Bi-parameter Potential Theory and Carleson Measures for the Dirichlet Space on the Bidisc

Nicola Arcozzi[*]   Pavel Mozolyako[†]   Karl-Mikael Perfekt[‡]

Giulia Sarfatti[§]

Abstract: We characterize the Carleson measures for the Dirichlet space on the bidisc, hence also its multiplier space. Following Maz'ya and Stegenga, the characterization is given in terms of a capacitary condition. We develop the foundations of a bi-parameter potential theory on the bidisc and prove a Strong Capacitary Inequality. In order to do so, we have to overcome the obstacle that the Maximum Principle fails in the bi-parameter theory.

1 Introduction

Notation. We denote by $\mathbb{D}$ the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ in the complex plane and by $\partial\mathbb{D}$ its boundary. We write $A \lesssim B$ $(A \gtrsim B)$ if there is a constant independent on the variables on which A and B depend (which might be numbers, variables, sets...) such that $A \leq CB$ ($CA \geq B$ respectively), and $A \approx B$, if $A \lesssim B$ and $A \gtrsim B$.

NICOLA ARCOZZI, PAVEL MOZOLYAKO, KARL-MIKAEL PERFEKT, AND GIULIA SARFATTI

In 1979, Alice Chang [16], extending a foundational result of Carleson [14] in one variable, characterized the Carleson measures for the bi-harmonic Hardy space of the bidisc, that is, those measures μ on $\mathbb{D}^2$ such that the identity operator boundedly maps $h^2(\mathbb{D}) \otimes h^2(\mathbb{D})$ into $L^2(\mu)$. In [15] Carleson had previously shown that the bi-parameter theory presents very peculiar features. Results in bi-parameter potential theory and harmonic analysis are scattered in the literature. See however Chapter 3 in the monograph [24], and the references mentioned at p. 124, and the lectures notes [17]. At the same time, Stegenga [27] characterized the Carleson measures for the holomorphic Dirichlet space on the unit disc. Following standard use in complex function theory, we say that a measure μ is a *Carleson measure for the Hilbert function space H* if H continuously embeds into $L^2(\mu)$.

Carleson measures proved to be a central notion in the analysis of holomorphic spaces, as they intervene in the characterization of multipliers, interpolating sequences, and Hankel-type forms, in Corona-type problems, in the characterization of exceptional sets at the boundary, and more. In this article we characterize the Carleson measures for the Dirichlet space on the bidisc, and we obtain as a consequence a characterization of its multiplier space.

As the Dirichlet space is defined by a Sobolev norm, it is not surprising that Stegenga's characterization is given in terms of a potential theoretic object, set capacity, and that the proof relies on deep results from Potential Theory, such as the Strong Capacitary Inequality. The main effort in this article is developing a bi-parameter potential theory which is rich enough to state and prove the characterization theorem. There are obstructions to doing so, which we will illustrate below.

Other approaches to similar problems have been suggested in the past. The closest result is Eric Sawyer's characterization of the weighted inequalities for the bi-parameter Hardy operator [25]. Sawyer's extremely clever combinatorial-geometric argument does not seem to work in our context, or at least we were not able to make it work. The difficulty lies in the fact that Sawyer deals with the product of two segments, while we work with the product of two hyperbolic discs. For similar reasons, we were not able to extend the *good-lambda* argument in [7] to the bi-parameter case. The simple approach via maximal functions in [8] could work, if knowledge concerning weighted maximal bi-parameter functions was more developed. The difficulty is that, contrary to the linear case, bi-parameter maximal functions do not always satisfy weighted L^2 inequalities. We refer to [13] and [17] for early, detailed accounts of two-parameter processes and their associated maximal functions, and [19] for seminal results on bi-parameter maximal functions. In the one parameter case, Carleson measures for the Dirichlet space can also be characterized using a Bellman function argument [5]. At the moment, however, the Bellman function technique, having at its heart stochastic optimization for martingales, does not work in the two time-parameter martingale theory underlying bi-parameter Potential Theory. The scheme of dualizing the embedding to translate the problem from bi-parameter–holomorphic to bi-parameter–dyadic has been borrowed from [8, 9].

We now state our results more precisely. Let $\mathcal{D}(\mathbb{D})$ be the Dirichlet space on the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$; that is, the space of the functions $f(z) = \sum_{n=0}^{\infty} \widehat{f}(n) z^n$, analytic in $\mathbb{D}$, such that the norm

$$\|f\|_{\mathcal{D}}^2 := \sum_{n=0}^{\infty} (n+1) \left|\widehat{f}(n)\right|^2 = \|f\|_{H^2(\mathbb{D})}^2 + \frac{1}{\pi} \int_{\mathbb{D}} \left|f'(z)\right|^2 dA(z) < \infty. \tag{1}$$

The Dirichlet space on the bidisc $\mathbb{D}^2$ can be temporarily defined as $\mathcal{D}(\mathbb{D}^2) := \mathcal{D}(\mathbb{D}) \otimes \mathcal{D}(\mathbb{D})$. The main aim of this article is proving the following.

Theorem 1.1 *Let $\mu \geq 0$ be a Borel measure on $\overline{\mathbb{D}}^2$, the closure of the bidisc. Then the following are equivalent.*

(I) *There is a constant $C_1 > 0$ such that*

$$\sup_{0 \leq r < 1} \int_{r\mathbb{D}^2} |f|^2 d\mu \leq C_1 \|f\|^2_{\mathcal{D}(\mathbb{D}^2)}, \quad f \in \mathcal{D}(\mathbb{D}^2); \tag{2}$$

(II) *There is a constant $C_2 > 0$ such that for all $n \geq 1$ and for all choices of arcs $J_1^1, \ldots, J_n^1$ and $J_1^2, \ldots, J_n^2$ on $\partial\mathbb{D}$, we have that*

$$\mu \left(\bigcup_{k=1}^{n} S\left(J_k^1 \times J_k^2\right) \right) \leq C_2 \operatorname{Cap}_{\left(\frac{1}{2}, \frac{1}{2}\right)} \left(\bigcup_{k=1}^{n} J_k^1 \times J_k^2 \right). \tag{3}$$

Moreover, the constants C_1 and C_2 are comparable independently of μ.

Here $S\left(J^1 \times J^2\right) = S\left(J^1\right) \times S\left(J^2\right)$ is the Carleson box in $\overline{\mathbb{D}}^2$ based on $J^1 \times J^2$ and $\operatorname{Cap}_{\left(\frac{1}{2}, \frac{1}{2}\right)}$ is the canonical extension of $\frac{1}{2}$-Bessel capacity from linear potential theory to bi-parameter potential theory, which will be defined later in the article. It can be estimated from above and below by the capacity which is naturally associated with the reproducing kernel of $\mathcal{D}(\mathbb{D}^2)$, and several other versions of capacity. In the one parameter case, Stegenga (see [27, Theorem 4.2]) exhibited examples of measures λ satisfying the appropriate version of (3) for single arcs, but not for arbitrary unions of arcs. It is easy to see that the measure $\mu = \lambda \otimes m$, where m is arclength, provides an example where (3) holds for rectangles, yet fails for arbitrary unions of them.

We can, for the moment informally, view (2) as the boundedness of the imbedding $Id : \mathcal{D}(\mathbb{D}^2) \hookrightarrow L^2(\mu)$. A measure satisfying (2) is a *Carleson measure* for $\mathcal{D}(\mathbb{D}^2)$ and we define

$$[\mu]_{CM} := \|Id\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2), L^2(\mu))}$$

as its *Carleson measure norm*. Actually, the result has a stronger version, which we will prove, where in the left hand side of (I), f is replaced by its *radial variation*, whose main contribution is given by

$$\operatorname{Var}_{12} f(\zeta, \xi) = \int_0^1 \int_0^1 |\partial_{zw} f(r\zeta, s\xi)| \, dr \, ds, \ (\zeta, \xi) \in \overline{\mathbb{D}}^2 \tag{4}$$

The full definition appears in Section 2.6. In particular, this allows us to recover, on the bidisc, Beurling's result [12] on exceptional sets for the radial variation of functions in the Dirichlet space.

A function b holomorphic in $\mathbb{D}^2$ is a *multiplier* of $\mathcal{D}(\mathbb{D}^2)$ if multiplication times b, $M_b : f \mapsto bf$, is bounded on $\mathcal{D}(\mathbb{D}^2)$. The operator norm of M_b is, by definition, the multiplier norm of b. From Theorem 1.1 we deduce a characterization of multipliers.

Theorem 1.2 *Let b be holomorphic in $\mathbb{D}^2$ and define the measure $d\mu_b := |\partial_{zw} b(z, w)|^2 dA(z) dA(w)$. There exist positive constants C_1, C_2 such that:*

$$\|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \approx [\mu_b]_{CM} + \|b\|^2_{H^\infty} + \sup_{w \in \mathbb{D}} [|\partial_z b(\cdot, w)|^2 dA(\cdot)]_{CM(\mathcal{D}(\mathbb{D}))} + \sup_{z \in \mathbb{D}} [|\partial_w b(z, \cdot)|^2 dA(\cdot)]_{CM(\mathcal{D}(\mathbb{D}))}, \tag{5}$$

where $[\cdot]_{CM(\mathcal{D}(\mathbb{D}))}$ denotes the Carleson measure norm with respect to the Dirichlet space in the disc.

NICOLA ARCOZZI, PAVEL MOZOLYAKO, KARL-MIKAEL PERFEKT, AND GIULIA SARFATTI

Up to this point the results exactly match those obtained by Stegenga in 1980. The proof, however, has to overcome a series of obstructions. The most prominent one is that the potentials of positive measures do not satisfy a maximum principle: the supremum of the potential $\mathbb{V}^\mu$ of a measure μ can be much larger than its supremum on the support of μ. Essentially, this is due to the fact that the product of distance functions on metric spaces X and Y is usually very far from itself being a distance on $X \times Y$. To oversimplify a large body of knowledge, a far reaching Potential Theory can be developed if the reciprocal of the defining kernel behaves like a distance, satisfying the triangle inequality up to a constant factor. The product of two copies of such kernels, however, fails to satisfy this property.

In order to isolate the essential difficulties, it is convenient to transfer the holomorphic problem to a dyadic one. In the sequel, T will denote a set of vertices labelling dyadic arcs in $\partial\mathbb{D}$, which might be seen as an oriented dyadic tree with respect to the relation $I \subset J$: we identify a dyadic arc with a vertex of T, and an oriented edge of T can be thought of as an inclusion relation $J \subset I$ of a a dyadic arc in the dyadic arc of twice its length. To each arc I we associate the usual Carleson box $S(I)$ and the Whitney square $Q(I)$ consisting of its upper half, $Q(I) = S(I) \setminus \cup_{J \subsetneq I} S(J)$.

It is convenient to set up a different notation for objects related to $\mathbb{D}$ and objects related to T: we will assign to a dyadic arc I a label α_I in T and, vice versa, each α in T will be the label of some dyadic arc $I_\alpha = I(\alpha)$. The *root* of T is o, where $I_o = \partial\mathbb{D}$. Also, T has a natural boundary ∂T, endowed with a metric which makes $\overline{T} = T \cup \partial T$ into a compact space. We can think of an element ζ in ∂T as the label for an infinite, decreasing sequence of dyadic arcs: $P(\zeta) := \{I_n\}_{n=0}^\infty$, with $I_n \supset I_{n+1}$, and $|I_n|/2\pi = 2^{-n}$. We write $\alpha \leq \beta$ if $I_\alpha \subseteq I_\beta$. This convention is different from the one used in [7]. To each α in $\overline{T}$ we associate the subset $P(\alpha) = \{\beta : \alpha \leq \beta \leq o\}$. To each α we also associate the region $S(\alpha) = \{\zeta \in \overline{T} : \alpha \in P(\zeta)\}$.

The geometric objects defined in the disc have natural counterparts in the bidisc: we define $S(J^1 \times J^2) = S(J^1) \times S(J^2)$, $Q(J^1 \times J^2) = Q(J^1) \times Q(J^2)$, and so on. The bitree $T^2 = T \times T$ labels dyadic rectangles. If $\alpha = (\alpha_z, \alpha_w) \in T \times T$, we associate to it $J(\alpha) = J(\alpha_z) \times J(\alpha_w)$. A basic fact is that T^2, identified with the set of the dyadic rectangles $J^1 \times J^2$, does not have a tree structure with respect to inclusion. We can move positive Borel measures from $\overline{\mathbb{D}}^2$ to $\overline{T}^2$, in a way which will be made precise later: to each Borel measure μ on $\overline{\mathbb{D}}^2$ we associate a unique Borel measure $\Lambda^*\mu$ on $\overline{T}^2$ and, vice-versa, to each Borel measure ν on $(\partial T)^2$ we can associate a unique Borel measure $\Lambda_*\nu$ on $(\partial\mathbb{D})^2$ (here we restrict ourselves to measures supported on $(\partial T)^2$, since we do not really need to transplant the measure on the rest of the bitree to the bidisc). Essentially, $\Lambda_*\nu$ associates in natural way a measure on the *distinguished boundary* $(\partial\mathbb{D})^2$ of $\mathbb{D}^2$ to the restriction of ν to the *distinguished boundary* $(\partial T)^2$ of T^2. Conversely, $\Lambda^*\mu$ concentrates the measure of $Q(I_\alpha) \subseteq \mathbb{D}^2$ into $\alpha \in T^2$, it essentially preserves the measures on the distinguished boundaries, and acts in a mixed way on the remaining parts of the boundaries. The precise definitions of Λ_* and Λ^* are slightly technical and will be given later.

We define a natural *bi-parameter Hardy operator* $\mathbb{I}$ acting on functions $\varphi : \overline{T}^2 \to \mathbb{R}$,

$$\mathbb{I}\varphi(\zeta) := \sum_{\alpha \in P(\zeta)} \varphi(\alpha), \tag{6}$$

provided the sum makes sense. This is certainly the case if $\varphi \geq 0$, which is what we will assume throughout the article. We define the operator $\mathbb{I}$ on real valued functions in order to have for its adjoint the usual definition. The operator $\mathbb{I}$ is analogous to the bi-parameter Hardy operator studied by Eric Sawyer in [25], and it is the bi-parameter version of the operator I introduced in [7]. Dually, we have the operator

$\mathbb{I}^*$ acting on (a priori, signed) Borel measures on $\overline{T}^2$, $\mathbb{I}^*\mu(\alpha) := \mu(S(\alpha))$. These simple operators encode all relevant information.

Theorem 1.3 *Let μ be a positive, Borel measure on $\overline{\mathbb{D}}^2$. Then*

$$C_1[\mu]_{CM} \leq \|\mathbb{I}\|^2_{\mathcal{B}(\ell^2(T^2), L^2(\Lambda^*\mu))} \leq C_2[\mu]_{CM}, \tag{7}$$

where C_1 and C_2 are universal constants. Moreover, we can replace $[\mu]_{CM}$ by the larger quantity

$$\sup_{\|f\|_{\mathcal{D}(\mathbb{D}^2)}=1} \|\mathrm{Var}\, f\|^2_{L^2(\overline{\mathbb{D}}^2, d\mu)},$$

*where the **radial variation** of f, $\mathrm{Var}\, f$, is defined in (29), and has (4) as its leading summand.*

With this theorem at hand, the characterization of the Carleson measures for $\mathcal{D}(\mathbb{D}^2)$ can be reduced to that of estimating the quantity $\|\mathbb{I}\|^2_{\mathcal{B}(\ell^2(T^2), L^2(\nu))}$, and it suffices to consider the case of nonnegative functions. In the linear case, in [7] this was done in terms of a Kerman–Sawyer [21] type testing condition. We will instead follow the capacitary path introduced by Maz'ya [23], then Adams [1], in proving sharp trace inequalities, which was transplanted by Stegenga [27] to the holomorphic world.

Following a general scheme [2, Sections 2.3-2.5], the operators $\mathbb{I}$ and $\mathbb{I}^*$ can interpreted in terms of a potential theory on $(\overline{T})^2 \times T^2$. More precisely, we consider the potential kernel $k(\zeta, \alpha) = \chi(\alpha \in P(\zeta)) = \chi(\zeta \in S(\alpha))$. Following [2], if $E \subseteq \overline{T}^2$, then we define *discrete bi-logarithmic capacity* by

$$\mathrm{Cap}(E) = \inf\left\{\|\varphi\|^2_{\ell^2(T^2)} : \mathbb{I}\varphi(\zeta) \geq 1 \text{ if } \zeta \in E\right\}. \tag{8}$$

The *trace inequality* we wish to prove is

Theorem 1.4 *There are positive constants C_1, C_2 such that, if ν is a Borel measure on $\overline{T}^2$, then*

$$C_1 \sup_{n \geq 1; \alpha_1, \ldots, \alpha_n \in T^2} \frac{\nu\left(\cup_{j=1}^n S(\alpha_j)\right)}{\mathrm{Cap}\left(\cup_{j=1}^n \partial S(\alpha_j)\right)} \leq \|\mathbb{I}\|^2_{\mathcal{B}(\ell^2(T^2), L^2(\nu))} \leq C_2 \sup_{n \geq 1; \alpha_1, \ldots, \alpha_n \in T^2} \frac{\nu\left(\cup_{j=1}^n S(\alpha_j)\right)}{\mathrm{Cap}\left(\cup_{j=1}^n \partial S(\alpha_j)\right)}. \tag{9}$$

Theorem 1.4 follows by a standard argument from a Strong Capacitary Inequality of Adams type [1].

Theorem 1.5 *There is a constant $C > 0$ such that, whenever $\varphi: T^2 \to [0, +\infty)$,*

$$\int_0^\infty \mathrm{Cap}(\{\zeta \in (\partial T)^2 : \mathbb{I}\varphi(\zeta) > \lambda\})d\lambda^2 \leq C\|\varphi\|^2_{\ell^2(T^2)}. \tag{10}$$

The Strong Capacitary Inequality is standard when Cap is the capacity associated to a radially decreasing kernel. This is the case, with sufficient approximation, in the linear case of a tree T, where Cap is associated with a Bessel-like kernel. See for example [2] for the general theory, [20] for the relation with semilinear equations, and [10] for case of trees and metric spaces. The literature on Bessel-like kernels is vast and we just mention a few titles. However, our capacity is associated with the tensor product of two Bessel-like kernels, which is itself very different from a Bessel kernel, in the same way that the tensor product of two distance functions is typically not a distance function. In particular, the Maximum Principle for potentials fails completely.

Proposition 1.1 *For any $\lambda \geq 1$ there exists a measure $\mu \geq 0$ on $(\partial T)^2$ such that $\mathbb{V}^\mu \leq 1$ on $\operatorname{supp}\mu$, but $\mathbb{V}^\mu(\zeta_0) \geq \lambda$ at some $\zeta_0 \in (\partial T)^2$. Moreover, we can take $\mu = \mu_E$ to be the equilibrium measure of some set $E \subseteq \partial T^2$.*

We go around this difficulty by proving an estimate showing that the set where the potential is large has small capacity. This is the main novelty of this article.

Theorem 1.6 *There is a positive constant C such that, if μ is the equilibrium measure of a Borel subset of $(\partial T)^2$ and $\lambda \geq 1$, one has:*

$$\operatorname{Cap}(\{\zeta \in (\partial T)^2 : \mathbb{V}^\mu(\zeta) > \lambda\}) \leq C\frac{\|\mathbb{I}^*\mu\|^2_{\ell^2(T^2)}}{\lambda^3}. \tag{11}$$

Theorem 1.6 is reduced to a new problem in linear Potential Theory on the tree T, which is solved. We have considered the Dirichlet space on the bidisc only. Some parts of our argument easily extend to more general environments; for instance, Theorem 1.2 extends to polydiscs. The dyadization scheme in Theorem 1.3 can be similarly extended to polytrees with any number of factors. We believe that the results can be extended using different powers in the definition of the Dirichlet norm: $1 < p < \infty$ is a natural choice. Weights could be taken into consideration. As this article enters unexplored territory, we have preferred to consider its most basic object: the unweighted Dirichlet space ($p = 2$) on the bidisc. Our results can also be used to prove trace inequalities in other contexts, using different dyadization schemes. We will return to this in other works. We have made an effort to provide all details of all proofs. We will point out, however, which parts of our arguments are, in our opinion, standard, and which are new.

Layout of the article

The paper is organized as follows. In Section 2 we prove Theorems 1.1 and 1.2 – modulo the Strong Capacitary Inequality – as well as several other statements mentioned above. We introduce the discrete model of the bidisc, move the problem there and solve the discrete version. The approach is adapted from one-dimensional techniques, and mostly follows [9] and [10]. We only present the general line of the argument, postponing technical details to the Appendix. Section 3 contains substantially new results. There we prove the Strong Capacitary Inequality on the bitree, which is a crucial part of our method. Section 4 contains some concluding reflections. In the Appendix, Section 5, we collect auxiliary results used or mentioned before, as well as some counterexamples.

2 Proof of Theorem 1: discretization and 'soft' argument

We start with introducing some notation and describe the properties of $\mathcal{D}(\mathbb{D}^2)$ that we use later.

Given a holomorphic function $f(z_1, z_2) = \sum_{m,n \geq 0} a_{mn} z_1^m z_2^n$ on $\mathbb{D}^2$ we let

$$\|f\|^2_{\mathcal{D}(\mathbb{D}^2)} = \sum_{m,n \geq 0} |a_{mn}|^2 (m+1)(n+1).$$

This norm can also be written as follows,

$$\|f\|_{\mathcal{D}(\mathbb{D}^2)}^2 = \frac{1}{\pi^2} \int_{\mathbb{D}^2} |\partial_{z_1,z_2} f(z_1,z_2)|^2 \, dA(z_1) \, dA(z_2) + \sup_{0 \le r_2 < 1} \frac{1}{2\pi^2} \int_{\partial\mathbb{D}} \int_{\mathbb{D}} |\partial_{z_1} f(z_1, r_2 e^{it})|^2 \, dA(z_1) \, dt +$$

$$\sup_{0 \le r_1 < 1} \frac{1}{2\pi^2} \int_{\mathbb{D}} \int_{\partial\mathbb{D}} |\partial_{z_2} f(r_1 e^{it}, z_2)|^2 \, ds \, dA(z_2) + \sup_{0 \le r_1, r_2 < 1} \frac{1}{4\pi^2} \int_{\partial\mathbb{D}} \int_{\partial\mathbb{D}} |f(r_1 e^{is}, r_2 e^{it})|^2 \, ds \, dt =$$

$$\|f\|_*^2 + \text{ other terms,}$$

where $\|f\|_*$ is a seminorm which is invariant under biholomorphisms of the bidisc. In what follows however we use an equivalent norm, arising from the representation $\mathcal{D}(\mathbb{D}^2) = \mathcal{D}(\mathbb{D}) \otimes \mathcal{D}(\mathbb{D})$. For $f \in \text{Hol}(\mathbb{D})$ let

$$\|f\|_{\mathcal{D}}^2 := \frac{1}{\pi} \int_{\mathbb{D}} |f'(z)|^2 \, dA(z) + C_0 |f(0)|^2, \tag{12}$$

where $C_0 > 0$ is a constant to be chosen shortly. It is classical fact that the Dirichlet space on the unit disc is a Reproducing Kernel Hilbert Space (RKHS) [3], and, consequently, $\mathcal{D}(\mathbb{D}^2)$ is one as well. The reproducing kernel K_z, $z \in \mathbb{D}^2$, generated by $\|\cdot\|_{\mathcal{D}}$, is

$$K_z(w) = \left(C_1 + \log \frac{1}{1 - \bar{z}_1 w_1} \right) \left(C_1 + \log \frac{1}{1 - \bar{z}_2 w_2} \right), \quad z, w \in \mathbb{D}^2 \tag{13}$$

a product of reproducing kernels for $\mathcal{D}(\mathbb{D})$ in respective variables. Here $C_1 = 1/C_0$. Hence K_z enjoys the following important property

$$\Re K_z(w) \approx |K_z(w)|, \quad z, w \in \mathbb{D}^2, \tag{14}$$

if we take C_1 to be large enough (see Lemma 5.1).

This Section is organized as follows. First we use duality arguments and the RKHS property of $\mathcal{D}(\mathbb{D}^2)$ to replace the Carleson Condition (2) with something more tangible, first doing so for measures supported strictly inside the bidisc. Then, in Section 2.2, we construct the discrete approximation of the bidisc – the bitree – and we introduce the basics of (logarithmic) Potential Theory there. Next we move all the objects from the bidisc to the bitree, obtaining an equivalent discrete characterization of Carleson measures for $\mathcal{D}(\mathbb{D}^2)$, see Theorem 1.3. Adopting the approach by Maz'ya we reduce the problem to verifying a certain property of the bilogarithmic potential — the Strong Capacitary Inequality. Assuming this inequality holds, see Theorem 1.5 and its proof in Section 3, we prove Theorem 1.1 for measures inside the bidisc in Section 2.5. In Section 2.6 we extend this result to measures supported on $\overline{\mathbb{D}}^2$ and replace the function f in (2) with its radial variation. Finally, in Section 2.7 we describe the multipliers of $\mathcal{D}(\mathbb{D}^2)$.

2.1 Duality approach

Let $\mu \ge 0$ be a finite Borel measure on $\overline{\mathbb{D}}^2$, and assume for a time being that $\mu(\partial\mathbb{D}^2) = 0$. We consider the general case later in Section 2.6.

To modify (2) we first observe that the embedding $Id : \mathcal{D}(\mathbb{D}^2) \to L^2(\mathbb{D}^2, d\mu)$ is bounded if and only if the adjoint operator

$$\Theta := Id^* : L^2(\mathbb{D}^2, d\mu) \to \mathcal{D}(\mathbb{D}^2)$$

is bounded as well. To proceed we need to know the action of Θ, which is provided by the RKHS property of $\mathcal{D}(\mathbb{D}^2)$. Indeed, given a function $g \in L^2(\mathbb{D}^2, d\mu)$ we have

$$(\Theta g)(z) = \langle \Theta g, K_z \rangle_{\mathcal{D}(\mathbb{D}^2)} = \langle g, K_z \rangle_{L^2(\mathbb{D}^2, d\mu)} = \int_{\mathbb{D}^2} g(w) \overline{K_z(w)} \, d\mu(w).$$

Then boundedness of Θ means that for any $g \in L^2(\mathbb{D}^2, d\mu)$ one has

$$\begin{aligned}
\|g\|^2_{L^2(\mathbb{D}^2, d\mu)} &\gtrsim \|\Theta g\|^2_{\mathcal{D}(\mathbb{D}^2)} = \langle \Theta g, \Theta g \rangle_{\mathcal{D}(\mathbb{D}^2)} = \langle g, \Theta g \rangle_{L^2(\mathbb{D}^2, d\mu)} = \\
&\int_{\mathbb{D}^2} \int_{\mathbb{D}^2} g(z) \overline{g(w)} K_z(w) \, d\mu(z) \, d\mu(w).
\end{aligned} \tag{15}$$

Taking g to be real and non-negative, we see that (15) becomes

$$\|g\|^2_{L^2(\mathbb{D}^2, d\mu)} \gtrsim \int_{\mathbb{D}^2} \int_{\mathbb{D}^2} g(z) g(w) \Re K_z(w) \, d\mu(z) \, d\mu(w), \tag{16}$$

since, clearly, $\int_{\mathbb{D}^2} \int_{\mathbb{D}^2} g(z) g(w) \Im K_z(w) \, d\mu(z) \, d\mu(w) = 0$. On the other hand, if g is an arbitrary function in $L^2(\mathbb{D}^2, d\mu)$, then (16) applied to $|g|$ gives

$$\|g\|^2_{L^2(\mathbb{D}^2, d\mu)} \gtrsim \int_{\mathbb{D}^2} \int_{\mathbb{D}^2} |g(z)| |g(w)| \Re K_z(w) \, d\mu(z) \, d\mu(w) \approx \int_{\mathbb{D}^2} \int_{\mathbb{D}^2} |g(z)| |g(w)| |K_z(w)| \, d\mu(z) \, d\mu(w),$$

and we are back at (15). To summarize, the measure μ on $\mathbb{D}^2$ is Carleson for $\mathcal{D}(\mathbb{D}^2)$ if and only if (16) holds for any non-negative function in $L^2(\mathbb{D}^2, d\mu)$. Observe that unlike (2), condition (16) does not mention the Sobolev norm of the Dirichlet space, nor its analytic structure, which makes it a much more viable candidate for discretization.

2.2 The bitree

Let T be a rooted directed uniform (each vertex has the same amount of children) infinite binary tree – in what follows we call such an object *a dyadic tree*. There is a natural order relation provided by the arborescent structure: given two points α, β in the vertex set $\mathcal{V}(T)$ we say that $\alpha \leq \beta$, if β lies on the unique geodesic connecting α and the root o, we also write $\beta \geq \alpha$, if $\alpha \leq \beta$. Here and when this is less cumbersome, we identify a geodesic with a sequence of vertices in the obvious way. Consider an infinite directed sequence ω starting at the root, i.e. $\omega = \{\omega^j\}_{j=0}^{\infty}$ with $\omega^0 = o$, $\omega^{j+1} \leq \omega^j$ and ω^j, ω^{j+1} are connected by an edge. The set of all such sequences is (as usual) called the *boundary* of T and denoted by ∂T, we also let $\overline{T} = T \cup \partial T$. For any $\omega = \{\omega^j\}_{j=0}^{\infty} \in \partial T$ we always say that $\omega \leq \omega^j$, $j \geq 0$.

If $\alpha, \beta \in \mathcal{V}(T) \cup \partial T$, then there there exists a unique point $\gamma \in \mathcal{V}(T) \cup \partial T$ that is *the least common ancestor* of α and β, we denote it by $\alpha \wedge \beta$. Namely, we have that $\gamma \geq \alpha$, $\gamma \geq \beta$, and if there is another point $\tilde{\gamma}$ satisfying these relations, then $\tilde{\gamma} \geq \gamma$. In other words, γ is the first intersection point of the geodesics connecting α and β to the root. The total amount of common ancestors of α and β is denoted by $d_T(\alpha \wedge \beta)$. Note that $d_T(\alpha \wedge \beta) = \text{dist}_T(\alpha \wedge \beta, o) + 1$, where dist_T is the usual graph distance on T. d_T may be infinite; for instance, $d_T(\omega \wedge \omega) = \infty$ whenever $\omega \in \partial T$. The *predecessor set* (with respect to the geometry of $\overline{T}$) of a point $\alpha \in \mathcal{V}(T) \cup \partial T$ is

$$\mathcal{P}(\alpha) = \{\beta \in \mathcal{V}(T) \cup \partial T : \beta \geq \alpha\}.$$

In particular, every point is its own predecessor. The *successor set* is

$$\mathcal{S}(\beta) := \{\alpha \in \mathcal{V}(T) \cup \partial T : \beta \in \mathcal{P}(\alpha)\}, \quad \beta \in \overline{T}.$$

Clearly $d_T(\alpha \wedge \beta) = |\mathcal{P}(\alpha \wedge \beta)|$.

A dyadic tree is a well known and often used way of discretizing the unit disc, so it is reasonable to assume that the discrete analogue of the bidisc $\mathbb{D} \times \mathbb{D}$ should look like a Cartesian product of two dyadic trees. With this observation in mind we define the graph T^2 which we call *bitree* as follows. Given two dyadic trees T_x and T_y we let

$$\mathcal{V}(T^2) := \mathcal{V}(T_x) \times \mathcal{V}(T_y),$$

to be the vertex set of T^2. In other words a vertex $\alpha \in \mathcal{V}(T^2)$ is a pair (α_x, α_y) of vertices of two (identical) coordinate trees. Given two vertices $\alpha, \beta \in \mathcal{V}(T^2)$ we connect them by an edge whenever $\alpha_x = \beta_x$ and α_y and β_y are connected by an edge in T_y, or, vice versa, $\alpha_y = \beta_y$ and α_x, β_x are neighbours. The order relation (and hence the direction of T^2) are induced from the coordinate trees, we say that $\alpha \leq \beta$ if and only if $\alpha_x \leq \beta_x$ *and* $\alpha_y \leq \beta_y$. As in one-dimensional case we define the boundary $\partial T^2 := T_x \times \partial T_y \bigcup \partial T_x \times T_y \bigcup \partial T_x \times \partial T_y$. The last part $\partial T_x \times \partial T_y$ we call the *distinguished boundary* of the bitree (similar to the bidisc setting) and denote by $(\partial T)^2$. We also let $\overline{T}^2 = T^2 \bigcup \partial T^2$. As before, we define predecessor and successor sets of a vertex $\alpha = (\alpha_x, \alpha_y)$ using the same notation

$$\mathcal{P}(\alpha) = \mathcal{P}(\alpha_x) \times \mathcal{P}(\alpha_y), \quad \mathcal{S}(\alpha) = \mathcal{S}(\alpha_x) \times \mathcal{S}(\alpha_y).$$

Sometimes, to avoid confusion, we specify the dimension by writing $\mathcal{S}_T(\alpha)$ for a point α in the tree T, and $\mathcal{S}_{T^2}(\alpha)$ for a point α in the bitree. We use the same convention for predecessor sets.

Similar to one-dimensional setting we denote the number, possibly infinite, of common ancestors of α and β by $d_{T^2}(\alpha \wedge \beta)$, where $\alpha \wedge \beta = (\alpha_x \wedge \beta_x, \alpha_y \wedge \beta_y)$ is the (unique) least common ancestor of α and β. We have

$$d_{T^2}(\alpha \wedge \beta) = d_T(\alpha_x \wedge \beta_x) \cdot d_T(\alpha_y \wedge \beta_y) = |\mathcal{P}(\alpha_x \wedge \beta_x)||\mathcal{P}(\alpha_y \wedge \beta_y)| = |\mathcal{P}(\alpha \wedge \beta)|.$$

In what follows we do not really need to consider the edges of T, since we are studying the unweighted Dirichlet space. From now on we do not distinguish the graph and its vertex set, in that we write $\alpha \in \overline{T}^2$ ($\overline{T}$) instead of $\alpha \in \mathcal{V}(T^2) \bigcup \partial T^2$ ($\mathcal{V}(T) \bigcup \partial T$). We also write $d_T(\alpha_x)$ and $d_{T^2}(\beta)$ instead of $d_T(\alpha_x \wedge \alpha_x)$ and $d_{T^2}(\beta \wedge \beta)$.

A natural way to interpret the dyadic tree is to identify its vertices with the approximating intervals for the classical Cantor set on the unit interval. Namely, consider the ternary Cantor set $E = \bigcap_{j=0}^{\infty} E_j$, where $E_0 = [0, 1]$, and E_k consists of 2^k closed intervals of length 3^{-k}. Then each point of T corresponds to a unique interval in E_j (or, more precisely, to its midpoint), and, similarly, ∂T maps to E. In the same vein the points of T^2 correspond to ternary rectangles (Cartesian products of centerpoints of intervals in E_j and E_k). In particular, $(\partial T)^2$ can be identified with E^2. Note that this means that $\overline{T}$ can be embedded into $\mathbb{R}^2$, and, consequently, $\overline{T}^2$ into $\mathbb{R}^4$. We will use this embedding to define a Potential Theory on the bitree. We also observe that T^2 no longer has unique geodesics; it is not acyclic like T. However, T^2 still *does not have directed cycles*.

2.3 Potential theory on the bitree

We start by defining a metric on $\overline{T}^2$: given $\alpha, \beta \in \overline{T}^2$ we let

$$\delta(\alpha, \beta) := 2^{-d_T(\alpha_x \wedge \beta_x)} + 2^{-d_T(\alpha_y \wedge \beta_y)} - \frac{1}{2}\left(2^{-d_T(\alpha_x)} + 2^{-d_T(\beta_x)} + 2^{-d_T(\alpha_y)} + 2^{-d_T(\beta_y)}\right). \quad (17)$$

This metric makes $\overline{T}^2$ into a compact space. The properties of a metric are easily verified for the function $\delta_x : \overline{T} \times \overline{T} \to [0, \infty)$, $\delta_x(\alpha_x, \beta_x) = 2^{-d_T(\alpha_x \wedge \beta_x)} - \frac{1}{2}\left(2^{-d_T(\alpha_x)} + 2^{-d_T(\beta_x)}\right)$, and δ is the sum of two copies of such distance, one on each factor of $\overline{T} \times \overline{T}$.

Denote by $\mathbb{M}$ the (open) bitree T^2 equipped with the counting measure v_c (so that $v_c(\{\alpha\}) = 1$, $\alpha \in T^2$). We define a kernel $G : \mathbb{R}^4 \times \mathbb{M} \to \mathbb{R}_+$ to be $G(\alpha, \beta) := \chi_{S_\beta}(\alpha)$, where $\alpha \in \overline{T}^2 \subset \mathbb{R}^4$ (here we consider $\overline{T}^2$ to be a compact subset of $\mathbb{R}^4$), $\beta \in T^2$ and $S_\beta := \{\gamma \in \overline{T}^2 : \gamma \leq \beta\}$ is the $\overline{T}^2$-successor set of β. It is easy to verify that G is lower semicontinuous on $\overline{T} \subset \mathbb{R}^4$ in first variable, and measurable on $\mathbb{M}$ in second variable. Extending kernels, functions, and measures from $\overline{T}^2$ to $\mathbb{R}^4$, by letting them be zero outside $\overline{T}^2$, we are squarely in the context of Adams and Hedberg [2, Chapter 2]. We thus have a well-defined potential theory on the bitree. We refer to [2, Chapter 2] for the general theory, while recalling some of its main features below.

Given a non-negative Borel measure μ on $\overline{T}^2$ (which by extension is Borel on $\mathbb{R}^4$) and a non-negative v_c-measurable function f on $\mathbb{M}$ we let

$$(\mathbb{I}f)(\alpha) := \int_{\mathbb{M}} G(\alpha, \beta)f(\beta)\,dv_c(\beta) = \sum_{\gamma \geq \alpha} f(\gamma), \quad (18a)$$

$$(\mathbb{I}^*\mu)(\beta) := \int_{\overline{T}^2} G(\alpha, \beta)\,d\mu(\alpha) = \int_{S(\beta)} d\mu(\alpha). \quad (18b)$$

Observe that a measure supported on T^2 and a non-negative function there are pretty much the same objects — a collection of masses assigned to the points of the bitree. The Potential Theory generated by these two operators leads us to the notions of bilogarithmic potential

$$\mathbb{V}^\mu := (\mathbb{I}\mathbb{I}^*)(\mu) \quad (19)$$

and capacity

$$\operatorname{Cap} E := \inf\left\{\int f^2\,dv_c : f \geq 0, (\mathbb{I}f)(\alpha) \geq 1, \alpha \in E\right\}, \quad (20)$$

for Borel set $E \subset \overline{T}^2$. Given two Borel measures $\mu, \nu \geq 0$ on $\overline{T}^2$ we define their mutual energy to be

$$\mathcal{E}[\mu, \nu] := \int_{\overline{T}^2} \mathbb{V}^\mu\,d\nu = \int_{\overline{T}^2} \mathbb{V}^\nu\,d\mu = \sum_{\alpha \in T^2} (\mathbb{I}^*\mu)(\alpha)(\mathbb{I}^*\nu)(\alpha), \quad (21)$$

the last two equalities following from Tonelli's theorem. When $\mu = \nu$ we write $\mathcal{E}[\mu]$ instead, and we call it *the energy of* μ. Given a Borel set $E \subset \overline{T}$ there exists a uniquely defined *equilibrium measure* $\mu_E \geq 0$ that generates the minimizer in (20), so that

$$\operatorname{Cap} E = \int_{T^2} (\mathbb{I}^*\mu_E)^2\,dv_C = \mathcal{E}[\mu_E] = \mu_E(E),$$

see [2]. If E is a compact set, then one also has $\operatorname{supp} \mu_E \subset E$.

2.4 From the bidisc to the bitree: measures supported inside

Now that the bitree has been defined, we can move all the objects from the bidisc here. In doing that we first assume that the measures in question are supported on $\mathbb{D}^2$ (so they have no mass on the boundary). We start with making a decomposition of the unit disc into dyadic Carleson boxes. For integer $j \geq 0$ and
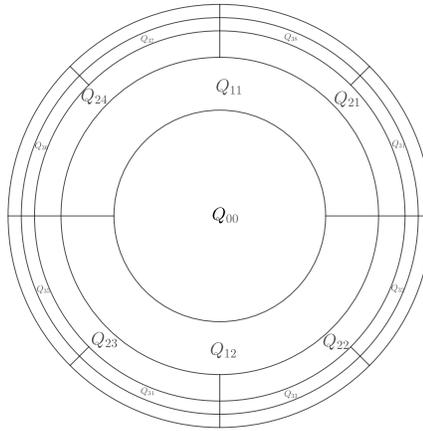

Figure 1: Discretized disc

$1 \leq l \leq 2^j$ let $z_{jl} = (1 - 2^{-j})e^{\frac{2\pi i(2l-1)}{2^{j+1}}}$, and for $z = re^{it} \in \mathbb{D}$ let $J(z) = \{e^{is} : t - (1-r)\pi \leq s \leq t + (1-r)\pi\}$, $S(z) = \{\rho e^{is} : e^{is} \in J(z); r \leq \rho \leq 1\}$, and let $Q(z) = \{\rho e^{is} \in S(z) : \frac{1-r}{2} \leq 1 - \rho \leq 1 - r\}$ be the 'upper half' of $S(z)$. We write $Q_{jl} := Q(z_{jl})$. Now we see that there is one-to-one map between points (vertices) of T and dyadic Carleson half-cubes Q_{jl}; Q_{00} corresponds to the root o, Q_{11} and Q_{12} to its two children etc. (see Fig. 1). In other words, for every $\alpha \in T$ there exists a unique half-cube Q_α, and vice versa, for every half-cube Q_{jl} there is exactly one point $\alpha^{jl} \in T$. The collection $\{Q_\alpha\}_{\alpha \in T}$ forms a covering of the unit disc. Note also that given a point $z \in \mathbb{D}$ it is possible to pick the half-box $Q_\alpha \ni z$ in a unique way. Though it can happen that there are several half-boxes Q_α containing z (up to four), we can still pick up one of them (say, whichever is closer to $\partial\mathbb{D}$ and/or with larger $\arg z_{jl}$), and we do this, wherever it is needed, in a consistent fashion throughout the whole paper.

Next we introduce an auxiliary graph $\mathfrak{G}$ in such a way that $\mathcal{V}(\mathfrak{G}) := \mathcal{V}(T)$, and $\{\alpha, \beta\}$ is an edge of $\mathfrak{G}$, if $\mathrm{cl}Q_\alpha \cap \mathrm{cl}Q_\beta \neq \emptyset$. Here cl means closure in Euclidean distance. Basically we take T and add extra edges connecting points corresponding to adjacent Carleson half-cubes, see Fig. 2. Given a vertex or leaf $\alpha \in \mathcal{V}(T) \bigcup \partial T$ we define the $\mathfrak{G}$-*extended predecessor set* to be $\mathcal{P}_\mathfrak{G}(\alpha) := \{\beta \in \mathcal{V}(T) = \mathcal{V}(\mathfrak{G}) : dist_\mathfrak{G}(\beta, \Gamma(o, \alpha)) \leq 1\}$, where $\Gamma(o, \alpha)$ is the (unique) geodesic in $\overline{T}$ connecting α and the root o. In other words, we take the $\overline{T}$-predecessor set $\mathcal{P}_T(\alpha)$ and add all the adjacent (in $\mathfrak{G}$) vertices (see Fig. 3). As
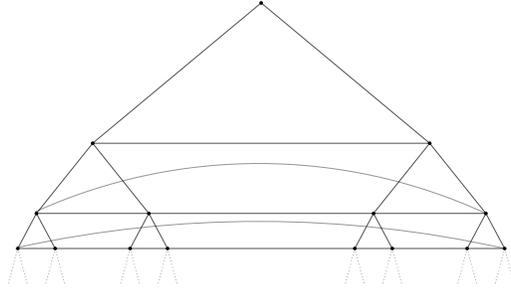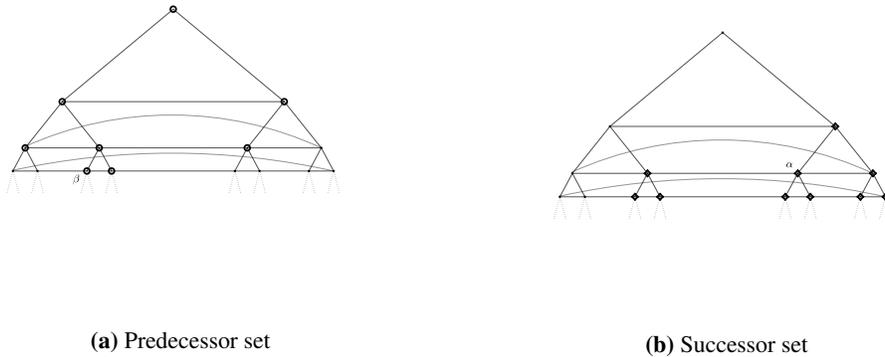
Figure 2: Graph $\mathfrak{G}$

(a) Predecessor set

(b) Successor set

Figure 3

before, we set $\mathcal{S}_{\mathfrak{G}}(\beta) := \{\alpha \in \overline{T} : \beta \in \mathcal{P}_{\mathfrak{G}}(\alpha)\}$ to be the $\mathfrak{G}$-*extended successor set.* Since (by definition) $\mathcal{P}(\alpha) \subset \mathcal{P}_{\mathfrak{G}}(\alpha)$ for any $\alpha \in \overline{T}$, we have the same inclusion for the successor sets, $\mathcal{S}(\alpha) \subset \mathcal{S}_{\mathfrak{G}}(\alpha)$, and this inclusion is proper unless the vertex in question is the root o. On the other hand, the successor sets are 'comparable on average'. To elaborate, let $N(\alpha)$ be the set of neighbours of α in $\mathfrak{G}$ (so α is connected by an edge in $\mathfrak{G}$ to the points in $N(\alpha)$). Then

$$\mathcal{S}_{\mathfrak{G}}(\alpha) \subset \bigcup_{\beta \in N(\alpha)} \mathcal{S}(\beta), \tag{22}$$

and $|N(\alpha)| \leq 5$ (in particular $\bigcup_{\alpha \in T} N(\alpha)$ covers each point at most 5 times). Another way to look at $\mathcal{S}_{\mathfrak{G}}(\alpha)$ is to consider the dyadic interval J_α and its two immediate neighbours of the same rank $J_\alpha^{\pm}$. Then

$$\mathcal{S}_{\mathfrak{G}}(\alpha) \cap T = \left\{ \beta \in T : J_\beta \subset J_\alpha \cup J_\alpha^- \cup J_\alpha^+ \right\}.$$

The correspondence between ∂T and $\partial \mathbb{D}$ will be explained later in Section 2.6. Finally, we let

$$d_{\mathfrak{G}}(\alpha \wedge \beta) := |\mathcal{P}_{\mathfrak{G}}(\alpha) \cap \mathcal{P}_{\mathfrak{G}}(\beta)|, \quad \alpha, \beta \in T.$$

As before, we keep the same notation for the bitree, namely given two points $\alpha = (\alpha_x, \alpha_y)$ and $\beta = (\beta_x, \beta_y)$ in T^2 we set

$$\mathcal{P}_{\mathfrak{G}^2}(\alpha) := \mathcal{P}_{\mathfrak{G}}(\alpha_x) \times \mathcal{P}_{\mathfrak{G}}(\alpha_y), \ \mathcal{S}_{\mathfrak{G}^2}(\alpha) := \mathcal{S}_{\mathfrak{G}}(\alpha_x) \times \mathcal{S}_{\mathfrak{G}}(\alpha_y)$$

$$d_{\mathfrak{G}^2}(\alpha \wedge \beta) := d_{\mathfrak{G}}(\alpha_x \wedge \beta_x) \cdot d_{\mathfrak{G}}(\alpha_y \wedge \beta_y)$$

The main reason to introduce this auxiliary graph $\mathfrak{G}$ is that the geometry of the tree T does not completely agree with the geometry of the unit disc $\mathbb{D}$. For instance, one can easily find a pair of points $z, w \in \mathbb{D}$, very close to each other, while the tree distance between α and β corresponding to these points (i.e. $z \in Q_\alpha$, $w \in Q_\beta$) is very large. It is a well-known (if somewhat minor) obstacle, and there are several ways to overcome it. We have chosen what we think is the simplest one, especially since we do not care about precise values of arising constants.

Taking two identical dyadic coordinate trees T_x, T_y we see that the collection $\{Q_\alpha\} := \{Q_{\alpha_x} \times Q_{\alpha_y}\}$, $\alpha = (\alpha_x, \alpha_y) \in T_x \times T_y = T^2$ gives almost a disjoint decomposition of the bidisc $\mathbb{D} \times \mathbb{D}$ (these Whitney cubes may intersect, but each point of the bidisc is counted at most 16 times). Assume that $\mu \geq 0$ is a Borel measure on $\mathbb{D}^2$ for which $\mu(\partial \mathbb{D}^2) = 0$ and that $g \in L^2(\mathbb{D}^2, d\mu)$ is a non-negative function. We then let

$$\tilde{\mu}(\alpha) := \mu(Q_\alpha),$$

$$\tilde{g}(\alpha) := \frac{1}{\mu(Q_\alpha)} \int_{Q_\alpha} g(z) \, d\mu(z), \quad \alpha = (\alpha_x, \alpha_y) \in T^2, \tag{23}$$

and we set $\tilde{g}(\alpha) := 0$, if $\mu(Q_\alpha) = 0$. Now we recall that μ is Carleson measure for $\mathcal{D}(\mathbb{D}^2)$ if and only if (16) holds for any g as above, namely

$$\|g\|_{L^2(\mathbb{D}^2, d\mu)}^2 = \int_{\mathbb{D}^2} g^2(z) \, d\mu(z) = \sum_{\alpha \in T^2} \int_{Q_\alpha} g^2(z) \, d\mu(z) \gtrsim \int_{\mathbb{D}^2} \int_{\mathbb{D}^2} g(z) g(w) \Re K_z(w) \, d\mu(z) \, d\mu(w) \approx$$

$$\sum_{\alpha \in T^2} \sum_{\beta \in T^2} \int_{Q_\alpha} \int_{Q_\beta} g(z) g(w) \Re K_z(w) \, d\mu(z) \, d\mu(w). \tag{24}$$

In order to proceed we need the following Lemma, the proof of which is given in Section 5.1.

Lemma 2.1 *For any $\alpha, \beta \in T^2$ we have*

$$\Re K_z(w) \approx d_{\mathfrak{G}^2}(\alpha \wedge \beta) = d_{\mathfrak{G}}(\alpha_x \wedge \beta_x) \cdot d_{\mathfrak{G}}(\alpha_y \wedge \beta_y), \quad z \in Q_\alpha, \ w \in Q_\beta.$$

.

Applying Lemma 2.1 to the right-hand side of (24) we get

$$\sum_{\alpha \in T^2} \sum_{\beta \in T^2} \int_{Q_\alpha} \int_{Q_\beta} g(z) g(w) \Re K_z(w) \, d\mu(z) \, d\mu(w) \approx$$

$$\sum_{\alpha \in T^2} \sum_{\beta \in T^2} \int_{Q_\alpha} \int_{Q_\beta} g(z) g(w) d_{\mathfrak{G}^2}(\alpha \wedge \beta) \, d\mu(z) \, d\mu(w) =$$

$$\sum_{\alpha \in T^2} \sum_{\beta \in T^2} \tilde{g}(\alpha) \tilde{g}(\beta) d_{\mathfrak{G}}(\alpha_x \wedge \beta_x) d_{\mathfrak{G}}(\alpha_y \wedge \beta_y) \tilde{\mu}(\alpha) \tilde{\mu}(\beta)$$

We attack the calculation from the end, letting $\sigma(\alpha) = \tilde{g}(\alpha)\tilde{\mu}(\alpha)$, $\alpha \in T^2$ (recall that measures and functions on T^2 are the same):

$$
\sum_{\gamma \in T^2} \left(\sum_{\alpha \in \mathcal{S}_{\mathfrak{G}^2}(\gamma)} \sigma(\alpha) \right)^2 = \sum_{\gamma \in T^2} \sum_{\alpha \in \mathcal{S}_{\mathfrak{G}^2}(\gamma)} \sum_{\beta \in \mathcal{S}_{\mathfrak{G}^2}(\gamma)} \sigma(\alpha)\sigma(\beta) = \sum_{\alpha \in T^2} \sum_{\beta \in T^2} \sum_{\gamma \in \mathcal{P}_{\mathfrak{G}^2}(\alpha) \cap \mathcal{P}_{\mathfrak{G}^2}(\beta)} \sigma(\alpha)\sigma(\beta) =
$$

$$
\sum_{\alpha \in T^2} \sum_{\beta \in T^2} |\mathcal{P}_{\mathfrak{G}^2}(\alpha) \cap \mathcal{P}_{\mathfrak{G}^2}(\beta)| \sigma(\alpha)\sigma(\beta) = \sum_{\alpha \in T^2} \sum_{\beta \in T^2} \tilde{g}(\alpha)\tilde{g}(\beta) d_{\mathfrak{G}}(\alpha_x \wedge \beta_x) d_{\mathfrak{G}}(\alpha_y \wedge \beta_y) \tilde{\mu}(\alpha)\tilde{\mu}(\beta).
$$

Repeating the calculation with $\mathcal{P}$ instead of $\mathcal{P}_{\mathfrak{G}^2}$ we obtain

$$
\sum_{\gamma \in T^2} \left(\sum_{\alpha \in \mathcal{S}(\gamma)} \sigma(\alpha) \right)^2 = \sum_{\alpha \in T^2} \sum_{\beta \in T^2} \tilde{g}(\alpha)\tilde{g}(\beta) d_{T^2}(\alpha \wedge \beta) \tilde{\mu}(\alpha)\tilde{\mu}(\beta).
$$

The successor set formula (22) implies that

$$
\sum_{\gamma \in T^2} \left(\sum_{\alpha \in \mathcal{S}_{\mathfrak{G}^2}(\gamma)} \sigma(\alpha) \right)^2 = \sum_{\gamma \in T^2} \sigma \left(\mathcal{S}_{\mathfrak{G}}(\gamma_x) \times \mathcal{S}_{\mathfrak{G}}(\gamma_y) \right)^2 \approx \sum_{\gamma \in T^2} \sigma \left(\mathcal{S}(\gamma_x) \times \mathcal{S}(\gamma_y) \right)^2 =
$$

$$
\sum_{\gamma \in T^2} \left(\sum_{\alpha \in \mathcal{S}(\gamma)} \sigma(\alpha) \right)^2.
$$

Combining the estimates above we see that (24) is equivalent to

$$
\sum_{\alpha \in T^2} \int_{Q_\alpha} g^2(z) \, d\mu(z) \gtrsim \sum_{\alpha \in T^2} \sum_{\beta \in T^2} \tilde{g}(\alpha)\tilde{g}(\beta) d_{T^2}(\alpha \wedge \beta) \tilde{\mu}(\alpha)\tilde{\mu}(\beta),
$$

where $\tilde{g}$ and $\tilde{\mu}$ are defined in (23). We see that if g is constant on the boxes Q_α and μ is Carleson measure for $\mathcal{D}(\mathbb{D}^2)$, then

$$
\|\tilde{g}\|^2_{L^2(T^2, d\tilde{\mu})} \gtrsim \sum_{\alpha \in T^2} \sum_{\beta \in T^2} \tilde{g}(\alpha)\tilde{g}(\beta) d_T(\alpha \wedge \beta) \tilde{\mu}(\alpha)\tilde{\mu}(\beta) = \sum_{\alpha \in T^2} \left(\sum_{\beta \leq \alpha} \tilde{g}(\beta)\tilde{\mu}(\beta) \right)^2. \tag{25}
$$

On the other hand, by Jensen's inequality,

$$
\sum_{\alpha \in T^2} \int_{Q_\alpha} g^2(z) \, d\mu(z) \geq \sum_{\alpha \in T^2} \tilde{g}^2(\alpha)\tilde{\mu}(\alpha),
$$

so if (25) holds for any non-negative $\tilde{g}$ in $L^2(T^2, d\tilde{\mu})$, then μ is Carleson.

Let $v \geq 0$ be a Borel measure on $\overline{T}^2$. Given a v-measurable function φ defined on $\overline{T}^2$, we let

$$
(\mathbb{I}_v^* \varphi)(\beta) := \int_{\mathcal{S}(\beta)} \varphi(\alpha) \, dv(\alpha), \quad \beta \in T^2. \tag{26}
$$

Since by our temporary assumption $\tilde{\mu}$ is supported on T^2, inequality (25) can be rewritten as

$$\|\mathbb{I}_{\tilde{\mu}}^* \tilde{g}\|_{\ell^2(T^2)}^2 \lesssim \|\tilde{g}\|_{L^2(T^2, d\tilde{\mu})}^2, \tag{27}$$

which means that the operator $\mathbb{I}_{\tilde{\mu}}^* : L^2(T^2, d\tilde{\mu}) \to \ell^2(T^2)$ is bounded (and its operator norm is comparable to the Carleson constant $[\mu]$). Given a pair of real functions $\varphi \in \ell^2(T^2)$ and $f \in L^2(T^2, d\nu)$ we see that

$$\langle \mathbb{I}_\nu^* \varphi, f \rangle_{\ell^2(T^2)} = \sum_{\beta \in T^2} \int_{S(\beta)} \varphi(\alpha) \, d\nu(\alpha) f(\beta) = \int_{\overline{T}^2} \varphi(\alpha) \sum_{\beta \in T^2} \chi_{S(\beta)}(\alpha) f(\beta) \, d\nu(\alpha) =$$

$$\int_{\overline{T}^2} \varphi(\alpha) \sum_{\beta \geq \alpha} f(\beta) \, d\nu(\alpha) = \langle \varphi, \mathbb{I}f \rangle_{L^2(\overline{T}^2, d\nu)}.$$

Hence $\mathbb{I}$, as an operator acting from $\ell^2(T^2)$ to $L^2(\overline{T}^2, d\nu)$, is adjoint to $\mathbb{I}_\nu^*$ and $\|\mathbb{I}\| = \|\mathbb{I}_\nu^*\|$. We arrive at the following statement.

Proposition 2.1 *Let $\mu \geq 0$ be a Borel measure on $\mathbb{D}^2$ and define $\tilde{\mu}$ as in (23). Then $\tilde{\mu}$ is a trace measure for discrete bi-parameter Hardy inequality,*

$$\int_{\overline{T}^2} (\mathbb{I}f)^2 \, d\tilde{\mu} \leq \tilde{C}_\mu \sum_{\alpha \in T^2} f^2(\alpha) \tag{28}$$

if and only if μ is Carleson for $\mathcal{D}(\mathbb{D}^2)$. The best possible constant in (28) is comparable to the Carleson constant of μ.

2.5   Proof of Theorem 1.1: Maz'ya approach

Here we show that trace measures for the bi-parameter Hardy operator admit a characterization via a discrete subcapacitary condition, as in Theorem 1.4. Then we translate this condition back to the continuous world, obtaining (3).

Let $\nu \geq 0$ be a Borel measure on $\overline{T}^2$ (note that now it might have non-zero mass on ∂T^2). We call it subcapacitary, if for any finite collection $\{\alpha^j\}_{j=1}^N \subset T^2$ one has

$$\nu \left(\bigcup_{j=1}^N S(\alpha^j) \right) \leq C \operatorname{Cap} \left(\bigcup_{j=1}^N S(\alpha^j) \right)$$

for some constant $C > 0$ the depends only on ν – the smallest such constant we denote by C_ν).

Assume now that ν is a trace measure for the Hardy operator,

$$\int_{\overline{T}^2} (\mathbb{I}f)^2 \, d\nu \lesssim \sum_{\alpha \in T^2} f^2(\alpha)$$

for any $f : T^2 \mapsto \mathbb{R}_+$. Given a Borel set $E \subset \overline{T}^2$ consider the family

$$\Omega_E = \{ f \in \ell^2(T^2) : f \geq 0, \, \mathbb{I}f \geq 1 \text{ on } E \}$$

of E-admissible functions. Then for any $f \in \Omega_E$ one has

$$\nu(E) = \int_{\overline{T}^2} \chi_E \, d\nu \leq \int_{\overline{T}^2} (\mathbb{I}f)^2 \, d\nu \lesssim \|f\|^2_{\ell^2(T^2)}.$$

Taking infimum over Ω_E we immediately get $\nu(E) \lesssim \operatorname{Cap} E$.

The other direction is more involved, and the argument follows the route pioneered by Maz'ya. Assume that $\nu \geq 0$ is a subcapacitary measure on $\overline{T}^2$ and that $f \in \ell^2(T^2)$, $f \geq 0$. By a distribution function argument

$$\int_{\overline{T}^2} f^2 \, d\nu \approx \sum_{k \in \mathbb{Z}} 2^{2k} \nu \left\{ \alpha \in \overline{T}^2 : \mathbb{I}f > 2^k \right\}.$$

Since $f \geq 0$, we have that if $(\mathbb{I}f)(\alpha) > 2^k$, then $(\mathbb{I}f)(\beta) > 2^k$ for any $\beta \leq \alpha$. Therefore for any $k \in \mathbb{Z}$ there exists a countable family $\{\alpha_k^j\}_{j=0}^\infty \subset T^2$ such that

$$\left\{ \alpha \in \overline{T}^2 : \mathbb{I}f > 2^k \right\} = \bigcup_{j=0}^\infty \mathcal{S}(\alpha_k^j),$$

with $(\mathbb{I}f)(\alpha_k^j) > 2^k$. It follows that

$$\nu \left\{ \alpha \in \overline{T}^2 : \mathbb{I}f > 2^k \right\} = \lim_{N \to \infty} \nu \left(\bigcup_{j=0}^N \mathcal{S}(\alpha_k^j) \right) \lesssim \lim_{N \to \infty} \operatorname{Cap} \left(\bigcup_{j=0}^N \mathcal{S}(\alpha_k^j) \right) = \operatorname{Cap} \left(\bigcup_{j=0}^\infty \mathcal{S}(\alpha_k^j) \right).$$

Now assume for a moment that the following inequality holds (see Theorem 1.5 and its proof in Section 3),

$$\sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap}\{\mathbb{I}f \geq 2^k\} \leq C \|f\|^2_{\ell^2(T^2)}$$

for some absolute constant $C > 0$. Then we immediately have

$$\sum_{k \in \mathbb{Z}} 2^{2k} \nu \left\{ \alpha \in \overline{T}^2 : \mathbb{I}f > 2^k \right\} \leq \sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap} \left\{ \alpha \in \overline{T}^2 : \mathbb{I}f \geq 2^k \right\} \leq C \|f\|^2_{\ell(T^2)},$$

for any $f \geq 0$ on T^2. Therefore ν is a trace measure for Hardy inequality. Theorem 1.4 is proven.

All that remains to finish the proof of Theorem 1.1 (for measures with zero mass on $\partial \mathbb{D}^2$, and still assuming the Strong Capacitary Inequality) is to go back to the bidisc. We start by defining a continuous version of capacity that is convenient for our purposes. The Riesz-Bessel kernel of order $(\frac{1}{2}, \frac{1}{2})$ on the torus $(\partial \mathbb{D})^2$ is

$$b_{(\frac{1}{2}, \frac{1}{2})}(z, \zeta) = |\theta_1 - \eta_1|^{-\frac{1}{2}} |\theta_2 - \eta_2|^{-\frac{1}{2}}, \quad z = (e^{i\theta_1}, e^{i\theta_2}), \, \zeta = (e^{i\eta_1}, e^{i\eta_2}) \in (\partial \mathbb{D})^2,$$

where the difference $\theta_i - \eta_i \in [-\pi, \pi)$ is taken modulo 2π. The kernel extends to a convolution operator on $(\partial \mathbb{D})^2$ acting on Borel measures supported there,

$$(B_{(\frac{1}{2}, \frac{1}{2})} \mu)(z) = \int_{(\partial \mathbb{D})^2} b_{(\frac{1}{2}, \frac{1}{2})}(z, \zeta) \, d\mu(\zeta).$$

Let $E \subset (\partial \mathbb{D})^2$ be a closed set. The $(\frac{1}{2}, \frac{1}{2})$-Bessel capacity of E is

$$\operatorname{Cap}_{(\frac{1}{2}, \frac{1}{2})}(E) = \inf\{\|h\|^2_{L^2((\partial \mathbb{D})^2, dm)} : h \geq 0 \text{ and } B_{(\frac{1}{2}, \frac{1}{2})} h \geq 1 \text{ on } E\},$$

and it is realized by an equilibrium measure μ_E:

$$\operatorname{Cap}_{(\frac{1}{2}, \frac{1}{2})}(E) = \mathcal{E}_{(\frac{1}{2}, \frac{1}{2})}[\mu_E] := \int_{(\partial \mathbb{D})^2} \left((B_{(\frac{1}{2}, \frac{1}{2})} \mu_E)(z) \right)^2 dm(z),$$

where m is normalized area measure on the torus $(\partial \mathbb{D})^2$.

Let $\{J_k\}_{k=0}^N$ be a finite collection of dyadic rectangles on $(\partial \mathbb{D})^2$, i.e. $J_k = J_k^1 \times J_k^2$, where J_k^i is a dyadic interval in $\partial \mathbb{D}$. For any such collection there exists a unique sequence $\{\alpha_k\}_{k=0}^N \subset T^2$ such that $J_k = S_{\alpha_k} \cap (\partial \mathbb{D})^2$ (here S_{α_k} is the Carleson box corresponding to α_k), and vice versa, any finite sequence $\{\alpha_k\}_{k=0}^N$ produces a family $\{J_k\}_{k=0}^N$ of dyadic rectangles. A standard argument shows that $(\frac{1}{2}, \frac{1}{2})$-Bessel capacity and discrete bilogarithmic capacity are comparable. For the proof, see Section 5.1.

Lemma 2.2 *For any finite collection $\{\alpha_k\}_{k=0}^N$ one has*

$$\operatorname{Cap}_{(\frac{1}{2}, \frac{1}{2})} \left(\bigcup_{k=0}^N J_k \right) \approx \operatorname{Cap} \left(\bigcup_{k=0}^N \mathcal{S}(\alpha_k) \right),$$

where α_k and J_k are related as above.

Theorem 1.1 follows immediately. Indeed, we have shown that $\mu \geq 0$ on $\mathbb{D}^2$ is Carleson if and only if its discrete image $\tilde{\mu}$ is a trace measure for Hardy operator, and that this happens if and only if $\tilde{\mu}$ is subcapacitary in $\overline{T}^2$. Since $\tilde{\mu}(\cup_{k=0}^N \mathcal{S}(\alpha_k)) = \mu(\cup_{k=0}^N S(J_k))$, (3) follows, and we are done.

2.6 From the bidisc to the bitree: general case

Up until now we assumed the measure μ to be supported inside the bidisc. Here we get rid of this restriction and prove Theorem 1.1 in full generality, still assuming the Strong Capacitary Inequality. We also show that $C[\mu]$ is comparable to $\sup_f \|\operatorname{Var} f\|^2_{L^2(\overline{\mathbb{D}}^2, d\mu)}$, as promised in Theorem 1.3. To do so we first need to define the discrete image of a measure with non-zero mass on the boundary $\partial \mathbb{D}^2$. We consider the case of the distinguished boundary $(\partial \mathbb{D})^2$ first, which is more interesting and contains the ingredients for the remaining part as well. The problem is that the boundaries of the complex disc and of the tree, and the measures supported on them, can not be identified without some care.

We introduce a map $\Lambda : (\partial T)^2 \to (\partial \mathbb{D})^2$, $\Lambda(\alpha) = (\Lambda_0(\alpha_x), \Lambda_0(\alpha_y))$, where $\Lambda_0 : \partial T \to \partial \mathbb{D}$ maps a geodesic $\omega = \{o = \omega^0, \omega^1, \dots\}$ to the point $\Lambda_0(\omega) = \bigcap_{n=0}^\infty S_{\omega^n} \in \partial \mathbb{D}$. We will use Λ to move measures back and forth from $(\partial T)^2$ to $(\partial \mathbb{D})^2$, in such a way corresponding measures have comparable mass and energy.

Consider on $\overline{T}$ the distance

$$\delta_0(\zeta, \xi) := 2^{-d_T(\zeta \wedge \xi)} - \frac{1}{2} \left(2^{-d_T(\zeta)} + 2^{-d_T(\xi)} \right)$$

It is clear that Λ_0 is a Lipschitz map with respect to the distance δ_0 on ∂T and Euclidean distance on $\partial \mathbb{D}$, and that Λ_0 is injective but for the set of the dyadic values $\frac{2\pi j}{2^n}$ with $1 \leqslant j \leqslant 2^n$, which have two preimages.

Then Λ is Lipschitz with respect to the distance δ defined in (17) on $(\partial T)^2$ and the usual distance of the torus $(\partial \mathbb{D})^2$. Given a positive, Borel measure ν on $(\partial T)^2$, let $\Lambda_* \nu(F) := \nu(\Lambda^{-1}(F))$ be its natural push-forward. We need to define an (unnatural) pull-back. Given a positive, Borel measure μ on $(\partial \mathbb{D})^2$, define $\Lambda^* \mu$ be the measure assigning to a Borel subset $E \subseteq (\partial T)^2$ the number

$$\Lambda^* \mu(E) = \int_{(\partial \mathbb{D})^2} \frac{\sharp(\Lambda^{-1}(\{z\}) \cap E)}{\sharp(\Lambda^{-1}(\{z\}))} d\mu(z),$$

that is,

$$\int_{(\partial T)^2} \varphi(x) d\Lambda^* \mu(x) = \int_{(\partial \mathbb{D})^2} \frac{\sum_{x \in \Lambda^{-1}(\{z\})} \varphi(x)}{\sharp(\Lambda^{-1}(\{z\}))} d\mu(z).$$

If it is well-defined, then $\Lambda^* \mu$ defines a countably additive, positive set function. But we have to show, first, that the function $z \mapsto \frac{\sharp(\Lambda^{-1}(\{z\}) \cap E)}{\sharp(\Lambda^{-1}(\{z\}))} = \varphi_E(z)$ is measurable on $(\partial \mathbb{D})^2$ (this is a simpler but slightly more technical version of the argument in [10]).

For each point α in T we denote its children by α_+ and α_-. We can split $\partial T = A_+ \cup A_- \cup A$ into the disjoint union of three Borel measurable sets: A_+ is the countable set of the geodesics $\omega = (\omega^n)_{n=0}^\infty$ such that $\omega^{n+1} = \omega_+^n$ definitely; A_- is defined similarly; $A = \partial T \setminus (A_+ \cup A_-)$. The map Λ_0 is injective on each set. Correspondingly, we split $(\partial T)^2$ into nine disjoint measurable sets $B_1, \ldots, B_9$, on each of which Λ is injective.

The map $z \mapsto \sharp(\Lambda^{-1}(\{z\}))$ takes on the value 1 on $\Lambda(A \times A)$, it takes on the value 2 on $\Lambda(A_\pm \times A \cup A \times A_\pm)$, it takes on the value 4 on $\Lambda(A_\pm \times A_\pm \cup A_\mp \times A_\pm)$; hence, it is Borel measurable on $(\partial \mathbb{D})^2$. Similarly, the map $z \mapsto \sharp(\Lambda^{-1}(\{z\}) \cap E)$ takes on the value 1 on $\Lambda(A \times A \cap E)$, etcetera; hence it is Borel measurable as well. Thus, φ_E is measurable, as desired.

Next we consider the measures on the rest of the bidisc. First we extend the map Λ on $\overline{T}^2$ by letting $\Lambda(\alpha_x, \alpha_y) = \Lambda_0(\alpha_x) \times \Lambda_0(\alpha_y)$, where $\Lambda_0(\alpha_x) := Q_{\alpha_x}$, $\Lambda_0(\alpha_y) := Q_{\alpha_y}$, if $\alpha = (\alpha_x, \alpha_y) \in T^2$. For a point (α_x, ω_y) on the mixed boundary $T \times \partial T$ we set $\Lambda((\alpha_x, \omega_y)) := Q_{\alpha_x} \times \{\Lambda_0(\omega_y)\}$, and we do the same for the other part of the boundary.

Assume μ to be a positive Borel measure on $\mathbb{D} \times \partial \mathbb{D}$ and let $\alpha_x \in T_x$, and E_y be a Borel subset of T_y. We define the pull-back to be

$$(\Lambda^* \mu)(\{\alpha_x\} \times E_y) := \int_{Q_{\alpha_x}} \int_{E_y} \frac{\sharp\{\Lambda_0^{-1}(\{z_2\}) \cap E_y\}}{\sharp\{\Lambda_0^{-1}(\{z_2\})\}} d\mu(z_1, z_2),$$

the integrand being measurable for the same reasons as above.
Any set $E \subset T \times \partial T$ is a disjoint countable union of the product sets, i.e. there exist families $\{\alpha_x^j\}, E_y^j$, $j = 0, \ldots$, such that

$$E = \bigvee_{j=0}^\infty \left(\{\alpha_x^j\} \times E_y^j \right).$$

Hence $\Lambda^* \mu$ admits a unique extension to Borel sets on $T \times \partial T$. The measures on $\partial \mathbb{D} \times \mathbb{D}$ are dealt with the same way.

Finally, for $\alpha \in T^2$ we put

$$(\Lambda^* \mu)(\alpha) := \mu(\Lambda(\alpha)) = \mu(Q_\alpha).$$

We also need the one-dimensional version of the pull-back. Consider a Borel measure $\mu \geq 0$ on the closed unit disc $\overline{\mathbb{D}}$, we define its pull-back to the tree T to be

$$(\Lambda_0^* \mu)(\alpha) = \mu(Q_\alpha), \ \Lambda_0^* \mu(E) = \int_{\partial \mathbb{D}} \frac{\sharp(\Lambda_0^{-1}(\{z\}) \cap E)}{\sharp(\Lambda_0^{-1}(\{z\}))} d\mu(z),$$

for a set $E \subset \partial T$ (it is much simpler in one dimension, since we do not need to take care of the mixed parts of the boundary).

There is no natural way to define a push-forward Λ_* of a measure on T^2 (or on $T \times \partial T$ for that matter), since a point mass on T^2 (a positive number attached to a point $\alpha \in T^2$) can be moved to Q_α in several different manners (for instance it could be spread uniformly over Q_α, or considered as a point mass, concentrated at the centerpoint of Q_α). On the other hand, in what follows we do not need to use a push-forward of such a measure anyway.

Now we can prove the following Theorem, which contains one half of Theorem 1.3.

Theorem 2.1 *Let μ be a Borel measure on $\overline{\mathbb{D}}^2$, then*

$$||\mathbb{I}||^2_{\mathcal{B}(\ell^2(T^2), L^2(\overline{T}^2, d\Lambda^* \mu))} \approx ||\operatorname{Id}||^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2), L^2(\overline{\mathbb{D}}^2, d\mu))}.$$

Moreover, $||\operatorname{Id}||^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2), L^2(\overline{\mathbb{D}}^2, d\mu))}$ *is also comparable with the best constant K^2 in the stronger inequality*

$$\int_{\overline{\mathbb{D}}^2} (\operatorname{Var} f)^2 d\mu \leqslant K^2 ||f||^2_{\mathcal{D}(\mathbb{D}^2)},$$

where Var *is the* radial variation *of f.*

Let f be holomorphic in $\mathbb{D}^2$. The *radial variation* of f at $(\zeta_1, \zeta_2) \in \overline{\mathbb{D}}^2$ is

$$\begin{aligned}
\operatorname{Var}(f)(\zeta_1, \zeta_2) &= \operatorname{Var}_{12}(f)(\zeta_1, \zeta_2) + \operatorname{Var}_1(f)(\zeta_1) + \operatorname{Var}_2(f)(\zeta_2) + |f(0,0)| = \\
&\int_0^{\zeta_1} \int_0^{\zeta_2} |\partial_{z_1, z_2} f(z_1, z_2)||dz_1||dz_2| + \int_0^{\zeta_1} |\partial_{z_1} f(z_1, 0)||dz_1| + \int_0^{\zeta_2} |\partial_{z_2} f(0, z_2)||dz_2| + |f(0,0)|.
\end{aligned} \tag{29}$$

Proof. We will prove the chain of implications $(A) \Rightarrow (B) \Rightarrow (C) \Rightarrow (A)$:

- (A) $\int_{\overline{\mathbb{D}}^2} \operatorname{Var}(f)^2 d\mu \leqslant K_0^2 ||f||^2_{\mathcal{D}(\mathbb{D}^2)}$.

- (B) $\int_{\overline{\mathbb{D}}^2} |f|^2 d\mu \leqslant K_1^2 ||f||^2_{\mathcal{D}(\mathbb{D}^2)}$.

- (C) $\int_{\overline{T}^2} |\mathbb{I}\varphi|^2 d\Lambda^* \mu \leqslant K_2^2 ||\varphi||^2_{\ell^2(T^2)}$.

We note that (B) means

$$\sup_{R<1} \int_{\mathbb{D}^2} |f(R\zeta_1, R\zeta_2)|^2 d\mu(\zeta_1, \zeta_2) \leqslant K_1^2 \|f\|_{\mathcal{D}(\mathbb{D}^2)}^2.$$

The implication $(A) \Rightarrow (B)$ is elementary: the inequality for the variation is a priori stronger than the inequality for $|f|$. For $(\zeta_1, \zeta_2) \in \mathbb{D}^2$:

$$
\begin{aligned}
|f(\zeta_1, \zeta_2)| &= \left| \int_0^{\zeta_1} \partial_{z_1} f(z_1, \zeta_2)\, dz_1 + \int_0^{\zeta_2} \partial_{z_2} f(0, z_2) dz_2 + f(0,0) \right| \\
&= \left| \int_0^{\zeta_1} \left(\int_0^{\zeta_2} \partial_{z_1 z_2} f(z_1, z_2)\, dz_2 + \partial_{z_1} f(z_1, 0) \right) dz_1 + \int_0^{\zeta_2} \partial_{z_2} f(0, z_2)\, dz_2 + f(0,0) \right| \\
&\leqslant \mathrm{Var}_{12}(f)(\zeta_1, \zeta_2) + \mathrm{Var}_1(f)(\zeta_1) + \mathrm{Var}_2(f)(\zeta_2) + |f(0,0)|.
\end{aligned}
$$

For $0 < r < 1$ let $f_r(z_1, z_2) = f(rz_1, rz_2)$. If μ satisfies (A), then it satisfies Carleson inequality for f_r with constant independent of r, and we are done.

The proof of the implications $(B) \Rightarrow (C)$ and $(C) \Rightarrow (A)$ are more involved, and before proceeding we need an additional smoothness property of Carleson measures: if a measure $\mu \geq 0$ on the closed bidisc is Carleson, then for any set $E \subset \overline{T}^2$ one has

$$(\Lambda^* \mu)(E) = \mu(\Lambda(E)). \tag{30}$$

It means essentially that Carleson measures have no singularities on coordinate slices of the torus (the dyadic grid $\{\partial \mathbb{D} \times \{2\pi j 2^{-n}\}\}$, $j, n \geq 0$ has no mass), see Lemma 5.2 for details.

We now prove $(B) \Rightarrow (C)$. Suppose μ satisfies

$$\int_{\mathbb{D}^2} |f(Rz_1, Rz_2)|^2 d\mu(z_1, z_2) =: \int_{\mathbb{D}^2} |f(z_1, z_2)|^2 d\mu_R(z_1, z_2) \leqslant K_1^2 \|f\|_{\mathcal{D}(\mathbb{D}^2)}^2$$

with K_1 independent of $R < 1$. Here $\mu_R(F) := \mu\left(\frac{1}{R}F\right)$, $F \subset \mathbb{C}^2$, where we consider the measure μ as a Borel measure on $\mathbb{C}^2$, supported on $\overline{\mathbb{D}}^2$. The measure μ_R has support in $\mathbb{D}^2$ and $\mu_R(\partial \mathbb{D}^2) = 0$. By Theorems 1.3 (already proven above for measures inside the bidisc) and 1.4 the measure $\nu_R := \Lambda^* \mu_R$ is subcapacitary on $\overline{T}^2$ for any $R \in (0, 1)$. It is enough to show that this implies the subcapacitary property of $\nu := \Lambda^* \mu$ as well.

Consider an arbitrary point $\alpha = (\alpha_1, \alpha_2) \in T^2$. We recall that it uniquely corresponds to the Carleson box $S_\alpha = S_{\alpha_1} \times S_{\alpha_2}$ with $S_{\alpha_k} = \{\rho_k e^{is_k} : e^{is_k} \in J_k; r_k \leq \rho_k < 1\}$, $k = 1, 2$. Here $r_k = 1 - 2^{-d_T(\alpha_k)+1}$ and J_k is a dyadic interval of generation $d_T(\alpha_k)$ on $\partial \mathbb{D}$ such that $J_k = \Lambda_0(\partial S(\alpha_k))$. Denote by $p(\alpha)$ the 'grandparent' of $\alpha \in T^2$, $p(\alpha) = (p_0(\alpha_1), p_0(\alpha_2))$, where $p_0(\alpha_k)$ is the immediate parent of α_k in respective coordinate tree. We claim that for $R \geq \max(r_1, r_2)$ one has

$$\nu(S(\alpha)) \leq \nu_R(S(p(\alpha))). \tag{31}$$

Indeed, for those values of R we immediately have $RS_\alpha \subset S_{p(\alpha)}$, since $Rr_k \geq 1 - 2(1 - r_k) = 1 - 2^{-d_T(p_0(\alpha_k))+1}$, $k = 1, 2$. The smoothness property (30) implies

$$
\begin{aligned}
\nu(S(\alpha)) &= (\Lambda^* \mu)(S(\alpha)) = \mu(\Lambda(S(\alpha))) = \mu(S_\alpha) = \mu_R(RS_\alpha) \leq \mu_R(S_{p(\alpha)}) = \\
&\mu_R(\Lambda(S(p(\alpha)))) = \nu_R(S(p(\alpha))),
\end{aligned}
$$

recalling $\Lambda(\mathcal{S}(\alpha)) = S_\alpha$.

Consider now any finite collection $\{\alpha^j\}_{j=1}^N$ of points in T^2. Taking R to be greater than $\max(r_k^j)$, $k = 1, 2$, $j = 1, \ldots, N$, we obtain

$$\nu\left(\bigcup_{j=1}^N \mathcal{S}(\alpha^j)\right) \leq \nu_R\left(\bigcup_{j=1}^N \mathcal{S}(p(\alpha^j))\right).$$

Since

$$\text{Cap}\left(\bigcup_{j=1}^N \mathcal{S}(\alpha^j)\right) \approx \text{Cap}\left(\bigcup_{j=1}^N \mathcal{S}(p(\alpha^j))\right),$$

see Lemma 5.3, it follows that $\nu = \Lambda^*\mu$ is a subcapacitary measure with constant comparable to K_1^2, hence we have (C).

Finally we show that $(C) \Rightarrow (A)$. We start with a local estimate for pieces of the (main term of) radial variation. Given a point $\zeta = (\zeta_1, \zeta_2) \in \overline{\mathbb{D}}^2$ let $P(\zeta) := \{z = (r_1\zeta_1, r_2\zeta_2), 0 \leq r_1, r_2 \leq 1\}$, so that $\text{Var}_{12}(\zeta) = \int_{P(\zeta)} |\partial_{z_1 z_2} f(z_1, z_2)| d|z_1| d|z_2|$. For $\zeta \in \overline{\mathbb{D}}^2$ and $\alpha \in T^2$ define

$$
\begin{aligned}
W(\zeta, \alpha) &:= \int_0^{\zeta_1} \int_0^{\zeta_2} \chi_{Q_\alpha \cap P(\zeta)}(z_1, z_2) |\partial_{z_1 z_2} f(z_1, z_2)| \, |dz_1| |dz_2| \\
&= \int_0^1 \int_0^1 \chi_{Q_\alpha}(p\zeta_1, q\zeta_2) |\partial_{z_1 z_2} f(p\zeta_1, q\zeta_2)| \cdot |\zeta_1||\zeta_2| \, dp \, dq \\
&\leq \max\{|\partial_{z_1 z_2} f(z_1, z_2)| : (z_1, z_2) \in Q_\alpha\} \cdot |Q_\alpha|^{1/2} \\
&= |\partial_{z_1, z_2} f(z_1(\alpha), z_2(\alpha))| \cdot |Q_\alpha|^{1/2} \\
&\qquad \text{for some point } z(\alpha) = (z_1(\alpha), z_2(\alpha)) \text{ in } Q_\alpha \\
&= \frac{|Q_\alpha|^{1/2}}{|D(z_1(\alpha), r) \times D(z_2(\alpha), s)|} \left| \int_{D(z_1(\alpha), r) \times D(z_2(\alpha), s)} \partial_{z_1 z_2} f(z_1, z_2) \, dA(z_1) \, dA(z_2) \right| \\
&\qquad \text{by the mean value principle with } r = \frac{1 - |z_1(\alpha)|}{2} \text{ and } s = \frac{1 - |z_2(\alpha)|}{2} \\
&\lesssim \left(\int_{D(z_1(\alpha), r) \times D(z_2(\alpha), s)} |\partial_{z_1 z_2} f(z_1, z_2)|^2 \, dA(z_1) \, dA(z_2) \right)^{1/2} \\
&\qquad \text{by Jensen's inequality.} \\
&=: (H((\alpha))^{1/2}.
\end{aligned}
$$

In other words, every piece of radial variation that passes through Q_α can be estimated by a single quantity $H(\alpha)$. Summing over α along the 'route' $P(\zeta)$ we get

$$
\begin{aligned}
\text{Var}_{12}(f)(\zeta) = \sum_{Q_\alpha \cap P(\zeta) \neq \emptyset} W(\zeta, \alpha) &\lesssim \sum_{Q_\alpha \cap P(\zeta) \neq \emptyset} (H(\alpha))^{1/2} \\
&\leq \sum_{\Lambda^{-1}(\zeta) = \beta} (\mathbb{I} H^{1/2})(\beta).
\end{aligned}
$$

To elaborate, if $\zeta \in \mathbb{D}^2$, then we can identify a Whitney box $Q_\beta \ni \zeta$ in a unique way, uniquely defining $\Lambda^{-1}(\zeta) = \beta$. If ζ lies on the boundary of the bidisc, then ζ has several (but boundedly many) Λ-preimages, and we just sum over all of them. Integrating:

$$
\begin{aligned}
\int_{\mathbb{D}^2} \mathrm{Var}_{12}(f)(\zeta)^2 d\mu(\zeta) &\lesssim \int_{\mathbb{D}^2} \sum_{\Lambda^{-1}(\zeta)=\beta} (\mathbb{I}H^{1/2}(\beta))^2 d\mu(\zeta) \\
&\approx \int_{\mathbb{D}^2} \frac{\sum_{\Lambda^{-1}(\zeta)=\beta} (\mathbb{I}H^{1/2}(\beta))^2}{\sharp\{\beta : \Lambda^{-1}(\zeta) = \beta\}} d\mu(\zeta) \\
&= \int_{T^2} (\mathbb{I}H^{1/2}(\beta))^2 d\Lambda^*\mu(\beta) \\
&\lesssim \sum_{\alpha \in T^2} H(\alpha) \\
&\quad \text{because } \Lambda^*\mu \text{ is a trace measure for the Hardy operator on the bitree} \\
&= \sum_{\alpha \in T^2} \int_{D(z_1(\alpha),r) \times D(z_2(\alpha),s)} |\partial_{z_1 z_2} f(z_1, z_2)|^2 dA(z_1) dA(z_2) \\
&= \int_{\mathbb{D}^2} \sharp\{\alpha : z \in D(z_1(\alpha),r) \times D(z_2(\alpha),s)\} |\partial_{z_1 z_2} f(z_1, z_2)|^2 \, dA(z_1) dA(z_2) \\
&\approx \int_{\mathbb{D}^2} |\partial_{z_1 z_2} f(z_1, z_2)|^2 dA(z_1) dA(z_2).
\end{aligned}
$$

Consider now $\mathrm{Var}_1(f)(\zeta_1) = \int_0^{\zeta_1} |\partial_{z_1} f(z_1, 0)| |dz_1|$, i.e. the radial variation of the function $z_1 \mapsto f(z_1, 0)$. By the one variable result (see [10]), we have that if μ_1 is a Carleson measure for $\mathcal{D}(\mathbb{D})$ then

$$\int_{\mathbb{D}} \mathrm{Var}_1(f)(\zeta_1)^2 d\mu_1(\zeta_1) \lesssim \frac{1}{\pi} \int_{\mathbb{D}} |\partial_{z_1} f(z_1, 0)|^2 dA(z_1). \tag{32}$$

Using the mean value property and Jensen's inequality we get

$$|\partial_{z_1} f(z_1, 0)|^2 = \left| \frac{1}{2\pi} \int_{\partial\mathbb{D}} \partial_{z_1} f(z_1, e^{is}) ds \right|^2 \leq \frac{1}{2\pi} \int_{\partial\mathbb{D}} |\partial_{z_1} f(z_1, e^{is})|^2 ds.$$

Therefore

$$\int_{\mathbb{D}} \mathrm{Var}_1(f)(\zeta_1)^2 d\mu_1(\zeta_1) \lesssim \frac{1}{2\pi^2} \int_{\mathbb{D}} \int_{\partial\mathbb{D}} |\partial_{z_1} f(z_1, e^{is})|^2 ds dx dy \lesssim \|f\|_{\mathcal{D}(\mathbb{D}^2)}^2.$$

We are left to show that if μ is a Carleson measure for $\mathcal{D}(\mathbb{D}^2)$, then the measure μ_1 defined on every subset $A \subseteq \overline{\mathbb{D}}$ by

$$\int_A d\mu_1(\zeta_1) := \int_A \int_{\overline{\mathbb{D}}} d\mu(\zeta_1, \zeta_2)$$

is a Carleson measure for $\mathcal{D}(\mathbb{D})$. Let us prove the implication in the discrete setting. By previous results, see [10], it is enough to show that $\Lambda_0^*\mu_1$ is a trace measure for the Hardy operator on the tree. In the one dimensional case we know that this is equivalent to requiring that

$$\sum_{\beta_1 \leq \alpha_1} (\Lambda_0^*\mu_1)(S_T(\beta_1))^2 \lesssim (\Lambda_0^*\mu_1)(S_T(\alpha_1)) \tag{33}$$

for any $\alpha_1 \in T$. Now

$$(\Lambda_0^* \mu_1)(\mathcal{S}_T(\alpha_1)) = \int_{\mathcal{S}_T(\alpha_1) \times \overline{T}} d\Lambda^* \mu = (\Lambda^* \mu)(\mathcal{S}_T(\alpha_1) \times \overline{T}).$$

Let $g := \chi_{\mathcal{S}_T(\alpha_1) \times \overline{T}}$. Then, for any $\beta = (\beta_1, \beta_2) \in T^2$

$$\mathbb{I}_{\Lambda^* \mu}^* g(\beta) = \int \int_{\mathcal{S}_{T^2}(\beta) \cap (\mathcal{S}_T(\alpha_1) \times \overline{T})} d\Lambda^* \mu = (\Lambda^* \mu)(\mathcal{S}_{T^2}(\beta) \cap (\mathcal{S}_T(\alpha_1) \times \overline{T})).$$

Hence, by the inequality for the adjoint operator we obtain

$$\Lambda^* \mu(\mathcal{S}_T(\alpha_1) \times \overline{T}) = \|g\|_{L^2(\Lambda^* \mu)}^2 \gtrsim \|\mathbb{I}_{\Lambda^* \mu}^* g\|_{\ell^2(T^2)}^2 = \sum_{\beta \in T^2} \Lambda^* \mu(\mathcal{S}_{T^2}(\beta) \cap (\mathcal{S}_T(\alpha_1) \times \overline{T}))^2$$

$$\text{by restricting to } \beta_1 \geq \alpha_1, \, \beta_2 = o$$

$$\geq \sum_{\beta_1 \geq \alpha_1} \Lambda^* \mu(\mathcal{S}_T(\beta_1) \times \overline{T})^2 = \sum_{\beta_1 \geq \alpha_1} \Lambda_0^* \mu_1(\mathcal{S}_T(\beta_1))^2,$$

thus proving inequality (33). Since $\Lambda_0^* \mu_1$ is a trace measure for the Hardy operator on T if and only if μ_1 is Carleson for $\mathcal{D}(\mathbb{D})$, (32) is proved.

The last term $|f(0,0)|$ is elementary to treat. By the subcapacitary property, $(\Lambda^* \mu)(\overline{T}^2) \leq \mathrm{Cap}(\overline{T}^2) C_\mu$, and therefore

$$\int_{\mathbb{D}^2} |f(0,0)|^2 d\mu = |f(0,0)|^2 \mu(\overline{\mathbb{D}}^2) = |f(0,0)|^2 (\Lambda^* \mu)(\overline{T}^2) \lesssim C_\mu |f(0,0)|^2 \leq C_\mu \|f\|_{\mathcal{D}(\mathbb{D}^2)}^2.$$

We are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

It follows immediately that

Proposition 2.2 *Suppose μ on $\partial \mathbb{D}^2$ is a Carleson measure for the Dirichlet space on the bidisc. If* $\|f\|_{\mathcal{D}(\mathbb{D}^2)}^2 < \infty$, *then*

$$\lim_{R \to 1} f(R\zeta_1, R\zeta_2) =: f(\zeta_1, \zeta_2) \text{ exists for } \mu\text{-a.e. } (\zeta_1, \zeta_2) \in (\partial \mathbb{D})^2$$

and

$$\lim_{R \to 1} \int_{(\partial \mathbb{D})^2} |f(R\zeta_1, R\zeta_2)|^2 d\mu(\zeta_1, \zeta_2) = \int_{(\partial \mathbb{D})^2} |f(\zeta)|^2 d\mu(\zeta) \leq \int_{(\partial \mathbb{D})^2} \mathrm{Var}(f)^2 d\mu \leq K_1^2 \|f\|_{\mathcal{D}(\mathbb{D}^2)}^2.$$

2.7 Characterization of multipliers: Theorem 1.2

We start by showing that the left-hand side of (5) dominates the right-hand side. A standard argument with reproducing kernels shows that $\|b\|_{H^\infty} \leq \|M_b\|_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))}$. Namely,

$$|b(z,w)| \|K_{(z,w)}\|_{\mathcal{D}(\mathbb{D}^2)}^2 = |b(z,w)| |K_{(z,w)}(z,w)| = \left| \langle bK_{(z,w)}, K_{(z,w)} \rangle_{\mathcal{D}(\mathbb{D}^2)} \right| \leq \|M_b\|_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \|K_{(z,w)}\|_{\mathcal{D}(\mathbb{D}^2)}^2.$$

On the other hand, we may view M_b as a vector-valued multiplier on the vector-valued Dirichlet space $\mathcal{D}(\mathbb{D}) \otimes \mathcal{D}(\mathbb{D}) \cong \mathcal{D}(\mathcal{D}(\mathbb{D}))$. That is, we identify the function $f(z,w) = \sum_{m,n} a_{m,n} z^m w^n$ with the function $F : \mathbb{D} \to \mathcal{D}(\mathbb{D})$, $F : z \mapsto F(z) = (w \mapsto f(z,w))$. Here we equip $\mathcal{D}(\mathbb{D})$ with the norm given in (1), and $\mathcal{D}(\mathcal{D}(\mathbb{D}))$ with the norm

$$\|F\|^2_{\mathcal{D}(\mathcal{D}(\mathbb{D}))} := \frac{1}{\pi} \int_{\mathbb{D}} \left\| \frac{dF}{dz}(z) \right\|^2_{\mathcal{D}(\mathbb{D})} dA(z) + \sup_{0<r<1} \frac{1}{2\pi} \int_0^{2\pi} \|F(re^{it})\|^2_{\mathcal{D}(\mathbb{D})} dt = \|f\|^2_{\mathcal{D}(\mathbb{D}^2)}.$$

The multiplicator with symbol $b = b(z,w)$ is then identified with

$$M_b \colon f(z,\cdot) \mapsto b(z,\cdot) f(z,\cdot).$$

For each fixed $z \in \mathbb{D}$, let $M_{b(z,\cdot)}$ denote multiplication by $b(z,\cdot)$ on $\mathcal{D}(\mathbb{D})$, and K_z the reproducing kernel of $\mathcal{D}(\mathbb{D})$ at z. Then, see [3, §2.5],

$$M_b^*(K_z \otimes h) = K_z \otimes M_{b(z,\cdot)}^* h, \quad h \in \mathcal{D}(\mathbb{D}),$$

and thus

$$\sup_{z \in \mathbb{D}} \|M_{b(z,\cdot)}\|_{\mathcal{B}(\mathcal{D}(\mathbb{D}))} \leq \|M_b\|_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))}. \tag{34}$$

We can apply Stegenga's characterization [27] of the multipliers of $\mathcal{D}(\mathbb{D})$ to the left hand side of (34), yielding that

$$\sup_{z \in \mathbb{D}} [|\partial_w b(z,\cdot)|^2 dA(\cdot)]_{CM(\mathcal{D}(\mathbb{D}))} \lesssim \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))}.$$

Hence,

$$\int_{\mathbb{D}} |\partial_w b(z,w)|^2 |f(z,w)|^2 dA(w) \lesssim \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \left(\int_{\partial\mathbb{D}} |f(z,e^{it})|^2 dt + \int_{\mathbb{D}} |\partial_w f(z,w)|^2 dA(w) \right), \tag{35}$$

which by integration, standard properties of $H^2(\mathbb{D})$, Fatou's Lemma, and dominated convergence, yields that

$$\int_{\partial\mathbb{D}} \int_{\mathbb{D}} |\partial_w b(e^{is},w)|^2 |f(e^{is},w)|^2 dA(w) \, ds = \int_{\mathbb{D}} \lim_{r \to 1} \int_{\partial\mathbb{D}} |\partial_w b(re^{is},w)|^2 |f(re^{is},w)|^2 ds \, dA(w)$$

$$\leq \lim_{r \to 1} \int_{\mathbb{D}} \int_{\partial\mathbb{D}} |\partial_w b(re^{is},w)|^2 |f(re^{is},w)|^2 ds \, dA(w)$$

$$\lesssim \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \lim_{r \to 1} \int_{\partial\mathbb{D}} \left(\int_{\partial\mathbb{D}} |f(re^{is},e^{it})|^2 dt + \int_{\mathbb{D}} |\partial_w f(re^{is},w)|^2 dA(w) \right) ds$$

$$\leq \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \|f\|^2_{\mathcal{D}(\mathbb{D}^2)}. \tag{36}$$

Applying (35) with $\partial_z f(z,\cdot)$ in place of f and integrating also yields that

$$\int_{\mathbb{D}} \int_{\mathbb{D}} |\partial_w b(z,w)|^2 |\partial_z f(z,w)|^2 dA(z) \, dA(w) \lesssim \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \|f\|^2_{\mathcal{D}(\mathbb{D}^2)}. \tag{37}$$

Similarly, the inequalities (34)-(37) also hold with the roles of z and w reversed.

Suppose M_b is bounded. Writing out the norm of bf, $f \in \mathcal{D}(\mathbb{D}^2)$, applying the triangle inequality, the fact that $b \in H^\infty(\mathbb{D}^2)$, and inequality (37) yields that

$$\int_{\mathbb{D}^2} |\partial_{zw}(bf)|^2 \, dA(z) \, dA(w)$$
$$\gtrsim \int_{\mathbb{D}^2} |(\partial_{zw}b)f|^2 dA(z) \, dA(w) - \int_{\mathbb{D}^2} \left[|(\partial_z b)\partial_w f|^2 + |(\partial_w b)\partial_z f|^2 + |b\partial_{zw}f|^2 \right] dA(z) \, dA(w)$$
$$\gtrsim \int_{\mathbb{D}^2} |(\partial_{zw}b)f|^2 dA(z) \, dA(w) - \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \|f\|^2_{\mathcal{D}(\mathbb{D}^2)} - \|b\|^2_{H^\infty} \|f\|^2_{\mathcal{D}(\mathbb{D}^2)}.$$

Hence,

$$\|M_b f\|^2_{\mathcal{D}(\mathbb{D}^2)} \gtrsim \int_\mathbb{D} \int_\mathbb{D} |f|^2 \, d\mu_b - \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))} \|f\|^2_{\mathcal{D}(\mathbb{D}^2)},$$

and thus

$$[\mu_b]_{CM} \lesssim \|M_b\|^2_{\mathcal{B}(\mathcal{D}(\mathbb{D}^2))}.$$

The computations thus far have shown that the left-hand side of (5) dominates the right-hand side. The converse inequality is also clear, using the triangle inequality, from the estimates we have made.

3 Strong Capacitary Inequality on the bitree

Here we prove Theorem 1.5. First we establish some extra notation. Similarly to the definitions in Section 2.3 we define the one-dimensional Hardy operator, its adjoint, and the logarithmic potential on the tree T:

$$(I\varphi)(\alpha) := \sum_{\gamma \geq \alpha} \varphi(\gamma);$$
$$(I^*\mu)(\beta) := \int_{\mathcal{S}_T(\beta)} d\mu(\alpha);$$
$$V^\mu := (II^*)(\mu),$$

where $\varphi \geq 0$ is a function on T and $\mu \geq 0$ is a Borel measure on $\overline{T}$. The (one-dimensional) logarithmic capacity is defined in the same way,

$$\operatorname{Cap} E = \inf\{\|\varphi\|^2_{\ell^2(T)} : I\varphi \geq 1 \text{ on } E\},$$

for a Borel set $E \subset \overline{T}$. We aim to prove the following result:

Theorem 3.1 *For any $f : T^2 \mapsto \mathbb{R}_+$ in $\ell^2(T^2)$ we have*

$$\sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap}\{\alpha \in \overline{T}^2 : \mathbb{I}f(\alpha) \geq 2^k\} \lesssim \sum_{\alpha \in T^2} f^2(\alpha) := \|f\|^2_{\ell^2(T^2)}. \tag{38}$$

Similar results were obtained by Adams [1], Maz'ya [23], and others. However, they were based on a certain property of the respective potential-theoretic kernels — that they were of 'radial nature'. In our context this roughly translates to the uniqueness of the geodesic between two different vertices of the

underlying graph. While this property is elementary for a uniform dyadic tree T (as well for a p-adic tree), the bitree T^2 does not enjoy it any more, and this is one of the main problems we have to overcome when we increase the dimension.

To highlight this difference we give a rough sketch of the proof for $d = 1$ (i.e. for a dyadic tree). Given $k \in \mathbb{Z}$ assume that $If(\alpha) \approx 2^{k+1}$ and $If(\beta) \approx 2^k$ for $\beta > \alpha$ in the tree. Then, since there exists a unique geodesic connecting α and β, we have that $\sum_{\beta > \tau \geq \alpha} f(\tau) \approx 2^k$. Therefore one can expect, if we set $f_k := \chi_{2^k < If \leq 2^{k+1}} f$, that $If_k \approx 2^k \approx If$ on $\{If \geq 2^{k+1}\}$, so that $2^{-k+1} f_k$ is admissible for this set, and $\sum_{k \in \mathbb{Z}} 2^{2k} \mathrm{Cap}\{If \geq 2^{k+1}\} \lesssim \sum_{k \in \mathbb{Z}} \|f_k\|^2 \approx \|f\|^2$, since the functions f_k have disjoint supports.

However we see that already for $d = 2$ there are many geodesics in T^2 with endpoints at α and β, and the above argument fails, since one can construct a function f_0 such that $\mathbb{I} f_0(\alpha) \leq 1$ for every $\alpha \in \mathrm{supp} f_0$, but for every $\lambda > 1$ there exists a point $\beta \in T^2$ with $\mathbb{I} f_0(\beta) > \lambda$ (see Proposition 5.2). In other words, the maximum principle does not hold for T^2. However, while it fails pointwise, a quantitative version of the maximum principle is still true — the set of 'bad' points has asymptotically small capacity, see Corollary 3.1. Therefore we can salvage enough of the argument to obtain Theorem 3.1.

The proof is based on the following rearrangement lemma, Lemma 3.1. We explain how it implies Theorem 3.1 in Section 3.1. Lemma 3.1 is proved in Section 3.3 by reduction to a one-dimensional statement, Lemma 3.4, which in turn is proved in Section 3.4.

Lemma 3.1 (Rearrangement Lemma) *Let $\mu \geq 0$ be any Borel measure with finite energy on $(\partial T)^2$. Given $\delta > 0$ we define the δ-level set of $\mathbb{V}^\mu$ by*

$$E^\delta := \{\alpha \in T^2 : \mathbb{V}^\mu(\alpha) \leq \delta\}.$$

We also define the δ-restricted potential and energy by setting

$$\mathbb{V}_\delta^\mu(\alpha) := \sum_{\beta \in E^\delta : \beta \geq \alpha} (\mathbb{I}^* \mu)(\beta); \quad \mathcal{E}_\delta[\mu] := \sum_{\alpha \in E^\delta} (\mathbb{I}^* \mu)^2(\alpha) = \int_{(\partial T)^2} \mathbb{V}_\delta^\mu \, d\mu.$$

For $\lambda \geq \delta$ let

$$E_{\delta, \lambda} := \{\alpha \in (\partial T)^2 : \mathbb{V}_\delta^\mu(\alpha) > \lambda\}.$$

Then there exists a function $\varphi : \overline{T}^2 \to \mathbb{R}_+$ supported on T^2 such that

$$\mathbb{I}\varphi(\alpha) = \sum_{\beta \geq \alpha} \varphi(\beta) > \lambda, \quad \alpha \in E_{\delta, \lambda}; \tag{39a}$$

$$\|\varphi\|_{\ell^2(T^2)}^2 = \sum_{\alpha \in T^2} (\varphi(\alpha))^2 \lesssim \frac{\delta}{\lambda} \mathcal{E}_\delta[\mu] = \frac{\delta}{\lambda} \sum_{\alpha \in E^\delta} (\mathbb{I}^* \mu(\alpha))^2. \tag{39b}$$

Remark. Observe that, by the maximum principle, $E_{\delta, \lambda} = \emptyset$ in the one-dimensional setting of a tree T.

Corollary 3.1 *Let $E \subset (\partial T)^2$ be a Borel set, and $\mu = \mu_E$ be the equilibrium measure for E. Given $\lambda > 1$ define*

$$E_\lambda := \{\omega \in (\partial T)^2 : \mathbb{V}^\mu(\omega) = \sum_{\beta \in \mathcal{P}(\omega)} \mathbb{I}^* \mu(\beta) > \lambda\}. \tag{40}$$

Then

$$\mathrm{Cap}\, E_\lambda \lesssim \frac{1}{\lambda^3} \mathrm{Cap}\, E. \tag{41}$$

Proof. Put $\delta = 1$. Since μ is equilibrium for E, we have $\{\alpha \in T^2 : (\mathbb{I}^* \mu)(\alpha) > 0\} \subset E^1$ and $E_{1, \lambda} = E_\lambda$. It remains to apply Lemma 3.1 with data $1, \lambda$. $\qquad\square$

3.1 Deducing Theorem 3.1 from Corollary 3.1

Here we mostly follow the argument from Adams and Hedberg [2, Chapter 7]. First we separate the ℓ^2-norm of f, reducing (38) to estimates of the level sets of $\mathbb{I}f$. We then prove that the energy scalar product of two equilibrium measures can be estimated by the capacities of the respective sets, Lemma 3.2. This is the key point of the argument, and it is here that we use the Rearrangement Lemma 3.1 (or, more precisely, Corollary 3.1). We finish the proof by showing that the mixed energy of the level sets (energy scalar product of their equilibrium measures) is concentrated on the diagonal (inequality (43)).

Removing $\|f\|_{\ell^2(T^2)}$.

Given $k \in \mathbb{Z}$ let $\tilde{E}_k$ be the k-th level set of $\mathbb{I}f$,

$$\tilde{E}_k = \{\alpha \in \overline{T}^2 : (\mathbb{I}f)(\alpha) > 2^k\}.$$

We then define E_k to be the boundary projection of $\tilde{E}_k$

$$E_k = \mathcal{S}(\tilde{E}_k) \cap (\partial T)^2,$$

where $\mathcal{S}(\tilde{E}_k) = \bigcup_{\beta \in \tilde{E}_k} \mathcal{S}(\beta)$. Corollary 5.1 (see Section 5.4) implies that

$$\operatorname{Cap} \tilde{E}_k \approx \operatorname{Cap} E_k, \quad k \in \mathbb{Z}.$$

Hence (38) is equivalent to

$$\sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap} E_k \lesssim \|f\|_{\ell^2(T^2)}^2.$$

Since $E_k \subset \mathcal{S}(\tilde{E}_k)$, we see that $\mathbb{I}f \geq 2^k$ on E_k as well. By its nature, E_k is a countable union of clopen rectangles on $(\partial T)^2$, hence $E_k = \bigcup_{j=1}^{\infty} E_k^j$ with $\{E_k^j\}$ is an increasing sequence of compact sets. Define μ_k and μ_k^j to be equilibrium measures for E_k and E_k^j respectively. Clearly $\lim_{j \to \infty} \operatorname{Cap} E_k^j = \operatorname{Cap} E_k$. For $k \in \mathbb{Z}$ and $j \geq 1$ we have

$$2^{2k} \operatorname{Cap} E_k^j = 2^k \int_{\overline{T}^2} 2^k \, d\mu_k^j \leq 2^k \int_{\overline{T}^2} \mathbb{I}f \, d\mu_k^j = 2^k \int_{T^2} f \, d(\mathbb{I}^* \mu_k^j) = 2^k \sum_{\alpha \in T^2} f(\alpha)(\mathbb{I}^* \mu_k^j)(\alpha)$$

by Tonelli's theorem. Since $\mathbb{I}^* \mu_k^j \to \mathbb{I}^* \mu_k$ in $\ell^2(T^2)$ [2, Proposition 2.3.12], we may pass to the limit as $j \to \infty$, obtaining

$$2^{2k} \operatorname{Cap} E_k \leq 2^k \int_{T^2} f \, d(\mathbb{I}^* \mu_k).$$

Summing this estimate over $k \in \mathbb{Z}$ and applying Cauchy-Schwartz we arrive at

$$\sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap} E_k \leq \int_{T^2} f \sum_{k \in \mathbb{Z}} 2^k \, d(\mathbb{I}^* \mu_k) \leq \|f\|_{\ell^2(T^2)} \|\sum_{k \in \mathbb{Z}} 2^k \mathbb{I}^* \mu_k\|_{\ell^2(T^2)}.$$

We conclude that (38) follows from

$$\|\sum_{k \in \mathbb{Z}} 2^k \mathbb{I}^* \mu_k\|_{\ell^2(T^2)}^2 \lesssim \sum_{k \in \mathbb{Z}} 2^{2k} \operatorname{Cap} E_k. \tag{42}$$

Equilibrium potential on the subset.

Expanding the left-hand side of (42) we obtain

$$\sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}}2^{j+k}\int_{\overline{T}^2}\mathbb{V}^{\mu_k}\,d\mu_j,$$

and this expression is symmetric over j and k. Therefore (42) is equivalent to

$$\sum_{k\in\mathbb{Z}}\sum_{j\leq k}2^{j+k}\int_{\overline{T}^2}\mathbb{V}^{\mu_k}\,d\mu_j\lesssim\sum_{k\in\mathbb{Z}}2^{2k}\int_{\overline{T}^2}\mathbb{V}^{\mu_k}\,d\mu_k,\tag{43}$$

since $\operatorname{Cap}E_k=|\mu_k|=\int_{\overline{T}^2}\mathbb{V}^{\mu_k}\,d\mu_k$. We see that in order to prove (38) we need to show that the sum on the left-hand side of (43) is dominated by its diagonal term. First we state the following lemma.

Lemma 3.2 *Let $F,E\subset(\partial T)^2$ be a pair of sets on the distinguished boundary of T^2, such that $\operatorname{Cap}F\leq\operatorname{Cap}E$, and let μ_F,μ_E be their equilibrium measures. Then there exists an absolute constant $C>1$ such that*

$$\int_{\overline{T}^2}\mathbb{V}^{\mu_E}\,d\mu_F\leq C|\mu_E|^{\frac{1}{3}}|\mu_F|^{\frac{2}{3}}=C(\operatorname{Cap}E)^{\frac{1}{3}}(\operatorname{Cap}F)^{\frac{2}{3}}.\tag{44}$$

Proof. If $\operatorname{Cap}F=0$, then (44) is trivial. Therefore we let

$$\lambda=\frac{\int_{\overline{T}^2}\mathbb{V}^{\mu_E}\,d\mu_F}{|\mu_F|},$$

and we assume that $\lambda\geq16$ (otherwise we just set C to be large enough). Define F_k, $k\in\mathbb{Z}_+$, to be the level sets of $\mathbb{V}^{\mu_E}$ on $\operatorname{cl}F$

$$\begin{aligned}F_k&:=\{\alpha\in\operatorname{cl}F:\ 2^{k-1}\leq\mathbb{V}^{\mu_E}(\alpha)<2^k\},\quad k\geq1\\F_0&:=\{\alpha\in\operatorname{cl}F:\ \mathbb{V}^{\mu_E}<1\}.\end{aligned}\tag{45}$$

Corollary 3.1 implies that

$$\operatorname{Cap}F_k\leq C2^{-3k}\operatorname{Cap}E$$

for some $C>0$. Clearly F_k, $k\geq0$, are disjoint, and $\operatorname{Cap}\left(\operatorname{cl}F\setminus\bigcup_{k=0}^{\infty}F_k\right)=0$ (since the potential of μ_E can be infinite only on a polar set). Hence we have

$$\lambda|\mu_F|=\int_{\overline{T}^2}\mathbb{V}^{\mu_E}\,d\mu_F=\sum_{k\geq0}\int_{F_k}\mathbb{V}^{\mu_E}\,d\mu_F\leq$$

$$\sum_{k\geq0}\int_{F_k}2^k\,d\mu_F=\sum_{k\geq0}2^k\mu_F(F_k).$$

Fix j such that $2^{j-1}\leq\lambda<2^j$. Since $\mu_F(F_k)=\mu_F(F_k\cap\operatorname{supp}\mu_F)\leq\operatorname{Cap}F_k$ (see Lemma 5.6), we get

$$\lambda|\mu_F|\leq\sum_{k=0}^{j-3}2^k\mu_F(F_k)+\sum_{k\geq j-2}2^k\operatorname{Cap}F_k\leq$$

$$2^{j-2}\mu_F(\operatorname{cl}F)+C\sum_{k\geq j-2}2^{-2k}\operatorname{Cap}E\leq\frac{\lambda}{2}\mu_F(\operatorname{cl}F)+C\lambda^{-2}\operatorname{Cap}E.$$

Hence

$$|\mu_F| \leq C\lambda^{-3}|\mu_E|,$$

which immediately implies (44). □

As we will see below, in order to prove the Strong Capacitary Inequality it actually suffices to show that

$$\int_{\overline{T}^2} \mathbb{V}^{\mu_E} \, d\mu_F \leq |\mu_E|^{\frac{1}{2}-\varepsilon}|\mu_F|^{\frac{1}{2}+\varepsilon},$$

for some $\varepsilon > 0$. Hölder's inequality gives $|\mu_E|^{\frac{1}{2}}|\mu_F|^{\frac{1}{2}}$ on the right-hand side (which is not good enough). On the other hand, in the tree setting, or, more generally, in any setting where the Maximum Principle holds, one has much better estimate

$$\int_{\overline{T}^2} \mathbb{V}^{\mu_E} \, d\mu_F \lesssim |\mu_F|.$$

Estimates of the potentials of the level sets.

Applying Lemma 3.2 to the sets E_k and E_j, $k \geq j$, yields

$$\int_{\overline{T}^2} \mathbb{V}^{\mu_j} \, d\mu_k \lesssim |\mu_k|^{\frac{2}{3}}|\mu_j|^{\frac{1}{3}},$$

since $E_k \subset E_j$. By Hölder's inequality

$$\sum_{k \in \mathbb{Z}} \sum_{j \leq k} 2^{k+j} \int_{\overline{T}^2} \mathbb{V}^{\mu_j} \, d\mu_k \lesssim \sum_{k \in \mathbb{Z}} \sum_{j \leq k} 2^{k+j} |\mu_k|^{\frac{2}{3}}|\mu_j|^{\frac{1}{3}} =$$

$$\sum_{k \in \mathbb{Z}} 2^{\frac{4}{3}k}|\mu_k|^{\frac{2}{3}} \cdot 2^{-\frac{1}{3}k} \sum_{j \leq k} 2^j |\mu_j|^{\frac{1}{3}} \leq \left(\sum_{k \in \mathbb{Z}} 2^{2k}|\mu_k| \right)^{\frac{2}{3}} \left(\sum_{k \in \mathbb{Z}} 2^{-k} \left(\sum_{j \leq k} 2^j |\mu_j|^{\frac{1}{3}} \right)^3 \right)^{\frac{1}{3}}.$$

Another application of Hölder's inequality to the second term on the right-hand side gives

$$\sum_{k \in \mathbb{Z}} 2^{-k} \left(\sum_{j \leq k} 2^j |\mu_j|^{\frac{1}{3}} \right)^3 = \sum_{k \in \mathbb{Z}} 2^{-k} \left(\sum_{j \leq k} 2^{\frac{1}{6}j} 2^{\frac{5}{6}j} |\mu_j|^{\frac{1}{3}} \right)^3 \lesssim$$

$$\sum_{k \in \mathbb{Z}} 2^{-k} \sum_{j \leq k} 2^{\frac{1}{2}j} \sum_{j \leq k} 2^{\frac{5}{2}j} |\mu_j| \lesssim \sum_{k \in \mathbb{Z}} 2^{2k}|\mu_k|.$$

Gathering the estimates we obtain

$$\sum_{k \in \mathbb{Z}} \sum_{j \leq k} 2^{k+j} \int_{\overline{T}^2} \mathbb{V}^{\mu_j} \, d\mu_k \lesssim \sum_{k \in \mathbb{Z}} 2^{2k}|\mu_k|,$$

which is (43).

We note that in the last part of the proof we did not use the fact that the sets E_k are generated by the function f. Indeed, (43) holds for any nested sequence $\{E_k\}$ of sets on the distinguished boundary.

3.2 Rearrangement and the energy decay

Before proceeding to the proof of Rearrangement Lemma we show how one can also deduce the energy decay rate of a measure outside its support. See e.g.[6] for some other applications.

Proposition 3.1 *Assume $\mu \geq 0$ is a Borel measure on $(\partial T)^2$ such that $\mathbb{V}^\mu \geq 1$ on $\operatorname{supp}\mu$. Then for any $0 < \delta \leq 1$ one has*

$$\mathcal{E}_\delta[\mu] \lesssim \delta^{\frac{1}{3}} \mathcal{E}[\mu].$$

In particular, there exists $\delta > 0$ such that

$$\mathcal{E}[\mu] - \mathcal{E}_\delta[\mu] = \sum_{\alpha:\, \mathbb{V}^\mu(\alpha) \geq \delta} (\mathbb{I}^*\mu)^2(\alpha) \geq \frac{9}{10} \mathcal{E}[\mu].$$

Proof. Fix $\delta \in (0,1]$, and let η_δ be such that

$$\eta_\delta := \frac{\mathcal{E}_\delta[\mu]}{\mathcal{E}[\mu]}.$$

Given $k \in \mathbb{Z}$ let $F_k := \{\alpha \in \operatorname{supp}\mu :\ 2^{-k} \geq \mathbb{V}_\delta(\alpha) > 2^{-k-1}\}$. Applying the rearrangement procedure to $\lambda = 2^{-k} \geq \delta$ we obtain a function φ_k that satisfies (39). In particular, $\mathbb{I}\varphi_k \geq 2^{-k}$ on F_k, therefore by Tonelli's theorem and (39b) one has

$$2^{-k}\mu(F_k) = \int_{F_k} 2^{-k}\, d\mu \leq \int_{F_k} \mathbb{I}\varphi_k\, d\mu = \int_{(\partial T)^2} \chi_{F_k}(\omega) \left(\sum_{\beta \geq \omega} \varphi_k(\beta) \right) d\mu(\omega) \leq$$

$$\int_{(\partial T)^2} \left(\sum_{\beta \geq \omega} \varphi_k(\beta) \right) d\mu(\omega) = \sum_{\beta \in T^2} \varphi_k(\beta)(\mathbb{I}^*\mu)(\beta) \leq \left(\sum_{\beta \in T^2} \varphi_k^2(\beta) \right)^{\frac{1}{2}} \left(\sum_{\beta \in T^2} (\mathbb{I}^*\mu)^2(\beta) \right)^{\frac{1}{2}} \leq$$

$$C_0 \left(2^k \delta \mathcal{E}_\delta[\mu] \right)^{\frac{1}{2}} \mathcal{E}^{\frac{1}{2}}[\mu] \leq C_0 2^{\frac{k}{2}} \delta^{\frac{1}{2}} \mathcal{E}[\mu]$$

for some absolute constant $C_0 > 1$. Set $N = N(\delta)$ to be such that $2^{-N+1} \geq \frac{1}{(10C_0)^2} \eta_\delta \geq 2^{-N}$. If $\eta_\delta \leq 2(10C_0^2)^2 \delta$, then the result follows immediately, hence we may assume that $2^{-N} \geq \delta$. Summing up over k we realize that

$$\mathcal{E}_\delta[\mu] = \int_{(\partial T)^2} \mathbb{V}_\delta^\mu\, d\mu \leq 2 \sum_{k=-\infty}^{+\infty} \int_{F_k} \mathbb{V}_\delta^\mu\, d\mu$$

$$\leq 2 \sum_{k=-\infty}^{N} 2^{-k}\mu(F_k) + 2^{-N+1}\mu\left(\left\{ \alpha \in \operatorname{supp}\mu :\ \mathbb{V}_\delta^\mu(\alpha) \leq 2^{-N} \right\} \right)$$

$$\leq 2 \sum_{k=-\infty}^{N} 2^{-k}\mu(F_k) + 2^{-N+1}|\mu| \leq 2 \sum_{k=-\infty}^{N} 2^{-k}\mu(F_k) + 2^{-N+1}\mathcal{E}[\mu],$$

since $\mathcal{E}[\mu] = \int_{(\partial T)^2} \mathbb{V}^\mu\, d\mu$ and $\mathbb{V}^\mu \geq 1$ on the support of μ. Therefore

$$\mathcal{E}_\delta[\mu] = \eta_\delta \mathcal{E}[\mu] \leq 2^{-N+1}\mathcal{E}[\mu] + 2C_0 \sum_{k \leq N} 2^{\frac{k}{2}} \delta^{\frac{1}{2}} \mathcal{E}[\mu],$$

hence

$$\eta_\delta \leq \frac{1}{5}\eta_\delta + 4C_0 2^{\frac{N}{2}}\delta^{\frac{1}{2}} \leq \frac{1}{5}\eta_\delta + \frac{4}{5}\eta_\delta^{-\frac{1}{2}}\delta^{\frac{1}{2}},$$

and we are done. □

3.3 Proof of Lemma 3.1. Reducing dimension.

We assume that $\delta = 1$ (otherwise we just rescale by replacing λ by λ/δ in the $\delta = 1$ statement), and from now on we write E_λ instead of $E_{1,\lambda}$.

We construct the function φ that satisfies (39). It is done separately on each layer of the form $T_x \times \{\alpha_y\}$, $\alpha_y \in T_y$ (although the statement we wish to prove is symmetric in the two components x and y, our argument is not, and it can be of course developed exchanging their roles). More precisely, for every $\alpha_y \in T_y$ we produce a function $\varphi_{\alpha_y} : \overline{T}^2 \to \mathbb{R}_+$ such that *(a)* it is supported on the layer $R = T_x \times \{\alpha_y\}$; *(b)* on a certain subset A_R of the set $E_\lambda \cap (\partial T_x \times \mathcal{S}(\alpha_y))$ it gives at least as much potential as μ restricted to this layer

$$(\mathbb{I}\varphi_{\alpha_y})(\varpi) = \sum_{\alpha \geq \varpi} \varphi_{\alpha_y}(\alpha) = \sum_{\alpha_x \geq \varpi_x} \varphi_{\alpha_y}(\alpha_x, \alpha_y) \geq \sum_{\alpha_x \geq \varpi_x : \alpha \in E^1} (\mathbb{I}^*\mu)(\alpha_x, \alpha_y), \quad \varpi \in E_R; \qquad (46)$$

(c) its ℓ^2-norm is much smaller than the energy of $\mu|_R$

$$\|\varphi_{\alpha_y}\|_{\ell^2(T^2)}^2 = \sum_{\alpha \in T^2} \varphi_{\alpha_y}^2(\alpha_x, \alpha_y) = \sum_{\alpha_x \in T_x} \varphi_{\alpha_y}^2(\alpha_x, \alpha_y) \lesssim \frac{1}{\lambda} \sum_{\alpha_x \in T_x : \alpha \in E^1} (\mathbb{I}^*\mu)^2(\alpha_x, \alpha_y), \qquad (47)$$

where we have recalled that φ_{α_y} is supported only on $T_x \times \{\alpha_y\}$. Each layer $T_x \times \{\alpha_y\}$ is essentially a dyadic tree, and the (restricted) potential of μ exhibits one-dimensional behaviour there, so we can consider the problem in the dyadic tree setting and use one-dimensional arguments.

Finally we set $\varphi = \frac{3}{2}\sum_{\alpha_y \in T_y} \varphi_{\alpha_y}$, and show that φ satisfies (39a) and (39b); the second inequality immediately follows from (47), since $T^2 = \bigcup_{\alpha_y \in T_y} T_x \times \{\alpha_y\}$ and the supports of φ_{α_y} are disjoint.

Construction of φ_{α_y}

Given $\gamma \in \overline{T}$ we define $\partial\mathcal{S}(\gamma) := \mathcal{S}(\gamma) \cap \partial T$ to be the boundary successor set of γ. Fix a point $\alpha_y \in T_y$, and let

$$A_R = \left\{ \varpi \in E_\lambda \cap (\partial T_x \times \partial\mathcal{S}(\alpha_y)) : \mathbb{V}_1^\mu(\varpi_x, \alpha_y) > \frac{\lambda}{3} \right\}.$$

In other words, $\varpi \in E_\lambda$ is in A_R, if $\varpi_y \leq \alpha_y$, and the (restricted) potential of μ at the 'fiber' (ϖ_x, α_y) is large enough. Define F_R to be the projection of A_R on the coordinate tree T_x,

$$F_R = \{\varpi_x \in \partial T_x : \text{ there exists } \varpi_y \leq \alpha_y \text{ such that } (\varpi_x, \varpi_y) \in A_R\}.$$

Observe that F_R is an open set in T_x.

We proceed by performing the dimension reduction argument — we restate our problem on the dyadic tree T_x. To do so we introduce auxiliary functions f_R and g_R supported on T_x. Let

$$f_R(\beta_x) := \int_{\partial\mathcal{S}(\beta_x)} \int_{\partial\mathcal{S}(\alpha_y)} d\mu(\varpi_x, \varpi_y) = (\mathbb{I}^*\mu)(\beta_x, \alpha_y),$$
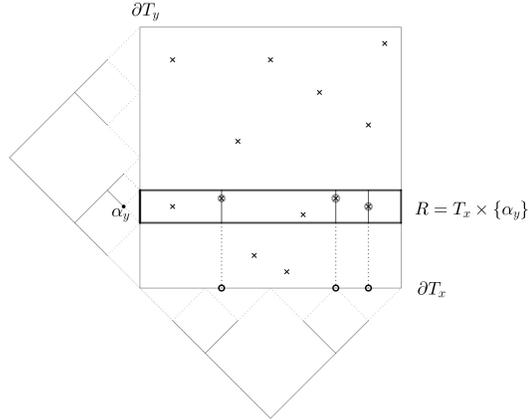
Figure 4:
$\times$: points in E_λ
$\otimes$: points in A_R
$\bigcirc$: points in F_R

if $(\beta_x, \alpha_y) \in E^1$ (i.e. if $\mathbb{V}^\mu(\beta_x, \alpha_y) \leq 1$), and $f_R(\beta_x) := 0$ otherwise. Next,

$$g_R(\beta_x) := \sum_{\beta_y \geq \alpha_y:\, (\beta_x, \beta_y) \in E^1} \int_{\partial \mathcal{S}(\beta_x)} \int_{\partial \mathcal{S}(\beta_y)} d\mu(\omega_x, \omega_y) = \sum_{\beta_y \geq \alpha_y:\, (\beta_x, \beta_y) \in E^1} (\mathbb{I}^* \mu)(\beta_x, \beta_y).$$

Therefore

$$(Ig_R)(\alpha_x) = \sum_{\beta_x \geq \alpha_x} g_R(\beta_x) = \sum_{\beta \geq \alpha:\, \beta \in E^1} (\mathbb{I}^* \mu)(\beta) = \mathbb{V}_1^\mu(\alpha_x, \alpha_y), \quad \alpha_x \in \overline{T}_x,\ \alpha = (\alpha_x, \alpha_y),$$

and, in particular,

$$(Ig_R)(\omega_x) = \mathbb{V}_1^\mu(\omega_x, \alpha_y) > \frac{\lambda}{3}, \tag{48}$$

if $\omega_x \in F_R$. On the other hand, if $\alpha_x \in \operatorname{supp} f_R$, then by definition of E^1 one has

$$(Ig_R)(\alpha_x) \leq 1. \tag{49}$$

Next we present some crucial properties of f_R and g_R.

Lemma 3.3 *Let f_R and g_R be as above. Given $\alpha_x \in T_x$ one always has $f_R(\alpha_x) \geq f_R(\alpha_x^-) + f_R(\alpha_x^+)$, where $\alpha_x^\pm$ are two children of α_x. In particular, the function $I f_R$ is positive superharmonic on T_x (i.e. $(I f_R)(\alpha_x) \geq \frac{1}{3}\left((I f_R)(\alpha_x^+) + (I f_R)(\alpha_x^-) + (I f_R)(p(\alpha_x))\right)$), and for any $\alpha_x \in T_x$ either $f_R(\alpha_x) = 0$, or $f(\beta_x) > 0$ for any $\beta_x \geq \alpha_x$. The same is true for g_R.*

Proof. All of these properties immediately follow from the definition of f_R an g_R, and the fact that if $(\beta_x, \beta_y) \in E^1$, then (γ_x, γ_y) is in E^1 as well for any $\gamma_x \geq \beta_x$, $\gamma_y \geq \beta_y$. $\qquad \square$

The inequalities (48) and (49) show that F_R is, in a sense, far away from the support of f_R, and we can express this property only in terms of F_R and supp f_R. Since F_R is open in T_x, we can exhaust it by compacts, i.e. there exists an increasing sequence of compact sets F_k such that $F = \bigcup_{k=1}^{\infty} F_k$. Define ρ_k and ρ_R to be equilibrium measures for F_k and F_R. We have $\frac{\lambda}{3} V^{\rho_k} \leq I g_R$ on F_k. By the Domination Principle, given in Lemma 5.6, it follows that

$$\frac{\lambda}{3} V^{\rho_k} \leq I g_R$$

everywhere on T_x. In particular,

$$V^{\rho_k}(\tau_x) \leq \frac{3}{\lambda}, \quad \tau_x \in \text{supp} f_R.$$

Since $V^{\rho_k} \to V^{\rho_R}$ pointwise on T_x, we have

$$V^{\rho_R}(\tau_x) \leq \frac{3}{\lambda}, \quad \tau_x \in \text{supp} f_R.$$

We have moved all pieces of our problem, constructing φ_{α_y}, to the dyadic tree T_x, and its solution is given by the following lemma, the proof of which will be given in the next subsection.

Lemma 3.4 (One-dimensional statement) *Let F be an open set on the boundary ∂T of the dyadic tree T and let ρ be its equilibrium measure. Assume that a function $f : T \to \mathbb{R}_+$ satisfies*

$$V^{\rho}(\alpha) \leq \delta, \quad \alpha \in \text{supp} f \tag{50}$$

with some $\delta \leq \frac{1}{3}$. Then there exists a non-negative measure σ such that

$$V^{\sigma}(\omega) \geq (If)(\omega), \quad \omega \in F; \tag{51a}$$
$$\mathcal{E}[\sigma] \lesssim \delta \|f\|_{\ell^2(T)}^2. \tag{51b}$$

Suppose for the moment that Lemma 3.4 holds. We apply it with $T = T_x$, $F = F_R$, $\delta = \frac{3}{\lambda}$ and $f = f_R$ to obtain a measure $\sigma = \sigma_R$ supported on $F_R \subset \partial T_x$ that satisfies (51). Now we define

$$\varphi_{\alpha_y}(\alpha_x, \alpha_y) := (I^* \sigma_R)(\alpha_x), \quad \alpha_x \in T_x,$$

and we set $\varphi_{\alpha_y} \equiv 0$ outside of $T_x \times \{\alpha_y\}$. We see that (51) implies (46) and (47). Finally we let

$$\varphi = \frac{3}{2} \sum_{\alpha_y \in T_y} \varphi_{\alpha_y}.$$

We are left to show that φ is the desired function, that is, it satisfies (39). The inequality (39b) follows immediately from (47)

$$\|\varphi\|_{\ell^2(T^2)}^2 = \frac{9}{4} \sum_{\alpha_y \in T_y} \|\varphi_{\alpha_y}\|_{\ell^2(T^2)}^2 = \frac{9}{4} \sum_{\alpha_y \in T_y} \|(I^* \sigma_R)\|_{\ell^2(T_x)}^2 \lesssim \sum_{\alpha_y \in T_y} \frac{1}{\lambda} \|f_R\|_{\ell^2(T_x)}^2 =$$

$$\frac{1}{\lambda} \sum_{\alpha = (\alpha_x, \alpha_y) : \alpha \in E^1} f_R^2(\alpha_x) = \frac{1}{\lambda} \sum_{\alpha \in E^1} (\mathbb{I}^* \mu)^2(\alpha).$$

To prove (39a) we use a stopping time argument. Fix a point $\omega \in E_\lambda$. We define $\alpha_y(\omega)$ to be the first (with respect to the natural order on T_y) point such that the (restricted) potential of μ on the fiber $(\omega_x, \alpha_y(\omega))$ exceeds $\frac{\lambda}{3}$. In other words,

$$\mathbb{V}_1^\mu(\omega_x, \alpha_y) > \frac{\lambda}{3}, \quad \alpha_y \leq \alpha_y(\omega),$$

and

$$\mathbb{V}_1^\mu(\omega_x, \alpha_y) \leq \frac{\lambda}{3}, \quad \alpha_y > \alpha_y(\omega)$$

(if $\mathbb{V}_1^\mu(\omega_x, o_y) \geq \frac{\lambda}{3}$, we set $\alpha_y(\omega) = o_y$, where o_y is the root of T_y). Clearly $\omega \in A_R$ with $R = T_x \times \{\alpha_y\}$ for $\alpha_y \leq \alpha_y(\omega)$ – remember that $\omega \in A_R$, if $\mathbb{V}_1^\mu(\omega_x, \alpha_y) > \frac{\lambda}{3}$. Therefore

$$V_1^\mu(\omega) = \sum_{\alpha \geq \omega:\, \alpha \in E^1} (\mathbb{I}^*\mu)(\alpha) = \sum_{\alpha_x \geq \omega_x} \sum_{\alpha_y \geq \omega_y} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y) =$$

$$\sum_{\alpha_x \geq \omega_x} \left(\sum_{\alpha_y(\omega) \geq \alpha_y \geq \omega_y} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y) + \sum_{\alpha_y > \alpha_y(\omega)} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y) \right),$$

where $\mathbb{I}_1^*\mu := \mathbb{I}^*\mu \cdot \chi_{E^1}$. By the definition of $\alpha_y(\omega)$,

$$\frac{\lambda}{3} > \sum_{\alpha_x \geq \omega_x} \sum_{\alpha_y > \alpha_y(\omega)} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y),$$

and therefore

$$\lambda \leq \mathbb{V}_1^\mu(\omega_x, \omega_y) \leq \sum_{\alpha_x \geq \omega_x} \sum_{\alpha_y(\omega) \geq \alpha_y \geq \omega_y} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y) + \frac{\lambda}{3}.$$

By (46)

$$\sum_{\alpha_x \geq \omega_x} \varphi_{\alpha_y}(\alpha_x, \alpha_y) \geq \sum_{\alpha_x \geq \omega_x} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y),$$

for $\alpha_y \leq \alpha_y(\omega)$. Therefore

$$\mathbb{I}\varphi(\omega) = \frac{3}{2} \sum_{\alpha_y \in T_y} \mathbb{I}\varphi_{\alpha_y}(\omega) = \frac{3}{2} \sum_{\alpha_y \geq \omega_y} \sum_{\alpha_x \geq \omega_x} \varphi_{\alpha_y}(\alpha_x, \alpha_y) \geq$$

$$\frac{3}{2} \sum_{\alpha_y(\omega) \geq \alpha_y \geq \omega_y} \sum_{\alpha_x \geq \omega_x} \varphi_{\alpha_y}(\alpha_x, \alpha_y) \geq \frac{3}{2} \sum_{\alpha_y(\omega) \geq \alpha_y \geq \omega_y} \sum_{\alpha_x \geq \omega_x} (\mathbb{I}_1^*\mu)(\alpha_x, \alpha_y) \geq$$

$$\frac{3}{2}\frac{2}{3}\lambda = \lambda,$$

proving (39a).

3.4 Proof of Lemma 3.1. One-dimensional argument: Lemma 3.4

As mentioned earlier, the condition (50) can be interpreted as a statement about the distance between F and supp f, in the sense that these two sets are far from each other. More precisely, if we want to find a function φ such that $I\varphi \geq If$ on F, there is a more effective, in terms of energy, solution than simply letting $\varphi = f$. A natural approach is to modify the equilibrium measure ρ of F, since $\varphi = I^*\rho$ provides the best way of acquiring unit potential on F.

The argument below goes as follows: first we split the set F into several parts in such a way that If is constant on each part. Then we modify the equilibrium measure ρ on each part according to the value of If there. Finally we show that the resulting measure satisfies (51).

Partition of F

We start with observing that $V^\rho(o) \leq \delta$. Indeed, since V^ρ is monotone on T (with respect to the natural order), we see that for any $\omega \in$ supp f

$$\delta \geq V^\rho(\omega) \geq V^\rho(o).$$

This allows us to define the δ-level sets of V^ρ,

$$F_\delta := \{\beta \in T : V^\rho(\beta) > \delta \text{ and } V^\rho(\alpha) \leq \delta, \text{ if } \alpha > \beta\}.$$

F_δ is essentially a stopping-time set for V^ρ. Define $\tilde{F} := \{\omega \in \text{cl} F : V^\rho = 1\}$. Since F is open, we have $V^\rho \equiv 1$ on F (see Lemma 5.6), hence $F \subset \tilde{F}$. Therefore $\tilde{F} \subset \mathcal{S}(F_\delta)$. Also we note that, if $\beta \in F_\delta$, then

$$V^\rho(\beta) = V^\rho(p(\beta)) + (I^*\rho)(\beta) \leq V^\rho(p(\beta)) + (I^*\rho)(o) =$$
$$V^\rho(p(\beta)) + V^\rho(o) \leq \delta + \delta = 2\delta,$$

where $p(\beta)$ denotes the immediate parent of β in T. In particular we see that supp f is outside $\mathcal{S}(F_\delta)$,

$$\text{supp} f \cap \mathcal{S}(F_\delta) = \emptyset,$$

where $\mathcal{S}(F_\delta) = \bigcup_{\beta \in F_\delta} \mathcal{S}(\beta)$. Also $\mathcal{S}(\beta_1) \cap \mathcal{S}(\beta_2) = \emptyset$ for any pair of (different) points $\beta_1, \beta_2 \in F_\delta$, so that the sets $\{\partial \mathcal{S}(\beta)\}_{\beta \in F_\delta}$ form a disjoint covering of the set $\tilde{F}$. Now we define the partition of ρ as follows

$$\rho_\beta = \rho|_{\partial \mathcal{S}(\beta)}, \quad \beta \in F_\delta.$$

Recall that supp $\rho = \text{cl} F \subset \partial T$ and that $\int_{\partial T} V^\rho d\rho = \text{Cap} F = |\rho|$, therefore $\rho(\text{cl} F \setminus \tilde{F}) = 0$. It follows immediately that

$$\rho = \sum_{\beta \in F_\delta} \rho_\beta.$$

We are ready to define the measure σ. Given $\beta \in F_\delta$ we set

$$\tilde{\sigma}_\beta := (If)(\beta)\rho_\beta,$$

and

$$\tilde{\sigma} = \sum_{\beta \in F_\delta} \tilde{\sigma}_\beta.$$

Finally we let

$$\sigma = (1-2\delta)^{-1}\tilde{\sigma}.$$

We are left to show that σ satisfies (51).

Inequality (51a)
Fix any $\omega \in F$. There exists exactly one point $\beta_\omega \in F_\delta$ such that $\beta_\omega > \omega$. By definition of ρ_β we have

$$(I^*\rho_{\beta_\omega})(\alpha) = (I^*\rho)(\alpha)$$

for any $\alpha \le \beta_\omega$. It follows that

$$
\begin{aligned}
V^{\rho_{\beta_\omega}}(\omega) = \sum_{\alpha \ge \omega}(I^*\rho_{\beta_\omega})(\alpha) \ge \sum_{\beta_\omega > \alpha \ge \omega}(I^*\rho_{\beta_\omega})(\alpha) = \\
\sum_{\beta_\omega > \alpha \ge \omega}(I^*\rho)(\alpha) = V^\rho(\omega) - V^\rho(\beta_\omega) \ge 1-2\delta,
\end{aligned}
\tag{52}
$$

since $V^\rho \equiv 1$ on F. Hence

$$
\begin{aligned}
V^\sigma(\omega) = (1-2\delta)^{-1}V^{\tilde{\sigma}}(\omega) \ge (1-2\delta)^{-1}V^{\tilde{\sigma}_{\beta_\omega}}(\omega) = \\
(1-2\delta)^{-1}(If)(\beta_\omega)V^{\rho_{\beta_\omega}}(\omega) \ge (If)(\beta_\omega) = (If)(\omega),
\end{aligned}
$$

since f has no mass on the set $\mathcal{S}(\beta_\omega)$.
Inequality (51b)
To prove this inequality we do a further partition of $\tilde{\sigma}$ and ρ with respect to the distribution of If. First we raise $\tilde{\sigma}$ and ρ to the set F_δ,

$$
\begin{aligned}
\tilde{\sigma}_\delta(\beta) &:= \chi_{F_\delta}(I^*\tilde{\sigma})(\beta), \\
\rho_\delta(\beta) &:= \chi_{F_\delta}(I^*\rho)(\beta), \quad \beta \in T.
\end{aligned}
$$

Clearly $\tilde{\sigma}_\delta, \rho_\delta$ are supported on F_δ, and $(I^*\tilde{\sigma}_\delta)(\alpha) = (I^*\tilde{\sigma})(\alpha)$ for any $\alpha \ge F_\delta$ (by this we mean that $\alpha \ge \theta$ for all $\theta \in F_\delta$). Also

$$\tilde{\sigma}_\delta(\beta) = (If)(\beta)\rho_\delta(\beta), \quad \beta \in F_\delta,\tag{53}$$

by the definition of $\tilde{\sigma}$.
Next we compute the energy of $\tilde{\sigma}$

$$
\begin{aligned}
\mathcal{E}[\tilde{\sigma}] = \int_{\overline{T}}V^{\tilde{\sigma}}d\tilde{\sigma} = \int_{\overline{T}}\sum_{\alpha \ge \omega}(I^*\tilde{\sigma})(\alpha)d\tilde{\sigma}(\omega) = \\
\sum_{\beta \in F_\delta}\int_{\mathcal{S}(\beta)}\sum_{\alpha \ge \beta}(I^*\tilde{\sigma})(\alpha)d\tilde{\sigma}(\omega) + \sum_{\beta \in F_\delta}\int_{\mathcal{S}(\beta)}\sum_{\beta > \alpha \ge \omega}(I^*\tilde{\sigma})(\alpha)d\tilde{\sigma}(\omega) \le \\
\sum_{\beta \in F_\delta}\sum_{\alpha \ge \beta}(I^*\tilde{\sigma}_\delta)(\alpha)\tilde{\sigma}_\delta(\beta) + \sum_{\beta \in F_\delta}\int_{\mathcal{S}(\beta)}\sum_{\beta \ge \alpha \ge \omega}(I^*\tilde{\sigma})(\alpha)d\tilde{\sigma}(\omega) := \\
(I) + (II).
\end{aligned}
$$

For the first term we have

$$(I) = \int_{F_\delta} V^{\tilde{\sigma}_\delta} \, d\tilde{\sigma}_\delta = \mathcal{E}[\tilde{\sigma}_\delta].$$

We expand the second term obtaining

$$(II) = \sum_{\beta \in F_\delta} \int_{\mathcal{S}(\beta)} \sum_{\beta \geq \alpha \geq \omega} (I^* \tilde{\sigma})(\alpha) \, d\tilde{\sigma}(\omega) = \sum_{\beta \in F_\delta} \int_{\mathcal{S}(\beta)} \sum_{\beta \geq \alpha \geq \omega} (I^* \tilde{\sigma}_\beta)(\alpha) \, d\tilde{\sigma}_\beta(\omega) =$$

$$\sum_{\beta \in F_\delta} (If)(\beta) \int_{\mathcal{S}(\beta)} \sum_{\beta \geq \alpha \geq \omega} (I^* \rho_\beta)(\alpha) \, d\tilde{\sigma}_\beta(\omega) \leq \sum_{\beta \in F_\delta} (If)(\beta)(I^* \tilde{\sigma})(\beta),$$

$$(54)$$

since

$$\sum_{\beta \geq \alpha \geq \omega} (I^* \rho_\beta)(\alpha) = \sum_{\beta \geq \alpha \geq \omega} (I^* \rho)(\alpha) \leq \sum_{\alpha \geq \omega} (I^* \rho)(\alpha) = V^\rho(\omega) \leq 1$$

for $\omega \in \mathrm{cl}F$. We see that

$$(II) \leq \int_{\overline{T}} If \, d\tilde{\sigma}_\delta = \sum_{\alpha \in T} f(\alpha)(I^* \tilde{\sigma}_\delta)(\alpha) =: \mathcal{E}[f, \tilde{\sigma}_\delta].$$

Hence

$$\mathcal{E}[\tilde{\sigma}] \leq \mathcal{E}[\tilde{\sigma}_\delta] + \mathcal{E}[f, \tilde{\sigma}_\delta].$$

We observe that in order to prove (51b) it is enough to show that

$$\mathcal{E}[\tilde{\sigma}_\delta] \leq C\delta \mathcal{E}[f, \tilde{\sigma}_\delta] \tag{55}$$

for some absolute constant $C > 0$. Indeed, by positivity of the energy integral we have

$$0 \leq \sum_{\alpha \in T} ((I^* \tilde{\sigma}_\delta)(\alpha) - C\delta f(\alpha))^2 = \mathcal{E}[\tilde{\sigma}_\delta] - 2C\delta \mathcal{E}[f, \tilde{\sigma}_\delta] + (C\delta)^2 \|f\|_{\ell^2(T)}^2 =$$

$$(\mathcal{E}[\tilde{\sigma}_\delta] - C\delta \mathcal{E}[f, \tilde{\sigma}_\delta]) + C\delta(C\delta \|f\|_{\ell^2(T)}^2 - \mathcal{E}[f, \tilde{\sigma}_\delta]),$$

so if the first term is negative, the second one must be positive. Hence $\mathcal{E}[f, \tilde{\sigma}_\delta] \leq C\delta \|f\|_{\ell^2(T)}^2$, and by (55),

$$\mathcal{E}[\tilde{\sigma}] \leq \mathcal{E}[\tilde{\sigma}_\delta] + \mathcal{E}[f, \tilde{\sigma}_\delta] \leq (C\delta)^2 \|f\|_{\ell^2(T)}^2 + C\delta \|f\|_{\ell^2(T)}^2,$$

which is (51b).

Now for any $k \in \mathbb{Z}$ we define

$$F_{\delta,k} := \{\beta \in F_\delta : 2^k \leq (If)(\beta) < 2^{k+1}\},$$

and we set $\tilde{\sigma}_{\delta,k} = \tilde{\sigma}_\delta|_{F_{\delta,k}}$, $\rho_{\delta,k} = \rho_\delta|_{F_{\delta,k}}$.
Clearly $\tilde{\sigma}_\delta = \sum_k \tilde{\sigma}_{\delta,k}$, and

$$\mathcal{E}[\tilde{\sigma}_\delta] = \int_{\overline{T}} V^{\tilde{\sigma}_\delta} \, d\tilde{\sigma}_\delta = \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \int_{\overline{T}} V^{\tilde{\sigma}_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} \leq 2 \sum_{k \in \mathbb{Z}} \sum_{j \leq k} \int_{\overline{T}} V^{\tilde{\sigma}_{\delta,j}} \, d\tilde{\sigma}_{\delta,k}.$$

For $j \leq k$, by (53), we have that $V^{\tilde{\sigma}_{\delta,j}} \leq 2^{j+1} V^{\rho_{\delta,j}}$ on $F_{\delta,k}$. Thus

$$\sum_{j \leq k} \int_{\overline{T}} V^{\tilde{\sigma}_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} \leq \sum_{j \leq k} 2^{j+1} \int_{\overline{T}} V^{\rho_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} \leq 2^{k+1} \int_{\overline{T}} \sum_{j \leq k} V^{\rho_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} =$$

$$2^{k+1} \int_{\overline{T}} V^{\sum_{j \leq k} \rho_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} \leq 2^{k+1} \int_{\overline{T}} V^{\rho_\delta} \, d\tilde{\sigma}_{\delta,k} \leq 2^{k+2} \delta |\tilde{\sigma}_{\delta,k}|,$$

since $V^{\rho_\delta} \leq 2\delta$ on F_δ. Summing this estimate over k we obtain

$$\sum_{k \in \mathbb{Z}} \sum_{j \leq k} \int_{\overline{T}} V^{\tilde{\sigma}_{\delta,j}} \, d\tilde{\sigma}_{\delta,k} \leq \sum_{k \in \mathbb{Z}} 2^{k+2} \delta |\tilde{\sigma}_{\delta,k}| =$$

$$\sum_{k \in \mathbb{Z}} \sum_{\beta \in F_{\delta,k}} 2^{k+2} \delta \tilde{\sigma}_{\delta,k}(\beta) \leq 4\delta \sum_{k \in \mathbb{Z}} \sum_{\beta \in F_{\delta,k}} (If)(\beta) \tilde{\sigma}_{\delta,k}(\beta) =$$

$$4\delta \sum_{\beta \in F_\delta} (If)(\beta) \tilde{\sigma}_\delta(\beta) = 4\delta \mathcal{E}[f, \tilde{\sigma}_\delta],$$

so we get (55), and therefore (51b).

4   Concluding remarks

We have characterized the Carleson measures for the Dirichlet space using, as in Stegenga's [27], a Strong Capacitary Inequality. In the one-parameter case, other characterizations can be given. In [7] and [22], the Carleson measures for $\mathcal{D}(\mathbb{D})$ are defined in terms of two, seemingly different, *one-box testing* conditions, in which the advantage is that they have to be verified for single Carleson boxes, and not unions thereof. The disadvantage is that the measure μ, unlike in the capacitary condition, appears on both sides of the testing inequality.

The bi-parameter non-linear case, $1 < p < \infty$, could also be considered; the space under scrutiny would be the tensor product of two copies of an analytic Besov space. The one dimensional case was considered for example in [28] and [7]. Here, we think that the needed tool is a bi-parameter version of Wolff's inequality [18], which could be considered as one half of the Muckenhoupt-Wheeden inequality. With that at hand, one could extend a sizeable portion of the potential theory we have developed here to the non-linear case.

The probabilistic theory underlying bi-parameter linear Potential Theory is that of two-parameter martingales [13, 17]. It would be interesting to make this relationship explicit, and to find a way to pass results from one theory to the other.

Much of the Potential Theory we have developed on bi-trees can be applied to yield a Potential Theory on product spaces much more general than $\partial \mathbb{D} \times \partial \mathbb{D}$, following, for example, the route taken in [10].

The Dirichlet space on the bidisc does not come with a Complete Nevanlinna–Pick kernel. In fact, no tensor product Hilbert space does [29]. If the kernel had the Complete Nevanlinna-Pick property, the characterization of Carleson measures for $\mathcal{D}(\mathbb{D}^2)$ would as consequence yield the characterization of its *universal interpolating sequences*, by a recent result of Aleman, Hartz, McCarthy, and Richter [4]. We think this is an interesting open problem, for which we have no guess. See [3, 26] for a deep and broad discussion of interpolating sequences for Hilbert function spaces.

5   Appendix

In this section we collect several results that were used or mentioned in the main text. First, in Section 5.1 we provide the proofs of the more technical results from Section 2 regarding the discretization procedure. Then we present some basic properties of bi-logarithmic potentials and equilibrium measures, see Lemma 5.6. In Section 5.3 we give counterexamples to the maximum and domination principles, in Propositions 5.2 and 5.3, respectively. Finally, in Theorem 5.1 we show that given a measure on $\overline{T}^2$, we can construct a measure supported on the distinguished boundary, equivalent to the original measure in the sense of potentials. From this we deduce that the capacity of a set is equivalent to the capacity of its boundary projection, see Corollary 5.1.

5.1

We start with providing some results justifying the discretization of the unit disc (bidisc) via the graphs $\mathfrak{G}$ and T. The graph $\mathfrak{G}$ serves as an intermediate point in the discretization scheme between the unit disc and the dyadic tree – see Section 2.4 for precise definitions. While it is more complicated and rather inconvenient to work with when compared to the tree T, its geometry is better suited for representing the unit disc, which is why we use it to justify passing from $\mathbb{D}$ to T (and therefore from $\mathbb{D}^2$ to T^2).

First we show that $\mathfrak{G}$ provides a model for the hyperbolic metric on $\overline{\mathbb{D}}$.

Lemma 5.1 *Given two points $z, w \in \overline{\mathbb{D}}^2$ one has*

$$\left| 10 + \log \frac{1}{1 - \bar{z}_1 w_1} \right| \left| 10 + \log \frac{1}{1 - \bar{z}_2 w_2} \right| \approx d_{\mathfrak{G}^2}(\alpha(z) \wedge \beta(w)) = d_{\mathfrak{G}}(\alpha_x(z) \wedge \beta_x(w)) d_{\mathfrak{G}}(\alpha_y(z) \wedge \beta_y(w)),$$

(56)

where $z = (z_1, z_2)$, $w = (w_1, w_2)$, and $\alpha, \beta \in \overline{T}$ are any of the preimages $\alpha(z) = (\alpha_x(z), \alpha_y(z)) \in \Lambda^{-1}(z)$, $\beta(w) = (\beta_x(w), \beta_y(w)) \in \Lambda^{-1}(w)$. We recall that the natural map Λ from $\overline{T}^2$ to $\overline{\mathbb{D}}^2$ was defined in Section 2.6.

Proof. Clearly it is enough to show (56) separately for each coordinate; we show that

$$\left| 10 + \log \frac{1}{1 - \bar{z}_1 w_1} \right| \approx d_{\mathfrak{G}}(\alpha_x \wedge \beta_x).$$

(57)

Note that

$$\left| 10 + \log \frac{1}{1 - \bar{z}_1 w_1} \right| \approx 10 + \log \frac{1}{|1 - \bar{z}_1 w_1|},$$

since $1 + \log \frac{1}{|1 - \bar{z}_1 w_1|} > 0$ for any pair of points $z_1, w_1 \in \overline{\mathbb{D}}$.

We start by assuming $z_1, w_1 \in \mathbb{D}$. Recall that there exist uniquely defined $\alpha_x, \beta_x \in T$ such that $z_1 \in Q_{\alpha_x}$, $w_1 \in Q_{\beta_x}$. Let J be the smallest interval, not necessarily dyadic, containing both J_{α_x} and J_{β_x}. We claim that $d_{\mathfrak{G}}(\alpha_x \wedge \beta_x) \approx \log |J|^{-1}$. Indeed, in order for γ to be a common point of the sets $\mathcal{P}_{\mathfrak{G}}(\alpha_x)$ and $\mathcal{P}_{\mathfrak{G}}(\beta_x)$, the dyadic interval J_γ has to be large, $|J_\gamma| \geq \frac{1}{2} \max(|J_{\alpha_x}|, |J_{\beta_x}|)$, and $3J_\gamma$ must have non-empty intersection with both J_{α_x} and J_{β_x}. The number of such intervals is approximately $\log \frac{1}{|J|}$. On the other hand, an elementary computation yields that

$$|1 - \bar{z}_1 w_1| \approx \max(1 - |z_1|, 1 - |w_1|, |\arg(\bar{z}_1 w_1)|) \approx |J|.$$

Now let $z_1, w_1 \in \partial \mathbb{D}$. If these two points coincide, then $d_{\mathfrak{G}} = \infty$ regardless of the choice of pre-images $\alpha_x \in \Lambda_0^{-1}(\{z_1\})$, $\beta_x \in \Lambda_0^{-1}(\{w_1\})$. Otherwise we let J to be the smallest interval containing z_1 and w_1, and repeat the above argument.

The cases $z_1 \in \mathbb{D}$, $w_1 \in \partial \mathbb{D}$ and $w_1 \in \partial \mathbb{D}$, $z_1 \in \mathbb{D}$ are dealt with similarly. $\qquad \square$

Next we investigate the properties of the map Λ, and the induced pull-backs and push-forwards of measures, as introduced in Section 2.6.

Lemma 5.2 *The Lipschitz map $\Lambda : \overline{T}^2 \to \overline{\mathbb{D}}^2$ induces maps $\Lambda_* : \mathrm{Meas}^+((\partial T)^2) \to \mathrm{Meas}^+((\partial \mathbb{D})^2))$ and $\Lambda^* : \mathrm{Meas}^+(\overline{\mathbb{D}}^2) \to \mathrm{Meas}^+(\overline{T}^2)$, Meas^+ denoting the space of non-negative Borel measures on the respective set, with the following properties:*

- $\Lambda_* \Lambda^* \mu = \mu$, *if μ is supported on $(\partial \mathbb{D})^2$.*

- *If $\mu \geq 0$ is a Borel measure on $(\partial \mathbb{D})^2$ with finite $(\frac{1}{2}, \frac{1}{2})$-energy, then $\Lambda^* \mu(E) = \mu(\Lambda(E))$ for all measurable sets E in $(\partial T)^2$. In particular, $\mu(\Delta) = 0$, where Δ is the dyadic grid on $(\partial \mathbb{D})^2$, the set of points with at least one dyadic coordinate.*

- *If $\nu \geq 0$ is a Borel measure on $(\partial T)^2$ with finite energy, then $\nu(\partial T_x \times \{\omega_y\} \bigcup \{\omega_x\} \times \partial T_y) = 0$ for any $\omega \in (\partial T)^2$.*

- *For such a measure ν, it holds that $\Lambda^* \Lambda_* \nu = \nu$.*

Proof. The first point is obvious.

Proof of the second point. It is enough to show that $\mu(\Delta) = 0$, since Δ precisely consists of the points where Λ^{-1} is not uniquely defined. In turn, one only has to prove that, say, $\mu(\{1\} \times \partial \mathbb{D}) = 0$, since Δ is a countable union of such sets.

Let us recall the dual definition of capacity: for any compact set $E \subset \overline{T}^2$ one has

$$\mathrm{Cap}^{\frac{1}{2}} E = \sup \left\{ \frac{\mu(E)}{\mathcal{E}^{\frac{1}{2}}[\mu]} : \ \mathrm{supp}\, \mu \subset E \right\},$$

and the maximizer is exactly the equilibrium measure μ_E. From here, is not difficult to see that the proof of the second point of the statement will follow if we show that $\{1\} \times \partial \mathbb{D}$ is a polar set, meaning that

$$\mathrm{Cap}_{(\frac{1}{2}, \frac{1}{2})}(\{1\} \times \partial \mathbb{D}) = 0.$$

This is almost a direct corollary of the one-dimensional fact that the Bessel $\frac{1}{2}$-capacity of a singleton on the unit circle is zero. To elaborate, let

$$h_K(e^{i\theta_1}, e^{i\theta_2}) := \frac{1}{K} \sum_{j=1}^{K} 2^{\frac{j}{2}} \chi_{[-2^{-j}, 2^{-j}]}(\theta_1).$$

Clearly $\int_{(\partial \mathbb{D})^2} h_K^2(z)\, dm(z) \lesssim \frac{1}{K}$, and an elementary computation shows that

$$(b_{(\frac{1}{2}, \frac{1}{2})} h_K)(e^{i0}, e^{i\theta_2}) \approx \frac{1}{K} \sum_{j=1}^{K} 2^{\frac{j}{2}} \int_0^{2^{-j}} \frac{d\theta_1}{\theta_1^{\frac{1}{2}}} \approx 1.$$

Hence Ch_K is an admissible function for some large $C > 1$. Letting K to infinity we immediately obtain the desired result.

Proof of the third point. Assume that $v(\{\omega_x\} \times \partial T_y) = \varepsilon > 0$ for some $\omega_x \in \partial T_x$. Then we immediately have $(\mathbb{I}^* v)(\alpha_x, o) \geq \varepsilon$ for any $\alpha_x > \omega_x$, and

$$\mathcal{E}[v] = \sum_{\alpha \in T^2} (\mathbb{I}^* v)^2(\alpha) \geq \sum_{\alpha_x > \omega_x} (\mathbb{I}^* v)^2(\alpha_x, o) = \infty.$$

The same argument shows that $\partial T_x \times \{\omega_y\}$ has measure zero.

Proof of the fourth point.

$$\Lambda^* \Lambda_* v(E) = \Lambda_* v(\Lambda(E)) = v(\Lambda^{-1}(\Lambda(E))) = v(E),$$

since Λ fails to be injective only on a set of vanishing v-measure. $\qquad \square$

Next we show that $\mathfrak{G}$ and T are similar in capacitary sense. In the one-parameter setting, much stronger results are available, see for example Lemma 2.14 in [11].

Lemma 5.3 *Given a finite family of points $\{\alpha^j\}_{j=1}^n \subset T^2$, one has*

$$\text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}(\alpha^j)\right) \approx \text{Cap}\left(\bigcup_{j=1}^n \partial \mathcal{S}(\alpha^j)\right) \approx \text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}_{\mathfrak{G}^2}(\alpha^j)\right) \approx \text{Cap}\left(\bigcup_{j=1}^n \partial \mathcal{S}_{\mathfrak{G}^2}(\alpha^j)\right). \qquad (58)$$

In particular,

$$\text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}(\alpha^j)\right) \approx \text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}(p(\alpha^j))\right),$$

where $p(\alpha^j) = (p(\alpha_x^j), p(\alpha_y^j))$ is the 'grandparent' of α^j in T^2, see the proof of Theorem 2.1.

Proof. The first and last equivalences of (58) come from the fact that the capacities of a set and its boundary projection are comparable, see Corollary 5.1. Since $\mathcal{S}_{\mathfrak{G}^2}(\alpha) \supset \mathcal{S}(\alpha)$ for any $\alpha \in T^2$, we have

$$\text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}(\alpha^j)\right) \leq \text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}_{\mathfrak{G}^2}(\alpha^j)\right).$$

To show the reverse inequality

$$\text{Cap}\, E := \text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}(\alpha^j)\right) \gtrsim \text{Cap}\left(\bigcup_{j=1}^n \mathcal{S}_{\mathfrak{G}^2}(\alpha^j)\right) =: \text{Cap}\, F,$$

we prove that the energies of μ_E and μ_F are comparable. We start by showing that the mixed energy of μ_E and μ_F dominates $\mathcal{E}[\mu_F]$, using an argument similar to the one in Section 2.4. We have

$$\mathcal{E}[\mu_E, \mu_F] = \sum_{\alpha \in T^2} (\mathbb{I}^* \mu_E)(\alpha)(\mathbb{I}^* \mu_F)(\alpha) = \sum_{\alpha \in T^2} \mu_E(\mathcal{S}(\alpha)) \mu_F(\mathcal{S}(\alpha)).$$

The successor set formula (22) implies that for any $\alpha \in T^2$ there exists a finite collection $G_\alpha = \{\beta_\alpha^j\}_{j=1}^N$, N independent of α, such that $\mathcal{S}_{\mathfrak{G}^2}(\alpha) \subset \bigcup_{j=1}^N \mathcal{S}_{T^2}(\beta_\alpha^j)$, and moreover, for any $\beta \in T^2$ there exist at most N points α such that $\beta \in G_\alpha$. It follows that

$$\sum_{\alpha \in T^2} \mu_E(\mathcal{S}(\alpha))\mu_F(\mathcal{S}(\alpha)) \gtrsim \sum_{\alpha \in T^2} \mu_E(\mathcal{S}(\alpha)) \sum_{\beta \in G_\alpha} \mu_F(\mathcal{S}_{T^2}(\beta)) \gtrsim \sum_{\alpha \in T^2} \mu_E(\mathcal{S}(\alpha))\mu_F(\mathcal{S}_{\mathfrak{G}^2}(\alpha)) =$$

$$\sum_{\alpha \in T^2} (\mathbb{I}^*\mu_E)(\alpha) \int_{\mathcal{S}_{\mathfrak{G}^2}(\alpha)} d\mu_F = \int_{\overline{T}^2} \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(\beta)} (\mathbb{I}^*\mu_E)(\alpha)\, d\mu_F(\beta).$$

Given $\beta \in F$ there must exist at least one α^j such that $\alpha^j \in \mathcal{P}_{\mathfrak{G}}^2(\beta)$. Since μ_E is the equilibrium measure of E, $\alpha^j \in E$, and $\alpha^j \in T^2$ (so that $\mathrm{Cap}\{\alpha^j\} > 0$) we have $1 \leq \mathbb{V}^{\mu_E}(\alpha^j) \leq \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(\alpha^j)} (\mathbb{I}^*\mu_E)(\alpha) \leq \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(\beta)} (\mathbb{I}^*\mu_E)(\alpha)$. It follows immediately that $\int_{\overline{T}^2} \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(\beta)} (\mathbb{I}^*\mu_E)(\alpha)\, d\mu_F(\beta) \geq \mu_F(F) = \mathcal{E}[\mu_F]$, therefore

$$\mathcal{E}[\mu_E, \mu_F] \geq \varepsilon \mathcal{E}[\mu_F]$$

for some $\varepsilon > 0$ that does not depend on E or F. By positivity of the energy integral

$$0 \leq \mathcal{E}[\mu_E - \varepsilon\mu_F] = \mathcal{E}[\mu_E] - 2\varepsilon\mathcal{E}[\mu_E, \mu_F] + \varepsilon^2 \mathcal{E}[\mu_F] = (\mathcal{E}[\mu_E] - \varepsilon\mathcal{E}[\mu_E, \mu_F]) + \varepsilon(\varepsilon\mathcal{E}[\mu_F] - \mathcal{E}[\mu_E, \mu_F]).$$

We have shown that the first term must be positive, which in turn implies that $\mathcal{E}[\mu_E] \geq \varepsilon^2 \mathcal{E}[\mu_F]$. We are done. $\qquad\square$

The next result compares the capacities of sets in $\overline{T}^2$ and $(\partial\mathbb{D})^2$. We refer to Sections 2.3 and 2.5 for the relevant definitions. By arguments in Section 5.4, we can always estimate the capacity of a set in $\overline{T}^2$ by the capacity of its boundary projection. Therefore we only consider sets on the distinguished boundaries of the bitree and bidisc. Moreover, it is sufficient to consider finite unions of 'rectangles'. The proof mostly consists of arguments from [10, Chapter 4], adapted to the two-parameter setting.

Lemma 5.4 *Let $\{\alpha^k\}_{j=1}^N$ be a finite collection of points in T^2. One then has*

$$\mathrm{Cap}\left(\bigcup_{k=1}^N \mathcal{S}(\alpha^k)\right) \approx \mathrm{Cap}\left(\bigcup_{k=1}^N \partial\mathcal{S}(\alpha^k)\right) \approx \mathrm{Cap}_{(\frac{1}{2},\frac{1}{2})}\left(\bigcup_{k=1}^N \Lambda(\partial\mathcal{S}(\alpha^k))\right) = \mathrm{Cap}_{(\frac{1}{2},\frac{1}{2})}\left(\bigcup_{k=1}^N J_{\alpha^k}\right), \quad (59)$$

where $J_\alpha = S_\alpha \cap (\partial\mathbb{D})^2$ is the intersection of the Carleson box S_α with the torus.

Proof. As before, the first equivalence follows from Corollary 5.1. Set $E := \bigcup_{k=1}^n \partial\mathcal{S}(\alpha^k)$ and $F := \bigcup_{k=1}^N J_{\alpha^k}$. Clearly both sets are compact in their respective topologies. Let ν_E, μ_F be the equilibrium measures for E and F, so that $\mathrm{supp}\,\nu_E \subset E$, $\mathrm{supp}\,\mu_F \subset F$. To compare the capacities of E and F we need to know how to move equilibrium measures between $(\partial T)^2$ and $(\partial\mathbb{D})^2$. Our first step in this direction is to show that

$$\mathrm{supp}\,\Lambda^*\mu_F \subset E, \ \mu_F(F) = (\Lambda^*\mu_F)(E);$$
$$\mathrm{supp}\,\Lambda_*\nu_E \subset F, \ \nu_E(E) = (\Lambda_*\nu_E)(F). \tag{60}$$

We start with μ_F. The mass conservation property can be shown separately on each rectangle. Fix any J_{α_k} and denote $\partial \mathcal{S}(\alpha^k)$ by R_{α^k}. Recalling the definition of Λ^* we see that

$$(\Lambda^* \mu_F)(R_{\alpha^k}) = \int_{(\partial \mathbb{D})^2} \psi(z) \, d\mu_F,$$

where $\psi(z) := \frac{\sharp\{\Lambda^{-1}(\{z\}) \cap R_{\alpha^k}\}}{\sharp\{\Lambda^{-1}(\{z\})\}}$, $z \in (\partial \mathbb{D})^2$. For any z in the (torus) interior of the rectangle J_{α^k} we clearly have $\psi(z) = 1$. Unfortunately $\Lambda^{-1}(J_{\alpha^k})$ is slightly larger than R_{α^k}, so $\psi(z)$ could be $\frac{1}{2}$ or $\frac{1}{4}$, depending on whether z is on the side of J_{α^k}, or is one of its corners. However, by Lemma 5.2 we see that $\mu_F(\partial J_{\alpha^k}) = 0$, since, clearly, $\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu_F] = \mu_F(F) < +\infty$. Here ∂J_{α^k} is the boundary of the rectangle J_{α^k} in the torus $(\partial \mathbb{D})^2$. It follows that $(\Lambda^* \mu_F)(R_{\alpha^k}) = \mu_F(J_{\alpha^k})$, hence $\mu_F(F) = (\Lambda^* \mu_F)(E)$. Arguing as above we obtain

$$(\Lambda^* \mu_F)((\partial T)^2 \setminus E) = \int_{(\partial \mathbb{D})^2} \frac{\sharp\{\Lambda^{-1}(\{z\}) \cap ((\partial T)^2 \setminus E)\}}{\sharp\{\Lambda^{-1}(\{z\})\}} \, d\mu_F = 0,$$

since μ_F has zero mass on the boundary of F in $(\partial \mathbb{D})^2$ and the set $\{\Lambda^{-1}(\{z\}) \cap ((\partial T)^2 \setminus E)\}$ is empty for any z in the interior of F. Therefore $\operatorname{supp} \Lambda^* \mu_F \subset E$, and we have the first part of (60).

The argument for ν_E is similar. Clearly $\Lambda^{-1}((\partial \mathbb{D})^2 \setminus F) \cap E = \emptyset$, hence $(\Lambda_* \nu_E)((\partial \mathbb{D})^2 \setminus F) = 0$, and $\operatorname{supp} \Lambda_* \nu_E \subset F$. Further, $\Lambda(E) = F$, and thus $E \subset \Lambda^{-1}(F)$, and by compactness of E

$$(\Lambda_* \nu_E)(F) = \nu_E(\Lambda^{-1}(F)) = \nu_E(E).$$

By the dual definition of capacity,

$$
\begin{aligned}
\operatorname{Cap} E &= \sup\left\{ \frac{(\nu(E))^2}{\mathcal{E}[\nu]} : \operatorname{supp} \nu \subset E \right\}, \\
\operatorname{Cap}_{(\frac{1}{2},\frac{1}{2})} F &= \sup\left\{ \frac{(\mu(F))^2}{\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu]} : \operatorname{supp} \mu \subset E \right\},
\end{aligned}
\tag{61}
$$

and ν_E, μ_F are the respective maximizers. Therefore

$$\operatorname{Cap} E = \frac{(\nu_E(E))^2}{\mathcal{E}[\nu_E]} = \frac{((\Lambda_* \nu_E)(F))^2}{\mathcal{E}[\nu_E]},$$

and

$$\operatorname{Cap}_{(\frac{1}{2},\frac{1}{2})} F = \frac{(\mu_F(F))^2}{\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu_F]} = \frac{((\Lambda^* \mu_F)(E))^2}{\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu_F]},$$

so in order to prove (59) it is enough to show that

$$\mathcal{E}[\nu_E] \approx \mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\Lambda_* \nu_E], \tag{62a}$$

$$\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu_F] \approx \mathcal{E}[\Lambda^* \mu_F]. \tag{62b}$$

Both of these equivalences follow from Lemma 5.5 below. Indeed, to obtain (62a) we apply this Lemma with $\nu = \nu_E$ and $\mu = \Lambda_* \nu_E$, similarly (62b) follows by assuming $\nu = \Lambda^* \mu_F$ and $\mu = \mu_F$. $\qquad\square$

Lemma 5.5 *Consider two Borel measures* $\mu, \nu \geq 0$ *on* $(\partial\mathbb{D})^2$ *and* $(\partial T)^2$ *respectively such that for any* $\alpha \in (\partial T)^2$ *they satisfy*

$$\mu(J_\alpha) = \nu(\Lambda^{-1}(J_\alpha)); \quad \nu(R_\alpha) = \mu(\Lambda(R_\alpha)), \tag{63}$$

and their respective energies are finite. Then

$$\mathcal{E}[\nu] \approx \mathcal{E}_{(\frac{1}{2}, \frac{1}{2})}[\mu]. \tag{64}$$

Proof. We start from the continuous side. Define $M := \Lambda^* m$, where m is the normalized area measure on the torus $(\partial\mathbb{D})^2$. Clearly, $M(\partial S(\alpha)) = 2^{-d_T(\alpha_x) - d_T(\alpha_y) + 2} = m(J_\alpha)$. Given a point $z \in (\partial\mathbb{D})^2$, set

$$\mathcal{P}_{\mathfrak{G}^2}(z) := \bigcup_{\omega \in \Lambda^{-1}(\{z\})} \mathcal{P}_{\mathfrak{G}^2}(\omega).$$

First we discretize the Bessel potential, namely, we show that for any $z \in (\partial\mathbb{D})^2$ one has

$$\int_{(\partial\mathbb{D})^2} b_{(\frac{1}{2}, \frac{1}{2})}(z, \zeta) \, d\mu(\zeta) \approx \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^* \nu)(\alpha)}{M^{\frac{1}{2}}(\partial S(\alpha))}. \tag{65}$$

Fix a point $z = (e^{i\theta_1}, e^{i\theta_2}) \in (\partial\mathbb{D})^2$. Let $\tilde{J}_z(\varepsilon, \delta)$ be a rectangle with centerpoint z, and the sidelengths $4\pi\varepsilon, 4\pi\delta$, that is,

$$\tilde{J}_z(\varepsilon, \delta) = \left\{ \zeta = (e^{i\eta_1}, e^{i\eta_2}) \in (\partial\mathbb{D})^2 : \frac{|\theta_1 - \eta_1|}{2\pi} \leq \varepsilon; \frac{|\theta_2 - \eta_2|}{2\pi} \leq \delta \right\}.$$

A simple computation gives

$$\int_0^{2\pi} \int_0^{2\pi} \frac{d\mu(e^{i\eta_1}, e^{i\eta_2})}{|\theta_1 - \eta_1|^{\frac{1}{2}} |\theta_2 - \eta_2|^{\frac{1}{2}}} \approx \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \int_{2^{-n_1 - 1} \leq \frac{|\theta_1 - \eta_1|}{2\pi} \leq 2^{-n_1}} \int_{2^{-n_2 - 1} \leq \frac{|\theta_2 - \eta_2|}{2\pi} \leq 2^{-n_2}} \frac{d\mu(e^{i\eta_1}, e^{i\eta_2})}{2^{-\frac{n_1}{2}} 2^{-\frac{n_2}{2}}} \approx$$

$$\sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \int_{\frac{|\theta_1 - \eta_1|}{2\pi} \leq 2^{-n_1}} \int_{\frac{|\theta_2 - \eta_2|}{2\pi} \leq 2^{-n_2}} \frac{d\mu(e^{i\eta_1}, e^{i\eta_2})}{2^{-\frac{n_1}{2}} 2^{-\frac{n_2}{2}}} = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \frac{\mu(\tilde{J}_z(2^{-n_1}, 2^{-n_2}))}{2^{-\frac{n_1}{2} - \frac{n_2}{2}}} \approx$$

$$\sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \frac{\mu(\tilde{J}_z(10 \cdot 2^{-n_1}, 10 \cdot 2^{-n_2}))}{2^{-\frac{n_1}{2} - \frac{n_2}{2}}}.$$

Fix some $n_1, n_2 \geq 0$ and consider $\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)$ such that $d_T(\alpha_x) = n_1 + 1$, $d_T(\alpha_y) = n_2 + 1$. Denote the collection of such points by $N_z(n_1, n_2)$. Then we have

$$\tilde{J}_z(2^{-n_1}, 2^{-n_2}) \subset \bigcup_{\alpha \in N_z(n_1, n_2)} J_\alpha \subset \tilde{J}_z(10 \cdot 2^{-n_1}, 10 \cdot 2^{-n_2}).$$

Indeed, if α is such a point, then J_α either contains z, or is one of the neighbouring rectangles of the same generation, and vice versa, all such rectangles correspond to some point in $\mathcal{P}_{\mathfrak{G}^2}(z)$. It follows that

$$\int_{(\partial\mathbb{D})^2} b_{(\frac{1}{2}, \frac{1}{2})}(z, \zeta) \, d\mu(\zeta) \approx \sum_{n_1, n_2 \geq 0} \frac{\mu\left(\bigcup_{\alpha \in N_z(n_1, n_2)} J_\alpha\right)}{2^{-\frac{n_1}{2} - \frac{n_2}{2}}}.$$

Since $\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu] < +\infty$, we have $\mu(\partial J_\alpha) = 0$ for any α by Lemma 5.2. Combined with (63) we obtain

$$\mu\left(\bigcup_{\alpha \in N_z(n_1,n_2)} J_\alpha\right) = \sum_{\alpha \in N_z(n_1,n_2)} \mu(J_\alpha) = \sum_{\alpha \in N_Z(n_1,n_2)} \nu(\Lambda^{-1}(J_\alpha)) = \sum_{\alpha \in N_z(n_1,n_2)} (\mathbb{I}^*\nu)(\alpha),$$

the last equality following from the fact that $\operatorname{supp} \nu \subset (\partial T)^2$ and $\Lambda(\partial \mathcal{S}(\alpha)) = J_\alpha$. Gathering the estimates we arrive at

$$\int_{(\partial \mathbb{D})^2} b_{(\frac{1}{2},\frac{1}{2})}(z,\zeta)\,d\mu(\zeta) \approx \sum_{n_1,n_2 \geq 0}\sum_{\alpha \in N_z(n_1,n_2)} \frac{(\mathbb{I}^*\nu)(\alpha)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}} = \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}},$$

which is (65).

The next part of the argument is actually a very special (linear) case of the well-known Wolff's inequality, that can be proven rather elementarily. We start by expanding the integrand,

$$\int_{(\partial \mathbb{D})^2}\left(\sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}}\right)^2 dm(z) = \int_{(\partial \mathbb{D})^2}\sum_{\alpha,\beta \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}(\beta))^{\frac{1}{2}}}\,dm(z).$$

Recall that $\alpha \in \mathcal{P}_{\mathfrak{G}^2}(\zeta)$ should be relatively close to the point ζ, namely

$$J_\alpha \subset \tilde{J}_\zeta(10 \cdot 2^{-d_T(\alpha_x)+1}, 10 \cdot 2^{-d_T(\alpha_y)+1}).$$

Therefore $\{z : \alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)\} \subset \tilde{J}_\alpha$, where $\tilde{J}_\alpha = \Lambda(\partial \mathcal{S}_{\mathfrak{G}^2}(p_2(\alpha)))$ and $p_2(\alpha) = (p_2(\alpha_x), p_2(\alpha_y))$ with $p_2(\alpha_x), p_2(\alpha_y)$ being the grandparents of α_x, α_y in the tree geometry (if one of the points is the root o or one of its children we assume the grandparent to be the root as well). It follows that

$$\int_{(\partial \mathbb{D})^2}\sum_{\alpha,\beta \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}(\beta))^{\frac{1}{2}}}\,dm(z) = \sum_{\alpha,\beta \in T^2}\int_{z:\alpha,\beta \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}(\beta))^{\frac{1}{2}}}\,dm(z) \leq$$

$$\sum_{\alpha,\beta \in T^2} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}(\beta))^{\frac{1}{2}}} \cdot m\left(\tilde{J}_\alpha \cap \tilde{J}_\beta\right) \lesssim$$

$$\sum_{\alpha,\beta \in T^2} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}(\beta))^{\frac{1}{2}}} \cdot M\left(\partial \mathcal{S}_{\mathfrak{G}^2}(p_2(\alpha)) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(p_2(\beta))\right) \lesssim$$

$$\sum_{\alpha,\beta \in T^2} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}}M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}}} \cdot M\left(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta)\right),$$

since $M(\partial \mathcal{S}_{\mathfrak{G}^2}(p_2(\alpha))) \approx M(\partial \mathcal{S}(\alpha)))$ for any $\alpha \in T^2$.

A point $\gamma \in T^2$ is called a proper $\mathfrak{G}^2$-descendant of $\tau \in T^2$ if $\gamma \in \mathcal{S}_{\mathfrak{G}^2}(\tau)$, and it lies strictly below τ, namely, either $d_T(\gamma_x) \geq d_T(\tau_x)$ and $d_T(\gamma_y) > d_T(\tau_y)$, or $d_T(\gamma_x) > d_T(\tau_x)$ and $d_T(\gamma_y) \geq d_T(\tau_y)$. Observe that for any two points $\alpha, \beta \in T^2$ there exists a (possibly non-unique) $\tau \in T^2$ such that α, β are proper

$\mathfrak{G}^2$-descendants of τ, and τ is minimal, that is, for any proper $\mathfrak{G}^2$-descendant γ of τ one of the points α, β is not a proper $\mathfrak{G}^2$-descendant of γ. Clearly,

$$\sum_{\alpha, \beta \in T^2} \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}} M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}}} \cdot M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta)) \le$$

$$\sum_{\tau \in T^2} \sum_{\alpha, \beta: \, \tau \text{ is minimal for } \alpha, \beta} \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}} M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}}} \cdot M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta)).$$

We aim to show that for any $\tau \in T^2$ one has

$$\sum_{\alpha, \beta: \, \tau \text{ is minimal for } \alpha, \beta} \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}} M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}}} \cdot M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta)) \lesssim (\nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)))^2. \quad (66)$$

Given $\alpha, \beta, \tau \in T^2$ let $a_x, a_y, b_x, b_y, t_x, t_y$ be their respective generation numbers, $a_x := d_T(\alpha_x) - 1, a_y := d_T(\alpha_y) - 1$ etc. First we note that $M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}} \approx 2^{-\frac{a_x + a_y}{2}}$, $M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}} \approx 2^{-\frac{b_x + b_y}{2}}$, and $M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha)) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta))) \lesssim 2^{-\max(a_x, b_x) - \max(a_y, b_y)}$, where the last equivalence is by a trivial estimate of $m(\tilde{J}_\alpha \cap \tilde{J}_\beta)$. The key observation here is as follows.

Proposition 5.1 *Assume that τ is minimal for α and β and either $a_x \ge t_x + 4$ and $b_x \ge t_x + 4$, or $a_y \ge t_y + 4$ and $b_y \ge t_y + 4$. Then $\mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \mathcal{S}_{\mathfrak{G}^2}(\beta) = \emptyset$. In other words, if α and β lie very 'deep' inside τ and are not 'perpendicular', then they must be 'far' from each other (see Fig. 5).*

Proof. Indeed, assume that $a_x \ge t_x + 4$, $b_x \ge t_x + 4$, and let $\gamma \in \mathcal{S}_{\mathfrak{G}^2}(\alpha) \cap \mathcal{S}_{\mathfrak{G}^2}(\beta)$, so that, in particular, $\gamma_x \in \mathcal{S}_{\mathfrak{G}}(\alpha_x) \cap \mathcal{S}_{\mathfrak{G}}(\beta_x)$. It follows immediately that both α_x, β_x belong either to $S_{\mathfrak{G}}(p_2(\alpha_x))$ (if $b_x \ge a_x$) or to $S_{\mathfrak{G}}(p_2(\beta_x))$ (if $b_x \le a_x$). Suppose we are in the first case. Then, since $a_x \ge t_x + 4$, we see that $p_2(\alpha_x) \in S_{\mathfrak{G}}(\tau_x)$ and $d_T(p_2(\alpha_x)) = a_x - 1 \ge t_x + 1$. Now define $\gamma = (\gamma_x, \gamma_y) := (p_2(\alpha_x), \tau_y)$. Clearly γ is a proper $\mathfrak{G}^2$-descendant of τ, and at the same time both α and β are proper $\mathfrak{G}^2$-descendants of γ. We have a contradiction. The case $a_y \ge t_y + 4$, $b_y \ge t_y + 4$ is done similarly. $\square$

Now we are ready to return to (66). Fix $\tau \in T^2$. By the observation above, if τ is minimal for α and β, then

$$0 \ne \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha)^{\frac{1}{2}} M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta)^{\frac{1}{2}}} \cdot M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha)) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta)))$$

$$\lesssim \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{2^{-\frac{a_x + a_y + b_x + b_y}{2}}} \cdot 2^{-\max(a_x, b_x) - \max(a_y, b_y)}$$

only if one of a_x, b_x and one of a_y, b_y are comparable to t_x and t_y, respectively. Note that we always have $a_x, b_x \ge t_x$ and $a_y, b_y \ge t_y$, since τ is minimal for α, β. Given a point $\alpha \in T^2$, assume that $a_x \le t_x + 4$ and $a_y \le t_y + 4$, and for $b_x \ge t_x$, $b_y \ge t_y$ denote by $A_\tau(\alpha, b_x, b_y)$ the set of all points β such that $d_T(\beta_x) = $
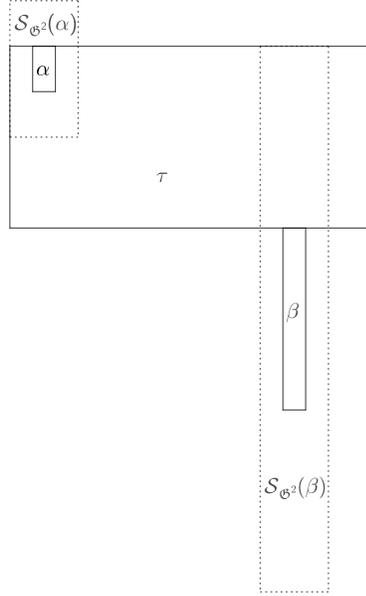
Figure 5

$b_x + 1, d_T(\beta_y) = b_y + 1$, and τ is minimal for α and β. Then

$$\sum_{b_x \geq t_x} \sum_{b_y \geq t_y} \sum_{\beta \in A_\tau(\alpha, b_x, b_y)} \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{2^{-\frac{a_x + a_y + b_x + b_y}{2}}} \cdot 2^{-\max(a_x, b_x) - \max(a_y, b_y)} \approx$$

$$(\mathbb{I}^* \nu)(\alpha) \cdot \sum_{b_x \geq t_x} \sum_{b_y \geq t_y} 2^{\frac{t_x + t_y + b_x + b_y}{2} - b_x - b_y} \sum_{\beta \in A_\tau(\alpha, b_x, b_y)} (\mathbb{I}^* \nu)(\beta) \approx (\mathbb{I}^* \nu)(\alpha) \sum_{\beta \in A_\tau(\alpha, b_x, b_y)} (\mathbb{I}^* \nu)(\beta) \lesssim$$

$$(\mathbb{I}^* \nu)(\alpha) \nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)),$$

since $\mathcal{S}(\beta^1) \cap \mathcal{S}(\beta^2) = \emptyset$ for any non-identical pair $\beta^1, \beta^2 \in A_\tau(\alpha, b_x, b_y)$, and $\bigcup_{\beta \in A_\tau(\alpha, b_x, b_y)} \subset \mathcal{S}_{\mathfrak{G}^2}(\tau)$. It follows that

$$\sum_{a_x = t_x}^{t_x + 4} \sum_{a_y = t_y}^{t_y + 4} \sum_{\alpha: \, \alpha \text{ is a proper } \mathfrak{G}^2\text{-descendant of } \tau, \, d_T(\alpha_x) = a_x + 1, d_T(\alpha_y) = a_y + 1} (\mathbb{I}^* \nu)(\alpha) \nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)) \approx (\nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)))^2.$$

The remaining cases (i.e. $a_x \leq t_x + 4$ and $b_y \leq t_y + 4$, $b_x \leq t_x + 4$ and $a_y \leq t_y + 4$, or $b_x \leq t_x + 4$ and $b_y \leq t_y + 4$) are dealt with similarly. Gathering all the estimates we arrive at (66), hence

$$\sum_{\tau \in T^2} \sum_{\alpha, \beta: \, \tau \text{ is minimal for } \alpha, \beta} \frac{(\mathbb{I}^* \nu)(\alpha)(\mathbb{I}^* \nu)(\beta)}{M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha))^{\frac{1}{2}} M(\partial \mathcal{S}_{\mathfrak{G}^2}(\beta))^{\frac{1}{2}}} \cdot M(\partial \mathcal{S}_{\mathfrak{G}^2}(\alpha)) \cap \partial \mathcal{S}_{\mathfrak{G}^2}(\beta))) \lesssim \sum_{\tau \in T^2} (\nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)))^2.$$

By the $\mathfrak{G}$-neighbours argument of Section 2.4 we have

$$\sum_{\tau \in T^2} (\nu(\mathcal{S}_{\mathfrak{G}^2}(\tau)))^2 \approx \sum_{\tau \in T^2} ((\mathbb{I}^*\nu)(\tau))^2 = \mathcal{E}[\nu],$$

and therefore

$$\int_{(\partial\mathbb{D})^2} \sum_{\alpha,\beta \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial\mathcal{S}(\alpha))^{\frac{1}{2}} M(\partial\mathcal{S}(\beta))^{\frac{1}{2}}} \, dm(z) \lesssim \mathcal{E}[\nu]. \tag{67}$$

On the other hand,

$$\int_{(\partial\mathbb{D})^2} \sum_{\alpha,\beta \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)(\mathbb{I}^*\nu)(\beta)}{M(\partial\mathcal{S}(\alpha))^{\frac{1}{2}} M(\partial\mathcal{S}(\beta))^{\frac{1}{2}}} \, dm(z) \geq \int_{(\partial\mathbb{D})^2} \sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{((\mathbb{I}^*\nu)(\alpha))^2}{M(\partial\mathcal{S}(\alpha))} \, dm(z) \geq$$

$$\sum_{\alpha \in T^2} \frac{((\mathbb{I}^*\nu)(\alpha))^2}{M(\partial\mathcal{S}(\alpha))} \cdot m\{z:\ \alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)\} \geq \sum_{\alpha \in T^2} \frac{((\mathbb{I}^*\nu)(\alpha))^2}{M(\partial\mathcal{S}(\alpha))} \cdot m(J_\alpha) = \sum_{\alpha \in T^2} ((\mathbb{I}^*\nu)(\alpha))^2 = \mathcal{E}[\nu]. \tag{68}$$

Combined with (65) these estimates give us

$$\mathcal{E}_{(\frac{1}{2},\frac{1}{2})}[\mu] = \int_{(\partial\mathbb{D})^2} \left(\int_{(\partial\mathbb{D})^2} b_{(\frac{1}{2},\frac{1}{2})}(z,\zeta) \, d\mu(\zeta) \right)^2 dm(z)$$

$$\approx \int_{(\partial\mathbb{D})^2} \left(\sum_{\alpha \in \mathcal{P}_{\mathfrak{G}^2}(z)} \frac{(\mathbb{I}^*\nu)(\alpha)}{M(\partial\mathcal{S}(\alpha))^{\frac{1}{2}}} \right)^2 dm(z) \approx \mathcal{E}[\nu],$$

which finishes the proof. $\qquad\square$

5.2

Lemma 5.6 *The following properties hold:*

1. *Let E be a Borel subset of $\overline{T}^2$ and let $\mu \geq 0$ be a Borel measure on $\overline{T}^2$ such that $\mathcal{E}[\mu] < \infty$ and $\mathbb{V}^\mu \leq 1$ quasi-everywhere on E. Then $\mathrm{Cap}\, E \geq \mu(E)$.*

2. *Let F be an open set in ∂T and ρ be its equilibrium measure. Then $\mathrm{supp}\,\rho = \mathrm{cl}F$, and $V^\rho \equiv 1$ on F.*

3. ***Maximum principle.*** *Let $\mu \geq 0$ be a measure on $\overline{T}$ with finite energy. Then*

$$\sup_{\alpha \in \mathrm{supp}\,\mu} V^\mu(\alpha) = \sup_{\alpha \in \overline{T}} V^\mu(\alpha). \tag{69}$$

4. ***Domination principle.*** *Let f be a non-negative function in $\ell^2(T)$ such that $f(\alpha) \geq f(\alpha^+) + f(\alpha^-)$ for any point $\alpha \in T$ and its two children $\alpha^\pm$. Suppose $\nu \geq 0$ is a Borel measure on $\overline{T}$ with finite energy such that*

$$(If)(\alpha) \geq V^\nu, \quad \alpha \in \mathrm{supp}\,\nu. \tag{70}$$

Then this inequality holds everywhere,

$$(If)(\alpha) \geq V^{\nu}(\alpha), \quad \alpha \in \overline{T}. \tag{71}$$

Proof. Property 1. Define, as usual, the restricted measure $\mu|_E$ by $\mu|_E(F) := \mu(E \cap F)$, and let μ_E be the equilibrium measure of E. Clearly, $\mathcal{E}[\mu|_E] < \infty$ and $\mathbb{V}^{\mu|_E} \leq \mathbb{V}^{\mu}$. We have

$$\mathcal{E}[\mu|_E] = \int_{\overline{T}^2} \mathbb{V}^{\mu|_E} d\mu|_E \leq \int_{\overline{T}^2} \mathbb{V}^{\mu_E} d\mu|_E = \mathcal{E}[\mu_E, \mu|_E],$$

since $\mathbb{V}^{\mu_E} \geq 1$ quasi-everywhere on E. Hence, since $\mathcal{E}[\mu_E - \mu|_E] \geq 0$,

$$\mu(E) \leq \mathcal{E}[\mu_E, \mu|_E] \leq \mathcal{E}[\mu_E] = \text{Cap}\, E.$$

Property 2. We first show that $V^{\rho} = 1$ on F. Fix a point $\omega \in F$. Since F is open, there is a point $\tau > \omega$ such that $\partial \mathcal{S}(\tau) \subset F$. Since ρ minimizes the energy, an elementary computation shows that for any $\alpha \leq \tau$ one must have $(I^*\rho)(\alpha) = 2(I^*\rho)(\alpha^+) = 2(I^*\rho)(\alpha^-)$, where $\alpha^{\pm}$ are the children of α. Therefore V^{ρ} is actually constant on $\partial \mathcal{S}(\tau)$, and thus $V^{\rho}(\omega) = 1$.

Next we show that $\text{supp}\,\rho = \text{cl}F$. Let $\omega \in \text{cl}F$, and consider an open neighbourhood U_{ω} of ω. If $\rho(U_{\omega}) = 0$ there is a smallest point $\alpha > \omega$ such that $\rho(\mathcal{S}(\alpha)) \neq 0$. Denote its two children by α^+ and α^-, and assume that $\omega \in \mathcal{S}(\alpha^+)$. Since F is open, $\mathcal{S}(\alpha^+) \cap F \neq \emptyset$. Therefore, $V^{\rho}(\alpha^+) = V^{\rho}(\alpha) \geq 1$, since $V^{\rho} = 1$ on F. On the other hand, $(I^*\rho)(\alpha^-) > 0$, by the minimality of α. Thus $V^{\rho}(\alpha^-) > 1$, which contradicts the fact that $V^{\rho} \leq 1$ on $\text{supp}\,\rho$. Therefore it must have been that $\rho(U_{\omega}) > 0$, and thus that $\text{cl}F \subset \text{supp}\,\rho$. The converse is elementary.

Property 3. It is enough to check (69) inside the tree (i.e. for $\alpha \in T$), since for any $\beta \in \partial T$ we have $V^{\mu}(\beta) = \sup_{\alpha > \beta} V^{\mu}(\alpha)$. Now assume that there exists a point $\beta \in T \setminus \text{supp}\,\mu$ such that

$$V^{\mu}(\beta) > V^{\mu}(\alpha), \quad \alpha \in \text{supp}\,\mu.$$

We see immediately that $\mathcal{S}(\beta) \cap \text{supp}\,\mu = \emptyset$, and hence there exists a unique point $\tau_{\beta} > \beta$ such that $\mathcal{S}(\tau_{\beta}) \cap \text{supp}\,\mu \neq \emptyset$, but $\mathcal{S}(\tau) \cap \text{supp}\,\mu = \emptyset$ for every $\tau < \tau_{\beta}$. Then $(I^*\mu)(\tau) = 0$ for such points τ, and

$$V^{\mu}(\beta) = V^{\mu}(\tau_{\beta}) + \sum_{\tau_{\beta} > \tau \geq \beta} (I^*\mu)(\beta) = V^{\mu}(\tau_{\beta}).$$

Monotonicity of V^{μ}, with respect to natural order on T, implies that $V^{\mu}(\tau_{\beta}) < V^{\mu}(\alpha)$ for any $\alpha \in \mathcal{S}(\tau_{\beta}) \cap \text{supp}\,\mu$, yielding a contradiction.

Property 4. As before, it is enough to show (71) only for points inside T. Now suppose there exists $\alpha^0 \in T$ such that

$$(If)(\alpha^0) < V^{\nu}(\alpha^0)$$

and

$$(If)(\tau) \geq V^{\nu}(\tau), \quad \tau > \alpha^0.$$

It follows immediately that $f(\alpha^0) < (I^*v)(\alpha^0)$. Hence one of the children of α^0, which we denote by α_1, satisfies $f(\alpha^1) < (I^*v)(\alpha^1)$. Continuing this argument we obtain a sequence $\{\alpha^k\}_0^\infty$ of nested points such that $f(\alpha^k) < (I^*v)(\alpha^k)$, $k = 0,1,\dots$. Denote the endpoint of this geodesic by $\omega = \bigcap_k S(\alpha^k)$. Clearly $\omega \in \operatorname{supp} v$. It follows that

$$(If)(\omega) = (If)(\alpha^0) + \sum_{k=1}^\infty f(\alpha^k) < V^v(\alpha^0) + \sum_{k=1}^\infty (I^*v)(\alpha^k) = V^v(\omega),$$

a contradiction. $\qquad\square$

5.3

Proposition 5.2 *For any $\lambda > 1$ there exists a point $\omega \in \overline{T}^2$ and a set $E \subset \overline{T}^2$ such that the equilibrium measure $\mu = \mu_E$ of this set satisfies*

$$\mathbb{V}^\mu(\omega) \geq \lambda. \tag{72}$$

Proof. Put $n = 20([\lambda] + 1)$ and $k = 20^n$. Now fix any point $\omega \in (\partial T)^2$ and for any $0 \leq i \leq n$ consider the unique point $\alpha^i = (\alpha_x^i, \alpha_y^i)$ that satisfies $\alpha^i \geq \omega$, $d_T(\alpha_x^i) = 20^{n-i}$, $d_T(\alpha_y^i) = 20^i$, so that $d_{T^2}(\alpha^i) = 20^n = k$. We have $\alpha_x^i \wedge \alpha_x^j = \alpha_x^i$, $\alpha_y^i \wedge \alpha_y^j = \alpha_y^j$ for $i > j$, hence

$$\sum_{i \neq j} d_{T^2}(\alpha^i \wedge \alpha^j) = \sum_{j < i} 20^{n+j-i} + \sum_{j > i} 20^{n+i-j} \leq \frac{2}{19} 20^n \leq \frac{k}{9}. \tag{73}$$

Now let $E := \{\alpha^i\}_{i=0}^n$ and $\mu = \mu_E$ to be the equilibrium measure of E. We claim that $\mathbb{V}^\mu(\omega) \geq \frac{n}{20} \geq \lambda$.

To show this we first note that the values of μ at α^i, $i = 0,\dots,n$ are more or less the same,

$$\sup_{0 \leq i \leq n} \mu(\alpha^i) \leq 5 \inf_{0 \leq i \leq n} \mu(\alpha^i). \tag{74}$$

Indeed, assume that i_1, i_2 are such that $\sup_{0 \leq i \leq n} \mu(\alpha^i) = \mu(\alpha^{i_1})$, $\inf_{0 \leq i \leq n} \mu(\alpha^i) = \mu(\alpha^{i_2})$, and

$$\mu(\alpha^{i_1}) > 5\mu(\alpha^{i_2}).$$

Since every element of E has non-zero capacity (actually $\operatorname{Cap}\{\alpha^i\} \equiv \frac{1}{k} > 0$), we have

$$1 \leq \mathbb{V}^\mu(\alpha^{i_2}) = \sum_{i \neq i_2} d_{T^2}(\alpha^i \wedge \alpha^{i_2})\mu(\alpha^i) + k\mu(\alpha^{i_2}) \leq$$

$$\sum_{i \neq i_2} d_{T^2}(\alpha^i \wedge \alpha^{i_2})\mu(\alpha^{i_1}) + k\mu(\alpha^{i_2}) \leq \frac{k}{9}\mu(\alpha^{i_1}) + \frac{k}{5}\mu(\alpha^{i_1}) \leq$$

$$\frac{1}{2}k\mu(\alpha^{i_1}) + \frac{1}{2}\sum_{i \neq i_1} d_{T^2}(\alpha^i \wedge \alpha^{i_1})\mu(\alpha^i) = \frac{1}{2}\mathbb{V}^\mu(\alpha^{i_1}).$$

On the other hand, $\mu(\alpha^{i_1}) > 0$, hence $\mathbb{V}^\mu(\alpha^{i_1}) = 1$, and we have a contradiction.

Furthermore,

$$1 = \mathbb{V}^{\mu}(\alpha^{i_1}) = k\mu(\alpha^{i_1}) + \sum_{i \neq i_1} d_{T^2}(\alpha^i \wedge \alpha^{i_1})\mu(\alpha^i) \leq$$

$$k\mu(\alpha^{i_1}) + \sum_{i \neq i_1} d_{T^2}(\alpha^i \wedge \alpha^{i_1})\mu(\alpha^i) \leq \frac{10}{9}k\mu(\alpha^{i_1}).$$

Therefore, for any $0 \leq i \leq n$,

$$\mu(\alpha^i) \geq \frac{1}{5}\mu(\alpha^{i_1}) \geq \frac{9}{50k}.$$

It follows immediately that

$$\mathbb{V}^{\mu}(\omega) = \sum_{i=0}^{n} d_{T^2}(\omega \wedge \alpha^i)\mu(\alpha^i) = \sum_{i=0}^{n} d_{T^2}(\alpha^i)\mu(\alpha^i) \geq kn\mu(\alpha^{i_2}) \geq \frac{9n}{50},$$

and we are done. $\qquad \square$

Proposition 5.3 *For any $\lambda > 0$ there exists a pair of measures $\mu, \nu \geq 0$ on $\overline{T}^2$ such that*

$$\mathbb{V}^{\nu}(\alpha) \geq \mathbb{V}^{\mu}(\alpha), \quad \alpha \in \operatorname{supp}\mu, \tag{75}$$

but

$$\sup_{\alpha \in \overline{T}^2} \mathbb{V}^{\nu}(\alpha) \leq \sup_{\alpha \in \overline{T}^2} \frac{1}{\lambda}\mathbb{V}^{\mu}(\alpha). \tag{76}$$

Proof. This is a direct corollary of Proposition 5.2 above. Indeed, given $\lambda > 0$ let $\mu = \mu_E$ be as in (72), and let $\nu := \chi_o$ be the unit point mass at the root. Then, clearly, $\mathbb{V}^{\nu} \equiv 1$ on $\overline{T}^2$, and in particular on $\operatorname{supp}\mu$, but $\sup_{\alpha \in \overline{T}^2} \mathbb{V}^{\mu} \geq \lambda$. $\qquad \square$

5.4

Let m_0 be normalized length on $\partial \mathbb{D}$. Define M_0 to be its natural pull-back on ∂T, $M_0 = \Lambda_0^* m_0$. In particular, we have

$$M_0(\partial \mathcal{S}(\beta)) = 2^{-d_T(\beta)+1}, \quad \beta \in T.$$

Similarly, as in Lemma 5.5, let $M = \Lambda^* m$, where m is normalized area measure on the torus $(\partial \mathbb{D})^2$. Clearly,

$$M(\partial \mathcal{S}(\beta)) = 2^{-d_T(\beta_x)-d_T(\beta_y)+2} = M_0(\partial \mathcal{S}(\beta_x))M_0(\partial \mathcal{S}(\beta_y)), \quad \beta = (\beta_x, \beta_y) \in T^2.$$

Let us show that d_{T^2} is almost a martingale with respect to the measure M.

Lemma 5.7 *Assume that $\alpha, \beta \in T^2$. Then*

$$d_{T^2}(\alpha \wedge \beta) \approx \frac{1}{M(\partial \mathcal{S}(\alpha))M(\partial \mathcal{S}(\beta))} \int_{\partial \mathcal{S}(\alpha)} \int_{\partial \mathcal{S}(\beta)} d_{T^2}(\xi \wedge \omega)dM(\xi)\,dM(\omega). \tag{77}$$

Proof. Due to multiplicativity it is enough to prove that, say,

$$d_T(\alpha_x \wedge \beta_x) \approx \frac{1}{M_0(\partial S(\alpha_x))M_0(\partial S(\beta_x))} \int_{\partial S(\alpha_x)} \int_{\partial S(\beta_x)} d_T(\xi_x \wedge \omega_x)dM(\xi_x)dM(\omega_x).$$

If $\xi_x \leq \alpha_x$ and $\omega_x \leq \beta_x$, then $d_T(\xi_x \wedge \omega_x) \geq d_T(\alpha_x \wedge \beta_x)$, hence

$$d_T(\alpha_x \wedge \beta_x) \leq \frac{1}{M_0(\partial S(\alpha_x))M_0(\partial S(\beta_x))} \int_{\partial S(\alpha_x)} \int_{\partial S(\beta_x)} d_T(\xi_x \wedge \omega_x)dM_0(\xi_x)dM_0(\omega_x).$$

To get the reverse inequality we first show that for any $\beta_x \in T_x$ and $\tau_x \in \overline{T}_x$ we have

$$\frac{1}{M_0(\partial S(\beta_x))} \int_{\partial S(\beta_x)} d_T(\tau_x \wedge \omega_x)dmM_0(\omega_x) \leq 3d_T(\tau_x \wedge \beta_x). \tag{78}$$

If $\tau_x \geq \beta_x$ or these two points are not comparable, then, clearly, $d_T(\tau_x \wedge \beta_x) = d_T(\tau_x \wedge \omega_x)$ for $\omega_x \leq \beta_x$, and (78) is trivial. Hence from now on we assume that $\tau_x < \beta_x$. Let $n := d_T(\beta_x)$ and $N := d_T(\tau_x)$. For every $n \leq k \leq N$ there exists exactly one point $\gamma_k \in T_x$ such that $\tau_x \leq \gamma_k \leq \beta_x$, and $d_T(\gamma_k) = k$ (in particular $\gamma_n = \beta_x$, $\gamma_N = \tau_x$). Define

$$S_k = \partial S(\gamma_k) \setminus \partial S(\gamma_{k+1}), \quad n \leq k \leq N-1,$$

and

$$S_N = \partial S(\tau_x).$$

If $\omega_x \in S_k$, then, clearly, $d_T(\tau_x \wedge \omega_x) = k$. Moreover, these sets are disjoint and form a covering of $\partial S(\beta_x)$. Also $M_0(S_k) = 2^{-k} - 2^{-k-1}, n \leq k \leq N-1$ and $M_0(S_N) = 2^{-N}$. We have

$$\frac{1}{M_0(\partial S(\beta_x))} \int_{\partial S(\beta_x)} d_T(\tau_x \wedge \omega_x)dM_0(\omega_x) =$$

$$2^{d_T(\beta_x)} \sum_{k=n}^{N} \int_{S_k} d_t(\tau_x \wedge \omega_x)dM_0(\omega_x) = 2^n \sum_{k=n}^{N} k \cdot M_0(S_k) \leq$$

$$2^n \sum_{k=n}^{N} k2^{-k} \leq 3n = 3d_T(\tau_x \wedge \beta_x),$$

and we arrive at (78). It follows immediately that

$$\frac{1}{M_0(\partial S(\alpha_x))M_0(\partial S(\beta_x))} \int_{\partial S(\alpha_x)} \int_{\partial S(\beta_x)} d_T(\xi_x \wedge \omega_x)dM_0(\xi_x)dM_0(\omega_x) \leq$$

$$3\frac{1}{M_0(\partial S(\alpha_x))} \int_{\partial S(\alpha_x)} d_T(\xi_x \wedge \beta_x)M_0(\xi_x) \leq 9d_T(\alpha_x \wedge \beta_x).$$

$\square$

Theorem 5.1 *Suppose that $\mu \geq 0$ is a Borel measure on $\overline{T}^2$ with finite energy. By the Disintegration Theorem we can define a measure μ_b supported on the $(\partial T)^2$ by*

$$d\mu_b(\omega_x, \omega_y) = \sum_{\beta > \omega} \frac{\mu(\beta)}{M(\partial \mathcal{S}(\beta))} dM(\omega) +$$

$$\sum_{\beta_y > \omega_y} \frac{d\mu(\omega_x, \beta_y)}{M_0(\partial \mathcal{S}(\beta_y))} dM_0(\omega_y) + \sum_{\beta_x > \omega_x} \frac{d\mu(\beta_x, \omega_y)}{M_0(\partial \mathcal{S}(\beta_x))} dM_0(\omega_x) + d\mu(\omega_x, \omega_y). \tag{79}$$

Then the potentials of μ and μ_b are equivalent,

$$\mathbb{V}^\mu(\alpha) \approx \mathbb{V}^{\mu_b}(\alpha), \quad \alpha \in \overline{T}^2. \tag{80}$$

Proof. Fix any point $\alpha \in \overline{T}^2$. We have

$$\mathbb{V}^{\mu_b}(\alpha) = \int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) \, d\mu_b(\omega) = \int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) \sum_{\beta > \omega} \frac{\mu(\beta)}{M(\partial \mathcal{S}(\beta))} dM(\omega) +$$

$$\int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) \sum_{\beta_y > \omega_y} \frac{d\mu(\omega_x, \beta_y)}{M_0(\partial \mathcal{S}(\beta_y))} dM_0(\omega_y) +$$

$$\int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) \sum_{\beta_x > \omega_x} \frac{d\mu(\beta_x, \omega_y)}{M_0(\partial \mathcal{S}(\beta_x))} dM_0(\omega_x) +$$

$$\int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) d\mu(\omega) := (I) + (II) + (III) + (IV).$$

By Tonelli's theorem and Lemma 5.7

$$(I) = \sum_{\beta \in T^2} \frac{\mu(\beta)}{M(\partial \mathcal{S}(\beta))} \int_{\partial \mathcal{S}(\beta)} d_{T^2}(\alpha \wedge \omega) \, dM(\omega) \approx \sum_{\beta \in T^2} \mu(\beta) d_{T^2}(\alpha \wedge \beta).$$

Similarly,

$$(II) =$$

$$\int_{\partial T_x} d_T(\alpha_x \wedge \omega_x) \sum_{\beta_y \in T_y} \left(\frac{1}{M_0(\partial \mathcal{S}(\beta_y))} \int_{\partial \mathcal{S}(\beta_y)} d_T(\alpha_y \wedge \omega_y) dM_0(\omega_y) \right) d\mu(\omega_x, \beta_y) \approx$$

$$\sum_{\beta_y \in T_y} d_T(\alpha_y \wedge \beta_y) \int_{\partial T_x} d_T(\alpha_x \wedge \omega_x) d\mu(\omega_x, \beta_y),$$

and

$$(III) \approx \sum_{\beta_x \in T_x} d_T(\alpha_x \wedge \beta_x) \int_{\partial T_y} d_T(\alpha_y \wedge \omega_y) d\mu(\beta_x, \omega_y).$$

We arrive at

$$V^{\mu_b}(\alpha) \approx \sum_{\beta \in T^2} \mu(\beta) d_{T^2}(\alpha \wedge \beta) + \sum_{\beta_y \in T_y} d_T(\alpha_y \wedge \beta_y) \int_{\partial T_x} d_T(\alpha_x \wedge \omega_x) d\mu(\omega_x, \beta_y) +$$

$$\sum_{\beta_x \in T_x} d_T(\alpha_x \wedge \beta_x) \int_{\partial T_y} d_T(\alpha_y \wedge \omega_y) d\mu(\beta_x, \omega_y) + \int_{(\partial T)^2} d_{T^2}(\alpha \wedge \omega) \, d\mu(\omega) =$$

$$\int_{T^2} d_{T^2}(\alpha \wedge \tau) \, d\mu(\tau) + \int_{\partial T_x \times T_y} d_{T^2}(\alpha \wedge \tau) \, d\mu(\tau) + \int_{T_x \times \partial T_y} d_{T^2}(\alpha \wedge \tau) \, d\mu(\tau) +$$

$$\int_{\partial T_x \times \partial T_y} d_{T^2}(\alpha \wedge \tau) \, d\mu(\tau) = \mathbb{V}^\mu(\alpha).$$

$\square$

Corollary 5.1 *Given a Borel set $E \subset \overline{T}^2$ define its boundary projection $\mathcal{S}_b(E) \subset (\partial T)^2$ to be*

$$\mathcal{S}_b(E) = \bigcup_{\beta \in E} \partial \mathcal{S}(\beta).$$

Then there exists a constant $C > 1$ such that

$$\operatorname{Cap} \mathcal{S}_b(E) \leq \operatorname{Cap} E \leq C \operatorname{Cap} \mathcal{S}_b(E). \tag{81}$$

Proof. We start by assuming that E is compact. The left inequality is trivial, since any function admissible for E is also admissible for $\mathcal{S}_b(E)$. Now let μ and ν be the equilibrium measures for E and $\mathcal{S}_b(E)$ respectively. By the definition of μ_b,

$$|\mu_b| = \int_{(\partial T)^2} d\mu_b = \int_{(\partial T)^2} \sum_{\beta > \omega} \frac{\mu(\beta)}{M(\partial \mathcal{S}(\beta))} dM(\omega) +$$

$$\int_{(\partial T)^2} \sum_{\beta_y > \omega_y} \frac{d\mu(\omega_x, \beta_y)}{M_0(\partial \mathcal{S}(\beta_y))} dM_0(\omega_y) + \int_{(\partial T)^2} \sum_{\beta_x > \omega_x} \frac{d\mu(\beta_x, \omega_y)}{M_0(\partial \mathcal{S}(\beta_x))} dM_0(\omega_x) +$$

$$\int_{(\partial T)^2} d\mu(\omega) = \sum_{\beta \in T^2} \mu(\beta) + \sum_{\beta_y \in T_y} \int_{\partial T_x} d\mu(\omega_x, \beta_y) + \sum_{\beta_x \in T_x} \int_{\partial T_y} d\mu(\beta_x, \omega_y) +$$

$$\int_{(\partial T)^2} d\mu = \int_{\overline{T}^2} d\mu = |\mu|.$$

By Theorem 5.1 and the fact that μ is an equilibrium measure,

$$|\mu_b| = |\mu| = \int_{\overline{T}^2} \mathbb{V}^\mu \, d\mu \approx \int_{\overline{T}^2} \mathbb{V}^{\mu_b} \, d\mu \approx \int_{\overline{T}^2} \mathbb{V}^{\mu_b} \, d\mu_b. \tag{82}$$

On the other hand, for every $C \in \mathbb{R}$ we have

$$0 \leq \int_{\overline{T}^2} \mathbb{V}^{\mu_b} \, d\mu_b - 2C \int_{\overline{T}^2} \mathbb{V}^\nu \, d\mu_b + C^2 \int_{\overline{T}^2} \mathbb{V}^\nu \, d\nu \leq \int_{\overline{T}^2} \mathbb{V}^{\mu_b} \, d\mu_b - 2C|\mu_b| + C^2|\nu|,$$

since ν is an equilibrium measure for $\mathcal{S}_b(E)$ and $\mathbb{V}^\nu \geq 1$ quasi-everywhere on $\mathcal{S}_b(E) \supset \operatorname{supp}\mu_b$. By (82) there is a $C > 1$, independent of E, such that

$$0 \leq \int_{\overline{T}^2} \mathbb{V}^{\mu_b}\, d\mu_b - C|\mu_b| + C\left(C|\nu| - |\mu_b|\right) \leq C\left(C|\nu| - |\mu_b|\right).$$

Therefore

$$C\operatorname{Cap}\mathcal{S}_b(E) = C|\nu| \geq |\mu_b| = |\mu| = \operatorname{Cap}E,$$

and we get the second half of (81).

Given a general set E we exhaust it by compact sets E_k from inside. Then $\lim_{k\to\infty}\operatorname{Cap}E_k = \operatorname{Cap}E$ and $\lim_{k\to\infty}\operatorname{Cap}\mathcal{S}_b(E_k) = \operatorname{Cap}\mathcal{S}_b(E)$, an we still have (81) by the argument above. $\qquad\square$

Acknowledgments

References

[1] Adams, David R. *On the existence of capacitary strong type estimates in* $\mathbb{R}^n$. Ark. Mat. 14 (1976), no. 1, 125-140. 5, 25

[2] Adams, David R.; Hedberg, Lars Inge. *Function spaces and potential theory.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 314. Springer-Verlag, Berlin, 1996. xii+366 pp. 5, 10, 27

[3] Agler, Jim; McCarthy, John E. *Pick Interpolation and Hilbert Function Spaces.* Graduate Studies in Mathematics, 44. AMS, 2002. 308 pp. 7, 24, 38

[4] Aleman, Alexandru; Hartz, Michael; McCarthy, John E.; Richter, Stefan. *Interpolating sequences in spaces with the complete Pick property.* Int. Math. Res. Not. (2017). 38

[5] Arcozzi, Nicola; Holmes, Irina; Mozolyako, Pavel; Volberg, Alexander. *Bellman function sitting on a tree.* Int. Math. Res. Not. (2019). 2

[6] Arcozzi, Nicola; Holmes, Irina; Mozolyako, Pavel; Volberg, Alexander. *Bi-parameter embedding and measures with restriction energy condition.* Math. Ann. 377 (2020), no. 1-2, 643-674.

[7] Arcozzi, Nicola; Rochberg, Richard; Sawyer, Eric. *Carleson measures for analytic Besov spaces.* Rev. Mat. Iberoamericana 18 (2002), no. 2, 443-510. 30

[8] Arcozzi, Nicola; Rochberg, Richard; Sawyer, Eric. *The characterization of the Carleson measures for analytic Besov spaces: a simple proof.* Complex and harmonic analysis, 167-177, DEStech Publ., Inc., Lancaster, PA, 2007. 2, 4, 5, 38

[9] Arcozzi, Nicola; Rochberg, Richard; Sawyer, Eric. *Carleson measures for the Drury-Arveson Hardy space and other Besov-Sobolev spaces on complex balls.* Adv. Math. 218 (2008), no. 4, 1107-1180. 2

[10] Arcozzi, Nicola; Rochberg, Richard; Sawyer, Eric T.; Wick, Brett D. *Potential theory on trees, graphs and Ahlfors-regular metric spaces.* Potential Anal. 41 (2014), no. 2, 317-366. 2, 6

[11] Bishop, Christopher. *Interpolating sequences for the Dirichlet space and its multipliers.* (1994), preprint https://www.math.stonybrook.edu/~bishop/papers/mult.pdf. 5, 6, 18, 22, 38, 42

[12] Beurling, Arne. *Ensembles exceptionnels.* (French) Acta Math. 72 (1940), 1–13. 41

[13] Cairoli, Renzo. *Une inégalité pour martingales à indices multiples et ses applications.* (French) Séminaire de Probabilités, IV (Univ. Strasbourg, 1968/69) 4 (1970), Lecture Notes in Mathematics, Vol. 124. Springer, Berlin, 1-27 3

[14] Carleson, Lennart. *Interpolations by bounded analytic functions and the corona problem.* Ann. of Math. 76 (1962), no. 3, 547-559. 2, 38

[15] Carleson, Lennart. *A counterexample for measures bounded on H^p for the bi-disc.* Mittag-Leffler Report (1974), no. 7. 2

[16] Chang, Sun-Yung A. *Carleson measure on the bi-disc.* Ann. of Math. 109 (1979), no. 3, 613-620. 2

[17] Gundy, R. F. Inégalités pour martingales à un et deux indices: l'espace H^p. (French) Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978), Lecture Notes in Math., 774, Springer, Berlin, 1980, 251–334. 2

[18] Hedberg, Lars Inge; Wolff, Thomas. H. *Thin sets in nonlinear potential theory.* Ann. Inst. Fourier (Grenoble) 33 (1983), no. 4, 161-187. 2, 38

[19] Jessen, B.; Marcinkiewicz, J., Zygmund, A. *Note on the differentiability of multiple integrals.* Fund. Math. 25 (1935), 217–234. 38

[20] Kalton, N. J.; Verbitsky, I. E. *Nonlinear equations and weighted norm inequalities.* Trans. Amer. Math. Soc. 351 (1999), no. 9, 3441-3497. 2

[21] Kerman, Ron; Sawyer, Eric T. *The trace inequality and eigenvalue estimates for Schrödinger operators.* Ann. Inst. Fourier (Grenoble) 36 (1986), no. 4, 207-228. 5

[22] Kerman, Ron; Sawyer, Eric T. *Carleson measures and multipliers of Dirichlet-type spaces.* Trans. Amer. Math. Soc. 309 (1988), no. 1, 87-98. 5

[23] Maz'ya, V. G. *Certain integral inequalities for functions of several variables.* (Russian) Problems of mathematical analysis, No. 3: Integral and differential operators, Differential equations. Izdat. Leningrad. Univ., Leningrad, 1972, 33-68. 38

[24] Muscalu, Camil; Schlag, Wilhelm *Classical and multilinear harmonic analysis.* Vol. II. Cambridge Studies in Advanced Mathematics, 138. Cambridge University Press, Cambridge, 2013. 5, 25

[25] Sawyer, Eric T. *Weighted inequalities for the two-dimensional Hardy operator.* Studia Math. 82 (1985), no. 1, 1-16. 2

[26] Seip, Kristian. *Interpolation and sampling in spaces of analytic functions.* University Lecture Series, 33. American Mathematical Society, Providence, RI, 2004. xii+139 pp. 2, 4

[27] Stegenga, David A. *Multipliers of the Dirichlet space.* Illinois J. Math. 24 (1980), no. 1, 113-139. 38

[28] Wu, Zhijian. *Carleson measures and multipliers for Dirichlet spaces*. (English summary) J. Funct. Anal. 169 (1999), no. 1, 148-163. 2, 3, 5, 24, 38

[29] Young, Nicholas. Personal communication. 38
    38

AUTHORS

Nicola Arcozzi
Dipartimento di Matematica
Università di Bologna "Alma Mater"
P.zza di P.ta S. Donato 5, 40126 Bologna, Italy
nicola.arcozzi@unibo.it

Pavel Mozolyako
Department of Mathematics and Computer Science
Saint-Petersburg University
14th line V.O., 29B, 199178 Saint-Petersburg, Russia
pmzlcroak@gmail [dot] com

Karl-Mikael Perfekt
Department of Mathematical Sciences
Norwegian University of Science and Technology
7491 Trondheim, Norway
karl-mikael [dot] perfekt@ntnu [dot] no

Giulia Sarfatti
Dipartimento di Ingegneria Industriale e Scienze Matematiche
Università Politecnica delle Marche
via Brecce Bianche 12, 60131 Ancona, Italy
g [dot] sarfatti@univpm [dot] it