Sine}_\infty$  is the so-called clock process: the set  $2\pi\mathbb{Z}$  shifted by a uniformly distributed random variable on  $[0, 2\pi]$ . The bound (5) in this case provides a quantitative description of the limit in distribution of  $\text{Sine}_\beta$  as  $\beta \rightarrow \infty$ .

The theorem requires a positive lower bound on  $\beta$ , hence it cannot describe the  $\beta \rightarrow 0$  behavior. It is known that in this case the limit in distribution of  $\text{Sine}_\beta$  is a homogeneous Poisson process; see [Allez and Dumaz \(2014\)](#).

*Structure of the paper.* The proof of [Theorem 1](#) relies on a precise coupling of a hyperbolic random walk and hyperbolic Brownian motion.

The starting point is a hyperbolic heat kernel bound. Consider the squared Euclidean norm for a hyperbolic Brownian motion started at the origin in the Poincaré disk model. We show

that this quantity at a small time is very close in total variation to a beta random variable, which also stochastically dominates it (Lemma 9 in Section 3.1).

This leads to a coupling of the hyperbolic random walk steps with hyperbolic Brownian increments at stopping times (Sections 3.2 and 3.3). For each single step, with high probability we stop at a fixed time. Otherwise, we just wait for the Brownian motion to hit the right distance. At the tail of the random walk, a slightly different coupling is implemented.

A modulus of continuity estimate for hyperbolic Brownian motion then implies that the random walk is close to the Brownian path (Section 3.4).

In Section 4, we show that if two paths are close and escape to the boundary of  $\mathbb{H}$  similarly as a geodesic then the corresponding operators are also close. Finally, in Section 5 we use the linear rate of escape for hyperbolic Brownian motion to complete the proof of Theorem 1. Section 6 proves Theorem 3. Some of the technical facts needed are collected in the Appendix.

*Historical background.* The modern history of random matrices originates from Wigner (1951), who used them to approximate the spectrum of self-adjoint operators from statistical physics point of view.

In the following decades, the scaling behavior of a number of random matrix models were derived. The point process limits of the random matrix spectra were described via the limiting joint densities, usually relying on some algebraic structure of the finite models. (See the monographs Mehta (2004), Anderson, Guionnet and Zeitouni (2010) and Forrester (2010) for an overview of the classical results.)

Dumitriu and Edelman (2002) constructed tridiagonal random matrix models with spectrum distributed as beta ensembles, one parameter extensions of classical random matrix models. Edelman and Sutton (2007) observed that under the appropriate scaling, these tridiagonal matrix models behave like approximate versions of random stochastic operators, and conjectured that scaling limits of beta ensembles can be described as the spectra of these objects.

These conjectures were confirmed in Ramírez, Rider and Virág (2011) and Ramírez and Rider (2009) for the soft and hard edge scaling limits of beta ensembles. The authors rigorously defined the stochastic differential operators that show up as limits, and proved the convergence of the finite ensembles to the spectrum of these operators.

In Valkó and Virág (2009) and Killip and Stoiciu (2009), the bulk scaling limit of the Gaussian and circular beta ensembles were derived, and the counting functions of the limit processes were characterized via coupled systems of SDEs. In Nakano (2014) and Valkó and Virág (2017), it was shown that the scaling limit of the circular beta ensemble is the same as  $\text{Sine}_\beta$ , the bulk limit of the Gaussian beta ensemble. Furthermore, Valkó and Virág (2017) constructed a stochastic differential operator with a spectrum given by  $\text{Sine}_\beta$  and showed that several random matrix limits can be described via differential operators parametrized by certain random walks or diffusions.

The coupling of the circular beta ensemble for  $\beta = 2$  (the circular unitary ensemble) to its limit, the  $\text{Sine}_2$  process has been recently studied in Bourgade, Najnudel and Nikeghbali (2013), Maples, Najnudel and Nikeghbali (2019) and Meckes and Meckes (2016). In Bourgade, Najnudel and Nikeghbali (2013), the circular unitary ensembles of various sizes are coupled together and it is shown that the scaled ensembles converge a.s. to a  $\text{Sine}_2$  process. Moreover, a bound of the form (4) is given with an exponent  $\varepsilon > 0$ . This coupling was further studied in Maples, Najnudel and Nikeghbali (2019) where a bound of the form (4) is given with an exponent  $1/3$ .

In Meckes and Meckes (2016), the total variation distance between the counting functions of the finite and the limiting process is considered. Denote by  $\mathcal{N}_n$  the counting function of

the appropriately scaled circular unitary ensemble of size  $n$ , and by  $\mathcal{N}$  the counting function of the Sine<sub>2</sub> process. It is shown that for any fixed interval  $I$  the following bound holds:

$$(6) \quad d_{\text{TV}}(\mathcal{N}_n(I), \mathcal{N}(I)) \leq 5 \frac{|I|^2}{n^{3/2}} \quad \text{for } n \geq n_0(I).$$

This provides a bound on the distance between the distributions for the number of points in a given interval, but does not seem to imply a process level result.

**2. Stochastic differential operators.** We review the framework introduced in Valkó and Virág (2017) to study random matrix ensembles via differential operators.

2.1. *Dirac operators.* We consider differential operators of the form

$$(7) \quad \tau : f \rightarrow R^{-1}(t)J \frac{d}{dt} f.$$

Here,  $f : [0, 1) \rightarrow \mathbb{R}^2$ , and

$$(8) \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad R = \frac{1}{2} X^t X, \quad X = \frac{1}{\sqrt{y}} \begin{pmatrix} 1 & -x \\ 0 & y \end{pmatrix},$$

with  $x : [0, 1) \rightarrow \mathbb{R}$  and  $y : [0, 1) \rightarrow (0, \infty)$ . We consider boundary conditions parametrized by nonzero vectors  $u_0, u_1 \in \mathbb{R}^2$  where we assume that  $u_0^t J u_1 = 1$ . We set the domain of the differential operator  $\tau$  as

$$(9) \quad \text{dom}(\tau) = \left\{ v \in L_R^2 \cap \text{AC} : \tau v \in L_R^2, v(0)^t J u_0 = 0, \lim_{s \rightarrow 1} v(s)^t J u_1 = 0 \right\}.$$

Here,  $L_R^2$  is the  $L^2$  space of functions  $f : [0, 1) \rightarrow \mathbb{R}^2$  with the  $L^2$  norm  $\|f\|_2^2 = \int_0^1 f^t R f ds$ , while AC is the set of absolutely continuous functions.

The function  $\gamma = x + iy$  is a path in the upper half-plane  $\{(x, y) : y > 0\}$ . In Valkó and Virág (2017), it was shown that various properties of  $\tau$  can be identified by treating  $\gamma$  as a path in the hyperbolic plane  $\mathbb{H}$  (using the upper half-plane representation) with  $u_0, u_1$  identified with boundary points  $\eta_0, \eta_1$  of  $\mathbb{H}$ . The set of boundary points of  $\mathbb{H}$  in the upper half-plane representation is  $\mathbb{R} \cup \{\infty\}$ . A nonzero vector  $v = (v_1, v_2)^t \in \mathbb{R}^2$  can be identified with the boundary point  $\mathcal{P}v \in \partial\mathbb{H}$  where  $\mathcal{P}v = \frac{v_1}{v_2}$  if  $v_2 \neq 0$  and  $\mathcal{P}v = \infty$  if  $v_2 = 0$ . To show the dependence on these parameters, we use the notation  $\tau = \text{Dir}(\gamma, \eta_0, \eta_1)$ .

For a given boundary point  $\eta \in \partial\mathbb{H}$ , the (signed) horocyclic distance of points  $a$  and  $b$  in  $\mathbb{H}$  with respect to  $\eta$  is defined as

$$d_\eta(a, b) = \lim_{z \rightarrow \eta} (d_{\mathbb{H}}(a, z) - d_{\mathbb{H}}(b, z)).$$

Here,  $d_{\mathbb{H}}$  is the hyperbolic distance and the limit is evaluated along a sequence of points in  $\mathbb{H}$  converging to  $\eta$ . We record the following formulas for the half-plane representation:

$$(10) \quad d_{\mathbb{H}}(x_1 + iy_1, x_2 + iy_2) = \text{arccosh} \left( 1 + \frac{(x_1 - x_2)^2 + (y_1 - y_2)^2}{2y_1 y_2} \right),$$

$$(11) \quad d_\eta(x + iy, i) = \begin{cases} \log\left(\frac{1}{y}\right) & \text{if } \eta = \infty, \\ \log\left(\frac{(x - q)^2 + y^2}{(1 + q^2)y}\right) & \text{if } \eta = q \in \mathbb{R}. \end{cases}$$

The following theorem gives a condition in terms of the parameters  $\gamma, \eta_0, \eta_1$  for  $\tau$  to be self-adjoint with a Hilbert–Schmidt inverse.

**THEOREM 4 (Valkó and Virág (2017)).** *Let  $\eta_0, \eta_1$  be distinct boundary points of  $\mathbb{H}$  and  $\gamma : [0, 1) \rightarrow \mathbb{H}$  be measurable and locally bounded. Assume that there is a  $\xi \in \mathbb{H}$  with*

$$(12) \quad \int_0^1 e^{d_{\eta_1}(\gamma(t), \xi)} dt < \infty \quad \text{and} \quad \int_0^1 \int_s^1 e^{d_{\eta_0}(\gamma(s), \xi) + d_{\eta_1}(\gamma(s), \xi)} ds dt < \infty.$$

*Then the operator  $\tau = \text{Dir}(\gamma, \eta_0, \eta_1)$  is self-adjoint on  $\text{dom}(\tau)$  and its inverse is Hilbert-Schmidt. The inverse  $\tau^{-1}$  is an integral operator on  $L^2_R$  with kernel function*

$$(13) \quad K(s, t) = (u_0 u_1^t \mathbf{1}(s < t) + u_1 u_0^t \mathbf{1}(s \geq t)).$$

*This means that if  $g \in L^2_R$  then  $(\tau^{-1}g)(x) = \int_0^1 K(x, y)R(y)g(y) dy$ .*

Suppose that  $\gamma, \eta_0, \eta_1$  satisfy the conditions of the theorem above and consider the operator  $\tau = \text{Dir}(\gamma, \eta_0, \eta_1)$ . Let  $\hat{\tau} = X\tau X^{-1}$ ; this means that  $(\hat{\tau}f)(x) = X(x)(\tau g)(x)$  where  $g(y) = X^{-1}(y)f(y)$ . Then  $\hat{\tau}$  is just  $\tau$  after a change of coordinates. In particular,  $\hat{\tau}$  is a self-adjoint differential operator on  $\{v : X^{-1}v \in \text{dom}(\tau)\} \subset L^2$ , with the same spectrum as  $\tau$ . We denote the inverse of  $\hat{\tau}$  by  $\mathfrak{r}\tau$  ( $\mathfrak{r}$  standing for resolvent). By Theorem 4, the operator  $\mathfrak{r}\tau$  is an integral operator acting on  $L^2$  functions with kernel

$$(14) \quad K_{\mathfrak{r}\tau}(s, t) = \frac{1}{2}(a(s)c(t)^t \mathbf{1}(s < t) + c(s)a(t)^t \mathbf{1}(s \geq t)),$$

where  $a(s) = X(s)u_0$  and  $c(s) = X(s)u_1$ . Thus for  $g \in L^2$  we have

$$(\mathfrak{r}\tau g)(x) = \int_0^1 K_{\mathfrak{r}\tau}(x, y)g(y) dy.$$

**2.2. Stochastic operators.** The Gaussian and circular  $\beta$  ensembles are defined via the following joint densities on  $\mathbb{R}^n$  and  $[0, 2\pi)^n$ , respectively,

$$(15) \quad p_{\beta,n}^g(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,\beta}^g} \prod_{1 \leq j < k \leq n} |\lambda_j - \lambda_k|^\beta e^{-\frac{\beta}{4} \sum_{j=1}^n \lambda_j^2},$$

$$(16) \quad p_{\beta,n}^c(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,\beta}^c} \prod_{1 \leq j < k \leq n} |e^{i\lambda_j} - e^{i\lambda_k}|^\beta.$$

For  $\beta = 2$ , these give the joint eigenvalue densities of the Gaussian and circular unitary ensemble. We use angles to represent the eigenvalues in the circular case.

The bulk scaling limits of these ensembles have been identified in Valkó and Virág (2009) and Killip and Stoiciu (2009). In Nakano (2014) and Valkó and Virág (2017), it was shown that the scaling limit of the circular beta ensemble is the same as the bulk limit of the Gaussian beta ensemble.

**THEOREM 5 (Valkó and Virág (2009)).** *Fix  $\beta > 0$  and  $|E| < 2$ . Let  $\Lambda_n^g$  be a finite point process with density (15). Then  $\sqrt{4 - E^2} \sqrt{n}(\Lambda_n^g - \sqrt{n}E)$  converges in distribution to a point process  $\text{Sine}_\beta$ .*

**THEOREM 6 (Killip and Stoiciu (2009), Nakano (2014), Valkó and Virág (2017)).** *Fix  $\beta > 0$  and let  $\Lambda_n^c$  be a finite point process with density (16). Then  $n\Lambda_n^c$  converges in distribution to the point process  $\text{Sine}_\beta$ .*

In Valkó and Virág (2017), the authors constructed random Dirac operators with spectrum given by  $\text{Sine}_\beta$  and the finite circular beta ensemble. Recall that the standard hyperbolic

Brownian motion  $x + iy$  in the upper half-plane representation started from  $i$  is the solution of the SDE

$$d(x + iy) = y(dB_1 + i dB_2), \quad x(0) + iy(0) = i,$$

where  $B_1, B_2$  are independent standard real Brownian motions.

**THEOREM 7 (Valkó and Virág (2017)).** Fix  $\beta > 0$  and let  $\mathcal{B}$  be a standard hyperbolic Brownian motion in the upper half-plane started from  $\mathcal{B}(0) = i$ . Set  $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta} \log(1 - t))$ ,  $t \in [0, 1)$ ,  $\eta_0 = \infty$  and  $\eta_1 = \lim_{t \rightarrow \infty} \mathcal{B}(t)$ . Then the operator

$$(17) \quad \text{Sine}_\beta = \text{Dir}(\tilde{\mathcal{B}}(t), \eta_0, \eta_1)$$

is a.s. self-adjoint with a Hilbert–Schmidt inverse and  $\text{spec}(\text{Sine}_\beta) \stackrel{d}{=} \text{Sine}_\beta$ .

The definition (17) can be extended to  $\beta = \infty$ . In this case, we set  $\tilde{\mathcal{B}}(t) = i$  for  $t \in [0, 1)$ . Then the corresponding operator is just  $2J \frac{d}{dt}$  on  $[0, 1)$  with boundary conditions  $u_0 = (-1, 0)^t$  and  $u_1 = (1, \eta_1)^t$ , with the same  $\eta_1$  as in Theorem 7. We call this operator  $\text{Sine}_\infty$ . A simple computation shows that  $\text{spec}(\text{Sine}_\infty) \stackrel{d}{=} 2\pi\mathbb{Z} + U$  where  $U$  is uniform on  $[0, 2\pi]$ . We denote this process  $\text{Sine}_\infty$ , this is sometimes referred to as the clock process.

Killip and Nenciu (2004) gave a construction for generating a random unitary matrix with eigenvalues distributed as the circular beta ensemble. Building on this result, Valkó and Virág (2017) produced a random Dirac operator representation for the circular beta ensemble using a random walk in  $\mathbb{H}$ .

Fix  $n \geq 1$  and let  $\zeta_0, \dots, \zeta_{n-2}$  be independent with  $\zeta_k$  distributed as  $\text{Beta}(1, \frac{\beta}{2}(n - k - 1))$ .<sup>1</sup> Set  $Y_k = \log(\frac{1 + \sqrt{\zeta_k}}{1 - \sqrt{\zeta_k}})$ . We define the random walk  $b_0, b_1, \dots, b_{n-1}$  in  $\mathbb{H}$  with a final boundary point  $b_n \in \partial\mathbb{H}$ . We set  $b_0 = i$ , and for  $0 \leq k \leq n - 2$  we choose  $b_{k+1}$  uniformly (according to the hyperbolic geometry) among the points in  $\mathbb{H}$  with hyperbolic distance  $Y_k$  from  $b_k$ , independently of the previous choices. The final point  $b_n$  is chosen uniformly on the boundary  $\partial\mathbb{H}$  as viewed from  $b_{n-1}$ , independently of the previous choices. Note that  $\zeta_k$  is the squared Euclidean norm of the random walk step in the Poincaré disk model with  $b_k$  at the origin.

**THEOREM 8 (Valkó and Virág (2017)).** Consider the random walk  $b_0, b_1, \dots, b_n$  defined above. Set  $\eta_0 = \infty$ ,  $\eta_1 = b_n$  and  $\mathcal{B}_n(t) = b_{\lfloor nt \rfloor}$  for  $t \in [0, 1)$ . The operator

$$\text{Circ}_{\beta,n} = \text{Dir}(\mathcal{B}_n(t), \eta_0, \eta_1)$$

is a.s. self-adjoint with a Hilbert–Schmidt inverse and  $\text{spec}(\text{Circ}_{\beta,n}) \stackrel{d}{=} n\Lambda_n^c + 2\pi n\mathbb{Z}$  where  $\Lambda_n^c$  is the finite point process with joint density (16).

Theorems 7 and 8 imply that zero is not an eigenvalue for the operators  $\text{Sine}_\beta$  and  $\text{Circ}_{\beta,n}$  with probability one.

**2.3. Coupling  $\text{Circ}_{\beta,n}$  and  $\text{Sine}_\beta$ .** Fix  $\beta > 0$ . Let  $\mathcal{B}$  be a hyperbolic Brownian motion,  $\eta_0 = \infty$ ,  $\eta_1 = \mathcal{B}(\infty)$ , and consider the operator  $\text{Sine}_\beta = \text{Dir}(\mathcal{B}(-\frac{4}{\beta} \log(1 - t)), \eta_0, \eta_1)$ .

Let  $U$  be a uniform random variable on  $[0, 1]$  independent from  $\mathcal{B}$ , and let  $\mathcal{F}_t, t \geq 0$  be the natural filtration of  $\mathcal{B}$  enlarged with  $U$ . In Proposition 12 below, we construct an

<sup>1</sup>A random variable has distribution  $\text{Beta}(a, b)$  with  $a, b > 0$  if it has density  $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbf{1}(x \in [0, 1])$ .

array of stopping times  $\tau_{n,k}$  with respect to  $\mathcal{F}_t$  so that for each  $n \geq 1$  the random variables  $\mathcal{B}(\tau_{n,n}), \mathcal{B}(\tau_{n,n-1}), \dots, \mathcal{B}(\tau_{n,0})$  have the same joint distribution as the random walk  $b_0, \dots, b_n$  from Theorem 8. (Note that  $\mathcal{B}(\tau_{n,k})$  corresponds to  $b_{n-k}$ , so the index  $k$  matches up with the parameter of the Beta distribution in the appropriate step of the random walk.) Setting  $\text{Cir}_{\mathcal{C}\beta,n} = \text{Dir}(\tilde{\mathcal{B}}_n(t), \eta_0, \eta_1)$  with  $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n, \lceil(1-t)n\rceil})$ ,  $t \in [0, 1)$  gives the appropriate coupling of  $\text{Sine}_{\beta}$  and the sequence  $\{\text{Cir}_{\mathcal{C}\beta,n}\}_{n \geq 1}$  that is used in Theorem 1.

**3. Coupling construction.** The goal of this section is to construct the coupling of the hyperbolic random walk and the hyperbolic Brownian motion that appears in Theorem 1. Since the steps in the random walk have rotationally invariant distributions and the same is true for the increments of the hyperbolic Brownian motion, it would be easy to embed the walk via simple hitting times. However, this “naive” embedding would not give enough control for us to obtain the error bound in Theorem 1. Instead we construct a coupling that also exploits the fact that the single step hyperbolic distance distributions in the random walk can be well approximated with the distance distribution of the hyperbolic Brownian motion at a certain fixed time.

3.1. *A heat kernel bound on the hyperbolic plane.* Our coupling relies on a careful estimate of the transition density of hyperbolic Brownian motion. Although there are a number of similar bounds in the literature (see, e.g., Davies and Mandouvalos (1988)), we could not find one that would be strong enough for our purposes. We show that the distribution of the distance of hyperbolic Brownian motion from its starting point at time  $t \leq 1$  can be well approximated estimated using a  $\text{Beta}(1, \frac{2}{t} - 1/2)$  random variable.

LEMMA 9. *Let  $\mathcal{B}(t)$  be standard hyperbolic Brownian motion and let  $t \in (0, 1]$ . Let  $Y = \log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$  where  $\xi$  has distribution  $\text{Beta}(1, \frac{2}{t} - 1/2)$  and set  $\zeta = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ . Then the following statements hold:*

- (a)  $P(Y > r) \geq P(\zeta > r)$  for all  $r > 0$ , in other words  $Y$  stochastically dominates  $\zeta$ .
- (b) The total variation distance of  $\zeta$  and  $Y$  is bounded by  $\frac{3}{2}t$ .

The proof of the lemma relies on a precise analysis of the explicit formula for the transition density. We leave it for Section A.1 in the Appendix.

3.2. *Single step coupling.* We first concentrate on a single step in the hyperbolic random walk corresponding to  $\text{Cir}_{\mathcal{C}\beta,n}$  and couple it to the hyperbolic Brownian motion.

PROPOSITION 10. *Fix  $\gamma \geq 3/2$ . Let  $\mathcal{B}$  be hyperbolic Brownian motion and  $U$  an independent uniform random variable on  $[0, 1]$ . Let  $\mathcal{F}_t, t \geq 0$  be the filtration of  $\mathcal{B}$  enlarged with  $U$ . Consider a Poincaré disk representation of the hyperbolic plane where  $\mathcal{B}(0) = 0$ .*

*There exists a finite random variable  $\sigma > 0$  so that the following hold:*

- 1.  $\sigma$  is a stopping time with respect to  $\mathcal{F}_t$ .
- 2.  $\mathcal{B}(\sigma)$  has rotationally invariant distribution and  $|\mathcal{B}(\sigma)|^2$  has  $\text{Beta}(1, \gamma)$  distribution.
- 3.  $P(\sigma \geq \frac{4}{2\gamma+1}) = 1$  and  $P(\sigma \neq \frac{4}{2\gamma+1}) \leq \frac{3}{\gamma}$ .
- 4. For  $r \geq 8$ , we have  $P(\sigma > r/\gamma) \leq 3e^{-\frac{1}{3}r^{1/3}}$ .

The proof of the proposition will rely on Lemma 9 and the Lemma 11 below. Lemma 11 is a standard coupling statement, we include its proof in the Appendix for completeness.

LEMMA 11. Assume that  $X_1$  and  $X_2$  are random variables so that  $X_2$  stochastically dominates  $X_1$  and the total variation distance of their distributions is  $\varepsilon$ . Then there exists a measurable function  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  so that if  $U$  is a uniform random variable on  $[0, 1]$ , independent of  $X_1$  then the following hold:

- (a)  $g(X_1, U)$  has the same distribution as  $X_2$ .
- (b)  $P(X_1 \leq g(X_1, U)) = 1$ .
- (c)  $P(X_1 = g(X_1, U)) = 1 - \varepsilon$ .

PROOF OF PROPOSITION 10. Set  $t = \frac{4}{2\gamma+1}$ , then  $0 < t < 1$ . Recall that if  $z$  is in the Poincaré disk with  $|z| = r < 1$  then  $d_{\mathbb{H}}(0, z) = \log(\frac{1+r}{1-r})$ .

Let  $\xi$  be a random variable with distribution  $\text{Beta}(1, \gamma)$ . Then by Lemma 9 the random variable  $\log(\frac{1+|\mathcal{B}(t)|}{1-|\mathcal{B}(t)|})$  is stochastically dominated by  $\log(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}})$  and their total variation distance is bounded by  $\frac{3}{2}t$ . Since  $\log(\frac{1+r}{1-r})$  is strictly increasing in  $r$ , we get that  $|\mathcal{B}(t)|$  is stochastically dominated by  $\sqrt{\xi}$  and their total variation distance is bounded by  $\frac{3}{2}t$ .

By Lemma 11, there exists a measurable function  $g$  so that almost surely  $g(|\mathcal{B}(t)|, U) \geq |\mathcal{B}(t)|$ ,  $g(|\mathcal{B}(t)|, U)$  has the same distribution as  $\sqrt{\xi}$ , and  $P(|\mathcal{B}(t)| \neq g(|\mathcal{B}(t)|, U)) \leq \frac{3}{2}t \leq 3/\gamma$ .

We set

$$\sigma = \inf\{s \geq t : |\mathcal{B}(s)| = g(|\mathcal{B}(t)|, U)\}.$$

Then  $\sigma$  is an a.s. finite stopping time with respect to  $\mathcal{F}_t$  and almost surely  $\sigma \geq t$ . Because  $\sigma$  only depends on  $|\mathcal{B}|$  and  $U$ , it follows that  $\mathcal{B}(\sigma)$  has rotationally invariant distribution. Finally, from our construction we get that  $|\mathcal{B}(\sigma)|^2 = g(|\mathcal{B}(t)|, U)^2$  has  $\text{Beta}(1, \gamma)$  distribution and  $P(\sigma \neq t) \leq \frac{3}{2}t \leq 3/\gamma$ .

The only thing left to prove is the tail bound for  $\sigma$ . We start with the bound

$$(18) \quad \begin{aligned} P(\sigma > r/\gamma) &\leq P\left(\sigma > r/\gamma \text{ and } d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) \leq \frac{r^{1/3}}{\gamma^{1/2}}\right) \\ &\quad + P\left(d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) > \frac{r^{1/3}}{\gamma^{1/2}}\right). \end{aligned}$$

For the rest of the proof, we assume  $r \geq 8$ . Then  $\sigma > r/\gamma > t$  and from the definition of  $\sigma$ , it follows that  $d_{\mathbb{H}}(0, \mathcal{B}(s)) < d_{\mathbb{H}}(0, \mathcal{B}(\sigma))$  for  $t \leq s < \sigma$ . Thus we can bound the first term on the right of (18) by writing

$$\begin{aligned} &P\left(\sigma > r/\gamma \text{ and } d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) \leq \frac{r^{1/3}}{\gamma^{1/2}}\right) \\ &\leq P\left(\max_{t \leq s \leq r/\gamma} d_{\mathbb{H}}(0, \mathcal{B}(s)) \leq \frac{r^{1/3}}{\gamma^{1/2}}\right) \\ &\leq P\left(\max_{t \leq s \leq r/\gamma} d_{\mathbb{H}}(\mathcal{B}(t), \mathcal{B}(s)) \leq 2\frac{r^{1/3}}{\gamma^{1/2}}\right) \\ &= P\left(\max_{0 \leq s \leq r/\gamma-t} d_{\mathbb{H}}(0, \mathcal{B}(s)) \leq 2\frac{r^{1/3}}{\gamma^{1/2}}\right). \end{aligned}$$

Since  $t \leq \frac{3}{2\gamma}$  and  $r \geq 8$ , we have  $r/\gamma - t \geq \frac{4}{5}r\gamma^{-1}$ . Using the bound (67) from Lemma 20 of the Appendix, we get

$$P\left(\max_{0 \leq s \leq r/\gamma - t} d_{\mathbb{H}}(0, \mathcal{B}(s)) \leq 2\frac{r^{1/3}}{\gamma^{1/2}}\right) \leq P\left(\max_{0 \leq s \leq \frac{4}{5}r\gamma^{-1}} d_{\mathbb{H}}(0, \mathcal{B}(s)) \leq 2\frac{r^{1/3}}{\gamma^{1/2}}\right) \leq \frac{4}{\pi}e^{-\frac{\pi^2 r^{1/3}}{40}}.$$

For the second term in (18), we recall that by construction  $d_{\mathbb{H}}(0, \mathcal{B}(\sigma))$  has the same distribution as  $\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right)$  where  $\xi$  has distribution Beta(1,  $\gamma$ ). By an explicit computation,

$$P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) = \operatorname{sech}^{2\gamma}\left(\frac{u}{2}\right).$$

We have  $\operatorname{sech}(x) = \frac{2}{e^x + e^{-x}} \leq 2e^{-x}$  for  $x > 0$ , which gives the following upper bound for  $u \geq 4 \log 2$ :

$$P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) \leq 2^{2\gamma}e^{-\gamma u} \leq e^{-\frac{\gamma u}{2}}.$$

For  $0 \leq u \leq 4 \log 2$ , we have  $\log \operatorname{sech}(u/2) \leq -\frac{u^2}{12}$  so for these values we get

$$(19) \quad P\left(\log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right) > u\right) \leq e^{-\frac{\gamma u^2}{6}}.$$

From this, we get

$$P\left(d_{\mathbb{H}}(0, \mathcal{B}(\sigma)) > \frac{r^{1/3}}{\gamma^{1/2}}\right) \leq \max\left(e^{-\frac{\sqrt{\gamma}r^{1/3}}{2}}, e^{-\frac{r^{2/3}}{6}}\right) \leq e^{-\frac{1}{3}r^{1/3}},$$

where in the last step we used  $\gamma \geq 3/2$  and  $r \geq 8$ . Collecting our estimates, we get

$$P(\sigma > r/\gamma) \leq \frac{4}{\pi}e^{-\frac{\pi^2 r^{1/3}}{40}} + e^{-\frac{1}{3}r^{1/3}} \leq 3e^{-\frac{1}{5}r^{1/3}},$$

which completes the proof of the proposition.  $\square$

3.3. *Path coupling.* Using Proposition 10 repeatedly, we can provide a coupling of the hyperbolic Brownian motion and the hyperbolic random walk appearing in the construction of  $\operatorname{Circ}_{\beta, n}$ .

PROPOSITION 12. *Let  $\mathcal{B}$  be hyperbolic Brownian motion,  $U_k, k \geq 1$  Uniform[0, 1] random variables, and  $\xi_k, k \geq 1$  random variables with distribution Beta(1,  $\frac{\beta}{2}k$ ), with  $\mathcal{B}, U_1, U_2, \dots, \xi_1, \xi_2, \dots$  all independent. Let  $\mathcal{F}_t, t \geq 0$  be the filtration of  $\mathcal{B}$  enlarged with the random variables  $U_k, \xi_k, k \geq 1$ .*

*There exists a collection of stopping times  $(\tau_{n,k}; 1 \leq n, 0 \leq k \leq n)$  with respect to  $\mathcal{F}_t, t \geq 0$  so that the following statements hold:*

1. *For each fixed  $n$ , we have  $0 = \tau_{n,n} < \tau_{n,n-1} < \dots < \tau_{n,0} = \infty$ , and the random variables  $\Delta\tau_{n,k} := \tau_{n,k} - \tau_{n,k+1}$  are independent for  $k = 1, 2, \dots, n - 1$ .*

2. For each  $n$ , the process  $(b_k^{(n)} = \mathcal{B}(\tau_{n,n-k}), k = 0, 1, \dots, n)$  is a hyperbolic random walk with the same distribution as the one given above Theorem 8.

3. Let  $t_{n,k} = \frac{4}{\beta} \log(\frac{n}{k})$  for  $1 \leq k \leq n$ . There exists a random integer  $N_0 > 0$  so that for  $n \geq N_0$  and  $\log^6 n \leq k \leq n$  we have almost surely

$$(20) \quad t_{n,k} - \frac{4}{\beta^2 k} \leq \tau_{n,k} \leq t_{n,k} + \frac{\log^{4+1/2} n}{k}.$$

4. For  $k$  fixed the hyperbolic distance  $d_{\mathbb{H}}(\mathcal{B}(\tau_{n,k}), \mathcal{B}(\tau_{n,k-1}))$  does not depend on  $n$  as long as  $k < \log^6 n$ .

PROOF. We first give the construction of the stopping times, then prove that they satisfy all the conditions.

For a fixed  $n \geq 1$ , we define  $\tau_{n,k}$  recursively, starting with  $\tau_{n,n} = 0$ . If for a certain  $k < n$ , we have already defined  $\tau_{n,k+1}$  then we define  $\tau_{n,k}$  as follows.

- In the case of  $k \geq \max(\log^6 n, \frac{3}{\beta})$ :

We apply Proposition 10 with  $\gamma = \frac{\beta}{2}k$  for the hyperbolic Brownian motion  $\tilde{\mathcal{B}}(t) = \mathcal{B}(t + \tau_{n,k+1}) - \mathcal{B}(\tau_{n,k+1}), t \geq 0$  and the independent uniform random variable  $U_k$ , and denote the constructed stopping time by  $\sigma_{n,k}$ . We set  $\tau_{n,k} = \tau_{n,k+1} + \sigma_{n,k}$ .

- In the case of  $1 \leq k < \max(\log^6 n, \frac{3}{\beta})$ :

We set

$$\tau_{n,k} = \inf \left\{ t \geq \tau_{n,k+1} : d_{\mathbb{H}}(\mathcal{B}(\tau_{n,k+1}), \mathcal{B}(t)) = \log \left( \frac{1 + \sqrt{\xi_k}}{1 - \sqrt{\xi_k}} \right) \right\}.$$

- For  $k = 0$ , we define  $\tau_{n,k} = \infty$ .

Note that we use the coupling given in Proposition 10 when  $k$  is not too small compared to  $n$ . This will enable us to prove the estimate (20), and allows us to control the distance between  $\mathcal{B}(t_{n,k})$  and  $b_k^{(n)}$  (see Proposition 13 below) in this regime. The bounds in Proposition 10 are not strong enough to use this approach for all  $k$ , that is why we need to use a different coupling construction for small values of  $k$ .

By construction, the random variables  $\tau_{n,k}, 1 \leq k \leq n$  are a.s. finite stopping times with respect to the filtration  $\mathcal{F}_t, t \geq 0$ , and they satisfy conditions 1, 2 and 4. To check Condition 3, we first choose  $n_0$  so that for  $n \geq n_0$  we have  $\log^6 n \geq \frac{3}{\beta}$  and  $\log^{3+3/8} n \geq 8$ . During the rest of this proof, we will assume  $1 \leq \log^6 n \leq k \leq n$ .

For  $n \geq n_0$ , by Proposition 10 we have almost surely

$$\tau_{n,k} \geq \sum_{j=k}^{n-1} \frac{4}{\beta j + 1} \geq \frac{4}{\beta} \log \left( \frac{n}{k} \right) - \frac{4}{\beta^2 k}.$$

This takes care of the lower bound in (20).

For the upper bound, recall the definition of  $\sigma_{n,k} = \tau_{n,k} - \tau_{n,k+1}$ . From Proposition 10, we have the following estimates:

$$(21) \quad \begin{aligned} P \left( \sigma_{n,k} \geq \frac{4}{\beta k + 1} \right) &= 1, & P \left( \sigma_{n,k} = \frac{4}{\beta k + 1} \right) &\geq 1 - \frac{6}{\beta k}, \\ P \left( \sigma_{n,k} > \frac{\log^{3+3/8} n}{\frac{\beta}{2} k} \right) &\leq 3e^{-\frac{1}{5} \log^{1+1/8} n}. \end{aligned}$$

Since  $\sum_{k,n} P(\sigma_{n,k} > \frac{\log^{3+3/8} n}{\frac{\beta}{2}k}) \leq 3 \sum_n n e^{-\frac{1}{5} \log^{1+1/8} n} < \infty$ , by the Borel–Cantelli lemma there is a random  $N_1 \geq n_0$  so that for  $n \geq N_1$  we have  $\sigma_{n,k} \leq \frac{\log^{3+3/8} n}{\frac{\beta}{2}k}$  a.s. Set

$$(22) \quad Z_{n,k} = \frac{4}{\beta k + 1} \mathbf{1}\left(\sigma_{n,k} = \frac{4}{\beta k + 1}\right) + \frac{2 \log^{3+3/8} n}{\beta k} \mathbf{1}\left(\sigma_{n,k} \neq \frac{4}{\beta k + 1}\right).$$

For  $n \geq N_1$ , we have  $\sigma_{n,k} \leq Z_{n,k}$  and  $\tau_{n,k} \leq \sum_{j=k}^{n-1} Z_{n,j}$  a.s.

Next, we will bound  $P(\sum_{j=k}^{n-1} (Z_{n,j} - \frac{4}{\beta j + 1}) \geq \frac{\log^{9/2} n}{k})$ . For  $\lambda > 0$  from (21) and (22), we get

$$(23) \quad E(e^{\sum_{j=k}^{n-1} \lambda(Z_{n,j} - \frac{4}{\beta k + 1})}) \leq \prod_{j=k}^{n-1} \left(1 - \frac{6}{\beta j} + \frac{6}{\beta j} e^{\lambda \frac{2 \log^{3+3/8} n}{\beta j}}\right).$$

Assuming that  $\lambda \frac{2 \log^{3+3/8} n}{\beta k} \leq 1$ , we can use that  $\frac{e^x - 1}{x} \leq 2$  for  $x \leq 1$  to bound the right-hand side of (23) as

$$\begin{aligned} \prod_{j=k}^{n-1} \left(1 - \frac{6}{\beta j} + \frac{6}{\beta j} e^{\lambda \frac{2 \log^{3+3/8} n}{\beta j}}\right) &\leq \prod_{j=k}^{n-1} \left(1 + \frac{6}{\beta j} \cdot 2\lambda \frac{2 \log^{3+3/8} n}{\beta j}\right) \\ &\leq \prod_{j=k}^{n-1} e^{\frac{24\lambda \log^{3+3/8} n}{\beta^2 j^2}} \\ &\leq e^{\frac{48\lambda \log^{3+3/8} n}{\beta^2 k}}. \end{aligned}$$

Setting now  $\lambda = \frac{\beta k}{2 \log^{3+3/8} n}$  and using the exponential Markov inequality, we obtain

$$\begin{aligned} P\left(\sum_{j=k}^{n-1} \left(Z_{n,j} - \frac{4}{\beta j + 1}\right) \geq \frac{\log^{4+1/2} n}{k}\right) \\ = E(e^{\sum_{j=k}^{n-1} \lambda(Z_{n,j} - \frac{4}{\beta k + 1})}) e^{-\lambda \frac{\log^{4+1/2} n}{k}} \\ \leq e^{\frac{48\lambda \log^{3+3/8} n}{\beta^2 k} - \lambda \frac{\log^{4+1/2} n}{k}} \leq e^{\frac{24}{\beta} - \frac{\beta}{2} \log^{1+1/8} n}. \end{aligned}$$

Since  $\sum_n n e^{\frac{24}{\beta} - \frac{\beta}{2} \log^{1+1/8} n} < \infty$ , the Borel–Cantelli lemma implies that there is a random  $N_0 \geq N_1$  so that for  $n \geq N_0$  and  $\log^6 n \leq k \leq n$  we have a.s.,

$$\sum_{j=k}^{n-1} Z_{n,j} \leq \sum_{j=k}^{n-1} \frac{4}{\beta j + 1} + \frac{\log^{4+1/2} n}{k} \leq \frac{4}{\beta} \log\left(\frac{n}{k}\right) + \frac{\log^{4+1/2} n}{k}.$$

Since  $\tau_{n,k} \leq \sum_{j=k}^{n-1} Z_{n,j}$ , the upper bound in (20) follows.  $\square$

3.4. *Path comparison.* Let  $\mathcal{B}$  be a hyperbolic Brownian motion and consider the stopping times  $\tau_{n,k}$  constructed in Proposition 12. Set  $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n, \lceil (1-t)n \rceil})$  and  $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta} \log(1-t))$ . The next proposition gives uniform bounds on  $d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t))$ . The estimates rely on path properties of the hyperbolic Brownian motion which are stated in Propositions 22 and 23, and proved in the Appendix.

PROPOSITION 13. *There is a random integer  $N^*$  so that for  $n \geq N^*$  we have the following a.s. inequalities:*

$$(24) \quad d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)) \leq \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}} \quad \text{if } 0 \leq t \leq 1 - \frac{\log^6 n}{n},$$

$$(25) \quad d_{\mathbb{H}}\left(\tilde{\mathcal{B}}_n\left(1 - \frac{\log^6 n}{n}\right), \tilde{\mathcal{B}}_n(t)\right) \leq (\log \log n)^4 \quad \text{if } 1 - \frac{\log^6 n}{n} \leq t < 1.$$

PROOF. Let  $T_n = 1 - \frac{\log^6 n}{n}$ . Consider  $N_0$  from the statement of Proposition 12. For  $n \geq N_0$  and  $0 \leq t \leq T_n$ , we have a.s.,

$$(26) \quad \begin{aligned} & \left| \tau_{n, \lceil (1-t)n \rceil} - \frac{4}{\beta} \log\left(\frac{1}{1-t}\right) \right| \\ & \leq \frac{4}{\beta^2 \lceil (1-t)n \rceil} + \frac{\log^{9/2} n}{\lceil (1-t)n \rceil} + \frac{4}{\beta} \left| \log\left(\frac{n}{\lceil (1-t)n \rceil}\right) - \log\left(\frac{1}{1-t}\right) \right| \\ & \leq \frac{\log^{4+5/8} n}{(1-t)n}, \end{aligned}$$

where for the second inequality we also assume  $n \geq n_0$  with an  $n_0$  only depending on  $\beta$ . Inequality (26) implies

$$d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)) \leq \max\{d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+u)) : |u| \leq h, 0 \leq s+u\},$$

with  $h = \frac{\log^{4+5/8} n}{(1-t)n}$  and  $s = \frac{4}{\beta} \log\left(\frac{1}{1-t}\right)$ . Note that  $h \leq \frac{\log^{4+5/8} n}{\log^6 n} \leq \frac{1}{\log n}$ .

Consider the random constant  $h_0$  from the statement of Proposition 22 of the Appendix. If  $n \geq e^{h_0^{-1}}$ , then  $h \leq h_0$  and we may apply Proposition 22 with  $s, s+u$  if  $0 \leq u \leq h$  and with  $s+u, s$  if  $0 \leq -u \leq \min(h, s)$ . Using the fact that  $h \log\left(2 + \frac{s+1}{h}\right)$  is monotone increasing in  $h$  and  $s$ , we get

$$\begin{aligned} & \max\{d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+u)) : |u| \leq h, 0 \leq s+u\} \\ & \leq 20 \sqrt{h \log\left(2 + \frac{s+1}{h}\right)} \leq \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}}, \end{aligned}$$

if  $0 \leq t \leq T_n$  and  $n \geq N_1$ , with a random integer  $N_1$ .

Next, we prove the estimate for the  $T_n \leq t < 1$  case. Recall the construction of the stopping times  $\tau_{n,k}$  from the proof of Proposition 12. From the construction, it follows that

$$(27) \quad \begin{aligned} & d_{\mathbb{H}}\left(\tilde{\mathcal{B}}_n\left(1 - \frac{\log^6 n}{n}\right), \tilde{\mathcal{B}}_n(t)\right) \\ & \leq \max\{d_{\mathbb{H}}(\mathcal{B}(\tau_{n, \lceil (1-T_n)n \rceil}), \mathcal{B}(s)) : \tau_{n, \lceil (1-T_n)n \rceil} \leq s \leq \tau_{n,1}\}. \end{aligned}$$

From Proposition 22 (and the comment after it), we get that there is a random constant  $C_B$  depending only on  $\mathcal{B}$  so that

$$(28) \quad \begin{aligned} & \max\{d_{\mathbb{H}}(\mathcal{B}(\tau_{n, \lceil (1-T_n)n \rceil}), \mathcal{B}(s)) : \tau_{n, \lceil (1-T_n)n \rceil} \leq s \leq \tau_{n,1}\} \\ & \leq C_B (\tau_{n,1} - \tau_{n, \lceil (1-T_n)n \rceil}) \sqrt{\log(3 + \tau_{n,1})}. \end{aligned}$$

We will show that there is a random constant  $0 < C < \infty$  (not depending on  $n$ ) so that a.s. for all  $n$

$$(29) \quad \tau_{n,1} - \tau_{n, \lceil (1-T_n)n \rceil} \leq C (\log \log n)^3, \quad \tau_{n,1} \leq C \log n.$$

This bound together with (27) and (28) implies the estimate (25) for  $n > N^*$ , where  $N^*$  is random.

Consider the Beta distributed random variables  $\xi_j$  used in the construction in the proof of Proposition 12. Setting  $Y_j = \log\left(\frac{1+\sqrt{\xi_j}}{1-\sqrt{\xi_j}}\right)$  and using the tail bound (19), we get the following bound for  $0 < r < 4 \log 2$ :

$$P(Y_j \geq r) = \operatorname{sech}^{\beta j}(r/2) \leq e^{-\frac{\beta jr^2}{12}}.$$

This implies  $\sum_j P(Y_j \geq \frac{4}{\sqrt{\beta}}\sqrt{\frac{\log 2j}{j}}) < \infty$ , and shows that there is a random constant  $C_\xi < \infty$  depending only on the sequence  $\{\xi_n\}$ , so that a.s.,

$$(30) \quad Y_j < C_\xi \sqrt{\frac{\log 2j}{j}}.$$

We will prove that there are random constants  $A$  and  $N_1$  so that for all  $n > N_1$  and  $1 \leq k < n(1 - T_n)$  we have a.s.,

$$(31) \quad \tau_{n,k} - \tau_{n,k+1} \leq A \frac{\log(2k) \log \log n}{k}.$$

The bound (26) applied for  $t = T_n$  shows that for  $n \geq N_0$ ,

$$(32) \quad \tau_{n, \lceil (1-T_n)n \rceil} \leq \frac{4}{\beta} \log n.$$

From (31) and (32), the bounds in (29) follow directly.

Let  $C_0$  be the random constant from the statement of Proposition 23. Choose  $A > 0$  so that  $\frac{C_0 A}{2} > C_\xi^2$  where  $C_\xi$  is the constant from (30), and set  $h_{n,k} = A \frac{\log(2k) \log \log n}{k}$ . For  $1 \leq k < n(1 - T_n)$ , we have

$$A \frac{\log \log n}{\log^6 n} \leq h_{k,n} \leq A \frac{\log(2 \log^6 n) \log \log n}{k}$$

for  $n \geq 3$ . From this, it follows that there is a constant  $N_2 \geq 3$  (depending only on  $C_0$  and  $C_\xi$ ) so that the following inequality holds for all  $N_2 \leq n$  and  $1 \leq k < n(1 - T_n) = \log^6 n$ :

$$(33) \quad \frac{C_0}{4} \frac{h_{n,k}}{\log(2 + \frac{8}{\beta} \log n + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})} \geq C_\xi^2 \frac{\log 2k}{k}.$$

Consider  $1 \leq k < n(1 - T_n)$  and the stopping time  $\tau_{n,k+1}$ . By (74) of Proposition 23, there exists  $0 < u < h_{k,n}$  so that

$$\begin{aligned} & d_{\mathbb{H}}(\mathcal{B}(\tau_{n,k+1}), \mathcal{B}(\tau_{n,k+1} + u))^2 \\ & \geq \frac{C_0}{4} \frac{h_{n,k}}{\log(2 + \tau_{n,k+1} + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})}. \end{aligned}$$

If

$$(34) \quad \frac{C}{4} \frac{h_{n,k}}{\log(2 + \tau_{n,k+1} + h_{n,k}) + \log(h_{n,k} + h_{n,k}^{-1})} \geq C_\xi^2 \frac{\log 2k}{k}$$

then by (30) and the construction of the stopping times  $\tau_{n,k}$  we had  $\tau_{n,k} \leq \tau_{n,k+1} + h_{n,k}$ . The inequality (34) holds for  $k + 1 = \lceil n(1 - T_n) \rceil$  if  $n \geq N_2$  by (32) and (33). From this (34), and hence  $\tau_{n,k} \leq \tau_{n,k+1} + h_{n,k}$  follows for all  $1 \leq k < n(1 - T_n)$  by induction as long as  $\sum_{k=1}^{n(1-T_n)} h_{k,n} \leq \frac{4}{\beta} \log n$ , which holds for  $n$  large enough. This completes the proof of (31) and also that of (25).  $\square$

**4. Hilbert–Schmidt bounds.** This section contains general bounds on Dirac operators in the case when the corresponding hyperbolic paths escape to the boundary of  $\mathbb{H}$  with a positive speed. Proposition 14 below shows that such a Dirac operators has a Hilbert–Schmidt inverse. The next result, Proposition 15 below, compares such a Dirac operator to its truncated version. Finally, Proposition 16 below compares two Dirac operators if their driving paths are close.

These propositions will be used in the next section to prove Theorems 1 and 3. In particular, for Theorem 1 we will bound  $\|rSine_\beta - rCirc_{\beta,n}\|_{HS}^2$  by replacing each integral operator with its truncated version using Proposition 15, and then use the path comparison in Proposition 13 together with Proposition 16 to estimate the norm difference of the truncated operators.

**PROPOSITION 14 (Hilbert–Schmidt property).** *Let  $\{\gamma(t), 0 \leq t\}$  be a measurable path in  $\mathbb{H}$ ,  $\eta_0, \eta_1 \in \partial\mathbb{H}$  distinct boundary points and  $z_0 \in \mathbb{H}$ . For a  $v > 0$ , set  $\tilde{\gamma}(t) = \gamma(v \log(\frac{1}{1-t}))$ . Let  $z(t)$  be the point moving with speed  $\alpha > 0$  on the geodesic connecting  $z_0$  to  $\eta_1$  with  $z(0) = z_0$ . Assume that there are constants  $b > 0$  and  $0 \leq \varepsilon < v^{-1}$  so that for all  $t \geq 0$  we have*

$$(35) \quad d_{\mathbb{H}}(\gamma(t), z(t)) \leq b + \varepsilon t.$$

Then the operator

$$(36) \quad \tau = \text{Dir}(\tilde{\gamma}(t), \eta_0, \eta_1)$$

is self-adjoint on the appropriate domain and  $\tau^{-1}$  is Hilbert–Schmidt.

Here, the “appropriate domain” is described in and around (9).

**PROOF.** We will check that the conditions of Theorem 4 are satisfied.

If  $Q$  is an isometry of  $\mathbb{H}$ , then  $d_{\mathbb{H}}(z_1, z_2) = d_{\mathbb{H}}(Qz_1, Qz_2)$  and  $d_\eta(z_1, z_2) = d_{Q\eta}(Qz_1, Qz_2)$ . Take an isometry  $Q$  for which  $Qz_0 = i$  and  $Q\eta_1 = \infty$  and denote  $Q\eta_0$  by  $q$ . The geodesic  $z(t)$  is mapped into the geodesic connecting  $i$  with  $\infty$  with speed  $\alpha$ , thus  $Qz(t) = ie^{\alpha t}$ . From (11), it follows that

$$\begin{aligned} d_{\eta_1}(z(t), z_0) &= d_\infty(ie^{\alpha t}, i) = -\alpha t, \\ d_{\eta_0}(z(t), z_0) &= d_q(ie^{\alpha t}, i) = \alpha t + \log\left(\frac{q^2 e^{-2\alpha t} + 1}{1 + q^2}\right) \leq \alpha t. \end{aligned}$$

From the triangle inequality, we get

$$(37) \quad \begin{aligned} d_{\eta_1}(\gamma(t), z_0) &\leq d_{\eta_1}(z(t), z_0) + d_{\mathbb{H}}(\gamma(t), z(t)) \leq -(\alpha - \varepsilon)t + b, \\ d_{\eta_0}(\gamma(t), z_0) &\leq d_{\eta_0}(z(t), z_0) + d_{\mathbb{H}}(\gamma(t), z(t)) \leq (\alpha + \varepsilon)t + b. \end{aligned}$$

The bounds in (12) now follow easily:

$$\begin{aligned} \int_0^1 e^{d_{\eta_1}(\tilde{\gamma}(t), z_0)} dt &\leq \int_0^1 e^{-(\alpha - \varepsilon)v \log(\frac{1}{1-t}) + b} dt \\ &= e^b \int_0^1 (1-t)^{(\alpha - \varepsilon)v} dt = \frac{e^b}{1 + (\alpha - \varepsilon)v} < \infty \end{aligned}$$

and

$$\begin{aligned} & \int_0^1 \int_0^t e^{d_{\eta_0}(\tilde{\gamma}(s), z_0) + d_{\eta_1}(\tilde{\gamma}(t), z_0)} ds dt \\ & \leq e^{2b} \int_0^1 \int_0^t (1-s)^{-(\alpha+\varepsilon)\nu} (1-t)^{(\alpha-\varepsilon)\nu} ds dt \\ & = \frac{e^{2b}}{2(1+(\alpha-\varepsilon)\nu)(1-\varepsilon\nu)} < \infty. \end{aligned} \quad \square$$

Recall from (14) that  $\mathfrak{r}\tau$  is an integral operator with kernel  $K_{\mathfrak{r}\tau}$ . For  $0 < T < 1$ , we denote by  $\mathfrak{r}_T\tau$  the integral operator with kernel

$$K_{\mathfrak{r}_T\tau}(x, y) \cdot \mathbf{1}(0 \leq x \leq T, 0 \leq y \leq T).$$

PROPOSITION 15 (Hilbert–Schmidt truncation). *Let  $\gamma, \eta_0, \eta_1, z_0, \alpha, \nu, \tilde{\gamma}, z(t)$  be as in Proposition 14, and define  $\tau$  according to (36).*

1. *Assume that for some  $0 < b$  and  $0 \leq \kappa < 1$  the following inequality holds for all  $0 \leq t$ :*

$$(38) \quad d_{\mathbb{H}}(\gamma(t), z(t)) \leq b + t^\kappa,$$

and that for some  $c_0 < \infty$  we have

$$(39) \quad \alpha\nu \leq c_0.$$

Then for any  $T \in (0, 1)$ , we have

$$(40) \quad \|\mathfrak{r}\tau - \mathfrak{r}_T\tau\|_{\text{HS}}^2 \leq C(1-T)^{1+\frac{1}{2}\min(\alpha\nu, 1)} \left(1 + \log \frac{1}{1-T}\right)$$

with  $C$  depending only on  $\eta_0, \eta_1, b, \alpha, \kappa$  and  $c_0$  from (39).

2. *Assume that for some  $0 < T < 1$  (38) holds for  $0 \leq t \leq \nu \log(\frac{1}{1-T})$  with some  $0 < b, 0 \leq \kappa < 1$ . Assume further that  $d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}(T)) \leq M$  for  $t \geq T$  and that (39) holds. Then we have*

$$(41) \quad \|\mathfrak{r}\tau - \mathfrak{r}_T\tau\|_{\text{HS}}^2 \leq Ce^{2M}(1-T)^{1+\frac{1}{2}\min(\alpha\nu, 1)} \left(1 + \log \frac{1}{1-T}\right),$$

with  $C$  depending only on  $\eta_0, \eta_1, b, \alpha, \kappa$  and  $c_0$  from (39).

PROOF. We denote the representation of  $\gamma$  in the half-plane by  $x + iy$  and use  $\tilde{x} + i\tilde{y}$  for the representation of  $\tilde{\gamma}$ . We represent  $\eta_0, \eta_1$  with nonzero vectors  $u_0, u_1$  that satisfy  $u_0^t J u_1 = 1$ . Recall the integral kernel of  $\mathfrak{r}\tau$  from (14). From the definition of  $\mathfrak{r}_T\tau$ , we get

$$2\|\mathfrak{r}\tau - \mathfrak{r}_T\tau\|_{\text{HS}}^2 = \int_T^1 \int_s^1 |a(s)|^2 |c(t)|^2 dt ds + \int_0^T \int_T^1 |a(s)|^2 |c(t)|^2 dt ds,$$

where  $a(s) = X(s)u_0$  and  $c(s) = X(s)u_1$  with  $X = \frac{1}{\sqrt{\tilde{y}}} \begin{pmatrix} 1 & -\tilde{x} \\ 0 & \tilde{y} \end{pmatrix}$ .

From (11), one can check that if  $u \in \mathbb{R}^2$  is a nonzero vector then  $e^{d_u(x+iy, i)} = \frac{|Xu|^2}{|u|^2}$ . Using the triangle inequality, we get

$$(42) \quad \begin{aligned} |a(s)|^2 &= |X(s)u_0|^2 = |u_0|^2 e^{d_{\eta_0}(\tilde{\gamma}, i)} \leq |u_0|^2 e^{d_{\eta_0}(\tilde{\gamma}, z_0) + d_{\mathbb{H}}(z_0, i)}, \\ |c(s)|^2 &= |X(s)u_1|^2 = |u_1|^2 e^{d_{\eta_1}(\tilde{\gamma}, i)} \leq |u_1|^2 e^{d_{\eta_1}(\tilde{\gamma}, z_0) + d_{\mathbb{H}}(z_0, i)}. \end{aligned}$$

Using now (37) and recalling  $\tilde{\gamma}(t) = \gamma(v \log(\frac{1}{1-t}))$ , we get that if (38) holds for all  $t$  then

$$(43) \quad |c(t)|^2 \leq C_0(1-t)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})}, \quad |a(t)|^2 \leq C_0(1-t)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})},$$

where  $C_0$  depends on  $u_0, u_1$  and  $b$ . This leads to

$$(44) \quad \begin{aligned} & 2\|r\tau - r_T\tau\|_{\text{HS}}^2 \\ & \leq C_0^2 \int_T^1 \int_s^1 (1-t)^{\alpha v} (1-s)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} e^{v^\kappa \log^\kappa(\frac{1}{1-s})} dt ds \\ & \quad + C_0^2 \int_0^T (1-s)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-s})} ds \int_T^1 (1-t)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} dt. \end{aligned}$$

Note that for  $0 < s < 1$  and  $0 < \varepsilon < 1$  we have

$$\begin{aligned} \int_s^1 (1-t)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} dt & \leq \int_s^1 (1-t)^{(1-\varepsilon)\alpha v} dt \max_{0 \leq t \leq 1} (1-t)^{\varepsilon\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} \\ & \leq (1-s)^{(1-\varepsilon)\alpha v + 1} C(\varepsilon\alpha, \kappa), \end{aligned}$$

where  $C(\varepsilon\alpha, \kappa) = \max_{0 \leq x \leq 1} e^{-\varepsilon\alpha x + x^\kappa}$  is a positive constant depending on  $\varepsilon\alpha$  and  $\kappa$ . This leads to

$$\begin{aligned} & \int_T^1 \int_s^1 (1-s)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-s})} (1-t)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} dt ds \\ & \leq C(\varepsilon\alpha, \kappa) \int_T^1 (1-s)^{1-\varepsilon\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-s})} ds \leq C(\varepsilon\alpha, \kappa)^2 (1-T)^{2-2\varepsilon\alpha v}, \end{aligned}$$

where for the validity of the last step we also assume  $2\varepsilon c_0 \leq 1$ . Moreover,

$$\begin{aligned} & \int_0^T (1-s)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-s})} ds \int_T^1 (1-t)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-t})} dt \\ & \leq C(\varepsilon\alpha, \kappa)^2 (1-T)^{1+(1-\varepsilon)\alpha v} \int_0^T (1-s)^{-(1+\varepsilon)\alpha v} ds \\ & \leq C(\varepsilon\alpha, \kappa)^2 \log \frac{1}{1-T} \cdot (1-T)^{1+\min(1-2\varepsilon\alpha v, (1-\varepsilon)\alpha v)}, \end{aligned}$$

where the last bound follows from the inequality

$$(45) \quad \int_0^T (1-s)^r ds \leq \log \frac{1}{1-T} \cdot (1-T)^{\min(1+r, 0)},$$

which holds for all  $0 \leq T < 1$  and  $r \in \mathbb{R}$ .

Now choose  $\varepsilon = \min(\frac{1}{2}, \frac{1}{4c_0})$  with  $c_0$  from (39). Collecting all of our bounds and returning to (44), we get the estimate in (40).

Now assume that (38) holds for  $0 \leq t \leq v \log(\frac{1}{1-T})$ , and  $d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}(T)) \leq M$  for  $t \geq T$ . Then for  $0 \leq t \leq T$ , we still have (43), while for  $t \geq T$  we can use  $d_\eta(x, y) \leq d_\eta(x, z) + d_{\mathbb{H}}(y, z)$  together with (42) to get

$$\begin{aligned} |c(t)|^2 & \leq C_0 e^M (1-T)^{\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-T})}, \\ |a(t)|^2 & \leq C_0 e^M (1-T)^{-\alpha v} e^{v^\kappa \log^\kappa(\frac{1}{1-T})}, \end{aligned}$$

for  $T \leq t < 1$ . This gives

$$\begin{aligned} & 2\|\mathfrak{r}\tau - \mathfrak{r}_T\tau\|_{\text{HS}}^2 \\ & \leq C_0^2 e^{2M} \int_T^1 \int_s^1 e^{2\nu^\kappa \log^\kappa(\frac{1}{1-T})} dt ds \\ & \quad + C_0^2 e^M \int_0^T (1-s)^{-\alpha\nu} e^{\nu^\kappa \log^\kappa(\frac{1}{1-s})} ds \int_T^1 (1-T)^{\alpha\nu} e^{\nu^\kappa \log^\kappa(\frac{1}{1-T})} dt. \end{aligned}$$

The bound (41) now follows by using similar estimates as in the proof of (40) from (44).  $\square$

PROPOSITION 16 (Hilbert–Schmidt approximation). *Let  $\gamma, \eta_0, \eta_1, z_0, \alpha, \nu, \tilde{\gamma}, z(t)$  be as in Proposition 14, and define  $\tau$  according to (36). Assume that (38) holds for  $0 \leq t \leq \nu \log(\frac{1}{1-T})$  for some  $T \in (0, 1)$ ,  $0 < b$  and  $0 \leq \kappa < 1$ .*

*Suppose that the path  $\tilde{\gamma}_1$  is measurable and for  $0 \leq t \leq \nu \log(\frac{1}{1-T})$  we have*

$$(46) \quad \sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))\right)^2 \leq \min(\delta(1-t)^{-1}, M)$$

for some  $M, \delta > 0$ .

Consider  $\tau_1 = \text{Dir}(\{\tilde{\gamma}_1(t), t \geq 0\}, \eta_0, \eta_1)$ , define  $\mathfrak{r}_T\tau$  as in Proposition 15, and  $\mathfrak{r}_T\tau_1$  similarly. Then

$$(47) \quad \|\mathfrak{r}_T\tau - \mathfrak{r}_T\tau_1\|_{\text{HS}}^2 \leq C(M+1)\delta,$$

with a constant  $C$  depending only on  $\eta_0, \eta_1, \kappa, b, \alpha$  and  $\nu$ .

PROOF. Denote the representation of  $\tilde{\gamma}_1$  in the half-plane by  $\tilde{x}_1 + i\tilde{y}_1$ .

The hyperbolic distance formula (10) in the upper half-plane representation gives

$$4 \sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{\gamma}, \tilde{\gamma}_1)\right)^2 = \frac{(\tilde{x} - \tilde{x}_1)^2}{\tilde{y}\tilde{y}_1} + \left(\sqrt{\frac{\tilde{y}}{\tilde{y}_1}} - \sqrt{\frac{\tilde{y}_1}{\tilde{y}}}\right)^2.$$

Consider  $a, c$ , defined as in the proof of Proposition 15 and the analogously defined  $a_1, c_1$ .

If  $u \in \mathbb{R}^2$  is a nonzero vector,  $X = \frac{1}{\sqrt{\tilde{y}}}\begin{pmatrix} 1 & -\tilde{x} \\ 0 & \tilde{y} \end{pmatrix}$  and  $X_1 = \frac{1}{\sqrt{\tilde{y}_1}}\begin{pmatrix} 1 & -\tilde{x}_1 \\ 0 & \tilde{y}_1 \end{pmatrix}$  then

$$\frac{|Xu - X_1u|}{|Xu|} = \frac{|(I - X_1X^{-1})Xu|}{|Xu|} \leq \|I - X_1X^{-1}\|_2.$$

An explicit computation gives

$$\begin{aligned} \|I - X_1X^{-1}\|_2^2 &= \left(\frac{x - x_1}{\sqrt{\tilde{y}\tilde{y}_1}}\right)^2 + \left(1 - \sqrt{\frac{\tilde{y}_1}{\tilde{y}}}\right)^2 + \left(1 - \sqrt{\frac{\tilde{y}}{\tilde{y}_1}}\right)^2 \\ &\leq \left(\frac{\tilde{x} - \tilde{x}_1}{\sqrt{\tilde{y}\tilde{y}_1}}\right)^2 + \left(\sqrt{\frac{\tilde{y}}{\tilde{y}_1}} - \sqrt{\frac{\tilde{y}_1}{\tilde{y}}}\right)^2 \\ &= 4 \sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{x} + i\tilde{y}, \tilde{x}_1 + i\tilde{y}_1)\right)^2. \end{aligned}$$

This yields

$$\frac{|c - c_1|^2}{|c|^2} \leq 4 \sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{\gamma}, \tilde{\gamma}_1)\right)^2, \quad \frac{|a - a_1|^2}{|a|^2} \leq 4 \sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{\gamma}, \tilde{\gamma}_1)\right)^2.$$

To compute  $\|\mathfrak{r}_T \tau - \mathfrak{r}_T \tau_1\|_{\text{HS}}^2$ , we need to estimate  $\text{tr}(\Delta K(s, t)\Delta K(s, t)^t)$  where  $\Delta K(s, t) = \frac{1}{2}(a(s)c(t)^t - a_1(s)c_1(t)^t)$ . Using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} &4\text{tr}(\Delta K(s, t)\Delta K(s, t)^t) \\ &\leq 3(|a(s)|^2|c(t) - c_1(t)|^2 + |a(s) - a_1(s)|^2|c(t)|^2 \\ &\quad + |a(s) - a_1(s)|^2|c(t) - c_1(t)|^2). \end{aligned}$$

The previous estimates with (46) yield

$$\begin{aligned} &\|\mathfrak{r}_T \tau - \mathfrak{r}_T \tau_1\|_{\text{HS}}^2 \\ &= 2 \int_0^T \int_0^t \text{tr}(\Delta K(s, t)\Delta K(s, t)^t) ds dt \\ &\leq 24 \int_0^T \int_0^t |a(s)|^2|c(t)|^2 \left( \sinh\left(\frac{d_{\mathbb{H}}(s)}{2}\right)^2 + \sinh\left(\frac{d_{\mathbb{H}}(t)}{2}\right)^2 \right. \\ &\quad \left. + \sinh\left(\frac{d_{\mathbb{H}}(s)}{2}\right)^2 \sinh\left(\frac{d_{\mathbb{H}}(t)}{2}\right)^2 \right) ds dt \\ &\leq 24 \int_0^T \int_0^t |a(s)|^2|c(t)|^2 \left( \delta(1-t)^{-1} + \delta(1-s)^{-1} \right. \\ &\quad \left. + \frac{1}{2}\delta M((1-t)^{-1} + (1-s)^{-1}) \right) ds dt, \end{aligned}$$

where we used the notation  $d_{\mathbb{H}}(t) = d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))$ . Using the arguments in the proof of Proposition 15, we get that (43) holds for  $0 \leq t \leq T$ , with a constant  $C_1$  depending on  $\eta_0, \eta_1$  and  $b$ . Using the temporary notation  $g(t) = \nu^\kappa \log^\kappa(\frac{1}{1-t})$ , this leads to

$$\begin{aligned} &\|\mathfrak{r}_T \tau - \mathfrak{r}_T \tau_1\|_{\text{HS}}^2 \\ &\leq 24C_1^2\delta \left(1 + \frac{M}{2}\right) \int_0^T \int_0^t e^{g(t)+g(s)}((1-s)^{-\alpha\nu-1}(1-t)^{\alpha\nu} \\ &\quad + (1-s)^{-\alpha\nu}(1-t)^{\alpha\nu-1}) ds dt \\ &\leq 24C_1^2\delta(2+M) \int_0^1 e^{2g(t)}(1-t)^{\alpha\nu-1} \int_0^t (1-s)^{-\alpha\nu} ds dt \\ &\leq 24C_1^2\delta(2+M) \int_0^1 e^{2g(t)} \log\left(\frac{1}{1-t}\right)(1-t)^{\min(\alpha\nu-1,0)} dt, \end{aligned}$$

where we used (45) in the last step.

Since the integral  $\int_0^1 e^{2\nu^\kappa \log^\kappa(\frac{1}{1-t})} \log(\frac{1}{1-t})(1-t)^{\min(\alpha\nu-1,0)} dt$  is finite for any given  $0 \leq \kappa < 1, \nu > 0$  and  $\alpha > 0$ , the bound (47) now follows.  $\square$

**5. Proof of Theorem 1.** We now return to the proof of our main theorem.

PROOF OF THEOREM 1. Let  $\mathcal{B}$  be a hyperbolic Brownian motion. Set  $\tilde{\mathcal{B}}(t) = \mathcal{B}(-\frac{4}{\beta} \log(1-t))$  for  $0 \leq t < 1$ . Set  $\eta_0 = \infty, \eta_1 = \mathcal{B}(\infty)$  and set  $\text{Sine}_\beta = \text{Dir}(\tilde{\mathcal{B}}, \eta_0, \eta_1)$ .

Let  $U_k, \xi_k, k \geq 1$  be random variables independent of each other and  $\mathcal{B}$  with distributions given in Proposition 12, and let  $\tau_{n,k}$  be the stopping times constructed there. According to the proposition, the path  $\tilde{\mathcal{B}}_n(t) = \mathcal{B}(\tau_{n, \lceil(1-t)n\rceil}), t \in [0, 1)$  has the same distribution as the path in the construction of  $\text{Circ}_{\beta,n}$  in Theorem 8, and we may write  $\text{Circ}_{\beta,n} = \text{Dir}(\tilde{\mathcal{B}}_n, \eta_0, \eta_1)$ .

To prove the bound (1), we set  $T_n = 1 - \frac{\log^6 n}{n}$ . Recall the definition of  $r_T$  from Section 4 and write

$$\begin{aligned}
 & \|r\text{Sine}_\beta - r\text{Circ}_{\beta,n}\|_{\text{HS}}^2 \\
 (48) \quad & \leq 3\|r_{T_n}\text{Sine}_\beta - r_{T_n}\text{Circ}_{\beta,n}\|_{\text{HS}}^2 \\
 & \quad + 3\|r\text{Sine}_\beta - r_{T_n}\text{Sine}_\beta\|_{\text{HS}}^2 \\
 & \quad + 3\|r\text{Circ}_{\beta,n} - r_{T_n}\text{Circ}_{\beta,n}\|_{\text{HS}}^2.
 \end{aligned}$$

We will use Propositions 13, 15 and 16 to estimate the three terms on the right. Let  $z(t)$  be the point moving with speed 1/2 on the geodesic connecting  $\mathcal{B}(0)$  to  $\mathcal{B}(\infty)$ . From Lemma 21 of the Appendix, it follows that for  $\kappa = 2/3$  there is a random  $b$  so that a.s. for all  $t \geq 0$  we have

$$(49) \quad d_{\mathbb{H}}(\mathcal{B}(t), z(t)) \leq b + t^{2/3}.$$

Applying the first statement of Proposition 15 for  $\text{Sine}_\beta = \text{Dir}(\tilde{\mathcal{B}}(t), \eta_0, \eta_1)$  with  $\gamma = \mathcal{B}$ ,  $v = \frac{4}{\beta}$ ,  $\alpha = 1/2$ ,  $\kappa = 2/3$  and  $c_0 = \frac{2}{\beta}$ , we get

$$\|r\text{Sine}_\beta - r_{T_n}\text{Sine}_\beta\|_{\text{HS}}^2 \leq C \left(\frac{\log^6 n}{n}\right)^{1+\frac{1}{2}\min(\frac{2}{\beta}, 1)} \left(1 + \log \frac{n}{\log^6 n}\right)$$

with a random  $C$  depending on  $\mathcal{B}$  and  $\beta$ .

Recall from Proposition 13 that for  $n \geq N^*$  we have

$$(50) \quad d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t)) \leq \frac{\log^{3-1/8} n}{\sqrt{(1-t)n}} \leq 1 \quad \text{if } 0 \leq t \leq T_n,$$

$$(51) \quad d_{\mathbb{H}}(\tilde{\mathcal{B}}_n(T_n), \tilde{\mathcal{B}}_n(t)) \leq (\log \log n)^4 \quad \text{if } T_n \leq t < 1.$$

Let  $\mathcal{B}_n(t) = \tilde{\mathcal{B}}_n(1 - e^{-\frac{\beta}{4}t})$ , then  $\tilde{\mathcal{B}}_n(t) = \mathcal{B}_n(\frac{4}{\beta} \log(\frac{1}{1-t}))$  for  $0 \leq t < 1$ . From (49), (50) and the triangle inequality, we get

$$d_{\mathbb{H}}(\mathcal{B}_n(t), z(t)) \leq b + 1 + t^{2/3} \quad \text{for } 0 \leq t \leq \frac{4}{\beta} \log\left(\frac{1}{1-T_n}\right).$$

Recall that  $\text{Circ}_{\beta,n} = \text{Dir}(\tilde{\mathcal{B}}_n(t), \eta_0, \eta_1)$ . Applying the second statement of Proposition 15 with  $\gamma = \mathcal{B}_n$ ,  $v = \frac{4}{\beta}$ ,  $\alpha = 1/2$ ,  $\kappa = 2/3$ ,  $c_0 = \frac{2}{\beta}$  and  $M = (\log \log n)^4$ ,  $T = T_n$  we get

$$\begin{aligned}
 & \|r\text{Circ}_{\beta,n} - r_{T_n}\text{Circ}_{\beta,n}\|_{\text{HS}}^2 \\
 & \leq C e^{2(\log \log n)^4} \left(\frac{\log^6 n}{n}\right)^{1+\frac{1}{2}\min(\frac{2}{\beta}, 1)} \left(1 + \log \frac{n}{\log^6 n}\right)
 \end{aligned}$$

if  $n \geq N^*$ , with a random  $C$  depending on  $\mathcal{B}$  and  $\beta$ .

Since  $\sinh(x/2)^2 \leq x^2$  for  $0 \leq x \leq 1$ , from (50) we get

$$\sinh\left(\frac{1}{2}d_{\mathbb{H}}(\tilde{\mathcal{B}}(t), \tilde{\mathcal{B}}_n(t))\right)^2 \leq \frac{\log^{6-1/4} n}{n}(1-t)^{-1} \leq 1 \quad \text{if } 0 \leq t \leq T_n.$$

Hence we may use Proposition 16 with  $\gamma = \mathcal{B}$ ,  $\gamma_1 = \mathcal{B}_n$ ,  $v = \frac{4}{\beta}$ ,  $\alpha = 1/2$ ,  $T = T_n$ ,  $\delta = \frac{\log^{6-1/4} n}{n}$ ,  $M = 1$ ,  $\kappa = 2/3$  to get

$$\|r_{T_n}\text{Sine}_\beta - r_{T_n}\text{Circ}_{\beta,n}\|_{\text{HS}}^2 \leq C \frac{\log^{6-1/4} n}{n} \quad \text{if } n \geq N^*,$$

with a random  $C$  depending on  $\mathcal{B}$  and  $\beta$ . Collecting our estimates, going back to (48) and modifying the random lower bound  $N^*$  appropriately, we get

$$\|rSine_\beta - rCirc_{\beta,n}\|_{HS}^2 \leq \frac{\log^6 n}{n} \quad \text{for } n \geq N^*.$$

This completes the proof of Theorem 1.  $\square$

Before proving the three statements of Corollary 2, we state a law of large numbers for the points of  $Sine_\beta$ .

PROPOSITION 17. *Suppose that the points of  $Sine_\beta$  are given by  $\lambda_k, k \in \mathbb{Z}$  with  $\lambda_0 < 0 \leq \lambda_1$ . Then with probability one, we have*

$$(52) \quad \lim_{k \rightarrow \infty} \frac{\lambda_k}{k} = \lim_{k \rightarrow -\infty} \frac{\lambda_k}{k} = 2\pi.$$

PROOF. In Holcomb and Valkó (2015), it was shown that  $\frac{1}{\lambda} \#\{\lambda_k : 0 \leq \lambda_k \leq \lambda\}$  satisfies a large deviation principle as  $\lambda \rightarrow \infty$  with scale  $\lambda^2$  and rate function  $\beta I(\rho)$ , where  $I(\frac{1}{2\pi}) = 0$  is the global minimum. From this, the statement follows for  $k \rightarrow \infty$  by a simple Borel–Cantelli argument. The  $k \rightarrow -\infty$  case follows similarly using the symmetry of the  $Sine_\beta$  process.  $\square$

PROOF OF COROLLARY 2. The bound (2) follows directly from (1) and the Hoffman–Wielandt inequality for compact integral operators (see, e.g., Bhatia and Elsner (1994)). Note that it might be possible to get a sharper bound by comparing the eigenfunctions as well.

From Proposition 17, we see that there is a random constant  $C$  so that a.s.  $|\lambda_k| \leq C(|k| + 1)$  for all  $k$ . Set  $a_n = \frac{n^{1/4}}{\log^2 n}$ . From (1), it follows that for large enough  $n$  we have

$$(53) \quad \sup_k |\lambda_k^{-1} - \lambda_{k,n}^{-1}| \leq \sqrt{\sum_k |\lambda_k^{-1} - \lambda_{k,n}^{-1}|^2} \leq \frac{\log^3 n}{n^{1/2}}.$$

Let  $b_{k,n} = \lambda_k^{-1} - \lambda_{k,n}^{-1}$ , then

$$(54) \quad (\lambda_{k,n} - \lambda_k)(1 - b_{k,n}\lambda_k) = b_{k,n}\lambda_k^2.$$

For large enough  $n$ , we have

$$\max_{|k| \leq a_n} |b_{k,n}\lambda_k| \leq \frac{\log^3 n}{n^{1/2}} C(a_n + 1) = C \left( \frac{\log n}{n^{1/4}} + \frac{\log^3 n}{n^{1/2}} \right),$$

which means that  $\lim_{n \rightarrow \infty} \max_{|k| \leq a_n} |b_{k,n}\lambda_k| = 0$  a.s. From (54), for large enough  $n$ , we have

$$\max_{|k| \leq a_n} |\lambda_{k,n} - \lambda_k| \leq 2C^2(a_n + 1)^2 \max_{|k| \leq a_n} |b_{k,n}| \leq 4C^2 a_n^2 \frac{\log^3 n}{n^{1/2}} = \frac{4C^2}{\log n}.$$

This completes the proof of (3).

By Proposition 17, we may choose a random  $C > 0$  so that  $|\lambda_k| \leq C\sqrt{1+k^2}$  for all  $k$ . Then we have  $\frac{1}{C\sqrt{1+k^2}} \leq \frac{1}{|\lambda_k|}$  and for  $n \geq N$ ,  $|k| \leq n^{1/2-\varepsilon}$  we have

$$\frac{1}{|\lambda_{k,n}|} \geq \frac{1}{|\lambda_k|} - \left| \frac{1}{\lambda_k} - \frac{1}{\lambda_{k,n}} \right| \geq \frac{1}{C\sqrt{1+k^2}} - \frac{\log^3 n}{\sqrt{n}} \geq \frac{1}{2C\sqrt{1+k^2}}.$$

(For the last inequality, one might need to change  $N$ .) Then for such  $k$  and  $n$ , we get

$$|\lambda_k - \lambda_{k,n}| = \left| \frac{1}{\lambda_k} - \frac{1}{\lambda_{k,n}} \right| |\lambda_k| |\lambda_{k,n}| \leq \frac{\log^3 n}{n^{1/2}} 2C^2(1+k^2) \leq \frac{1+k^2}{n^{1/2-\varepsilon}},$$

again by setting a large enough random lower bound on  $n$ .  $\square$

**6. Beta dependence in the  $\text{Sine}_\beta$  operator.** The construction of the  $\text{Sine}_\beta$  operator provides a natural coupling of this operator for all values of  $\beta$ . The techniques we developed for the proof of Theorem 1 can be used to estimate how the operator  $\text{Sine}_\beta$  (and the process  $\text{Sine}_\beta$ ) depends on the value of  $\beta$  in this coupling.

PROOF OF THEOREM 3. We start with the case  $\beta' < \infty$ . We have

$$\text{Sine}_\beta = \text{Dir}(\tilde{\gamma}, \eta_0, \eta_1), \quad \text{Sine}_{\beta'} = \text{Dir}(\tilde{\gamma}_1, \eta_0, \eta_1),$$

where  $\tilde{\gamma}(t) = \mathcal{B}(\frac{4}{\beta} \log(\frac{1}{1-t}))$ ,  $\tilde{\gamma}_1(t) = \mathcal{B}(\frac{4}{\beta'} \log(\frac{1}{1-t}))$ ,  $\eta_0 = \infty$  and  $\eta_1 = \mathcal{B}(\infty)$ . We will estimate  $\|\text{rSine}_\beta - \text{rSine}_{\beta'}\|_{\text{HS}}$  using Propositions 15 and 16. Note that with probability one  $\eta_0 \neq \eta_1$ , and hence a.s. 0 is not an eigenvalue for any of the  $\text{Sine}_\beta$  operators.

Set  $T = 1 - \delta < 1$ . Cutting off  $\text{Sine}_\beta$  and  $\text{Sine}_{\beta'}$  at  $T$  and using the first statement of Proposition 15 with  $\alpha = 1/2$ ,  $\kappa = 2/3$ ,  $c_0 = \frac{2}{\theta}$ , and  $\nu = \frac{4}{\beta}$  and  $\frac{4}{\beta'}$ , respectively, as in the proof of Theorem 1) gives

$$\begin{aligned} \|\text{rSine}_\beta - \text{r}_T \text{Sine}_\beta\|_{\text{HS}}^2 &\leq C_0 \delta^{1+\frac{1}{2} \min(\frac{2}{\beta}, 1)} (1 + \log \delta^{-1}), \\ \|\text{rSine}_{\beta'} - \text{r}_T \text{Sine}_{\beta'}\|_{\text{HS}}^2 &\leq C_0 \delta^{1+\frac{1}{2} \min(\frac{2}{\beta'}, 1)} (1 + \log \delta^{-1}) \end{aligned}$$

with a random  $C_0$  depending only on  $\mathcal{B}$ .

To estimate  $\|\text{r}_T \text{Sine}_\beta - \text{r}_T \text{Sine}_{\beta'}\|_{\text{HS}}^2$ , we first bound  $d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))$  in  $[0, T]$ . For  $0 < t \leq T$ , we have  $(\frac{4}{\beta'} - \frac{4}{\beta}) \log(\frac{1}{1-t}) \leq \delta \log(\delta^{-1}) < 1$  and by Proposition 22 (and the comment following the proposition), we get the bound

$$\begin{aligned} (55) \quad &d_{\mathbb{H}}\left(\mathcal{B}\left(\frac{4}{\beta} \log\left(\frac{1}{1-t}\right)\right), \mathcal{B}\left(\frac{4}{\beta'} \log\left(\frac{1}{1-t}\right)\right)\right)^2 \\ &< C^2 \delta \log\left(\frac{1}{1-t}\right) \log\left(2 + \frac{\frac{4}{\beta'} \log(\frac{1}{1-t}) + 1}{\delta \log(\frac{1}{1-t})}\right) \end{aligned}$$

for  $0 \leq t \leq T$  with a random  $C$  depending only on  $\mathcal{B}$ .

Since  $0 \leq t \leq T = 1 - \delta$ , the right-hand side of (55) can be bounded as

$$C^2 \delta \log\left(\frac{1}{1-t}\right) \left(\log\left(2 + \frac{4}{\theta} \delta^{-1}\right) + \log\left(1 + \frac{1}{\delta \log(\frac{1}{1-t})}\right)\right) \leq C^2 c_1,$$

with a constant  $c_1$  depending only on  $\theta$ .

Using  $\log(\frac{1}{1-t}) \leq (1-t)^{-1/2}$  (which holds for  $0 \leq t < 1$ ), we also get the bound

$$d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))^2 < \min\left(C_1 \delta \log\left(\frac{1}{\delta}\right) (1-t)^{-1/2}, C_1\right), \quad t \in [0, T]$$

with a constant  $C_1$  depending on  $\mathcal{B}$  and  $\theta$ . We can turn this into an upper bound of the form

$$\begin{aligned} (56) \quad &\sinh\left(\frac{1}{2} d_{\mathbb{H}}(\tilde{\gamma}(t), \tilde{\gamma}_1(t))\right)^2 \\ &\leq \min\left(C_2 \delta \log\left(\frac{1}{\delta}\right) (1-t)^{-1/2}, C_2\right), \quad t \in [0, T], \end{aligned}$$

with  $C_2$  depending on  $\mathcal{B}$  and  $\theta$ .

Using the arguments in the proof of Proposition 16, we get the bounds

$$\begin{aligned} &\|\text{r}_T \text{Sine}_\beta - \text{r}_T \text{Sine}_{\beta'}\|_{\text{HS}}^2 \\ &\leq C_3 \delta \log\left(\frac{1}{\delta}\right) \int_0^T \int_0^t e^{g(t)+g(s)} ((1-s)^{-\frac{2}{\beta}-1/2} (1-t)^{\frac{2}{\beta}} \end{aligned}$$

$$\begin{aligned}
 &+ (1-s)^{-\frac{2}{\beta}}(1-t)^{\frac{2}{\beta}-1/2} ds dt \\
 &\leq 2C_3\delta \log\left(\frac{1}{\delta}\right) \int_0^1 e^{2g(t)}(1-t)^{\frac{2}{\beta}-1/2} \int_0^t (1-s)^{-\frac{2}{\beta}} ds dt \\
 &\leq 2C_3\delta \log\left(\frac{1}{\delta}\right) \int_0^1 e^{2g(t)} \log\left(\frac{1}{1-t}\right) (1-t)^{\min(\frac{2}{\beta}-1/2, 1/2)} dt,
 \end{aligned}$$

with  $g(t) = (4/\beta)^{2/3} \log^{2/3}(\frac{1}{1-t})$  and  $C_3$  still only depending on  $\mathcal{B}$  and  $\theta$ . Since  $\theta \leq \beta$ , we have

$$\begin{aligned}
 &\int_0^1 e^{2g(t)} \log\left(\frac{1}{1-t}\right) (1-t)^{\min(\frac{2}{\beta}-1/2, 1/2)} dt \\
 &\leq \int_0^1 e^{2g(t)} \log\left(\frac{1}{1-t}\right) (1-t)^{\frac{1}{\beta} \min(2, \theta) - 1/2} dt \\
 &\leq \int_0^1 \log\left(\frac{1}{1-t}\right) (1-t)^{-1/2} dt \max_{x>0} e^{2x^{2/3} - \frac{1}{4} \min(2, \theta)x} \leq C_4,
 \end{aligned}$$

with  $C_4$  depending only on  $\theta$ . This gives

$$\|r_T \text{Sine}_\beta - r_T \text{Sine}_{\beta'}\|_{\text{HS}}^2 \leq C_5\delta \log\left(\frac{1}{\delta}\right),$$

with  $C_5$  depending only on  $\mathcal{B}$  and  $\theta$ . (Note that we needed the bound (56) that was slightly better than the assumption (46) in Proposition 16 to get an upper bound here that does not depend on  $\beta, \beta'$ .)

Collecting all the terms gives

$$\|r \text{Sine}_\beta - r \text{Sine}_{\beta'}\|_{\text{HS}}^2 \leq C\delta \log\left(\frac{1}{\delta}\right),$$

with a  $C$  depending only on  $\mathcal{B}$  and  $\theta$ . The Hoffman–Wielandt inequality completes the proof of (5).

To treat the  $\beta' = \infty$  case, we note that here  $\tilde{\gamma}_1(t) = \mathcal{B}(0)$ , hence the integral kernel of  $r \text{Sine}_{\beta'}$  is constant on the sets  $\{(s, t) : 0 \leq s < t < 1\}$  and  $\{(s, t) : 0 \leq t \leq s < 1\}$ , with the constants depending only on  $\eta_1 = \mathcal{B}(\infty)$ . Setting  $\delta = \frac{4}{\beta} \leq 1/3$  and  $T = 1 - \delta$ , we have

$$\|r \text{Sine}_\infty - r_T \text{Sine}_\infty\|_{\text{HS}}^2 \leq C_0\delta$$

with  $C_0$  depending only  $\mathcal{B}(\infty)$ .

Using (56), we have

$$\|r \text{Sine}_\beta - r_T \text{Sine}_\beta\|_{\text{HS}}^2 \leq C_1\delta \log\left(\frac{1}{\delta}\right)$$

with  $C_1$  depending only on  $\mathcal{B}$  and  $\theta$ .

The term  $\|r_T \text{Sine}_\beta - r_T \text{Sine}_\infty\|_{\text{HS}}^2$  can be bounded similarly as in the case  $\beta' < \infty$ , leading to the same upper bound. This completes the proof of (5) in the  $\beta' = \infty$  case.  $\square$

### APPENDIX

In the first part of the Appendix, we collect the proofs of Lemmas 9 and 11 that were used in the single step coupling of Proposition 10. In the second part, we collect some estimates on the hyperbolic Brownian motion.

**A.1. Proof of the coupling statements.** We now return to the proof of Lemma 9. Let  $\mathcal{B}$  be a hyperbolic Brownian motion and set  $\zeta = \zeta_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$ .

The circumference of a hyperbolic circle of radius  $r$  is  $2\pi \sinh(r)$  and the density function of the hyperbolic BM at time  $t$  and distance  $r$  is given by the following formula (see, e.g., Karpelevič, Tutubalin and Šur (1959)):

$$g(r, t) = \frac{\sqrt{2}e^{-t/8}}{(2\pi t)^{3/2}} \int_r^\infty \frac{se^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh r}} ds.$$

From this, the density of  $\zeta_t$  is

$$(57) \quad p_\zeta(r, t) = 2\pi \sinh(r)g(r, t) = \frac{e^{-t/8} \sinh(r)}{\sqrt{\pi t^3/2}} \int_r^\infty \frac{se^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh r}} ds,$$

and the tail of the cumulative distribution function is

$$(58) \quad \begin{aligned} 1 - F_\zeta(r) &= \int_r^\infty \sinh(u) \frac{e^{-t/8}}{\sqrt{\pi t^3/2}} \int_u^\infty \frac{se^{-\frac{s^2}{2t}}}{\sqrt{\cosh s - \cosh u}} ds du \\ &= \int_r^\infty \frac{e^{-t/8} se^{-\frac{s^2}{2t}}}{\sqrt{\pi t^3/2}} \int_r^s \frac{\sinh(u)}{\sqrt{\cosh s - \cosh u}} du ds \\ &= \int_r^\infty \frac{2e^{-t/8} se^{-\frac{s^2}{2t}}}{\sqrt{\pi t^3/2}} \sqrt{\cosh(s) - \cosh(r)} ds. \end{aligned}$$

Let  $Y = Y_\gamma = \log\left(\frac{1+\sqrt{\xi}}{1-\sqrt{\xi}}\right)$  where  $\xi$  has distribution Beta(1,  $\gamma$ ). We record the cumulative distribution function  $F_Y$  and the probability density function  $p_Y$  of  $Y$ , which follow from direct computation with the Beta distribution:

$$(59) \quad F_Y(r) = 1 - \operatorname{sech}^{2\gamma}\left(\frac{r}{2}\right), \quad p_Y(r) = \gamma \sinh\left(\frac{r}{2}\right) \operatorname{sech}^{2\gamma+1}\left(\frac{r}{2}\right).$$

We start with a simple estimate.

LEMMA 18. For  $0 \leq r \leq s$ , we have

$$(60) \quad \frac{1}{\sqrt{2}}\sqrt{s^2 - r^2} \leq \sqrt{\cosh(s) - \cosh(r)} \leq \frac{1}{\sqrt{2}}\sqrt{s^2 - r^2} \exp\left(\frac{r^2 + s^2}{24}\right).$$

PROOF. The statement follows from the bound  $1 \leq \frac{\sinh(x)}{x} \leq e^{x^2/6}$  and

$$\frac{2(\cosh(s) - \cosh(r))}{s^2 - r^2} = \frac{\sinh((s - r)/2)}{(s - r)/2} \cdot \frac{\sinh((s + r)/2)}{(s + r)/2}. \quad \square$$

Applying Lemma 18 to (57) and (58) and computing the resulting integrals directly, we get the following bounds:

$$(61) \quad p_\zeta(r, t) \geq \left(1 + \frac{t}{12}\right)^{-1/2} \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{r^2}{12} - \frac{t}{8}} \sinh(r) =: p_-(r, t),$$

$$(62) \quad p_\zeta(r, t) \leq \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{t}{8}} \sinh(r) =: p_+(r, t),$$

$$(63) \quad 1 - F_\zeta(r) \leq \left(1 - \frac{t}{12}\right)^{-3/2} e^{-\frac{r^2}{2t} + \frac{r^2}{12} - \frac{t}{8}},$$

where the last bound is valid for  $0 < t < 12$ .

We also record the following bounds on  $\log \cosh(r)$  which can be readily checked by differentiation and Taylor expansion:

$$(64) \quad \log \cosh(x) \leq \frac{x^2}{2}, \quad \text{for all } x,$$

$$(65) \quad \log \cosh(x) \geq \frac{x^2}{2} - \frac{x^4}{12}, \quad \text{for } 0 \leq x \leq 1.$$

PROOF OF (A) IN LEMMA 9. From (59) and (64), we get

$$1 - F_Y(r) = \exp(-2\gamma \log \cosh(r/2)) \geq \exp\left(-\gamma \frac{r^2}{4}\right) = \exp\left(-\frac{r^2}{2t} + \frac{r^2}{8}\right).$$

One can check that if  $0 < t \leq 1$  and  $t \leq r$  then

$$\left(1 - \frac{t}{12}\right)^{-3/2} e^{-\frac{r^2}{2t} + \frac{r^2}{12} - \frac{t}{8}} \leq \exp\left(-\frac{r^2}{2t} + \frac{r^2}{8}\right).$$

Together with (63), this proves the statement for  $t \leq r$ .

To prove the statement in the  $0 \leq r < t$  case, we will show

$$\int_0^r p_\zeta(u, t) du \geq \int_0^r p_Y(u) du \quad \text{for } 0 \leq r < t \leq 1$$

using the lower bound (61) on  $p_\zeta(r, t)$ . We will prove that for  $0 < r < t \leq 1$  we have

$$\begin{aligned} p_-(r, t) &= \left(1 + \frac{t}{12}\right)^{-1/2} \frac{1}{t} e^{-\frac{r^2}{2t} - \frac{r^2}{12} - \frac{t}{8}} \sinh(r) \\ &\geq \gamma \sinh\left(\frac{r}{2}\right) \operatorname{sech}^{2\gamma+1}\left(\frac{r}{2}\right) = p_Y(r). \end{aligned}$$

The last inequality is equivalent to

$$(66) \quad e^{-\frac{r^2}{2t} - \frac{r^2}{12}} \geq e^{t/8} \sqrt{1 + \frac{t}{12}} \left(1 - \frac{t}{4}\right) \exp\left(-\left(\frac{4}{t} + 1\right) \log \cosh(r/2)\right).$$

For  $0 \leq t \leq 1$ , we have  $e^{t/8} \sqrt{1 + \frac{t}{12}} \left(1 - \frac{t}{4}\right) \leq 1$ . Then by the bound (65), we get

$$\begin{aligned} e^{t/8} \sqrt{1 + \frac{t}{12}} \left(1 - \frac{t}{4}\right) \exp\left(-\left(\frac{4}{t} + 1\right) \log \cosh(r/2)\right) \\ \leq \exp\left(-\left(\frac{4}{t} + 1\right) \left(\frac{r^2}{8} - \frac{r^4}{192}\right)\right). \end{aligned}$$

To complete the proof of (66), we need to show that for  $0 < r < t \leq 1$  we have

$$\left(\frac{4}{t} + 1\right) \left(\frac{r^2}{8} - \frac{r^4}{192}\right) \geq \frac{r^2}{2t} + \frac{r^2}{12},$$

which follows from direct computation.  $\square$

Now we turn to the proof of the total variation bound.

PROOF OF (B) IN LEMMA 9. We need to show that  $\int_0^\infty |p_Y(r) - p_\zeta(r, t)| dr \leq 3t$ . By part (a), we have  $P(\zeta > r) \leq P(Y_\gamma > r)$  which leads to

$$\int_K^\infty |p_Y(r) - p_\zeta(r, t)| dr \leq 2(1 - F_Y(K)) = 2\operatorname{sech}^{2\gamma}\left(\frac{K}{2}\right), \quad \text{for } K > 0.$$

Setting  $K = 2\sqrt{t \log(2/t)} \leq 2$ , we get

$$\begin{aligned} \int_K^\infty |p_Y(r) - p_\zeta(r, t)| dr &\leq 2 \operatorname{sech}^{2\gamma} \left( \frac{K}{2} \right) \leq 2 \exp \left( -2\gamma \cdot \frac{5}{12} (K/2)^2 \right) \\ &\leq 2(t/2)^{\frac{5}{4}} \leq t, \end{aligned}$$

where we used (65) in the second inequality.

Using the triangle inequality and the bound (62), we get

$$\int_0^K |p_Y(r) - p_\zeta(r, t)| dr \leq \int_0^K |p_Y(r) - p_+(r, t)| dr + \int_0^K (p_+(r, t) - p_\zeta(r, t)) dr.$$

We can bound the second integral explicitly:

$$\begin{aligned} \int_0^K (p_+(r, t) - p_\zeta(r, t)) dr &\leq \int_0^\infty (p_+(r, t) - p(r, t)) dr \\ &= \frac{\sqrt{\frac{\pi}{2}} e^{\frac{3t}{8}} (2\Phi(\sqrt{t}) - 1)}{\sqrt{t}} - 1 \\ &\leq e^{\frac{3t}{8}} - 1 \leq t/2. \end{aligned}$$

Introduce

$$p_0(r, t) = \frac{2e^{-\frac{r^2}{2t}} \sinh\left(\frac{r}{2}\right)}{t}.$$

Then

$$\frac{p_+}{p_0} = e^{-t/8} \cosh(r/2), \quad \frac{p_Y}{p_0} = (1 - t/4) e^{\frac{r^2}{2t}} \operatorname{sech}^4\left(\frac{r}{2}\right).$$

We have, for  $0 \leq t \leq 1, 0 \leq r \leq 2$ :

$$\left| \frac{p_+}{p_0} - 1 \right| \leq \frac{t}{8} + \frac{r^2}{4}.$$

Using the bounds (64) and (65), we get that for  $0 \leq r \leq K < 2$  we have

$$1 = \exp\left(\frac{r^2}{2t} - \frac{4r^2}{t} \frac{1}{8}\right) \leq e^{\frac{r^2}{2t}} \operatorname{sech}^4\left(\frac{r}{2}\right) \leq \exp\left(\frac{r^2}{2t} - \frac{4}{t} \left(\frac{r^2}{8} - \frac{r^4}{192}\right)\right) = e^{\frac{r^4}{48t}}.$$

This leads to

$$\left| \frac{p_Y}{p_0} - 1 \right| \leq \frac{t}{4} + |e^{\frac{r^4}{48t}} - 1| \leq \frac{t}{4} + \frac{r^4}{24t},$$

where we used that  $\frac{r^4}{t} \leq t \log(2/t)^2 \leq 1$  if  $t \leq 1$ . Note also that  $\sinh(x) \leq \frac{6}{5}x$  for  $x \leq 1$ . From this,

$$\begin{aligned} \int_0^K |p_Y(r) - p_+(r, t)| dr &\leq \int_0^K \left( \frac{3}{8}t + \frac{1}{4}r^2 + \frac{r^4}{24t} \right) p_0(r) dr \\ &\leq \frac{6}{5} \int_0^\infty \left( \frac{3}{8}t + \frac{1}{4}r^2 + \frac{r^4}{24t} \right) \frac{e^{-\frac{r^2}{2t}}}{t} dr = \frac{29}{20}t. \end{aligned}$$

Collecting all our estimates gives

$$\int_0^\infty |p_Y(r) - p_\zeta(r, t)| dr \leq 3t. \quad \square$$

The proof of Lemma 11 follows a standard coupling construction.

PROOF OF LEMMA 11. Denote the distributions of  $X_1, X_2$  by  $\mu_1, \mu_2$ , and let  $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ . Denote that density function of  $X_i$  with respect to  $\mu$  by  $f_i$ . From our assumptions, it follows that

$$\begin{aligned} \tilde{f}_0(x) &= \frac{1}{1-\varepsilon} \min(f_1(x), f_2(x)), & \tilde{f}_1(x) &= \frac{1}{\varepsilon} |f_1(x) - f_2(x)|^+, \\ \tilde{f}_2(x) &= \frac{1}{\varepsilon} |f_2(x) - f_1(x)|^+, \end{aligned}$$

are also density functions with respect to  $\mu$ , and the distributions corresponding to  $\tilde{f}_1, \tilde{f}_2$  are stochastically ordered just as  $X_1$  and  $X_2$ . Moreover,  $(1 - \varepsilon)\tilde{f}_0 + \varepsilon\tilde{f}_i = f_i$  for  $i = 1, 2$ .

Recall that if  $F$  is a cumulative distribution function,  $F^{-1}(x) = \sup\{y : F(y) < x\}$  is its generalized inverse, and  $U$  is uniform on  $[0, 1]$  then  $F^{-1}(U)$  has cumulative distribution function given by  $F$ .

Denote the cumulative distribution function corresponding to  $\tilde{f}_i$  by  $\tilde{F}_i$ . Let  $U_1, U_2$  be independent uniform random variables on  $[0, 1]$  and consider the pair of random variables

$$\begin{aligned} (Y_1, Y_2) &= \mathbf{1}(U_1 \leq 1 - \varepsilon)(\tilde{F}_0^{-1}(U_2), \tilde{F}_0^{-1}(U_2)) \\ &\quad + \mathbf{1}(U_1 > 1 - \varepsilon)(\tilde{F}_1^{-1}(U_2), \tilde{F}_2^{-1}(U_2)). \end{aligned}$$

In plain words, with probability  $1 - \varepsilon$ , we generate  $(Z, Z)$  where  $Z$  has density  $\tilde{f}_0$ , and with probability  $\varepsilon$  we generate  $(\tilde{X}_1, \tilde{X}_2)$  where  $\tilde{X}_i$  has density  $\tilde{f}_i$  and  $\tilde{X}_1 \leq \tilde{X}_2$  a.s. Then  $Y_1 \leq Y_2$  a.s.,  $P(Y_1 = Y_2) = 1 - \varepsilon$  and  $Y_i$  has the same distribution as  $X_i$ . Consider the regular conditional distribution of  $Y_2$  given  $Y_1$ , and let  $g(x, u)$  be the generalized inverse of the conditional cumulative distribution function of  $Y_2$  given  $Y_1 = x$ . Then  $(X_1, g(X_1, U))$  has the same joint distribution as  $(Y_1, Y_2)$ , and thus  $g$  satisfies the requirements of the lemma. □

**A.2. Hyperbolic Brownian motion estimates.** The first two lemmas give estimates on the behavior of the process  $d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$  where  $\mathcal{B}$  is a (standard) hyperbolic Brownian motion.

LEMMA 19. *There is a coupling of a hyperbolic Brownian motion  $\mathcal{B}$  and a 2-dimensional standard Brownian motion  $W$  so that almost surely for all  $t \geq 0$  we have*

$$|W(t)| \leq d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t)) \leq |W(t)| + t/2.$$

PROOF. The process  $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$  satisfies the SDE

$$dq = db + \frac{\coth q}{2} dt, \quad q(0) = 0,$$

where  $b$  is a standard Brownian motion. This follows, for example, from (10) and the fact that the half-plane representation  $x + iy$  of  $\mathcal{B}$  satisfies the SDE  $d(x + iy) = y(dB_1 + i dB_2)$  for  $\mathcal{B}$ .

Consider the following diffusions with the same driving Brownian motion  $b$  as in  $q$ :

$$dq_1 = db + \frac{1}{2q_1} dt, \quad dq_2 = db + \left(\frac{1}{2q_2} + \frac{1}{2}\right) dt, \quad q_1(0) = q_2(0) = 0.$$

Since  $\frac{1}{x} \leq \coth x \leq \frac{1}{x} + 1$  for  $x > 0$ , we have a.s.  $q_1 \leq q \leq q_2$ . (This follows from standard comparison theorems; see, e.g., Ikeda and Watanabe (1989).) The process  $q_1$  is a 2-dimensional Bessel process, and it can be written as the absolute value of a 2-dimensional Brownian motion  $W$ .

For the upper bound, note that we have  $q_1 \leq q_2$ , and taking the difference of the SDEs for  $q_2$  and  $q_1$  we get

$$(q_2 - q_1)' = \frac{1}{2q_2} - \frac{1}{2q_1} + \frac{1}{2} \leq \frac{1}{2}.$$

Thus  $q(t) \leq q_2(t) \leq q_1(t) + t/2 = |W(t)| + t/2$ , which completes the proof.  $\square$

LEMMA 20. *Let  $\mathcal{B}$  be a standard hyperbolic Brownian motion. Then for  $t > 0, a > 0$  we have*

$$(67) \quad P\left(\max_{0 \leq s \leq t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \leq a\right) \leq \frac{4}{\pi} e^{-\frac{\pi^2 t}{8a^2}}.$$

If  $0 < t \leq a$ , then

$$(68) \quad P\left(\max_{0 \leq s \leq t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \geq a\right) \leq \frac{16\sqrt{t}}{a\sqrt{\pi}} e^{-\frac{a^2}{16t}}.$$

PROOF. By Lemma 19, the process  $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$  stochastically dominates  $|W(t)|$  where  $W = (B_1, B_2)$  is a 2-dimensional Brownian motion. Then

$$\begin{aligned} P\left(\max_{0 \leq s \leq t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \leq a\right) &\leq P\left(\max_{0 \leq s \leq t} |W(s)| \leq a\right) \\ &\leq P\left(\max_{0 \leq s \leq t} |B_1(s)| \leq a\right) \leq \frac{4}{\pi} e^{-\frac{\pi^2 t}{8a^2}}, \end{aligned}$$

which proves (67). The last step follows from the following identity for the standard Brownian motion  $B_1$  (see, e.g., Section 7.4 of Mörters and Peres (2010)):

$$P\left(\max_{0 \leq s \leq t} |B_1(s)| \leq u\right) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-\frac{(2k+1)^2 \pi^2 t}{8u^2}}.$$

From Lemma 19, it follows that  $q_t = d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t))$  is stochastically dominated by the process  $|W(t)| + t/2$  where  $W$  is a 2-dimensional standard Brownian motion. Thus

$$\begin{aligned} &P\left(\max_{0 \leq s \leq t} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(s)) \geq a\right) \\ &\leq P\left(\max_{0 \leq s \leq t} (|W(s)| + s/2) \geq a\right) \leq P\left(\max_{0 \leq s \leq t} |W(s)| \geq a - t/2\right) \\ &\leq 4P\left(\max_{0 \leq s \leq t} B(s) \geq \frac{1}{\sqrt{2}}(a - t/2)\right) = 4P\left(|B(t)| \geq \frac{1}{\sqrt{2}}(a - t/2)\right) \\ &\leq \frac{8\sqrt{t}}{(a - t/2)\sqrt{\pi}} e^{-\frac{(a-t/2)^2}{4t}} \leq \frac{16\sqrt{t}}{a\sqrt{\pi}} e^{-\frac{a^2}{16t}}. \end{aligned}$$

In the second line, we used that  $W(s) \geq x$  implies that one of the coordinates of  $W$  is at least  $\frac{1}{\sqrt{2}}x$  in absolute value, then used the reflection principle. Finally, we used the well-known tail bound for the normal distribution and  $t \leq a$ .  $\square$

The next lemma shows that  $\mathcal{B}$  approaches its limit point with speed  $1/2$ . (Note that proved bound is not optimal.)

LEMMA 21. Let  $\mathcal{B}$  be a hyperbolic Brownian motion and let  $z(t)$  be the point moving with speed  $1/2$  on the geodesic connecting  $\mathcal{B}(0)$  to  $\mathcal{B}(\infty) = \lim_{t \rightarrow \infty} \mathcal{B}(t)$ . Then there is a random  $C < \infty$  so that almost surely

$$d_{\mathbb{H}}(\mathcal{B}(t), z(t)) \leq C + t^{1/2} \log(1 + t)$$

for all  $t \geq 0$ .

PROOF. Consider the half-plane representation of  $\mathbb{H}$  where  $\mathcal{B}(0) = i$  and  $\mathcal{B}(\infty) = \infty$ , and denote the representation of  $\mathcal{B}$  by  $x + iy$ . Then  $x + iy$  is a hyperbolic Brownian motion conditioned to hit  $\infty$ , in particular, it satisfies

$$dy = y dB_1 + dt, \quad dx = y dB_2, \quad y(0) = 1, \quad x(0) = 0,$$

where  $B_1, B_2$  are independent standard Brownian motions. (See, e.g., Valkó and Virág (2017).) The geodesic connecting  $i$  and  $\infty$  is  $\{ie^t, t \geq 0\}$ , and the point moving with speed  $1/2$  is  $z(t) = ie^{t/2}$ . By the triangle inequality,

$$\begin{aligned} d_{\mathbb{H}}(\mathcal{B}(t), z(t)) &\leq d_{\mathbb{H}}(x + iy, iy) + d_{\mathbb{H}}(iy, z(t)) \\ (69) \qquad \qquad \qquad &= \operatorname{arccosh}\left(1 + \frac{x^2}{2y^2}(t)\right) + |\log(y(t)e^{-t/2})| \\ &\leq \log\left(2 + \frac{x^2}{y^2}(t)\right) + |\log(y(t)e^{-t/2})|. \end{aligned}$$

We can explicitly solve for  $y$  and  $x$  from the SDE:

$$y(t) = e^{B_1(t)+t/2}, \quad x(t) = \int_0^t e^{B_1(s)+s/2} dB_2(s).$$

We have  $|\log(y(t)e^{-t/2})| = |B_1(t)|$ , and using the law of iterated logarithm we get the bound

$$(70) \qquad \qquad \qquad |\log(ye^{-t/2})| = |B_1(t)| \leq C_0 + \sqrt{t \log(1 + t)}$$

for all  $t \geq 0$  with some random  $C_0$  depending on  $B_1$ .

To bound  $\log(2 + \frac{x^2}{y^2})$ , we start with the observation that there is a standard Brownian motion  $B$  so that  $x(t) = B(\int_0^t y(s)^2 ds)$ . Thus, using the law of iterated logarithm again,

$$\begin{aligned} \left(\frac{x(t)}{y(t)}\right)^2 &\leq y(t)^{-2} \left(2C_0^2 + 2 \int_0^t y(s)^2 ds \log\left(1 + \int_0^t y(s)^2 ds\right)\right) \\ &= 2C_0 e^{-2B(t)-t} + 2 \int_0^t e^{2(B(s)-B(t))-t+s} ds \log\left(1 + \int_0^t e^{2B(s)+s} ds\right). \end{aligned}$$

Using the bound  $|B(s)| \leq C_0 + \sqrt{s \log(1 + s)}$  for  $0 \leq s \leq t$  leads to

$$\left(\frac{x(t)}{y(t)}\right)^2 \leq C_1(1 + te^{4\sqrt{t \log(1+t)}})$$

with a random  $C_1$  depending only on  $\mathcal{B}(\cdot)$ . Using (69) and (70), the statement follows.  $\square$

The next statement gives an estimate on the modulus of continuity of the hyperbolic Brownian motion. The proof follows that of the analogous statement for standard Brownian motion, we include it for completeness. The constants have not been optimized.

PROPOSITION 22. *Let  $\mathcal{B}$  be a hyperbolic Brownian motion. Then there is a random constant  $0 < h_0 \leq 1$  so that a.s.,*

$$(71) \quad d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+h)) \leq 20\sqrt{h \log\left(2 + \frac{s+1}{h}\right)},$$

for all  $0 < h \leq h_0$  and  $0 \leq s$ .

Note that the proof below also shows that there is a random constant  $C$  so that (71) holds for all  $0 < h \leq 1$  with  $C$  in place of 20.

PROOF. Let  $I_{m,n} = [m2^{-n}, (m+1)2^{-n}]$  and  $\Delta_{m,n} = \max_{t \in I_{m,n}} d_{\mathbb{H}}(\mathcal{B}(m2^{-n}), \mathcal{B}(t))$  for  $m, n \geq 0$ . If  $2^{-n/2} < u$ , then by Lemma 20 we have

$$P(\Delta_{m,n} \geq 2^{-n/2}u) = P\left(\max_{0 \leq t \leq 2^{-n}} d_{\mathbb{H}}(\mathcal{B}(0), \mathcal{B}(t)) \geq 2^{-n/2}u\right) \leq \frac{16}{u\sqrt{\pi}}e^{-\frac{u^2}{16}}.$$

Thus for  $n \geq 0, m \geq 0$  we get

$$P\left(\Delta_{m,n} \geq \frac{9}{2} \cdot 2^{-n/2} \sqrt{\log(2^n + m + 1)}\right) \leq \frac{3(2^n + m + 1)^{-5/4}}{\sqrt{\log(2^n + m + 1)}}.$$

We have  $\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} (2^n + m + 1)^{-5/4} < \infty$ . By the Borel–Cantelli lemma, there is a random  $N_0 \geq 1$  so that if  $n \geq N_0$  then

$$(72) \quad \Delta_{m,n} \leq \frac{9}{2} \sqrt{2^{-n} \log(2^n + m + 1)}.$$

We will show that (71) holds with  $0 < h \leq h_0 = 2^{-N_0}$  and  $0 \leq s$ . Let  $m, n$  be nonnegative integers with  $2^{-n-1} < h \leq 2^{-n}$  and  $m2^{-n} \leq s \leq (m+1)2^{-n}$ . Then we have  $n \geq N_0$ , and using the triangle inequality and (72) we get

$$\begin{aligned} d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(s+h)) &\leq 2\Delta_{m,n} + \Delta_{m+1,n} \leq 3 \cdot \frac{9}{2} \sqrt{2^{-n} \log(2^n + m + 2)} \\ &\leq 20\sqrt{h \log\left(2 + \frac{s+1}{h}\right)}, \end{aligned}$$

which completes the proof.  $\square$

Our next proposition gives a lower bound on the fluctuations of the hyperbolic Brownian motion. The proof again follows that of the analogous statement for standard Brownian motion.

PROPOSITION 23. *Let  $\mathcal{B}$  be a hyperbolic Brownian motion. Then there is a random constant  $C_0 < \infty$  so that a.s. for any  $0 \leq s < t$  there exists  $u, v \in [s, t]$  with*

$$(73) \quad d_{\mathbb{H}}(\mathcal{B}(u), \mathcal{B}(v))^2 \geq C_0 \frac{t-s}{\log(2+t) + \log(t-s + \frac{1}{t-s})}.$$

Moreover, with the same constant  $C_0$  a.s. for any  $0 \leq s < t$  there exists  $u \in [s, t]$  with

$$(74) \quad d_{\mathbb{H}}(\mathcal{B}(s), \mathcal{B}(v))^2 \geq \frac{1}{4} C_0 \frac{t-s}{\log(2+t) + \log(t-s + \frac{1}{t-s})}.$$

PROOF. The second part of the statement follows from the first part using the triangle inequality.

To prove (73), it is enough to show the statement for pairs of the form  $s = m2^n, t = (m + 1)2^n$  with  $m, n \in \mathbb{Z}$  and  $m > 0$ . We partition the interval  $[s, t]$  into  $k = \lceil 1 + |n| + 5 \log(2 + m) \rceil$  subintervals  $[a_i, a_{i+1})$  of size  $\frac{2^n}{k}$ . For a given  $0 \leq i < k$ , we have

$$\begin{aligned} P\left(d_{\mathbb{H}}(\mathcal{B}(a_i), \mathcal{B}(a_{i+1})) \leq \frac{2^{n/2}}{\sqrt{k}}\right) &= P\left(d_{\mathbb{H}}\left(\mathcal{B}(0), \mathcal{B}\left(\frac{2^n}{k}\right)\right) \leq \frac{2^{n/2}}{\sqrt{k}}\right) \\ &\leq P\left(\left|B\left(\frac{2^n}{k}\right)\right| \leq \frac{2^{n/2}}{\sqrt{k}}\right) \leq P(|B(1)| \leq 1) \leq \frac{4}{5}, \end{aligned}$$

where  $B(t)$  is a standard Brownian motion and we used Lemma 19. Using the Markov property of the hyperbolic Brownian motion, we get that

$$P\left(d_{\mathbb{H}}(\mathcal{B}(a_i), \mathcal{B}(a_{i+1})) \leq \frac{2^{n/2}}{\sqrt{k}} \text{ for all } 0 \leq i < k\right) \leq \left(\frac{4}{5}\right)^k \leq \left(\frac{4}{5}\right)^{|n|+5 \log(2+m)}.$$

Since  $\sum_{m=1}^{\infty} \sum_{n \in \mathbb{Z}} \left(\frac{4}{5}\right)^{|n|+5 \log(2+m)} < \infty$ , there are a.s. finitely many pairs  $m, n$  for which we cannot find  $u, v \in [m2^n, (m + 1)2^n]$  with

$$d_{\mathbb{H}}(\mathcal{B}(u), \mathcal{B}(v))^2 \leq \frac{2^n}{5 \log(2 + m) + |n|}.$$

Now (73) follows with a random  $C_0$  for pairs of the form  $s = m2^n, t = (m + 1)2^n$ , and from this (73) follows with a modified  $C_0$  for all  $s, t$ .  $\square$

**Acknowledgments.** We thank Paul Bourgade, Yun Li and Bálint Vető for interesting discussions and comments. We thank the anonymous referee for valuable comments and suggestions, especially for suggesting to investigate the  $\beta' = \infty$  case in Theorem 3.

The first author was supported in part by NSF CAREER award DMS-1053280, NSF award DMS-1712551 and the Simons Foundation.

The second author was supported by the Canada Research Chair program the NSERC Discovery Accelerator grant, the MTA Momentum Random Spectra research group and the ERC consolidated grant 648017 (Abért).

REFERENCES

ALLEZ, R. and DUMAZ, L. (2014). From sine kernel to Poisson statistics. *Electron. J. Probab.* **19** no. 114, 25. MR3296530 <https://doi.org/10.1214/EJP.v19-3742>

ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

BHATIA, R. and ELSNER, L. (1994). The Hoffman–Wielandt inequality in infinite dimensions. *Proc. Indian Acad. Sci. Math. Sci.* **104** 483–494. MR1314392 <https://doi.org/10.1007/BF02867116>

BOURGADE, P., NAJNUDEL, J. and NIKEGHBALI, A. (2013). A unitary extension of virtual permutations. *Int. Math. Res. Not. IMRN* **18** 4101–4134. MR3106884 <https://doi.org/10.1093/imrn/rns167>

DAVIES, E. B. and MANDOUVALOS, N. (1988). Heat kernel bounds on hyperbolic space and Kleinian groups. *Proc. Lond. Math. Soc.* (3) **57** 182–208. MR0940434 <https://doi.org/10.1112/plms/s3-57.1.182>

DUMITRIU, I. and EDELMAN, A. (2002). Matrix models for beta ensembles. *J. Math. Phys.* **43** 5830–5847. MR1936554 <https://doi.org/10.1063/1.1507823>

EDELMAN, A. and SUTTON, B. D. (2007). From random matrices to stochastic operators. *J. Stat. Phys.* **127** 1121–1165. MR2331033 <https://doi.org/10.1007/s10955-006-9226-4>

FORRESTER, P. J. (2010). *Log-Gases and Random Matrices. London Mathematical Society Monographs Series* **34**. Princeton Univ. Press, Princeton, NJ. MR2641363 <https://doi.org/10.1515/9781400835416>

HOLCOMB, D. and VALKÓ, B. (2015). Large deviations for the  $\text{Sine}_\beta$  and  $\text{Sch}_\tau$  processes. *Probab. Theory Related Fields* **163** 339–378. MR3405620 <https://doi.org/10.1007/s00440-014-0594-4>

- IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. [MR1011252](#)
- KARPELEVIČ, F. I., TUTUBALIN, V. N. and ŠUR, M. G. (1959). Limit theorems for compositions of distributions in the Lobačevskiĭ plane and space. *Theory Probab. Appl.* **4** 399–402.
- KILLIP, R. and NENCIU, I. (2004). Matrix models for circular ensembles. *Int. Math. Res. Not.* **50** 2665–2701. [MR2127367](#) <https://doi.org/10.1155/S1073792804141597>
- KILLIP, R. and STOICIU, M. (2009). Eigenvalue statistics for CMV matrices: From Poisson to clock via random matrix ensembles. *Duke Math. J.* **146** 361–399. [MR2484278](#) <https://doi.org/10.1215/00127094-2009-001>
- MAPLES, K., NAJNUDEL, J. and NIKEGBALI, A. (2019). Strong convergence of eigenangles and eigenvectors for the circular unitary ensemble. *Ann. Probab.* **47** 2417–2458. [MR3980924](#) <https://doi.org/10.1214/18-AOP1311>
- MECKES, E. S. and MECKES, M. W. (2016). Self-similarity in the circular unitary ensemble. *Discrete Anal.* Paper No. 9, 14. [MR3555192](#) <https://doi.org/10.19086/da.736>
- MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics (Amsterdam)* **142**. Elsevier/Academic Press, Amsterdam. [MR2129906](#)
- MÖRTERS, P. and PERES, Y. (2010). *Brownian Motion. Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge Univ. Press, Cambridge. [MR2604525](#) <https://doi.org/10.1017/CBO9780511750489>
- NAKANO, F. (2014). Level statistics for one-dimensional Schrödinger operators and Gaussian beta ensemble. *J. Stat. Phys.* **156** 66–93. [MR3215116](#) <https://doi.org/10.1007/s10955-014-0987-x>
- RAMÍREZ, J. A. and RIDER, B. (2009). Diffusion at the random matrix hard edge. *Comm. Math. Phys.* **288** 887–906. [MR2504858](#) <https://doi.org/10.1007/s00220-008-0712-1>
- RAMÍREZ, J. A., RIDER, B. and VIRÁG, B. (2011). Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.* **24** 919–944. [MR2813333](#) <https://doi.org/10.1090/S0894-0347-2011-00703-0>
- VALKÓ, B. and VIRÁG, B. (2009). Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** 463–508. [MR2534097](#) <https://doi.org/10.1007/s00222-009-0180-z>
- VALKÓ, B. and VIRÁG, B. (2017). The  $\text{Sine}_\beta$  operator. *Invent. Math.* **209** 275–327. [MR3660310](#) <https://doi.org/10.1007/s00222-016-0709-x>
- WIGNER, E. P. (1951). On a class of analytic functions from the quantum theory of collisions. *Ann. of Math.* (2) **53** 36–67. [MR0039059](#) <https://doi.org/10.2307/1969342>