

STOCHASTIC EQUIDISTRIBUTION AND GENERALIZED ADELIC MEASURES

JOHN R. DOYLE, PAUL FILI, AND BELLA TOBIN

ABSTRACT. We study the dynamics of stochastic families of rational maps on the projective line. As such families can be infinite and may not typically be defined over a single number field, we introduce the concept of generalized adelic measures, generalizing previous notions introduced by Favre and Rivera-Letelier and Mavraki and Ye. Generalized adelic measures are defined over the measure space of places of an algebraic closure of the rationals, using a framework established by Allcock and Vaaler. This turns our heights from sums over places into integrals. We prove an equidistribution result for generalized adelic measures, and use this result to prove an equidistribution result for random backwards orbits in stochastic arithmetic dynamics.

CONTENTS

1. Introduction	1
Acknowledgments	5
2. Generalized Adelic Measures	5
2.1. Definitions	5
2.2. Existence results for finite height measures	9
3. Equidistribution for generalized adelic measures	17
3.1. Heights associated to generalized adelic measures are Weil heights	17
3.2. Measures defined over a single number field	21
3.3. Equidistribution results	25
4. Stochastic Dynamics	36
4.1. Stochastic dynamical heights	36
4.2. An example with places without continuous potentials	41
4.3. Pullback formula for stochastic heights	43
4.4. Stochastic dynamical equidistribution	44
4.5. Stochastic Julia and Fatou sets	51
References	53

1. INTRODUCTION

Let S be a countable set of rational maps defined over $\overline{\mathbb{Q}}$. If we endow S with a probability measure ν_1 , we think of the pair (S, ν_1) as a *stochastic dynamical system*, where for a map $\varphi \in S$, the quantity $\nu_1(\varphi)$ represents the probability

Date: November 16, 2021.

2010 Mathematics Subject Classification. 11G50, 37P30, 37P50, 37P05.

Key words and phrases. Equidistribution, stochastic dynamical systems, adelic measure.

of applying φ in a random walk. Recently, Healey and Hindes [HH19] defined a stochastic height function and proved that, under certain assumptions, it exists and defines a Weil height.¹ (See also earlier work and an equidistribution theorem for random sequences of morphisms defined over a single number field in Kawaguchi [Kaw07].) We briefly recall their construction: We associate to an element $\gamma_n = (\varphi_1, \dots, \varphi_n) \in S^n$ its natural composition map, and think of it as a function on the projective line $\mathbb{P}^1(\overline{\mathbb{Q}})$, that is, we will write $\gamma_n(\alpha) = \varphi_n \circ \dots \circ \varphi_1(\alpha)$ and take the degree of γ_n to be $\deg(\gamma_n) = \deg(\varphi_1) \cdots \deg(\varphi_n)$. We endow S^n with the natural product measure $\nu_n = \nu_1 \times \dots \times \nu_1$, so that $\nu_n((\varphi_1, \dots, \varphi_n)) = \nu_1(\varphi_1) \cdots \nu_1(\varphi_n)$. Then the *stochastic height* associated to (S, ν_1) , when it exists, is the function $h_S : \mathbb{P}^1(\overline{\mathbb{Q}}) \rightarrow [0, \infty)$ defined by

$$h_S(\alpha) = \lim_{n \rightarrow \infty} \mathbb{E}_{S^n} \frac{1}{\deg(\gamma_n)} h(\gamma_n(\alpha)),$$

where h is the absolute logarithmic Weil height on $\mathbb{P}^1(\overline{\mathbb{Q}})$, and where the expectation is taken over $\gamma_n \in S^n$ with the probability measure ν_n . Notice that if $S = \{\varphi\}$ where φ is rational map of degree at least 2, and $\nu_1(\varphi) = 1$, this reduces to the usual (dynamical) canonical height associated to φ .

It is well-known in arithmetic dynamics that points of low height play a special role: the points of height zero are precisely the preperiodic points of φ , and it is known (see for example [FRL06, BR06, CL06, CLT09, Thu05, Yua08]) that sequences of points of small height equidistribute according to certain canonical measures determined by φ at every place of the number field over which φ is defined. A typical example of this is the sequence of probability measures supported equally on the set of preimages $\varphi^{-n}(\alpha)$ for each $n \in \mathbb{N}$, where α is a point which is not exceptional for the map φ . These preimage sets consist of points of geometrically decreasing height, and equidistribute according to the canonical measure at every place. This canonical measure is naturally supported on the Julia set of the map and thus is also tied to the dynamics of the map.

However, the situation for stochastic dynamical systems is a little bit more complicated: Healey and Hindes prove that while having stochastic height zero does guarantee that the forward orbit under the entire system S is finite, they also prove that there are only finitely many points of height zero, unless the maps all share the same set of preperiodic points, or equivalently, the same Julia set at every place; see [HH19, Corollaries 1.4-5]. (We note in passing that typically two maps that have the same Julia sets share a common iterate, though there are exceptions; see [Ye15].) Of course, with only finitely many points of height zero, it is not immediately clear that there is any equidistribution theorem to prove.

The goal of this paper is to prove that, in fact, a natural stochastic analogue of the equidistribution of preimages is true under certain natural restrictions on the stochastic family of rational maps (S, ν_1) . Our main theorem, proven as Theorem 4.8 below, is the following:

Theorem (Equidistribution of random backwards orbits). *Let S be a countable set of rational maps defined over an algebraic closure $\overline{\mathbb{Q}}$, with each map being of degree at least 2, and let ν_1 be a probability measure on S with respect to which the maps in S are L^1 height controlled. Then for almost every place y of $\overline{\mathbb{Q}}$ there exists a*

¹In fact, they proved their result for endomorphisms of projective varieties. However, in this paper, we will restrict our attention to families of rational maps on the projective line.

certain canonical stochastic dynamical measure ρ_y such that the following is true: Let $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$ be any point which is not in the exceptional set of the stochastic system S . Then the backwards orbit measures $\Delta_{n,\alpha}$ under S converge weakly to ρ_y as $n \rightarrow \infty$.

The precise definitions used above can be found in Section 4 below, but here we give an informal summary of the conditions of the main theorem:

- (1) For each rational map, it is known that the canonical height h_φ and the standard height differ by at most an explicit constant C_φ depending on φ . Then expected value of this constant over the stochastic family is assumed to be finite.
- (2) The backwards orbit measures $\Delta_{n,\alpha}$ are defined recursively by writing each measure $\Delta_{n,\alpha}$ as a countable sum of weighted point masses in $\mathbb{P}^1(\overline{\mathbb{Q}})$ (starting with $\Delta_{0,\alpha} = \delta_\alpha$, the unit point mass at α), then taking all possible preimages of the points in the support of $\Delta_{n,\alpha}$ under the maps in S , weighted according to the probability $\nu_1(\varphi)$.

The assumption that α is not exceptional simply means that the grand orbit of α is infinite.

In order to prove (or even precisely state) our theorem, we need to extend the current framework for talking about heights on the projective line. In order to explain what is new in our framework, let us briefly recall the concept of adelic measures due to Favre and Rivera-Letelier [FRL06] and its relation to height.

The notion of an *adelic measure* was introduced by Favre and Rivera-Letelier [FRL06] in order to generalize the idea of the canonical Weil height associated to the dynamics of rational map, based on its canonical measure at every place. Adelic measures are families of measures, defined over a number field, which are the ‘standard measure’ (associated to the local Weil height) at almost all places of the base field and which differ from the standard measure in a controlled fashion at the remaining places.

This notion was later generalized by Mavraki and Ye [MY] to the notion of a *quasi-adelic measure*, which allowed infinitely many ‘bad’ places where the measure was not the standard measure, provided that a global summability condition was met in order to ensure that the resulting height was still a Weil height. Mavraki and Ye then proved an equidistribution result for quasi-adelic measures. Their generalization was inspired by the study of dynamical heights in parameterized families, where it turned out that many of the resulting height functions failed to be adelic in the original sense of Favre and Rivera-Letelier.

The heights associated to adelic and quasi-adelic measures have one common restriction: they are all defined over a single base global field which satisfies the product formula, for example, a number field. Heights are then computed as averages of Galois orbits of points over this base field. As we wish to study the stochastic dynamics of families of maps which may not be defined over a single base field, we will need a more general framework. The framework for this more general definition of the height was developed by Allcock and Vaaler [AV09] and will allow us to define the notion of a *generalized adelic measure* below.

As we are interested in heights associated to stochastic dynamical systems, we will give a construction that, under the assumptions on the family S of rational maps with probability measure ν_1 given above, allows us to define a generalized

adelic measure ρ associated to S such that $h_S = h_\rho$. We note that there are three basic levels of definition for the generalized adelic measure ρ :

- (1) If S is finite, then ρ is an adelic measure in the sense of Favre and Rivera-Letelier. This is because there is a single number field over which all maps in S are defined, and a finite number of places of bad reduction in total, so the resulting measure for the stochastic dynamical system will also be defined over a single number field and equal to the standard measure at all but finitely many places.
- (2) If S is infinite, but all $\varphi \in S$ are defined over a single number field K , then ρ may not be an adelic measure, as there may be infinitely many places of bad reduction. Our L^1 height control assumption guarantees, however, that the resulting measure is still a quasi-adelic measure in the sense of Mavraki and Ye.
- (3) If S is infinite and the field of definition is of infinite degree, then ρ may not be an adelic measure or quasi-adelic measure, but the L^1 height control assumption guarantees that it is a generalized adelic measure.

In Theorem 3.17, we prove analogues of the main theorems of Favre and Rivera-Letelier [FRL06, Théorèmes 1-2], namely, that the heights associated to generalized adelic measures are essentially-positive Weil heights, and that points of small height equidistribute according to the generalized adelic measure; however, this equidistribution result only holds outside a set of places of measure zero, according to a natural measure on the space of places. This measure is the natural extension of the notion of ‘local degree over global degree’ at a place of a number field, and in particular, the set of places of \mathbb{Q} which lie over a given place of a number field always has positive measure, so that the equidistribution theorem, when applied to a generalized adelic measure defined over a single number field, applies at every place.

Proving results in this context has required several innovative definitions and proofs. For example, heights are no longer defined in terms of average Galois orbits of a point over the base field, and some arguments that relied on this concept have changed. We define our heights more generally for discrete probability measures on $\mathbb{P}^1(\mathbb{Q})$, possibly with infinite support, and we classify which discrete probability measures have finite height. In fact, we prove that finiteness of the height of Δ depends only on Δ and not on which generalized adelic measure is being used (Corollary 3.2).

The structure of this paper is as follows. In Section 2 we define the notion of a generalized adelic measure, introduce the notion of heights for discrete probability measures on $\mathbb{P}^1(\mathbb{Q})$ and introduce the notion for finite height for a discrete probability measure.

In Section 3, we prove the main theorems on the existence of Weil heights associated to generalized adelic measures. We then introduce a notion of what it means for a sequence of discrete probability measures to be *well-distributed* in Definition 3.12. This notion replaces the classical notion of the degree of our points tending to infinity in the equidistribution theorem, which may no longer apply as even one step in a backwards orbit for an infinite family of maps may result in a discrete probability measure concentrated on a countably infinite set. We then prove a local equidistribution theorem for such measures in Proposition 3.16. The global theorem then follows in Theorem 3.17, the main result of this section. We note that

even when our generalized adelic measure is an adelic measure in the original sense of Favre and Rivera-Letelier, these equidistribution theorems are still an expansion of the original equidistribution theorems, as we prove equidistribution for a much wider class of probability measures on $\mathbb{P}^1(\overline{\mathbb{Q}})$.

Finally, in Section 4, we prove that for an L^1 height controlled stochastic family of rational maps, the stochastic height is given by a generalized stochastic measure associated to the family. We prove a stochastic pullback formula for the stochastic dynamical system akin to the classical pullback formula in Theorem 4.6. We note that an analogue of the pushforward formula does not, in general, hold for stochastic families. Finally, we define the notion of a *random backwards orbit measure* associated to a starting point $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$, and we prove the stochastic equidistribution for random backwards orbits of non-exceptional points in Theorem 4.8. We conclude with a definition of the stochastic Julia and Fatou sets associated to the stochastic system.

In a forthcoming paper [DFT], the authors will prove analogues of classical results about the dynamics of the Julia and Fatou sets for stochastic dynamical system.

Acknowledgments. The authors would like to thank Joseph Silverman for helpful suggestions, Xander Faber for suggesting a reference to [Gub97], and Paul Nguyen for computational assistance.

2. GENERALIZED ADELIC MEASURES

2.1. Definitions. Let us start by recalling the basic notation and measures associated to the space of places of $\overline{\mathbb{Q}}$ from [AV09]. This construction makes $\overline{\mathbb{Q}}$ into an M -field in the sense of Gubler [Gub97]. We note that this definition works more generally for a global base field in place of \mathbb{Q} ; however, we will focus on the case of \mathbb{Q} as the prime field here. For a number field K , let M_K denote its set of places. Notice that the collection of sets M_K as K ranges over finite extensions of \mathbb{Q} forms a projective system with the natural maps $M_L \rightarrow M_K$ for $L \supset K$ given by restriction of places. The set of places of $\overline{\mathbb{Q}}$ is then the projective limit

$$Y = \varprojlim_K M_K,$$

where K ranges over finite extensions of \mathbb{Q} . If we endow each M_K with the discrete topology, this induces a natural projective topology on Y . Notice that each $y \in Y$ has an associated absolute value $|\cdot|_y$, and that for any place $v \in M_K$, we have a natural notion of $y \in Y$ extending v or not, so we can still ask if y lies above v , denoted by $y \mid v$ as usual. Further, each $y \mid p$ corresponds to a choice of embedding $\overline{\mathbb{Q}} \hookrightarrow \mathbb{C}_p$, and by making a choice for one embedding at each prime p , the remaining choices are all determined. We shall assume that we have made such a choice for each rational prime p .

Let us denote, for each number field K and place $v \in M_K$, the set

$$Y(K, v) = \{y \in Y : y \mid v\}.$$

Notice that with the subspace topology induced by Y , this set is compact open (in fact, it has the profinite topology, as a projective limit of the finite sets of places above v for each number field containing K). Allcock and Vaaler demonstrated [AV09, §4] that the absolute Galois group acts naturally on each set $Y(\mathbb{Q}, p)$ for

$p \in M_{\mathbb{Q}}$, and that its Haar measure induces a natural Borel probability measure $\mu = \mu|_{Y(\mathbb{Q}, p)}$ such that for each $v \mid p$ a place of K/\mathbb{Q} ,

$$\mu(Y(K, v)) = \frac{[K_v : \mathbb{Q}_p]}{[K : \mathbb{Q}]}.$$

This defines a Borel measure μ on Y given by using this local measure on every (disjoint) set $Y(\mathbb{Q}, p)$. For an algebraic number $\alpha \in K$, the usual absolute logarithmic Weil height can then be expressed in two ways:

$$h(\alpha) = \sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log^+ |\alpha|_v = \int_Y \log^+ |\alpha|_y d\mu(y),$$

where the single bar notation $|\cdot|_y$ for $y \mid p$ agrees with the usual normalization of the absolute value of \mathbb{C}_p . Notice that in the above equation, we are thinking of $f(y) = \log^+ |\alpha|_y$ as a real-valued function of the places Y of $\overline{\mathbb{Q}}$. Indeed, it is easy to see that $f(y)$ is locally constant on the sets $Y(K, v)$ as v ranges over M_K , making it continuous, and further that it is compactly supported, as it is nonzero at only a finite number of places v of K .

Define the *standard measure* at a place v to be

$$(1) \quad \lambda_v = \begin{cases} d\theta/2\pi & \text{if } v \mid \infty, \\ \delta_{\zeta_{0,1}} & \text{if } v \nmid \infty, \end{cases}$$

where $d\theta/2\pi$ is the usual normalized arc-length measure of the unit circle in $\mathbb{A}_v^1 = \mathbb{C}$ if $v \mid \infty$, and $\delta_{\zeta_{0,1}}$ denotes the point mass at the Gauss point $\zeta_{0,1}$ if $\mathbb{A}_v^1 = \mathbb{A}_{\text{Berk}}^1(\mathbb{C}_v)$ if $v \nmid \infty$. Notice that this definition makes sense regardless of whether v is a place of a number field K or a place of $\overline{\mathbb{Q}}$.

Favre and Rivera-Letelier define an *adelic measure* ρ , defined over a base number field K , as a collection of Borel probability measures $\rho = (\rho_v)_{v \in M_K}$ where each ρ_v is a Borel probability measure on \mathbb{P}_v^1 , the Berkovich projective line over \mathbb{C}_v , with the following conditions:

- (1) At each $v \in M_K$, ρ_v admits a continuous potential with respect to the standard measure, that is, there is a continuous function $g_v : \mathbb{P}_v^1 \rightarrow \mathbb{R}$ such that $\Delta g_v = \rho_v - \lambda_v$, where Δ denotes the normalized Laplacian² on the Berkovich projective line \mathbb{P}_v^1 .
- (2) At all but finitely many places $v \in M_K$, we have $\rho_v = \lambda_v$.

We recall that $\mathbb{P}_v^1 = \mathbb{P}^1(\mathbb{C}_v)$ is the Berkovich projective line, a Hausdorff, compact, and uniquely path-connected space in which $\mathbb{P}^1(\mathbb{C}_v)$ is a dense subspace. For more details on Berkovich space, we refer the reader to, for example, [FRL06, BR10, Bak08].

They then define a height function h_ρ associated to any adelic measure ρ , and prove [FRL06, Théorème 1] that the height function h_ρ satisfies:

- (1) $h_\rho(\alpha) = h(\alpha) + O(1)$ where the big- O constant is independent of α and depends only on ρ , and
- (2) h_ρ is essentially nonnegative, in the sense that for any $\epsilon > 0$, the set $\{\alpha \in \overline{K} : h_\rho(\alpha) < -\epsilon\}$ is finite.

²When v is archimedean, $\mathbb{P}_v^1 = \mathbb{P}^1(\mathbb{C})$, and $\Delta = (1/2\pi)(\partial^2/\partial x^2 + \partial^2/\partial y^2)$. For the definition in the case of a non-archimedean v we refer the reader to [FRL06, §4].

We will need to generalize our notion of adelic measure to deal with the fact that, when we define a stochastic height for an infinite family of maps, first, we may no longer have a single number field over which all of the maps are defined, and second, we may have bad reduction at an infinite number of places.

Definition 2.1. For each $y \in Y$ a place of $\overline{\mathbb{Q}}$, let \mathcal{B}_y denote the space of Borel probability measures on the Berkovich projective line \mathbb{P}_y^1 over \mathbb{C}_y . A *generalized adelic measure* ρ is a function $Y \rightarrow \mathcal{B}_y$ that satisfies the following conditions:

- (1) For every rational prime p , there exists a measurable function

$$g : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow [-\infty, \infty]$$

such that, for μ -almost all $\overline{\mathbb{Q}}$ -places $y \in Y(\mathbb{Q}, p)$, the function $g_y(z) = g(y, z)$ is a continuous function $\mathbb{P}^1(\mathbb{C}_p) \rightarrow \mathbb{R}$, normalized so that $g_y(\infty) = 0$, which satisfies $\Delta g_y = \rho_y - \lambda_y$. Note that for any given $y \in Y(\mathbb{Q}, p)$, we identify \mathbb{P}_y^1 with $\mathbb{P}^1(\mathbb{C}_p)$, where the embedding is determined by our choice of place y .

- (2) If $C : Y \rightarrow [0, \infty]$ is the function defined, for the μ -almost everywhere set where g_y defined above is continuous, by

$$C(y) = \sup_{z \in \mathbb{P}_y^1} |g_y(z)|,$$

and $C(y) = \infty$ on the remaining μ -measure zero set of places y , then $C \in L^1(Y)$, that is,

$$\int_Y C(y) d\mu(y) < \infty.$$

It is important to note that a generalized adelic measure may have a nonempty set of places of μ -measure zero where the measure ρ_y does not admit a continuous potential with respect to the standard measure. At these places, equidistribution results may not hold. An example of such a stochastic dynamical system is given in Example 4.5.

Definition 2.2. Let ρ be a generalized adelic measure, and for each rational prime p let $g : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow [-\infty, \infty]$ be as in Definition 2.1. If there exists a number field K such that $g_{y_1} = g_{y_2}$ whenever y_1 and y_2 lie over the same place of K , then we say that ρ is *defined over K* .

As we noted in the introduction, we will show below in Section 3.2 that when a generalized adelic measure is defined over a single number field K , the concept reduces to that of a quasi-adelic measure or an adelic measure, depending on whether the number of places where the measure is not equal to the standard measure is finite or infinite. Our goal in introducing this generalization is to naturally capture appropriate limits of such measures as the base fields of definition K grow. But first, we will define the height associated to a generalized adelic measure and prove some of its properties.

First, for any place $y \in Y$ and Borel probability measures ρ_y, σ_y on \mathbb{P}_y^1 , define (when it exists) the *local energy pairing* to be:

$$(2) \quad (\rho_y, \sigma_y)_y = \iint_{\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y} -\log|z - w|_y d\rho_y(z) d\sigma_y(w),$$

where $A_y^1 = A^1(\mathbb{C}_y)$ denotes the Berkovich affine line over \mathbb{C}_y and $\text{Diag}_y = \{(z, z) : z \in \mathbb{C}_y\}$. (Note that we are only excluding the *classical*, that is, type I, points of the diagonal.)

One key fact which we will have occasion to use several times in this paper is that if ρ_y, σ_y are two probability measures on P_y^1 for some place $y \in Y$ which admit continuous potentials, then

$$(3) \quad (\rho_y - \sigma_y, \rho_y - \sigma_y)_y \geq 0,$$

with equality if and only if $\rho_y = \sigma_y$. Proofs of this fact can be found in [FRL06, Prop. 2.6, Prop. 4.5]; see also [Fil, Theorem 1], which among other things proves that the closely related Arakelov-Zhang pairing is the square of a metric on the space of adelic measures.

It is important that the reader be aware that, whenever y is a non-archimedean place, the kernel $|z - w|_y$ in (2) and throughout this paper must be read as the natural extension to the Berkovich line A_y^1 of the usual distance, which is denoted by $\sup\{z, w\}$ in the article of Favre and Rivera-Letelier [FRL06, §3.3] and as the *Hsia kernel* $\delta(z, w)_\infty$ in the book of Baker and Rumely [BR10, §4]. As the appropriate extension is canonically defined, there is no danger in keeping the ‘classical’ notation, so long as the reader is well aware that this is no longer actually a distance function on P_y^1 ; as an important example, if ζ is the Gauss point of P_y^1 , then $\delta(\zeta, \zeta)_\infty = \sup\{\zeta, \zeta\} = 1 > 0$. (In fact, all non-classical points ζ of the Berkovich line over \mathbb{C}_y have positive Hsia kernel $\delta(\zeta, \zeta)_\infty > 0$.)

For a generalized adelic measure ρ and a discrete probability measure Δ on $\mathbb{P}^1(\overline{\mathbb{Q}})$, we define the *height* of Δ associated to ρ , when it exists, to be:

$$(4) \quad h_\rho(\Delta) = \frac{1}{2} \int_Y (\rho_y - \Delta, \rho_y - \Delta)_y d\mu(y).$$

When $\rho_y = \lambda_y$ is the standard measure for every $y \in Y$, then we will show that this definition extends the notion of the Weil height to discrete probability measures on $\mathbb{P}^1(\overline{\mathbb{Q}})$. In fact, this extends linearly, as we will show in Proposition 3.4 below; however, this fact is not immediate, as we must prove the vanishing of discriminant-type terms over discrete probability measures which may have infinite support, and existence and finiteness of these quantities is nontrivial.

Notice that Δ defines a well-defined discrete probability measure in P_y^1 for every place $y \in Y$, as each y defines an embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}_y = \mathbb{C}_p$ where $y \in Y(\mathbb{Q}, p)$. We note one key difference between our definition and that of Favre and Rivera-Letelier: Favre and Rivera-Letelier define the height as a sum over the places of K where the adelic measure ρ is defined over K , and for any $\alpha \in \overline{\mathbb{Q}}$ for which one wants to compute the height, one first forms the measure:

$$[\alpha]_K = \frac{1}{\#(G_K \cdot \alpha)} \sum_{z \in G_K \cdot \alpha} \delta_z,$$

where $G_K = \text{Gal}(\overline{K}/K)$ denotes the absolute Galois group over K and δ_z denotes the point mass at z . In this fashion, the measure $[\alpha]_K$ is a probability measure on $\overline{\mathbb{Q}}$ which is in fact invariant under G_K , and therefore defines for any single place $v \in M_K$ the same measure for every $y \in Y(K, v)$. They then define $h_\rho(\alpha)$ to be given by what we denote as $h_\rho([\alpha]_K)$. This has the convenience of not requiring them to change the places over which they are summing by using $K(\alpha)$ instead of K , however, as we will be dealing with generalized adelic measures, which are limits of

certain adelic measures that can be defined over increasing towers of number fields, it is neither practical nor helpful to use this trick. So we will define instead, for ρ a generalized adelic measure and for any $\alpha \in \overline{\mathbb{Q}}$,

$$(5) \quad h_\rho(\alpha) = h_\rho(\delta_\alpha),$$

where δ_α is the point mass at α . As we compute the height via (4), we note that the embedding of δ_α as a measure in \mathbf{A}_y^1 will change depending on the place y of $\overline{\mathbb{Q}}$.

We now want to show that h_ρ for generalized adelic measures ρ shares largely the same properties as the heights h_ρ for adelic measures ρ . We start with a definition which will characterize the discrete measures which have well-defined heights.

Definition 2.3. Let Δ be a discrete probability measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$. We say that Δ has finite height if

$$\int_Y -(\lambda_y, \Delta)_y d\mu(y) = \int_Y \int_{\mathbf{A}_y^1} \log^+ |z|_y d\Delta(z) d\mu(y) < \infty.$$

Notice that the integrand is nonnegative, and further, we have explicitly excluded any support at $\infty \in \mathbb{P}^1(\overline{\mathbb{Q}})$ from the integral.

Remark 2.4. We remind the reader that a discrete measure is one which can be written as a weighted sum of at most countably many Dirac measures, however, in any given completion, the support set of the measure need not be a discrete set in the usual topological sense.

As the integrand $\log^+ |z|_y$ in Definition 2.3 is a compactly supported locally constant function when $\Delta = \delta_\alpha$ (locally constant on the sets $Y(\mathbb{Q}(\alpha), v)$, for v a place of $\mathbb{Q}(\alpha)$), this is also true more generally when Δ is finite supported. In these cases the integral becomes a finite sum. It follows that we have the following lemma:

Lemma 2.5. *If Δ is a finitely supported probability measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$, then Δ has finite height.*

2.2. Existence results for finite height measures. Our goal is now to justify the definition of ‘finite height’ by proving:

Proposition 2.6. *The height $h_\rho(\Delta)$ defined in (4) exists and is finite for every generalized adelic measure ρ and every discrete probability measure Δ on $\mathbb{P}^1(\overline{\mathbb{Q}})$ which has finite height.*

Before proving Proposition 2.6, we need some auxiliary results, which we will use later.

Lemma 2.7. *Suppose y is a place of $\overline{\mathbb{Q}}$ and $g : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ is a continuous function such that $g(\infty) = 0$ and $\Delta g = \rho_y - \lambda_y$. Then the potential function*

$$(6) \quad p_{\rho_y}(z) = \int_{\mathbb{P}_y^1} \log |z - w|_y d\rho_y(w)$$

exists and is given by

$$p_{\rho_y}(z) = g(z) + \log^+ |z|_y.$$

Proof. Existence of the potential p_{ρ_y} follows from Lemmas 2.4 and 4.3 of [FRL06] in the archimedean and non-archimedean cases, respectively. To see the equality, observe that

$$\Delta p_{\rho_y}(z) = \rho_y - \delta_\infty \quad \text{and} \quad \Delta \log^+ |z|_y = \lambda_y - \delta_\infty,$$

where δ_∞ denotes the point mass at ∞ . Thus $\Delta(p_{\rho_y}(z) - g(z) - \log^+ |z|_y) = 0$, and so the function $p_{\rho_y}(z) - g(z) - \log^+ |z|_y$ is harmonic on all of \mathbb{P}_y^1 , but this means it is constant. By our choice of normalization $g(\infty) = 0$, it follows by considering the asymptotics as $z \rightarrow \infty$ that $p_{\rho_y}(z) = g(z) + \log^+ |z|_y$, as claimed. \square

Lemma 2.8. *Let Δ be a discrete probability measure on $\overline{\mathbb{Q}}$. Then for every finite place y of $\overline{\mathbb{Q}}$, we have*

$$(\lambda_y - \Delta, \lambda_y - \Delta)_y \geq 0.$$

For every infinite place y of $\overline{\mathbb{Q}}$, we have

$$(\lambda_y - \Delta, \lambda_y - \Delta)_y \geq -\log 2.$$

Proof. Write

$$\Delta = \sum_{i=1}^{\infty} t_i \delta_{\alpha_i},$$

where α_i are distinct points of $\overline{\mathbb{Q}}$, which, under the embedding given by our choice of place y we now think of as points of $\mathbb{P}^1(\mathbb{C}_y)$. Since we assume Δ is a probability measure, we have $t_i \geq 0$ for each i , and $\sum_i t_i = 1$. By the linearity of the pairing,

$$(\lambda_y - \Delta, \lambda_y - \Delta)_y = (\lambda_y, \lambda_y)_y - 2(\lambda_y, \Delta) + (\Delta, \Delta)_y$$

provided the terms on the right-hand side exist. We will show that the terms on the right side must be a sum of nonnegative terms, which is enough to prove our result. First note that $(\lambda_y, \lambda_y)_y = 0$, so we can ignore that term. The second term is

$$-2(\lambda_y, \Delta) = 2 \sum_{i=1}^{\infty} t_i \log^+ |\alpha_i|_y,$$

and

$$(\Delta, \Delta)_y = - \sum_{i \neq j} t_i t_j \log |\alpha_i - \alpha_j|_y.$$

Using $\sum_i t_i = 1$, we write

$$2 \sum_{i=1}^{\infty} t_i \log^+ |\alpha_i|_y = 2 \sum_{i=1}^{\infty} t_i \left(\sum_{j=1}^{\infty} t_j \right) \log^+ |\alpha_i|_y = 2 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} t_i t_j \log^+ |\alpha_i|_y.$$

Now, for any $z, w \in \overline{\mathbb{Q}}$,

$$|z - w|_y \leq \begin{cases} 2 \max\{|z|_y, |w|_y\} & \text{if } y \mid \infty, \\ \max\{|z|_y, |w|_y\} & \text{if } y \nmid \infty, \end{cases}$$

so

$$-\log |z - w|_y \geq \begin{cases} -\log 2 - \log^+ |z|_y - \log^+ |w|_y & \text{if } y \mid \infty, \\ -\log^+ |z|_y - \log^+ |w|_y & \text{if } y \nmid \infty, \end{cases}$$

or more simply, if we let $\chi_\infty(y)$ be the characteristic function of $Y(\mathbb{Q}, \infty)$,

$$(7) \quad \log^+ |z|_y + \log^+ |w|_y - \log |z - w|_y \geq -\chi_\infty(y) \log 2.$$

It follows that

$$\begin{aligned}
-2(\lambda_y, \Delta) + (\Delta, \Delta)_y &= \sum_{i \neq j} t_i t_j (\log^+ |\alpha_i|_y + \log^+ |\alpha_j|_y - \log |\alpha_i - \alpha_j|_y) \\
&\quad + \sum_{i=1}^{\infty} 2t_i^2 \log^+ |\alpha_i|_y \\
&\geq - \sum_{i \neq j} t_i t_j \chi_{\infty}(y) \log 2 \\
&\geq -\chi_{\infty}(y) \log 2,
\end{aligned}$$

and the conclusion follows. \square

Remark 2.9. We note that it is possible to improve the lower bound in the preceding lemma in the archimedean case; in particular, in the case where Δ is equally supported on a finite set, we could apply Mahler's inequality [Mah64] and Baker's sweeping generalization of Mahler's result to lower bounds for averages of Green's functions [Bak06]. However, these results are proven in a more limited context, and as we will not need a stronger result, we will not make any effort to improve the result here with those ideas.

We now prove a generalization of the product formula for the integral of discriminant-type terms over all places of $\overline{\mathbb{Q}}$:

Lemma 2.10. *Suppose that Γ, Δ are discrete probability measures on $\mathbb{P}^1(\overline{\mathbb{Q}})$ which have finite height. Then*

$$\int_Y (\Gamma, \Delta)_y d\mu(y) = 0.$$

Proof. The proof is essentially the product formula applied term by term; the main difficulty lies in justifying the interchange of the sum and the integral in order to apply the product formula to each individual difference of terms which appears. To do this, we will show that there exists a function in $L^1(Y)$ which we can add to the function $(\Gamma, \Delta)_y$ to make it nonnegative as a function of the place y .

We start by writing, as Γ and Δ are discrete probability measures,

$$\Gamma = \sum_{n=1}^{\infty} s_n \delta_{z_n} \quad \text{and} \quad \Delta = \sum_{n=1}^{\infty} t_n \delta_{w_n},$$

where $z_n, w_n \in \overline{\mathbb{Q}}$, δ_z denotes the Dirac point mass at $z \in \overline{\mathbb{Q}}$, and $\sum_n s_n = \sum_n t_n = 1$. By the definition of the energy pairing,

$$(\Gamma, \Delta)_y = \int_{\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y} -\log |z - w|_y d\Gamma(z) d\Delta(w) = \sum_{\substack{m, n \\ z_m \neq w_n}} -s_m t_n \log |z_m - w_n|_y,$$

where $\text{Diag}_y = \{(z, z) \in \mathbb{A}_y^1 \times \mathbb{A}_y^1 : z \in \mathbb{C}_y\}$ denotes the diagonal of classical points. (This makes no difference for an archimedean place y , but for the Berkovich line \mathbb{A}_y^1 over a nonarchimedean place, we are only excluding the diagonal of Type I points.) By the assumption that Γ and Δ have finite height, we have that

$$\int_Y \sum_n s_n \log^+ |z_n|_y d\mu(y) < \infty \quad \text{and} \quad \int_Y \sum_n t_n \log^+ |w_n|_y d\mu(y) < \infty.$$

Let χ_∞ be the characteristic function of $Y(\mathbb{Q}, \infty)$. We define the functions

$$f(y) = \sum_n s_n \log^+ |z_n|_y \quad \text{and} \quad g(y) = \sum_n t_n \log^+ |w_n|_y.$$

Note that $f, g \geq 0$ and $f, g \in L^1(Y)$ by our assumption that Γ and Δ have finite height. We also have $\chi_\infty \in L^1(Y)$, as χ_∞ is a constant function on its support, which has finite measure $\mu(Y(\mathbb{Q}, \infty)) = 1$. Notice that

$$\begin{aligned} & \chi_\infty(y) \log 2 + f(y) + g(y) + (\Gamma, \Delta)_y \\ &= \chi_\infty(y) \log 2 + \sum_m s_m \log^+ |z_m|_y \\ & \quad + \sum_n t_n \log^+ |w_n|_y - \sum_{\substack{m,n \\ z_m \neq w_n}} s_m t_n \log |z_m - w_n|_y \\ &= \sum_{m,n} s_m t_n \chi_\infty(y) \log 2 + \sum_{m,n} s_m t_n \log^+ |z_m|_y \\ & \quad + \sum_{m,n} s_m t_n \log^+ |w_n|_y - \sum_{\substack{m,n \\ z_m \neq w_n}} s_m t_n \log |z_m - w_n|_y, \end{aligned}$$

where we have used the fact that $\sum_m s_m = \sum_n t_n = 1$ above to insert the extra copies of s_m and t_n . Thus, applying Lemma 2.8, we have

$$\begin{aligned} & \chi_\infty(y) \log 2 + f(y) + g(y) + (\Gamma, \Delta)_y \\ &= \sum_{\substack{m,n \\ z_m = w_n}} s_m t_n (2 \log^+ |z_m|_y + \chi_\infty(y) \log 2) \\ & \quad + \sum_{\substack{m,n \\ z_m \neq w_n}} s_m t_n (\chi_\infty(y) \log 2 + \log^+ |z_m|_y + \log^+ |w_n|_y - \log |z_m - w_n|_y) \\ & \geq 0, \end{aligned}$$

where nonnegativity is clear for the first term and follows from (7) for the second term. As we have added L^1 -functions and the resulting function is nonnegative, and since our measure spaces are σ -finite, we can apply Tonelli's theorem to justify the interchange of summation and integration in the original integral we wished to compute:

$$\begin{aligned} (8) \quad & \int_Y (\Gamma, \Delta)_y d\mu(y) = \int_Y - \sum_{\substack{m,n \\ z_m \neq w_n}} s_m t_n \log |z_m - w_n|_y d\mu(y) \\ &= - \sum_{\substack{m,n \\ z_m \neq w_n}} s_m t_n \int_Y \log |z_m - w_n|_y d\mu(y). \end{aligned}$$

Notice however that for every $\alpha \in \overline{\mathbb{Q}}$, if we let $K = \mathbb{Q}(\alpha)$, then

$$\int_Y \log |\alpha|_y d\mu(y) = \sum_{v \in M_K} \mu(Y(K, v)) \log |\alpha|_v = \sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log |\alpha|_v = 0$$

by the product formula. It follows that

$$\int_Y \log |z_m - w_n|_y d\mu(y) = 0$$

for every m, n with $z_m \neq w_n$, and our proof is complete. \square

Using Lemma 2.10 we can now prove a result that justifies the use of the term ‘finite height’ in Definition 2.3:

Proposition 2.11. *Let Δ be a discrete probability measure on $\overline{\mathbb{Q}}$. Then Δ has finite height in the sense of Definition 2.3 if and only if $h(\Delta)$ exists and is finite.*

Proof. First, assume that Δ has finite height. Then by the bilinearity of the energy pairing,

$$h(\Delta) = \frac{1}{2} \int_Y (\Delta - \lambda_y, \Delta - \lambda_y)_y d\mu(y) = \frac{1}{2} \int_Y (\lambda_y, \lambda_y)_y - 2(\Delta, \lambda_y)_y + (\Delta, \Delta)_y d\mu(y),$$

presuming these quantities exist. But $(\lambda_y, \lambda_y)_y = 0$ for all y ,

$$-2 \int_Y (\Delta, \lambda_y)_y d\mu(y) = 2 \int_Y \int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta(z) d\mu(y) < \infty$$

by the assumption that Δ has finite height, and by Lemma 2.10,

$$\int_Y (\Delta, \Delta)_y d\mu(y) = 0.$$

For the other direction, we now suppose that $h(\Delta) < \infty$. If Δ is finitely supported, the result is trivial, so assume that Δ has countably infinite support. Write

$$\Delta = \sum_{i=1}^{\infty} t_i \delta_{\alpha_i}.$$

Then by Lemma 2.8, we have

$$(\Delta - \lambda_y, \Delta - \lambda_y)_y \geq -\chi_{\infty}(y) \log 2,$$

where χ_{∞} is the characteristic function of $Y(\mathbb{Q}, \infty)$. Since $\chi_{\infty} \in L^1(Y)$, Tonelli’s theorem applies to the integral defining the height $h(\Delta)$, and we can exchange the order of integration, and in particular, as we noted in the proof of Lemma 2.10, the product formula applied to the discriminant-type term yields

$$\int_Y \log |\alpha_i - \alpha_j|_y d\mu(y) = 0.$$

It follows, since $h(\Delta) < \infty$, that

$$(9) \quad h(\Delta) = \frac{1}{2} \int_Y (\Delta - \lambda_y, \Delta - \lambda_y)_y d\mu(y) = \int_Y \int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta(z) d\mu(y) < \infty,$$

and so Δ meets the finite height criterion. \square

It will follow from Proposition 2.11 and Theorem 3.1 below, which shows that for every generalized adelic measure ρ , $h_{\rho} = h + O(1)$ with big- O constant depending only on ρ , that the result of Proposition 2.11 is in fact true for all heights h_{ρ} associated to a generalized adelic measure.

Lemma 2.12. *Suppose y is a place of $\overline{\mathbb{Q}}$ and ρ_y is a Borel probability measure on \mathbb{P}_y^1 for which there exists a continuous potential $g_y : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ such that $g_y(\infty) = 0$ and that, with respect to the standard measure λ_y , we have $\Delta g_y = \rho_y - \lambda_y$. Then*

$$-2 \sup_{z \in \mathbb{P}_y^1} |g_y(z)| \leq (\rho_y, \rho_y)_y \leq 4 \sup_{z \in \mathbb{P}_y^1} |g_y(z)|.$$

In particular, for an adelic measure ρ , $|(\rho_y, \rho_y)_y| \leq 4C(y)$ for μ -almost every $y \in Y$.

Proof. Let

$$C(y) = \sup_{z \in \mathbb{P}_y^1} |g_y(z)|,$$

as in the definition of an adelic measure. It follows from [FRL06, Prop. 2.6] when y is archimedean and [FRL06, Prop 4.5] when y is non-archimedean that

$$0 \leq (\rho_y - \lambda_y, \rho_y - \lambda_y)_y = \int_{\mathbb{P}_y^1} -g_y(z) d(\rho_y - \lambda_y)(z).$$

It follows that

$$(10) \quad 0 \leq (\rho_y - \lambda_y, \rho_y - \lambda_y)_y \leq |\rho_y - \lambda_y|(\mathbb{P}_y^1) \cdot \sup_{z \in \mathbb{P}_y^1} |g_y(z)| \leq 2C(y).$$

By the linearity and symmetry of the energy pairing,

$$(\rho_y - \lambda_y, \rho_y - \lambda_y)_y = (\rho_y, \rho_y)_y - 2(\rho_y, \lambda_y)_y + (\lambda_y, \lambda_y)_y.$$

As before, we have $(\lambda_y, \lambda_y)_y = 0$. If we can bound the mixed term $(\rho_y, \lambda_y)_y$, our result will follow. By Lemma 2.7, the potential function

$$p_{\rho_y}(z) = \int_{\mathbb{P}_y^1} \log |z - w| d\rho_y(w)$$

of ρ_y is $p_{\rho_y}(z) = g_y(z) + \log^+ |z|_y$. We now apply Lemmas 2.4 and 2.5 of [FRL06] to say that

$$(11) \quad |(\rho_y, \lambda_y)| \leq \int_{\mathbb{P}_y^1} |p_{\rho_y}(z)| d\lambda_y(z) \leq \int_{\mathbb{P}_y^1} (\log^+ |z|_y + C(y)) d\lambda_y(z) = C(y).$$

Combining equations (10) and (11) the bounds follow. \square

We are now ready to prove Proposition 2.6, which states the $h_\rho(\Delta)$ exists and is finite for every generalized adelic measure ρ and every discrete probability measure Δ on $\mathbb{P}^1(\overline{\mathbb{Q}})$ which has finite height in the sense of Definition 2.3.

Proof of Proposition 2.6. By the linearity of the energy pairing, write:

$$(12) \quad \begin{aligned} h_\rho(\Delta) &= \frac{1}{2} \int_Y (\rho_y - \Delta, \rho_y - \Delta)_y d\mu(y) \\ &= \frac{1}{2} \int_Y (\rho_y, \rho_y) - 2(\Delta, \rho_y) + (\Delta, \Delta)_y d\mu(y). \end{aligned}$$

We will prove first that each of the three integrands in (12) above is an integrable function of the place y , and then that each of the three terms above is bounded above and below by an integrable function on Y , so that the integral exists, as claimed.

We start with the first term. Measurability of (ρ_y, ρ_y) as a function of y follows from the measurability of the function $g : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow \mathbb{R}$ for every rational prime p . To see that the first term has finite integral, we use Lemma 2.12 to say that

$$\left| \frac{1}{2} \int_Y (\rho_y, \rho_y)_y d\mu(y) \right| \leq \frac{1}{2} \int_Y 4C(y) d\mu(y) < \infty$$

by our assumption that $C(y) \in L^1(Y)$.

Consider now the second term in the integral,

$$\int_Y -(\rho_y, \Delta)_y d\mu(y) = \int_Y \int_{\mathbb{A}_y^1} p_{\rho_y}(z) d\Delta(z) d\mu(y),$$

where p_{ρ_y} denotes the potential function of ρ_y given by

$$p_{\rho_y}(z) = \int_{\mathbb{A}_y^1} \log|z - w|_y d\rho_y(w).$$

Since each point in the support of Δ is a single algebraic number, its y -adic absolute value defines a locally constant function of y , so measurability is trivial. Here we make use of the assumption that Δ has finite height, which gives

$$0 \leq \int_Y \int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta(z) d\mu(y) < \infty,$$

and now we need to bound the integral with $p_{\rho_y}(z)$ in place of $\log^+ |z|_y$. By Lemma 2.7, we know that

$$p_{\rho_y}(z) = g_y(z) + \log^+ |z|_y,$$

so the difference between the two integrals is bounded by

$$(13) \quad \int_Y \int_{\mathbb{A}_y^1} |g_y(z)| d\Delta(z) d\mu(y) \leq \int_Y \sup_{z \in \mathbb{P}_y^1} |g_y(z)| d\mu(y) = \int_Y C(y) d\mu(y) < \infty$$

as $C(y) \in L^1(Y)$.

The third term in the integrand, namely

$$\frac{1}{2} \int_Y (\Delta, \Delta)_y d\mu(y),$$

is equal to 0 by Lemma 2.10, since Δ is a discrete probability measure of finite height. We have shown that all three terms are integrable, so we are done. \square

We note that, in the case that ρ is defined over a single number field K/\mathbb{Q} , then we have the usual Galois invariance of the height:

Proposition 2.13. *Suppose that ρ is defined over a single number field K/\mathbb{Q} . Then the height h_ρ is $\text{Gal}(\overline{K}/K)$ -invariant in the sense that if we define $\sigma\Delta$ by linearly extending the action $\sigma(\delta_\alpha) = \delta_{\sigma\alpha}$ for all $\alpha \in \overline{\mathbb{Q}}$, then*

$$(14) \quad h_\rho(\sigma\Delta) = h_\rho(\Delta) \quad \text{for all } \sigma \in \text{Gal}(\overline{K}/K).$$

In particular, if $[\alpha]_K$ denotes the probability measure on $\overline{\mathbb{Q}}$ equally distributed on each K -Galois conjugate of α , then

$$(15) \quad h_\rho(\alpha) = h_\rho([\alpha]_K).$$

We remind the reader that the action of the absolute Galois group on the space of places Y is defined by $|\alpha|_{\sigma y} = |\sigma^{-1}(\alpha)|_y$, so that the action on Y is compatible with the action on $\overline{\mathbb{Q}}$; we use the inverse to ensure that

$$|\alpha|_{\tau\sigma y} = |\tau^{-1}(\alpha)|_{\sigma y} = |\sigma^{-1}\tau^{-1}(\alpha)|_y = |(\tau\sigma)^{-1}(\alpha)|_y.$$

Proof. By the linearity of the energy pairing, we have

$$\begin{aligned} h_\rho(\sigma\Delta) &= \frac{1}{2} \int_Y (\rho_y - \sigma\Delta, \rho_y - \sigma\Delta)_y d\mu(y) \\ &= \frac{1}{2} \int_Y (\rho_y, \rho_y) - 2(\sigma\Delta, \rho_y) + (\sigma\Delta, \sigma\Delta)_y d\mu(y). \end{aligned}$$

By Lemma 2.10, $\int_Y (\sigma\Delta, \sigma\Delta)_y d\mu(y) = 0$ so it suffices to show that

$$\int_Y (\sigma\Delta, \rho_y)_y d\mu(y) = \int_Y (\Delta, \rho_y)_y d\mu(y).$$

Writing $\Delta = \sum_i t_i \delta_{\alpha_i}$ where $\sum_i t_i = 1$, and using the fact that ρ is defined over K , we have

$$\begin{aligned} \int_Y (\sigma\Delta, \rho_y)_y d\mu(y) &= \int_{Y_K} \int_{\mathbb{A}^1 \times \mathbb{A}^1 \setminus \text{Diag}_y} -\log |z - w|_y d\sigma\Delta(z) d\rho_y(w) d\mu(y) \\ &= \int_{Y_K} \int_{\mathbb{A}_y^1} -p_{\rho_y}(z) \sigma\Delta(z) d\mu(y) \\ &= \int_{Y_K} \sum_i -t_i p_{\rho_y}(\sigma\alpha_i) d\mu(y). \end{aligned}$$

Since $p_{\rho_y}(\sigma\alpha_i) = p_{\rho_{\sigma^{-1}y}}(\alpha_i)$, then we have that

$$\begin{aligned} \int_Y (\sigma\Delta, \rho_y)_y d\mu(y) &= \int_{Y_K} \sum_i t_i p_{\rho_y}(\alpha_i) d\mu(y) \\ &= \int_Y (\Delta, \rho_y)_y d\mu(y), \end{aligned}$$

as desired. \square

We conclude this section with a lemma which will be needed to justify applications of Tonelli's theorem below:

Lemma 2.14. *Let Δ be a discrete probability measure on $\overline{\mathbb{Q}}$ and ρ a generalized adelic measure. Let $C(y)$ be the associated local bound on the Green's function associated to ρ at the place y . Then*

$$(\rho_y - \Delta, \rho_y - \Delta)_y \geq -4C(y) - \chi_\infty(y) \log 2,$$

where $\chi_\infty(y)$ denotes the characteristic function of $Y(\mathbb{Q}, \infty)$.

Proof. Our goal will be to give a lower bound for

$$(16) \quad (\rho_y - \Delta, \rho_y - \Delta)_y - (\lambda_y - \Delta, \lambda_y - \Delta)_y = (\rho_y, \rho_y)_y - 2(\rho_y - \lambda_y, \Delta)_y$$

and use Lemma 2.8 to bound the pairing of Δ with λ_y . Now, Lemma 2.12 gives that

$$(17) \quad (\rho_y, \rho_y)_y \geq -2C(y),$$

and by Lemmas 2.5 and 4.4 of [FRL06],

$$(18) \quad -2(\rho_y - \lambda_y, \Delta)_y = 2 \int_{\mathbb{A}_y^1} g_y(z) d\Delta(z),$$

so

$$(19) \quad |-2(\rho_y - \lambda_y, \Delta)_y| \leq 2 \sup_{z \in \mathbb{A}_y^1} |g_y(z)| = 2C(y).$$

Now, Lemma 2.8 yields

$$(\lambda_y - \Delta, \lambda_y - \Delta)_y \geq -\chi_\infty(y) \log 2,$$

so combining this with the results of (17) and (19) in equation (16) gives the result. \square

3. EQUIDISTRIBUTION FOR GENERALIZED ADELIC MEASURES

3.1. Heights associated to generalized adelic measures are Weil heights.

We now extend [FRL06, Théorème 1] to generalized adelic measures:

Theorem 3.1. *Let ρ be a generalized adelic measure. Then:*

- (1) $h_\rho(\Delta) = h(\Delta) + O(1)$ for every discrete probability measure Δ on $\mathbb{P}^1(\overline{\mathbb{Q}})$ with finite height, where the big- O constant is independent of Δ and depends only on ρ , and
- (2) h_ρ is essentially nonnegative, in the sense that for any $\epsilon > 0$, the set

$$\{\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}) : h_\rho(\alpha) < -\epsilon\}$$

is finite.

Proof. We start with the proof of the first statement. This is quite similar in spirit to the proof of boundedness in Proposition 2.6 above. We observe that:

$$h_\rho(\Delta) = \frac{1}{2} \int_Y (\rho_y - \Delta, \rho_y - \Delta)_y d\mu(y)$$

and

$$h(\Delta) = \frac{1}{2} \int_Y (\lambda_y - \Delta, \lambda_y - \Delta)_y d\mu(y).$$

Further, by Proposition 2.11, $h(\Delta)$ exists and is finite by our assumption on Δ . Since $(\lambda_y, \lambda_y)_y = 0$ at every place,

$$(20) \quad |h_\rho(\Delta) - h(\Delta)| \leq \left| \frac{1}{2} \int_Y (\rho_y, \rho_y)_y d\mu(y) \right| + \left| \int_Y (\rho_y - \lambda_y, \Delta)_y d\mu(y) \right|.$$

Now, by our assumption that ρ is a generalized adelic measure, we know that there exists a measurable function $g : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow \mathbb{R}$ at every rational prime p such that, if $g_y(z) = g(y, z)$, then for μ -almost all y , we have that g_y is continuous and $\Delta g_y(z) = \rho_y - \lambda_y$, and further, the function

$$C(y) = \sup_{z \in \mathbb{P}_y^1} |g_y(z)|$$

satisfies $C \in L^1(Y)$. Now, by Lemmas 2.5 and 4.4 of [FRL06],

$$(\rho_y - \lambda_y, \Delta)_y = \int_{\mathbb{A}_y^1} g_y(z) d\Delta(z),$$

so it follows that $|(\rho_y - \lambda_y, \Delta)_y| \leq C(y)$, independent of Δ . It follows that $h = h_\rho + O(1)$, and indeed, using (20) and applying Lemma 2.12 to bound the first term in (20), we can explicitly compute a big- O constant, which is independent of Δ :

$$(21) \quad |h_\rho(\Delta) - h(\Delta)| \leq 3 \int_Y C(y) d\mu(y).$$

In order to prove our height function is essentially nonnegative, assume for the sake of contradiction that, for some $\epsilon > 0$, the set

$$\{\alpha \in \overline{\mathbb{Q}} : h_\rho(\alpha) < -\epsilon\}$$

is instead infinite. Let $(\alpha_n)_{n=1}^\infty$ be sequence of distinct elements in this set and let Δ_N be the sequence of discrete probability measures on $\overline{\mathbb{Q}}$ given by

$$\Delta_N = \frac{1}{N} \sum_{n=1}^N \delta_{\alpha_n}.$$

As the sums involved in computing the height are finite, it follows from applying the product formula to cancel the discriminant terms that

$$h_\rho(\Delta_N) = \frac{1}{N} \sum_{n=1}^N h_\rho(\alpha_n),$$

and so $h_\rho(\Delta_N) < -\epsilon$ as well. However, for every place y of $\overline{\mathbb{Q}}$, it follows from applying Propositions 2.8 and 4.9 of [FRL06] with $\epsilon = 1/N$ that

$$(22) \quad \liminf_{N \rightarrow \infty} (\rho_y - \Delta_N, \rho_y - \Delta_N)_y \geq 0.$$

Now notice that by Lemma 2.14, the functions

$$f_n(y) = (\rho_y - \Delta_N, \rho_y - \Delta_N)_y$$

are almost nonnegative in the sense that $f_n(y) \geq -4C(y) + \chi_\infty(y) \log 2$, where χ_∞ denotes the characteristic function of $Y(\mathbb{Q}, \infty)$. As C and χ_∞ are both in $L^1(Y)$, it follows that Fatou's lemma applies to the function $f_n(y)$ and we can say that

$$(23) \quad \begin{aligned} 0 &\leq \frac{1}{2} \int_Y \liminf_{N \rightarrow \infty} (\rho_y - \Delta_N, \rho_y - \Delta_N)_y d\mu(y) \\ &\leq \liminf_{N \rightarrow \infty} \frac{1}{2} \int_Y (\rho_y - \Delta_N, \rho_y - \Delta_N)_y d\mu(y) \\ &= \liminf_{N \rightarrow \infty} h_\rho(\Delta_N) \leq -\epsilon < 0, \end{aligned}$$

which is a contradiction. It follows that our height is essentially nonnegative. \square

We can now state the corollary of Proposition 2.11 above, which we stated would follow from proving Theorem 3.1:

Corollary 3.2. *Let Δ be a discrete probability measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$ and ρ be a generalized adelic measure. Then Δ has finite height in the sense of Definition 2.3 if and only if $h_\rho(\Delta)$ exists and is finite.*

Remark 3.3. Note that Corollary 3.2 means that if $h_\rho(\Delta)$ exists and is finite for one generalized adelic measure, then it exists and is finite for all generalized adelic measures.

The reader may wonder what use we have for allowing arbitrary discrete probability measures on $\overline{\mathbb{Q}}$ rather than looking at point masses on algebraic numbers, or finite probability measures equally supported on the Galois conjugates of an algebraic number of a base field – especially given that we have gone to some trouble in order to define heights associated to this more general class of probability measures. There are several reasons for this, one being that the equidistribution theorem for stochastic dynamical systems, which we will prove below, naturally is phrased in terms of random backwards orbits of a point under the stochastic system, and when we have infinitely many maps in this system, our backwards orbits measures will generally have infinite support.

However, there are other reasons why we may wish to do this. One can be observed in the proof of the essential nonnegativity of the height associated to a generalized adelic measure, where, rather than taking Galois orbits over a base number field (which may not exist for a generalized adelic measure!) and applying Northcott's theorem, we proved essential nonnegativity by forming discrete measures that averaged over the algebraic numbers of negative height. The fact that this

proof avoided invoking Northcott's theorem seems to indicate it has the potential for broader applications.

Yet another reason that one might be interested in the heights of discrete measures on $\mathbb{P}^1(\overline{\mathbb{Q}})$, including the Weil height, is that they behave well in forming averages. The following proposition, which demonstrates this fact, will be a key result in proving our results, and is of, we believe, independent interest:

Proposition 3.4. *Let ρ be a generalized adelic measure and suppose that we are given a (possibly finite) sequence of discrete probability measures $\Delta_1, \Delta_2, \dots$ on $\mathbb{P}^1(\overline{\mathbb{Q}})$, each of finite height, and real numbers $t_1, t_2, \dots \geq 0$ with $\sum_n t_n = 1$. Then*

$$h_\rho\left(\sum_n t_n \Delta_n\right) = \sum_n t_n h_\rho(\Delta_n).$$

Remark 3.5. We note that in Proposition 3.4 above, we did not need to assume that our measures Δ_n above have finite height. However, if $h_\rho(\Delta_n) = \infty$ for some n and $t_n > 0$, it will follow that $h_\rho(\sum_n t_n \Delta_n) = \infty$ as well, as we will demonstrate that the integrands are all bounded below for each place y by an $L^1(Y)$ function.

Proof of Proposition 3.4. We recall that by (4),

$$\begin{aligned} h_\rho(\Delta_n) &= \frac{1}{2} \int_Y (\rho_y - \Delta_n, \rho_y - \Delta_n)_y d\mu(y) \\ &= \int_Y \frac{1}{2} (\rho_y, \rho_y)_y - (\Delta_n, \rho_y)_y + \frac{1}{2} (\Delta_n, \Delta_n)_y d\mu(y) \end{aligned}$$

for each $n \geq 1$. The key observation is that, by Lemma 2.14 above, our integrand is within an L^1 function of being nonnegative, so we will be able to apply Tonelli's theorem to bring the integrals over all places into the sums involved. This will eliminate the discriminant-type terms – the only terms in the height which do not *prima facie* appear to linearly average – and allow us to derive the result. This is less trivial than it may seem at first glance, as we are *not* assuming that all of the heights are finite.

We start by showing that, regardless of whether the height is finite or infinite, we can disregard the discriminant term in its computation. Let Δ be a given discrete probability measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$. As the result is trivial if Δ is supported on a finite set, we assume that it does not. Write Δ as

$$\Delta = \sum_{i=1}^{\infty} t_i \delta_{\alpha_i},$$

where $\alpha_i \in \mathbb{P}^1(\overline{\mathbb{Q}})$ are distinct, and let

$$\frac{1}{m} = \int_{\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y} 1 d\Delta(z) d\Delta(w),$$

where again $\text{Diag}_y = \{(z, z) \in \mathbb{A}_y^1 \times \mathbb{A}_y^1 : z \in \mathbb{C}_y\}$ is the diagonal of classical points. Notice that, since Δ is not a point mass as we assumed it is not supported on a finite set, $0 < m < 1$. Let $G(z, w)$ be the function given by

$$G(z, w) = m(p_{\rho_y}(z) + p_{\rho_y}(w)) - \log |z - w|_y + m(\rho_y, \rho_y)_y.$$

Notice that G is a slightly different normalization of the usual Arakelov-Green's function for ρ_y , where $p_{\rho_y}(z)$ is the associated potential function. Note that as Δ is supported on $\mathbb{P}^1(\overline{\mathbb{Q}})$, instead of our usual local integral over $\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y$,

we will integrate over $\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D$ where $D = \{(\alpha, \alpha) : \alpha \in \overline{\mathbb{Q}}\}$. with this normalization,

$$\int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} G(z, w) d\Delta(z) d\Delta(w) = (\rho_y - \Delta, \rho_y - \Delta)_y,$$

so

$$h_\rho(\Delta) = \frac{1}{2} \int_Y \int_{\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y} G(z, w) d\Delta(z) d\Delta(w) d\mu(y).$$

Now, by Lemma 2.14,

$$\begin{aligned} & (\rho_y - \Delta, \rho_y - \Delta)_y + 4C(y) + \chi_\infty(y) \log 2 \\ &= \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} G(z, w) d\Delta(z) d\Delta(w) + 4C(y) + \chi_\infty(y) \log 2 \\ &= \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} G(z, w) + m(4C(y) + \chi_\infty(y) \log 2) d\Delta(z) d\Delta(w) \\ &\geq 0, \end{aligned}$$

and since $4C(y) + \chi_\infty(y) \log 2$ is a nonnegative integrand with

$$\int_Y \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} (4C(y) + \chi_\infty(y) \log 2) d\Delta(z) d\Delta(w) d\mu(y) < \infty,$$

we can apply Tonelli's theorem to the integrals defining $h_\rho(\Delta)$ and interchange the order of integration:

$$h_\rho(\Delta) = \frac{1}{2} \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} \int_Y G(z, w) d\mu(y) d\Delta(z) d\Delta(w).$$

However, for any fixed $z \neq w$ in the support of Δ , we clearly have

$$\int_Y G(z, w) d\mu(y) = \int_Y m(p_{\rho_y}(z) + p_{\rho_y}(w)) - \log |z - w|_y + m(\rho_y, \rho_y)_y d\mu(y).$$

We write the integrals

$$\int_Y p_{\rho_y}(z) d\mu(y) = - \int_Y (\rho_y, \delta_z)_y d\mu(y), \quad \int_Y p_{\rho_y}(w) d\mu(y) = - \int_Y (\rho_y, \delta_w)_y d\mu(y),$$

and note that the integral of the third term is

$$\int_Y -\log |z - w|_y d\mu(y) = 0$$

by the product formula, as $z - w \in \overline{\mathbb{Q}}$ is a single algebraic number by our assumptions on Δ . It follows that

(24)

$$\begin{aligned} h_\rho(\Delta) &= \frac{1}{2} \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} \int_Y G(z, w) d\mu(y) d\Delta(z) d\Delta(w) \\ &= \frac{1}{2} \int_{\mathbb{A}^1(\overline{\mathbb{Q}}) \times \mathbb{A}^1(\overline{\mathbb{Q}}) \setminus D} m \int_Y (\rho_y, \rho_y)_y - (\rho_y, \delta_z)_y - (\rho_y, \delta_w)_y d\mu(y) d\Delta(z) d\Delta(w). \end{aligned}$$

Now again, we wish to show that we can add a nonnegative integrable function and make the integrand nonnegative, so we can apply Tonelli's theorem. By Lemma 2.12,

$$(\rho_y, \rho_y)_y \geq -2C(y),$$

and, as we noted in the proof of Lemma 2.14 above,

$$-(\rho_y - \lambda_y, \delta_z)_y = g_y(z).$$

By Lemma 2.7, $g_y(z) = p_{\rho_y}(z) - \log^+ |z|_y$, so the bound

$$\sup_{z \in \mathbb{A}_y^1} |g_y(z)| = C(y)$$

yields that

$$-(\rho_y, \delta_z)_y \geq -C(y).$$

Thus we can apply Tonelli's theorem and interchange the integrals once more:

(25)

$$\begin{aligned} h_\rho(\Delta) &= \frac{m}{2} \int_Y \int_{\mathbb{A}_y^1 \times \mathbb{A}_y^1 \setminus \text{Diag}_y} (\rho_y, \rho_y)_y - (\rho_y, \delta_z)_y - (\rho_y, \delta_w)_y d\Delta(z) d\Delta(w) d\mu(y) \\ &= \frac{1}{2} \int_Y (\rho_y, \rho_y) - 2(\rho_y, \Delta)_y d\mu(y). \end{aligned}$$

Using this, we see that

$$\begin{aligned} \sum_n t_n h_\rho(\Delta_n) &= \sum_n t_n \frac{1}{2} \int_Y (\rho_y, \rho_y) - 2(\rho_y, \Delta_n)_y d\mu(y) \\ &= \sum_n \frac{1}{2} \int_Y t_n (\rho_y, \rho_y) - 2t_n (\rho_y, \Delta_n)_y d\mu(y) \\ &= \sum_n \frac{1}{2} \int_Y t_n (\rho_y, \rho_y) - 2(\rho_y, t_n \Delta_n)_y d\mu(y). \end{aligned}$$

Again, using the same bounds as above,

$$(\rho_y, \rho_y) - 2(\rho_y, t_n \Delta_n)_y \geq -6C(y),$$

so again, Tonelli's theorem applies and we can interchange the sums and use the linearity of the pairing to say that

$$\begin{aligned} \sum_n t_n h_\rho(\Delta_n) &= \sum_n \frac{1}{2} \int_Y t_n (\rho_y, \rho_y) - 2(\rho_y, t_n \Delta_n)_y d\mu(y) \\ &= \frac{1}{2} \int_Y \sum_n t_n (\rho_y, \rho_y) - 2 \sum_n (\rho_y, t_n \Delta_n)_y d\mu(y) \\ &= \frac{1}{2} \int_Y (\rho_y, \rho_y) - 2 \left(\rho_y, \sum_n t_n \Delta_n \right)_y d\mu(y) = h_\rho \left(\sum_n t_n \Delta_n \right), \end{aligned}$$

which is what we wanted to prove. \square

3.2. Measures defined over a single number field. In this section we will compare our notion of a generalized adelic measure to the earlier notions of an *adelic measure* in Favre and Rivera-Letelier [FRL06] and of a *quasi-adelic measure* in [MY]. The main focus of this section is proving that, when our generalized adelic measure is defined over a single number field, the finiteness conditions given in Mavraki and Ye are in fact equivalent to those here. This section can be safely skipped if the reader is interested only in the new results of this paper.

First, observe that if K is a number field and $\rho = (\rho_v)_{v \in M_K}$ is an adelic measure in the sense of Favre and Rivera-Letelier, then ρ extends to a generalized adelic measure defined over K . Indeed, for a place $y \in Y$, if we let $v \in M_K$ be the place

such that $y \mid v$, we can define $\rho_y = \rho_v$ and $g_y = g_v$. For places $v \in M_K$ lying above $p \in M_{\mathbb{Q}}$, the functions $g_v(z) = g(v, z)$ are continuous on $\mathbb{P}^1(\mathbb{C}_p)$, normalized so that $g_v(\infty) = 0$, and they satisfy $\Delta g_v = \rho_v - \lambda_v$, hence the same is true for g_y for all $y \in Y(\mathbb{Q}, p)$. Moreover, since ρ_v must be the standard measure for all but finitely many places $v \in M_K$, we have $C(v) := \sup_{z \in \mathbb{P}^1(\mathbb{C}_v)} |g_v(z)| = 0$ for all but finitely many places $v \in M_K$, hence the integral $\int_Y C(y) d\mu(y)$ is really the finite sum $\sum_{v \in M_K} C(v)$, which yields a finite value. Thus, in the case that our generalized adelic measure is defined over a number field K and only differs from the standard measure at a finite number of places, our measure can be identified with an adelic measure.

Now suppose that our measure is defined over a single number field, but perhaps has an infinite number of places where it is not the standard measure. We wish to show that it defines a quasi-adelic measure in the sense of Mavraki and Ye. Let us recall their construction. Given measures $(\rho_v)_{v \in M_K}$, they define the *inner* and *outer radii* in the following fashion. First, they assume that each ρ_v admits a continuous potential, that is, that there is a continuous function $g : \mathbb{P}_v^1 \rightarrow \mathbb{R}$ such that

$$\Delta g = \rho_v - \lambda_v.$$

(We note in passing that their normalized Laplacian has the opposite sign as ours, so this equation is the reverse of what appears in their paper.) As any two choices of g differ only by a constant, g is uniquely determined by its value at ∞ . Define

$$p_v(z) = g(z) + \log^+ |z|_v.$$

They normalize so that if we define $G : \mathbb{C}_v^2 \rightarrow \mathbb{R} \cup \{-\infty\}$ by

$$G(z, w) = \begin{cases} p_v(z/w) + \log |w|_v & \text{if } z, w \in \mathbb{C}_v \text{ and } w \neq 0, \\ \log |z|_v + g(\infty) & \text{if } z, w \in \mathbb{C}_v \text{ and } z \neq 0, w = 0, \\ -\infty & \text{if } z = w = 0, \end{cases}$$

then G is a homogeneous logarithmic potential. (Note that the opposite sign in front of $g(\infty)$ versus equation (2.3) in [MY] is a result of our opposite choice of sign for the Laplacian.) In particular, $G(\alpha z, \alpha w) = \log |\alpha|_v + G(z, w)$ for all $z, w \in \mathbb{C}_v$ and $\alpha \in \mathbb{C}_v^\times$. For any $z \in \mathbb{P}^1(\mathbb{C}_v)$, we let $\tilde{z} = (z_0, z_1)$ denote any homogeneous lift of z to \mathbb{C}_v^2 , and for $z = (z_0, z_1)$ and $w = (w_0, w_1) \in \mathbb{C}_v^2$, we let $z \wedge w = z_0 w_1 - z_1 w_0$. If we then define the Arakelov-Green's function $g_{\rho_v} : \mathbb{P}_v^1 \times \mathbb{P}_v^1 \rightarrow \mathbb{R} \cup \{\infty\}$ by defining it for $z, w \in \mathbb{P}^1(\mathbb{C}_v)$ by

$$g_{\rho_v}(z, w) = -\log |\tilde{z} \wedge \tilde{w}|_v + G(\tilde{z}) + G(\tilde{w}),$$

then this extends naturally to a function on $\mathbb{P}_v^1 \times \mathbb{P}_v^1$ which is continuous and finite off of the diagonal. (We refer the reader to [BR06, Section 3.4] for more details on these Arakelov-Green's functions.) Notice that G , and hence g_{ρ_v} , are only defined up to the choice of constant $g(\infty)$. This is chosen so that the Arakelov-Green's function satisfies the normalization condition

$$(26) \quad \iint_{\mathbb{P}_v^1 \times \mathbb{P}_v^1} g_{\rho_v}(z, w) d\rho_v(z) d\rho_v(w) = 0.$$

As observed in [BR06, Section 3.5], this normalization is in fact equivalent to the (seemingly stronger) condition that

$$(27) \quad \int_{\mathbb{P}_v^1} g_{\rho_v}(z, w) d\rho_v(z) \equiv 0.$$

With this normalization, we then define the adelic set $(M_v)_{v \in M_K}$ by

$$M_v = \{(z, w) \in \mathbb{C}_v^2 : G(z, w) \leq 0\}.$$

In fact, M_v is a compact, circled, and pseudoconvex set in \mathbb{C}_v^2 , and by our choice of normalization, it has homogeneous capacity $\text{cap}(M_v) = 1$. (The notion of homogeneous capacity was introduced by DeMarco [DeM03] in the archimedean setting, and extended to the non-archimedean setting by Baker and Rumely [BR06, Section 3.3].) Endow \mathbb{C}_v^2 with the norm $\|(z, w)\| = \max\{|z|_v, |w|_v\}$ and let $B(r) = \{(z, w) \in \mathbb{C}_v^2 : \|(z, w)\| \leq r\}$. Then the *inner radius* of M_v is defined to be

$$r_{\text{in},v} = \sup\{r > 0 : B(r) \subset M_v\},$$

and the outer radius of M_v is defined to be

$$r_{\text{out},v} = \inf\{r > 0 : M_v \subset B(r)\}.$$

For a measure μ_v which admits a continuous potential we define the inner and outer radii of the measure to the inner and outer radii of the associated set $M_v \subset \mathbb{C}_v^2$ constructed with the above normalization.

Then Mavraki and Ye define for a number field K the sequence of measures $(\rho_v)_{v \in M_K}$ to be a *quasi-adelic measure* if the following conditions are met:

- (1) Each ρ_v is a probability measure on \mathbb{P}_v^1 which admits a continuous potential with respect to the standard measure.
- (2) We have the finiteness condition:

$$(28) \quad \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \cdot |\log r_{\text{in},v}| < \infty \quad \text{and} \quad \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \cdot |\log r_{\text{out},v}| < \infty.$$

If ρ is a generalized adelic measure defined over K and we identify it with $(\rho_v)_{v \in M_K}$, then clearly it meets condition (1). What must be shown is that our finiteness condition (2) from Definition 2.1 is equivalent to the finiteness condition (2) stated above for the inner and outer radii. It is not immediately obvious that this geometric condition on the adelic set M_v is related to the bounds of the potential function $g(z)$, however, we will see that this is indeed the case, and while our normalization differs slightly, the two conditions are indeed equivalent.

In order to see this, we first observe that with the notation above, for any $z \in \mathbb{P}_v^1$ and $\tilde{z} \in \mathbb{C}_v^2$ a lift, $\log \|\tilde{z}\|_v - G(\tilde{z})$ does not depend on the choice of lift, as each factor scales logarithmically with the absolute value of α if we replace \tilde{z} with $\alpha\tilde{z}$. Thus in fact, for $z \in \mathbb{C}_v$,

$$\log \|\tilde{z}\|_v - G(\tilde{z}) = \log^+ |z|_v - p_v(z) = -g(z).$$

Now, observe that for any \tilde{z} , by scaling by some $\alpha \in \mathbb{C}_v^\times$, we can arrange that $G(\alpha\tilde{z}) = 0$, so $\tilde{z} \in M_v$. By the definition of $r_{\text{in},v}$ and $r_{\text{out},v}$, it follows that

$$r_{\text{in},v} \leq \|\alpha\tilde{z}\| \leq r_{\text{out},v},$$

so $-g(z) = \log \|\tilde{z}\|_v - G(\tilde{z}) = \log \|\alpha\tilde{z}\|$, and thus

$$\log r_{\text{in},v} \leq -g(z) \leq \log r_{\text{out},v}.$$

In fact, by continuity and the density of \mathbb{C}_v in \mathbb{P}_v^1 , we have

$$(29) \quad \log r_{\text{in},v} = \inf_{z \in \mathbb{P}_v^1} -g(z) \quad \text{and} \quad \log r_{\text{out},v} = \sup_{z \in \mathbb{P}_v^1} -g(z).$$

Thus we see that the inner and outer radii are in fact bounds on the potential function g for which $\Delta g = \rho_v - \lambda_v$. However, Mavraki and Ye's choice of potential function differs from ours by a constant. We chose to normalize so that $g(\infty) = 0$, however, Mavraki and Ye normalized so that the double integral of the Arakelov-Green's function vanishes as in (26) above. To determine the value of $g(\infty)$ in their normalization, we use the fact that the function $F : \mathbb{P}_v^1 \rightarrow \mathbb{R}$ given by

$$F(w) = \int_{\mathbb{P}_v^1} g_{\rho_v}(z, w) d\rho_v(z)$$

is in fact constant as a function of w , as one can show that the function is harmonic on \mathbb{P}_v^1 ; see the discussion in [BR06, Section 3.5, Lemma 5.14] for more details. By the normalization in (27), $F \equiv 0$, and we will evaluate $F(\infty)$ in order to determine the constant:

$$\begin{aligned} F(\infty) &= \int_{\mathbb{P}_v^1} G(\tilde{z}) + G(1, 0) - \log |\tilde{z} \wedge (1, 0)|_v d\rho_v(z) \\ &= \int_{\mathbb{P}_y^1} p_v(z) + g(\infty) - \log |0 \cdot z - 1 \cdot 1|_v d\rho_v(z) \\ &= g(\infty) + \int_{\mathbb{P}_v^1} p_v(z) d\rho_v(z). \end{aligned}$$

Now let

$$u(z) = \int_{\mathbb{P}_v^1} \log |z - w|_v d\rho_v(w),$$

and observe that, by considering asymptotics as $z \rightarrow \infty$, $p_v(z) = g(\infty) + u(z)$. Further, by [FRL06, Lemmes 2.5, 4.4],

$$\int_{\mathbb{P}_v^1} u(z) d\rho_v(z) = -(\rho_v, \rho_v)_v.$$

It follows that

$$F(\infty) = 2g(\infty) - (\rho_v, \rho_v)_v = 0, \quad \text{so} \quad g(\infty) = \frac{1}{2}(\rho_v, \rho_v)_v.$$

Now let us define the analogue of our bounds on the potential function for this measure. Let $g_v : \mathbb{P}_v^1 \rightarrow \mathbb{R}$ be the potential function with $\Delta g_v = \rho_v - \lambda_v$ and normalized according to Mavraki and Ye's normalization $g_v(\infty) = \frac{1}{2}(\rho_v, \rho_v)_v$. Then if we define

$$C_v = \sup_{z \in \mathbb{P}_v^1} |g_v(z) - g_v(\infty)|,$$

then for the generalized measure $(\rho_y)_{y \in Y}$ defined by $\rho_y = \rho_v$ where $y \mid v$, the bounding function $C : Y \rightarrow [0, \infty]$ will be given by

$$(30) \quad C(y) = \sum_{v \in M_K} C_v \chi_{Y(K,v)}(y).$$

Then

$$\int_Y C(y) d\mu(y) = \sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} C_v.$$

Now, we wish to show that $C \in L^1(Y)$ if and only if equation (28) holds. First, suppose $C \in L^1(Y)$. Then

$$\sup_{z \in \mathbb{P}_v^1} |g(z)| \leq C_v + |g(\infty)| = C_v + \frac{1}{2}|(\rho_v, \rho_v)_v| \leq 3C_v,$$

where we have applied Lemma 2.12 to bound $|(\rho_v, \rho_v)_v| \leq 4C_v$. It follows that

$$-3C_v \leq \log r_{\text{in},v} \leq \log r_{\text{out},v} \leq 3C_v$$

and so

$$\sum_{v \in M_K} [K_v : \mathbb{Q}] |\log r_{\text{in},v}| \leq [K : \mathbb{Q}] \cdot 3 \int_Y C(y) d\mu(y) < \infty,$$

and likewise for the outer radius.

Now assume that $(\rho_v)_{v \in M_K}$ meets the condition (28). By equation (29),

$$\sup_{z \in \mathbb{P}_v^1} |g_v(z)| = \max\{|\log r_{\text{in},v}|, |\log r_{\text{out},v}|\}.$$

It follows that for our normalization,

$$C_v = \sup_{z \in \mathbb{P}_v^1} |g_v(z) - g_v(\infty)| \leq 2 \sup_{z \in \mathbb{P}_v^1} |g_v(z)| = 2 \max\{|\log r_{\text{in},v}|, |\log r_{\text{out},v}|\}.$$

In particular, this now guarantees that for $C(y)$ defined as in (30),

$$\int_Y C(y) d\mu(y) = 2 \sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \max\{|\log r_{\text{in},v}|, |\log r_{\text{out},v}|\} < \infty.$$

Thus, the two finiteness conditions are equivalent.

3.3. Equidistribution results. In order to define the regularization techniques used in the following section, we start by recalling the Favre and Rivera-Letelier retraction-by- ϵ map on $\mathbb{P}^1(\mathbb{C}_p)$, for non-archimedean primes p . Let $\epsilon \geq 0$ be given, and we define $\pi_\epsilon(\zeta)$ to be the unique minimal point in \mathbb{P}_p^1 on the arc between ζ and ∞ which has diameter at least ϵ . We will want to vary the function in ϵ , so we start by proving $\pi_\epsilon(\zeta)$ is continuous as a function of two variables:

Lemma 3.6. *Define the map*

$$\begin{aligned} \pi : \mathbb{R}_{\geq 0} \times \mathbb{P}^1(\mathbb{C}_p) &\rightarrow \mathbb{P}^1(\mathbb{C}_p) \\ (\epsilon, \zeta) &\mapsto \pi_\epsilon(\zeta). \end{aligned}$$

Then π is a continuous function, where $\mathbb{P}^1(\mathbb{C}_p)$ has the Berkovich topology, $\mathbb{R}_{\geq 0}$ has the usual real subspace topology, and $\mathbb{R}_{\geq 0} \times \mathbb{P}^1(\mathbb{C}_p)$ has the associated product topology.

Proof. As in the proof of the continuity of π_ϵ in [FRL06, Lemme 4.7], we take advantage of the fact that a subbase for the Berkovich topology is given by sets of the form

$$U(z, r) = \{\zeta \in \mathbb{P}^1(\mathbb{C}_p) : \sup\{\zeta, z\} < r\} \quad \text{where } z \in \mathbb{C}_p, r > 0,$$

and complements of closed discs:

$$\mathbb{P}^1(\mathbb{C}_p) \setminus \overline{U(z, r)} = \{\zeta \in \mathbb{P}^1(\mathbb{C}_p) : \sup\{\zeta, z\} > r\} \quad \text{where } z \in \mathbb{C}_p, r > 0.$$

We start by proving that $\pi^{-1}(U(z, r))$ is open for every $z \in \mathbb{C}_p$ and $r > 0$. Notice that, by [FRL06, Lemme 4.7], $\pi_\epsilon^{-1}(U(z, r)) = U(z, r)$ if $\epsilon < r$, but $\pi_\epsilon^{-1}(U(z, r)) = \emptyset$ if $\epsilon \geq r$. It follows that

$$\pi^{-1}(U(z, r)) = [0, r) \times U(z, r),$$

which is clearly open in the product topology. On the other hand, if we let $B = \mathbb{P}^1(\mathbb{C}_p) \setminus \overline{U(z, r)}$ for some $z \in \mathbb{C}_p$ and $r > 0$, we see again by [FRL06, Lemme 4.7] that $\pi_\epsilon^{-1}(B) = B$ if $0 \leq \epsilon \leq r$ and $\pi_\epsilon^{-1}(B) = \mathbb{P}^1(\mathbb{C}_p)$ if $\epsilon > r$. It follows that

$$\pi^{-1}(B) = ([0, r] \times B) \cup ((r, \infty) \times \mathbb{P}^1(\mathbb{C}_p)),$$

which is again an open set (note that it equals the union of two open sets, $[0, \infty) \times B$ and $(r, \infty) \times \mathbb{P}^1(\mathbb{C}_p)$). It follows that π is continuous. \square

As we will use it in the next section, we introduce the notion of ϵ -regularization of a measure.

Definition 3.7. Let ρ be a measure on \mathbb{P}_y^1 for a place y of $\overline{\mathbb{Q}}$ and let $\epsilon > 0$ be given. If $y \mid \infty$, we define the ϵ -regularization to be the measure ρ_ϵ given by requiring that for every continuous $f : \mathbb{P}_y^1 \rightarrow \mathbb{R}$,

$$\int f d\rho_\epsilon = \int \left(\int_0^1 f(z + \epsilon e^{2\pi it}) dt \right) d\rho(z).$$

If $y \nmid \infty$, then, following Favre and Rivera-Letelier [FRL06, §4.6], we define $\rho_\epsilon = (\pi_\epsilon)_*(\rho)$.

Remark 3.8. It will be clear to the reader that both of our regularizations are inspired by [FRL06], however, the archimedean regularization differs in that we avoided convolving with a smooth function. As a result, our potential function will not be smooth but will still be continuous, which is all we require; further, when approximating discrete measures, we will avoid an extraneous term which appears in the estimates of [FRL06]. We refer the reader to the discussion in [FP19, §2.1] for more details.

We also note that our regularization fails to smooth the potentials of point masses at $\infty \in \mathbb{P}^1(\mathbb{C}_y)$. Indeed, it is clear from the definition of our normalization that $(\delta_\infty)_\epsilon = \delta_\infty$. This was also true of the regularization in [FRL06], but just as in their work, in our equidistribution results, the resulting measures will never charge a single classical point (that is, a point of $\mathbb{P}^1(\mathbb{C}_y)$), so we will be able to ignore the point at infinity.

We state the basic convergence lemma:

Lemma 3.9. *Let ρ be a signed Borel measure on \mathbb{P}_y^1 . Then $\rho_\epsilon \rightarrow \rho$ as $\epsilon \rightarrow 0^+$ in the sense of weak convergence of measures. Further, if ρ_n is a sequence of signed Borel measures such that $\rho_n \rightarrow \rho$ weakly, then $(\rho_n)_\epsilon \rightarrow \rho_\epsilon$ weakly as $\epsilon \rightarrow 0^+$ as well.*

Proof. In the non-archimedean case, this is exactly [FRL06, Lemme 4.8], so we will not reprove it here. The archimedean case is, similarly to [FRL06, Lemme 2.7], trivial, but we will write it out for the sake of completeness. We identify \mathbb{P}_y^1 with $\mathbb{P}^1(\mathbb{C})$ and let a continuous real-valued function $f \in C(\mathbb{P}^1(\mathbb{C}))$ be given. We will show that

$$\int f d\rho_\epsilon \rightarrow \int f d\rho.$$

Note that by construction,

$$\int f d\rho_\epsilon = \int \left(\int_0^1 f(z + \epsilon e^{2\pi t}) dt \right) d\rho(z) \rightarrow \int f d\rho$$

as $\epsilon \rightarrow 0^+$, since the function f is continuous on a compact space and therefore we can bring the limit inside the integral. Since this holds for each continuous function, $\rho_\epsilon \rightarrow \rho$ in the weak sense of measures. Now suppose that $\rho_n \rightarrow \rho$ weakly, and fix $\epsilon > 0$. We wish to show that $(\rho_n)_\epsilon \rightarrow \rho_\epsilon$. Again, it suffices to show that for a given $f \in C(\mathbb{P}^1(\mathbb{C}))$,

$$\int f d(\rho_n)_\epsilon \rightarrow \int f d\rho_\epsilon \quad \text{as } n \rightarrow \infty.$$

But

$$\int f d(\rho_n)_\epsilon = \int \left(\int_0^1 f(z + \epsilon e^{2\pi t}) dt \right) d\rho_n(z) = \int f_\epsilon d\rho_n,$$

where

$$f_\epsilon(z) = \int_0^1 f(z + \epsilon e^{2\pi t}) dt.$$

But f_ϵ is again a continuous function, so since $\rho_n \rightarrow \rho$ weakly,

$$\int f_\epsilon d\rho_n \rightarrow \int f_\epsilon d\rho = \int f d\rho_\epsilon,$$

where the last equality follows by the definition of ρ_ϵ . \square

Lemma 3.10. *Let y be a place of $\overline{\mathbb{Q}}$ and Δ be a discrete probability measure on A_y^1 , given by*

$$\Delta = \sum_i t_i \delta_{\alpha_i},$$

where $\sum_i t_i = 1$ and $\alpha_i \in \mathbb{C}_y$, satisfying

$$(31) \quad \sum_i t_i \log^+ |\alpha_i|_y < \infty.$$

Then for any $0 < \epsilon \leq 1$, the ϵ -regularization Δ_ϵ admits a continuous potential with respect to the standard measure λ_y , that is, there is a continuous function $g_y : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ such that $g_y(\infty) = 0$ and

$$\Delta g_y = \Delta_\epsilon - \lambda_y.$$

Notice that we have assumed that Δ does not charge $\infty \in \mathbb{P}^1(\mathbb{C}_y)$. This assumption is primarily for our convenience at the moment and we will see that it is not relevant in the proof of the equidistribution theorem, as measures of small height will not converge to measures that charge a single point like $\infty \in \mathbb{P}^1(\mathbb{C}_y)$. It is possible to introduce a regularization that allows for continuous potentials when Δ has support at ∞ as well, however, this would complicate our analysis and is unnecessary, so we do not do it here.

Proof. First we recall for the reader's convenience that, for any $\epsilon > 0$, the ϵ -regularization of Δ is given by

$$\Delta_\epsilon = \sum_i t_i \delta_{\alpha_i, \epsilon},$$

where if $y \nmid \infty$, $\delta_{\alpha_i, \epsilon} = (\delta_{\alpha_i})_\epsilon$ denotes the point mass at $\pi_\epsilon(\alpha_i) = \zeta_{\alpha_i, \epsilon}$ and is the Favre-Rivera-Letelier smoothing of the measure, while if $y \mid \infty$, we let $\delta_{\alpha_i, \epsilon} = (\delta_{\alpha_i})_\epsilon$

denote the normalized arc length measure around the circle $\partial D(\alpha_i, \epsilon)$ in \mathbb{C} . We note that each individual measure $\delta_{\alpha_i, \epsilon}$ admits a continuous potential with respect to the standard measure λ_y given by

$$g_i(z) = \log \max\{|z - \alpha_i|_y, \epsilon\} - \log^+ |z|_y,$$

where in the case of y non-archimedean, we recall that $|z - \alpha_i|_y$ is really the Hsia kernel $\delta_\infty(z, \alpha_i)$, or $\sup\{z, \alpha_i\}$ in the notation of Favre and Rivera-Letelier. Notice that

$$\sup_{z \in \mathbb{P}_y^1} |g_i(z)| \leq \log(1 + |\alpha_i|_y) \leq \log 2 + \log^+ |\alpha_i|_y.$$

Define

$$g(z) = \sum_i t_i g_i(z) = \sum_i t_i \log \max\{|z - \alpha_i|_y, \epsilon\} - \log^+ |z|_y.$$

Then

$$\sum_i t_i |g_i(z)| \leq \log 2 + \sum_i t_i \log^+ |\alpha_i|_y < \infty$$

by our assumption on Δ . It follows by the Weierstrass M -test that $g \in C(\mathbb{P}_y^1)$, as claimed. \square

Lemma 3.11. *Let ρ be a generalized adelic measure and Δ be a discrete probability measure on $\overline{\mathbb{Q}}$ with finite height given by*

$$\Delta = \sum_{n=1}^{\infty} t_n \delta_{\alpha_n},$$

where $\alpha_n \in \mathbb{A}^1(\overline{\mathbb{Q}})$ are all distinct and $\sum_n t_n = 1$. Then for μ -almost every $y \in Y$, there exists a continuous function $\eta_y : [0, \infty) \rightarrow [0, \infty)$ such that $\eta(0) = 0$, and that for all $\epsilon \in (0, 1]$, the ϵ -regularization Δ_ϵ of Δ satisfies

$$(32) \quad (\rho_y - \Delta, \rho_y - \Delta)_y - (\rho_y - \Delta_\epsilon, \rho_y - \Delta_\epsilon)_y \geq -2\eta_y(\epsilon) + \sum_{n=1}^{\infty} t_n^2 \log \epsilon$$

and in particular,

$$(33) \quad (\rho_y - \Delta, \rho_y - \Delta)_y \geq -2\eta_y(\epsilon) + \sum_{n=1}^{\infty} t_n^2 \log \epsilon$$

for almost all $y \in Y(\mathbb{Q}, p)$. Further, the functions η_y depend on y and on ρ , but not on Δ .

We note that our lemma is inspired by Propositions 5 and 6 in Section 2 of [FP19], which itself was based closely on Propositions 2.8 and 4.9 of [FRL06] but uses a slightly different regularization of the discrete measures at the archimedean places.

Proof. Let $N = \{y \in Y : C(y) = \infty\}$, where $C(y)$ is the bounding function associated to our generalized adelic measure ρ in Definition 2.1. We have by assumption that $C(y) \in L^1(Y)$, so in particular, $\mu(N) = 0$. We note that by the condition that Δ has finite height,

$$\int_Y \int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta(z) d\mu(y) < \infty,$$

and it follows that for almost all $y \in Y$,

$$\int_{\mathbb{A}^1} \log^+ |z|_y d\Delta(z) < \infty.$$

This condition will be necessary to apply 3.10 and conclude that Δ_ϵ admits a continuous potential function, which will be vital to our proof.

Let $Y' \subset Y$ denote the set of places y where the above integral is finite and $C(y) < \infty$. Notice that $\mu(Y \setminus Y') = 0$. We will prove the result for all $y \in Y'$. Let $y \in Y'$ be given. For each $\epsilon \in (0, 1]$, let Δ_ϵ be the ϵ -regularization of the discrete measure Δ as defined above. We write Δ_ϵ explicitly as

$$(34) \quad \Delta_\epsilon = \sum_{n=1}^{\infty} t_n \delta_{\alpha_n, \epsilon},$$

where, if $y \nmid \infty$, $\delta_{\alpha_n, \epsilon}$ denotes the point mass at $\pi_\epsilon(\alpha_n) = \zeta_{\alpha_n, \epsilon}$ and is the Favre and Rivera-Rivera smoothing of the measure, while if $y \mid \infty$, we let $\delta_{\alpha_n, \epsilon}$ denote the normalized arc length measure around the circle $\partial D(\alpha_n, \epsilon)$ in \mathbb{C} . (The restriction that $\epsilon \leq 1$ is only relevant in the non-archimedean setting.) We then proceed by writing

$$(35) \quad (\rho_y - \Delta, \rho_y - \Delta)_y - (\rho_y - \Delta_\epsilon, \rho_y - \Delta_\epsilon)_y \\ = -2(\rho_y, \Delta - \Delta_\epsilon)_y + (\Delta, \Delta)_y - (\Delta_\epsilon, \Delta_\epsilon)_y.$$

We now proceed by finding lower bounds for the terms on the right hand side. We start with the difference $(\Delta, \Delta)_y - (\Delta_\epsilon, \Delta_\epsilon)_y$, largely following the proofs of [FP19, Prop. 5] and [FRL06, Lemme 4.11]. We write

$$(\Delta, \Delta)_y = \sum_{i \neq j} -t_i t_j \log |\alpha_i - \alpha_j|_y,$$

and by the linearity of the energy pairing,

$$(\Delta_\epsilon, \Delta_\epsilon)_y = \sum_{i, j} t_i t_j (\delta_{\alpha_i, \epsilon}, \delta_{\alpha_j, \epsilon})_y.$$

If $y \mid \infty$, then, following the idea proof of the proof of [FRL06, Lemme 2.10], then for $z \neq w \in \mathbb{C}$,

$$-(\delta_{z, \epsilon}, \delta_{w, \epsilon}) = \int_0^1 \int_0^1 \log |z + \epsilon \cdot e^{2\pi i r} - (w + \epsilon \cdot e^{2\pi i s})| dr ds \\ = \int_0^1 \max\{\log |z - (w + \epsilon \cdot e^{2\pi i s})|, \log \epsilon\} ds \\ \geq \max\left\{\int_0^1 \log |z - (w + \epsilon \cdot e^{2\pi i s})| ds, \log \epsilon\right\} \\ \geq \max\{\log |z - w|, \log \epsilon\} \geq \log |z - w| = -(\delta_z, \delta_w),$$

while if $y \nmid \infty$,

$$\log \sup\{\zeta_{\alpha_i, \epsilon}, \zeta_{\alpha_j, \epsilon}\} = \log \max\{\epsilon, |\alpha_i - \alpha_j|_y\} \geq \log |\alpha_i - \alpha_j|_y,$$

so that for all $i \neq j$ in the double sums,

$$(\Delta, \Delta)_y - (\Delta_\epsilon, \Delta_\epsilon)_y \geq 0.$$

Further, for the $i = j$ terms, $(\delta_{z, \epsilon}, \delta_{z, \epsilon})_y$ is the Robin constant for the disc of radius ϵ for both non-archimedean and archimedean places, and in both cases, one obtains $(\delta_{z, \epsilon}, \delta_{z, \epsilon})_y = -\log \epsilon$, so that

$$(36) \quad (\Delta, \Delta)_y - (\Delta_\epsilon, \Delta_\epsilon)_y \geq \sum_{i=1}^{\infty} t_i^2 \log \epsilon.$$

It remains to bound the term $-2(\rho_y, \Delta - \Delta_\epsilon)_y$. Our proof now breaks down into two cases. First, assume that $p \nmid \infty$ is a non-archimedean prime. Note that

$$(\rho_y, \Delta - \Delta_\epsilon)_y = \sum_{i=1}^{\infty} t_i (p_{\rho_y}(\alpha_i) - p_{\rho_y}(\zeta_{\alpha_i, \epsilon})),$$

where $p_{\rho_y}(z)$ is the potential associated to ρ_y . By Lemma 2.7, we know that $p_{\rho_y}(z) = g_y(z) + \log^+ |z|_p$. Now, since $g_y(z) = g(y, z)$ is assumed to be continuous as a map $Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow \mathbb{R}$, we can define

$$\eta_y(\epsilon) = \sup_{\zeta \in \mathbb{P}_p^1} |g_y(\pi_\epsilon(\zeta)) - g_y(\zeta)|,$$

and it follows by continuity of all of the variables involved and the compactness of the spaces that these suprema exist and are attained, that the resulting function is continuous in ϵ , and, further, that $\eta(0) = 0$. Further,

$$\begin{aligned} |p_{\rho_y}(\alpha_i) - p_{\rho_y}(\zeta_{\alpha_i, \epsilon})| &\leq |g_y(\alpha_i) - g_y(\zeta_{\alpha_i, \epsilon})| + |\log^+ |\alpha_i|_p - \log^+ |\zeta_{\alpha_i, \epsilon}|| \\ &\leq \eta_y(\epsilon), \end{aligned}$$

since $\log^+ |\zeta_{z, \epsilon}|_p = \log^+ |z|_p$ whenever $\epsilon \leq 1$. It follows that

$$(37) \quad |(\rho_y, \Delta - \Delta_\epsilon)_y| \leq \sum_{i=1}^{\infty} t_i \eta_y(\epsilon) = \eta_y(\epsilon).$$

For the second case, suppose that $y \mid \infty$. Let $\hat{\eta}_y : [0, \infty) \rightarrow [0, \infty)$ be the function given by

$$\hat{\eta}_y(\epsilon) = \sup_{\substack{z, w \in \mathbb{P}^1(\mathbb{C}) \\ \sigma(z, w) \leq \epsilon}} |g_y(z) - g_y(w)|,$$

where σ denotes the spherical metric on $\mathbb{P}^1(\mathbb{C})$. Again, by the continuity of g_y and compactness of the space \mathbb{P}_y^1 , $\hat{\eta}_y(\epsilon)$ is a continuous function and works as a uniform modulus of continuity for the potential functions over all places $y \mid \infty$. Let $\eta_y(\epsilon) = \hat{\eta}_y(\epsilon) + \epsilon$. Then for $z \in \mathbb{C}$,

$$\begin{aligned} \left| p_{\rho_y}(z) - \int p_{\rho_y} d\delta_{z, \epsilon} \right| &\leq \left| g_y(z) - \int g_y d\delta_{z, \epsilon} \right| + \left| \int \log^+ |w| d\delta_{z, \epsilon}(w) - \log^+ |z| \right| \\ &\leq \hat{\eta}_y(\epsilon) + \epsilon = \eta_y(\epsilon), \end{aligned}$$

where we have used

$$\left| \int_0^1 \log^+ |z + \epsilon \cdot e^{2\pi i t}| dt - \log^+ |z| \right| \leq \epsilon$$

to bound the second term, from which we again derive (37), this time in the archimedean case. Combining equation (35) with (36) and (37) gives the first part of the conclusion, equation (32) above.

For the second part of the conclusion, we note that by [FRL06, Prop. 2.6, Prop. 4.5], as Δ_ϵ admits a continuous potential by Lemma 3.10, we have

$$(\rho_y - \Delta_\epsilon, \rho_y - \Delta_\epsilon)_y \geq 0.$$

Therefore (35) yields

$$(38) \quad (\rho_y - \Delta, \rho_y - \Delta)_y \geq -2(\rho_y, \Delta - \Delta_\epsilon)_y + (\Delta, \Delta)_y - (\Delta_\epsilon, \Delta_\epsilon)_y,$$

which, combined with (36) and (37) gives the second conclusion (33). \square

The previous lemma defines a condition for the local pairings to be nonnegative in the limit, to which we will give a name:

Definition 3.12. A sequence of discrete probability measures $(\Delta_n)_{n \in \mathbb{N}}$ on a topological space is *well-distributed* if

$$(39) \quad \sum_{z \in \text{supp}(\Delta_n)} \Delta_n(\{z\})^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Notice that if the Δ_n are well-distributed, then we must have $|\text{supp}(\Delta_n)| \rightarrow \infty$, and further, if the Δ_n are measures on $\overline{\mathbb{Q}}$, this condition is independent of the particular embedding into \mathbb{P}_y^1 for any place y of $\overline{\mathbb{Q}}$.

Example 3.13. Let K/\mathbb{Q} be a field, and for $n \in \mathbb{N}$ let $\alpha_n \in \overline{K}$ and let $\Delta_n = [\alpha_n]_K$ be the probability measure equally supported on each $\text{Gal}(\overline{K}/K)$ -conjugate of α_n . If $[K(\alpha_n) : K] \rightarrow \infty$, then the sequence $(\Delta_n)_{n \in \mathbb{N}}$ is well-distributed, as

$$\sum_{z \in \text{supp}(\Delta_n)} \Delta_n(\{z\})^2 = \sum_{i=1}^{[K(\alpha_n):K]} \frac{1}{[K(\alpha_n) : K]^2} = \frac{1}{[K(\alpha_n) : K]} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

We will not use the definition of the well-distributed property directly in our proof of the equidistribution results, but rather we will use it in an equivalent but slightly more technical form:

Lemma 3.14. *If a sequence of discrete probability measures $(\Delta_n)_{n \in \mathbb{N}}$ on a topological space is well-distributed, then there exists a sequence of real numbers $(\epsilon_n)_{n \in \mathbb{N}}$ such that*

$$0 < \epsilon_n \leq 1 \quad \text{for each } n \in \mathbb{N}, \quad \epsilon_n \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

and

$$(40) \quad \sum_{z \in \text{supp}(\Delta_n)} \Delta_n(\{z\})^2 \log \epsilon_n \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Proof. For each $n \geq 1$ let

$$x_n = \sum_{z \in \text{supp}(\Delta_n)} \Delta_n(\{z\})^2.$$

The well-distributed condition says that $x_n \rightarrow 0$ as $n \rightarrow \infty$. Note that as Δ_n is a discrete probability measure, we must have $0 < x_n \leq 1$. So we may simply take $\epsilon_n = x_n$, then

$$\sum_{z \in \text{supp}(\Delta_n)} \Delta_n(\{z\})^2 \log \epsilon_n = x_n \log x_n \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad \square$$

Proposition 3.15. *Let y be a place of $\overline{\mathbb{Q}}$, let ρ_n be a sequence of Borel probability measures on \mathbb{P}_y^1 each of which admits a continuous potential with respect to the standard measure λ_y , and let ρ be a probability measure meeting the same conditions. If*

$$(\rho_n - \rho, \rho_n - \rho)_y \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

then $\rho_n \rightarrow \rho$ as $n \rightarrow \infty$ in the sense of weak convergence of measures.

Proof. Let $g_\rho : \mathbb{P}_y^1 \times \mathbb{P}_y^1 \rightarrow \mathbb{R} \cup \{\infty\}$ be the function given by

$$(41) \quad g_\rho(z, w) = p_\rho(z) + p_\rho(w) - \log |z - w|_y + (\rho, \rho)_y,$$

where $p_\rho(z)$ is the potential function of ρ , as defined in Lemma 2.7 above. We note that $g_\rho(z, w)$ is the Arakelov-Zhang function defined in [BR06]. It is easy to see that $g_\rho(z, w)$ is continuous off of the diagonal, and lower-semicontinuous on the diagonal, where it is equal to ∞ . It follows from [FRL06, Lemmes 2.4, 2.5, 4.3, 4.4] that we can ignore the diagonal (of classical points) in integrating $g_\rho(z, w)$ against ρ_n , so

$$(\rho_n - \rho, \rho_n - \rho)_y = \iint_{\mathbb{P}_y^1 \times \mathbb{P}_y^1} g_\rho(z, w) d\rho_n(z) d\rho_n(w).$$

It follows from [FRL06, Prop. 2.6 and 4.5] and the slightly stronger [BR06, Theorem 3.45] that this pairing is nonnegative and vanishes if and only if $\rho_n = \rho$. (The theorem in [BR06] is stronger in that it removes the hypothesis that ρ_n also admits a continuous potential.)

Let ρ' be a weak limit of some subsequence of ρ_n , which must exist by Prokhorov's theorem. Since our space \mathbb{P}_y^1 is metrizable (we refer the reader to the discussion of [BR10, Chapter 1.5] for details), we can apply the Portmanteau theorem for weak convergence of measures to the lower-semicontinuous function $g_\rho(z, w)$ to conclude that

$$(42) \quad 0 = \liminf_{n \rightarrow \infty} \iint_{\mathbb{P}_y^1 \times \mathbb{P}_y^1} g_\rho(z, w) d\rho_n(z) d\rho_n(w) \geq \iint_{\mathbb{P}_y^1 \times \mathbb{P}_y^1} g_\rho(z, w) d\rho'(z) d\rho'(w) \\ = (\rho - \rho', \rho - \rho')_y \geq 0,$$

where the nonnegativity follows from the fact that our measures both have continuous potentials. But then $(\rho - \rho', \rho - \rho')_y = 0$ so $\rho = \rho'$, and since every subsequence of ρ_n has a further subsequence which converges to this weak limit, the entire sequence has weakly limits to ρ . \square

We are now ready to prove the local equidistribution result. We note that our result is very much inspired by Propositions 2.11 and 4.12 of [FRL06], however, there are several differences, most notably, that our discrete measures may in fact be infinitely supported.

Proposition 3.16 (Local equidistribution). *Let y be a place of $\overline{\mathbb{Q}}$ and ρ be a Borel probability measure on \mathbb{P}_y^1 which admits a continuous potential in the sense that there exists a continuous function $g : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ such that $g(\infty) = 0$ and $\Delta g = \rho - \lambda_y$. Suppose that Δ_n is a sequence of discrete probability measures on \mathbb{P}_y^1 which are well-distributed, which satisfy the local finiteness condition*

$$(43) \quad \int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta_n(z) < \infty$$

for each $n \in \mathbb{N}$, and for which

$$(44) \quad \limsup_{n \rightarrow \infty} (\rho - \Delta_n, \rho - \Delta_n)_y \leq 0.$$

Then the measures Δ_n converge weakly to ρ in the sense of weak convergence of measures.

Proof. First, we note that we may assume that no $\Delta_n = \delta_\infty$, as the well-distributed condition requires that the support of Δ_n must grow, so we would be removing at most finitely many terms. Second, let us show that we may assume that no Δ_n measure charges $\infty \in \mathbb{P}^1(\overline{\mathbb{Q}})$. To see this, note that if $\Delta_n(\{\infty\}) \not\rightarrow 0$, then there exists a subsequence Δ_{n_k} with $\Delta_{n_k}(\{\infty\}) \geq c > 0$ for all $k \in \mathbb{N}$. But then the Δ_n are no longer well-distributed, as for these indices n_k we have

$$\sum_{z \in \text{supp}(\Delta_{n_k})} \Delta(\{z\})^2 \geq c^2 > 0.$$

Define a new sequence of probability measures by

$$\Delta'_n = \frac{1}{1 - \Delta_n(\{\infty\})} \Delta_n|_{\mathbb{A}_y^1}.$$

Notice that, since $\Delta_n(\{\infty\}) \rightarrow 0$, any weak limit of any subsequence of Δ'_n is the same as the weak limit of any subsequence of Δ_n . Therefore, by replacing Δ_n by Δ'_n if necessary, we can assume that no Δ_n measure charges $\infty \in \mathbb{P}^1(\mathbb{C}_y)$.

Write each Δ_n as

$$\Delta_n = \sum_i t_i^{(n)} \delta_{z_i^{(n)}},$$

where $z_i^{(n)} \in \mathbb{A}^1(\mathbb{C}_y)$ and $\sum_i t_i^{(n)} = 1$. By equation (32) from Lemma 3.11, there exists a function $\eta = \eta_y$, depending on ρ and the place y , but not on Δ_n , such that

$$(\rho_y - \Delta_n, \rho_y - \Delta_n)_y - (\rho_y - (\Delta_n)_\epsilon, \rho_y - (\Delta_n)_\epsilon)_y \geq -2\eta(\epsilon) + \sum_{i=1}^{\infty} (t_i^{(n)})^2 \log \epsilon,$$

or

$$(\rho_y - (\Delta_n)_\epsilon, \rho_y - (\Delta_n)_\epsilon)_y \leq (\rho_y - \Delta_n, \rho_y - \Delta_n)_y + 2\eta(\epsilon) - \sum_{i=1}^{\infty} (t_i^{(n)})^2 \log \epsilon.$$

Now by Lemma 3.10, $(\Delta_n)_\epsilon$ admits a continuous potential for each n and $\epsilon \in (0, 1]$, so we also know that

$$0 \leq (\rho_y - (\Delta_n)_\epsilon, \rho_y - (\Delta_n)_\epsilon)_y.$$

Let (ϵ_n) be a sequence of real numbers in $(0, 1]$ for which the conclusion of Lemma 3.14 holds. For this sequence,

$$2\eta(\epsilon_n) - \sum_{i=1}^{\infty} (t_i^{(n)})^2 \log \epsilon_n \rightarrow 0$$

as $n \rightarrow \infty$. It follows that

$$\lim_{n \rightarrow \infty} (\rho_y - (\Delta_n)_{\epsilon_n}, \rho_y - (\Delta_n)_{\epsilon_n})_y = 0.$$

Since our measures admit continuous potentials, it follows from Proposition 3.15 that we have the weak convergence of measures

$$\lim_{n \rightarrow \infty} (\Delta_n)_{\epsilon_n} = \rho_y.$$

Now, for any subsequence of the Δ_n , we know by Prokhorov's theorem that there is a further subsequence, which we denote $(\Delta_{n_k})_{k=1}^{\infty}$, that has a weak limit, say ρ' . Then by Lemma 3.9, for any ϵ_m we have

$$\lim_{k \rightarrow \infty} (\Delta_{n_k})_{\epsilon_m} = \rho'_{\epsilon_m},$$

which implies

$$\lim_{m \rightarrow \infty} \lim_{k \rightarrow \infty} (\Delta_{n_k})_{\epsilon_m} = \lim_{m \rightarrow \infty} \rho'_{\epsilon_m} = \rho'.$$

Since the subsequence with $m = n_k$ converges to ρ , so we must have $\rho' = \rho$, and it follows that every convergent subsequence of Δ_n converges to ρ , so $\Delta_n \rightarrow \rho$, as claimed. \square

Theorem 3.17 (Equidistribution for generalized adelic heights). *Let Δ_n be a well-distributed sequence of probability measures on $\overline{\mathbb{Q}}$ and ρ be a generalized adelic measure. If $h_\rho(\Delta_n) \rightarrow 0$ as $n \rightarrow \infty$, then $\Delta_n \rightarrow \rho_y$ in the weak sense of convergence of measures at μ -almost every place $y \in Y$.*

Remark 3.18. We note that if our generalized adelic measure is in fact defined over a single number field, then the condition of μ -almost every place becomes every place of that number field, as all places of any number field have positive μ measure. We also remind the reader that weak convergence of measures here refers to weak-* convergence of the measures as positive linear operators on the space of (compactly supported) continuous functions of the Berkovich line \mathbb{P}_y^1 .

Proof. Let $f_n : Y \rightarrow \mathbb{R}$ be the functions given by the local energy pairings:

$$f_n(y) = \frac{1}{2}(\rho_y - \Delta_n, \rho_y - \Delta_n)_y,$$

so that

$$h_\rho(\Delta_n) = \int_Y f_n(y) d\mu(y).$$

Note that since $h_\rho(\Delta_n) \rightarrow 0$, we may assume without loss of generality by tossing out at most finitely many terms of the sequence that $h_\rho(\Delta_n) < \infty$ for all n , hence that Δ_n meets the finite height condition by Corollary 3.2. Our first goal is to prove that $f_n(y) \rightarrow 0$ for almost all places $y \in Y$. We will then show that Δ_n meets the local finiteness condition needed in Proposition 3.16 for almost all $y \in Y$, so at the intersection of these two sets of places, we can apply our local equidistribution result Proposition 3.16 to get the desired result.

In order to prove that $f_n(y) \rightarrow 0$ for almost all y , we start by breaking $f_n(y)$ into its positive and negative parts, that is, we let

$$f_n(y) = f_n^+(y) - f_n^-(y),$$

where $f_n^+, f_n^- \geq 0$. Write

$$\Delta_n = \sum_i t_i^{(n)} \delta_{\alpha_i^{(n)}},$$

where the $\alpha_i^{(n)} \in \overline{\mathbb{Q}}$, $t_i^{(n)} \geq 0$, and $\sum_i t_i^{(n)} = 1$. (Note that the sum for Δ_n is countably infinite or finite.) Applying Lemma 3.11, we know that for μ -almost every place y , there is a continuous nonnegative function $\eta_y : [0, \infty) \rightarrow [0, \infty)$ depending on ρ and the y such that $\eta_y(0) = 0$ and such that

$$f_n^-(y) \leq \eta_y(\epsilon) + \frac{1}{2} \sum_i t_{n,i}^2 \log \epsilon^{-1} \quad \text{for all } \epsilon \in (0, 1].$$

By Lemma 3.14 and our assumption that the Δ_n are well-distributed, there exists a sequence of choices $\epsilon = \epsilon_n$ with $0 < \epsilon_n \leq 1$, $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$, and such that the right hand side of the above equation must vanish as $n \rightarrow \infty$, where the

first term vanishes by the continuity of η_y and $\eta_y(0) = 0$, and the second vanishes by (40). It follows that $f_n^-(y) \rightarrow 0$ pointwise for μ -almost every $y \in Y$.

Our goal is now to prove that in fact f_n^- is bounded by an integrable function, so that we can apply dominated convergence and conclude that $\int f_n^-(y) d\mu(y) \rightarrow 0$. We start by writing:

$$f_n(y) - \frac{1}{2}(\lambda_y - \Delta_n, \lambda_y - \Delta_n)_y = \frac{1}{2}(\rho_y, \rho_y)_y - (\rho_y - \lambda_y, \Delta_n)_y.$$

If $y \notin Y(\mathbb{Q}, \infty)$, then by Lemma 2.8, we have $(\lambda_y - \Delta_n, \lambda_y - \Delta_n)_y \geq 0$, so in fact,

$$f_n(y) \geq \frac{1}{2}(\rho_y, \rho_y)_y - (\rho_y - \lambda_y, \Delta_n)_y.$$

We recall that by Lemma 2.12,

$$(\rho_y, \rho_y)_y \geq -2C(y),$$

and

$$-(\rho_y - \lambda_y, \Delta_n)_y = \int_{A_y^1} g_y(z) d\Delta_n(z),$$

so $-(\rho_y - \lambda_y, \Delta_n)_y \geq -C(y)$. It follows that for $y \notin Y(\mathbb{Q}, \infty)$,

$$f_n(y) \geq -3C(y), \quad \text{or} \quad f_n^-(y) \leq 3C(y).$$

For $y \in Y(\mathbb{Q}, \infty)$, let η_y be the function from Lemma 3.11 where it exists. Recall from the proof of Lemma 3.11 that for $y \mid \infty$, $\eta_y(\epsilon) = \hat{\eta}_y(\epsilon) + \epsilon$ where

$$\hat{\eta}_y(\epsilon) = \sup_{\substack{z, w \in \mathbb{P}^1(\mathbb{C}) \\ \sigma(z, w) \leq r}} |g_y(z) - g_y(w)|.$$

Now observe that as $C(y) = \sup_{z \in \mathbb{P}_y^1} |g_y(z)|$, we have $\hat{\eta}_y(\epsilon) \leq 2C(y)$ for all y away from a set of μ -measure zero. Then Lemma 3.11 gives $f_n(y) \geq -2\eta_\infty(\epsilon)$, that is, taking $\epsilon = 1$,

$$f_n^-(y) \leq 2\eta_y(1) \leq 2C(y) + 1 \quad \text{for } \mu\text{-a.e. } y \in Y(\mathbb{Q}, \infty).$$

Define the function $F : Y \rightarrow [0, \infty)$ by

$$F(y) = \begin{cases} 2C(y) + 1 & \text{if } y \in Y(\mathbb{Q}, \infty), \\ 3C(y) & \text{if } y \in Y \setminus Y(\mathbb{Q}, \infty). \end{cases}$$

Clearly, $\int_Y F d\mu \leq 1 + 3 \int C d\mu < \infty$, so F is integrable. Lebesgue's dominated convergence theorem then applies and tells us that $\int_Y f_n^- d\mu \rightarrow 0$, which, together with the assumption that

$$h_\rho(\Delta_n) = \int_Y f_n d\mu = \int_Y f_n^+ d\mu - \int_Y f_n^- d\mu \rightarrow 0,$$

tells us that $\int f_n^+ d\mu \rightarrow 0$ as well. It follows that $\int |f_n| d\mu \rightarrow 0$, so we conclude that $f_n(y) \rightarrow 0$ for μ -almost all places $y \in Y$.

Now we show that for almost all $y \in Y$, we have that Δ_n meets the two conditions of Proposition 3.16, so we can conclude that $\Delta_n \rightarrow \rho$ for μ -a.e. place $y \in Y$. To see this, notice that as each Δ_n has finite height,

$$\int_Y \int_{A_y^1} \log^+ |z|_y d\Delta_n(z) d\mu(y) < \infty.$$

It follows that for a set of μ -almost all places which we denote as Y_n ,

$$\int_{\mathbb{A}_y^1} \log^+ |z|_y d\Delta_n(z) < \infty \quad \text{for all } y \in Y_n.$$

Since $\mu(Y \setminus Y_n) = 0$, we form the set

$$Y' = \bigcap_{n=1}^{\infty} Y_n,$$

and we can conclude that $\mu(Y \setminus Y') = 0$ as well. Let $Y'' = \{y \in Y : f_n(y) \rightarrow 0\}$. As we argued above, $\mu(Y \setminus Y'') = 0$, so if we let $Z = Y' \cap Y''$, then again, $\mu(Y \setminus Z) = 0$. For all $y \in Z$, Δ_n meets the conditions of Proposition 3.16, and so we can conclude that the set of y for which we get the desired equidistribution result has full μ -measure. \square

4. STOCHASTIC DYNAMICS

4.1. Stochastic dynamical heights. The main goal of this section is to establish that the stochastic height is in fact the height associated to a certain generalized adelic measure which depends on the family of maps S and the probability measure ν_1 associated to this family.

We will recall here the main results of Healey and Hindes [HH19] in the context of arithmetic dynamics. Let S be a finite or countably infinite set of rational maps defined over an algebraic closure $\overline{\mathbb{Q}}$, with each map being of degree at least 2. We view these as functions of the projective line over $\overline{\mathbb{Q}}$ to itself. Let ν_1 be a probability measure on S . For $n \in \mathbb{N}$, let ν_n be the product measure induced by ν_1 on S^n ; that is, for $\gamma_n = (\varphi_1, \dots, \varphi_n) \in S^n$, we have $\nu_n(\gamma_n) = \nu_1(\varphi_1) \cdots \nu_1(\varphi_n)$. For convenience, for each $\gamma_n = (\varphi_1, \dots, \varphi_n) \in S^n$ we will write $\gamma_n(\alpha)$ for the composition $\varphi_n \circ \cdots \circ \varphi_1(\alpha)$. Furthermore, we let ν denote the unique product measure on $S^{\mathbb{N}}$ which satisfies, for all $k \in \mathbb{N}$ and $A_1, \dots, A_k \subset S$, $\nu(A_1 \times A_2 \times \cdots \times A_k \times S \times S \times \cdots) = \prod_{i=1}^k \nu_1(A_i)$ defined on the smallest σ -algebra generated by sets of the form $A_1 \times \cdots \times A_k \times S \times S \times \cdots$. With this σ -algebra and measure we view $S^{\mathbb{N}}$ as the space of all independent identically distributed (i.i.d.) sequences in S . Given a sequence $\gamma = (\varphi_i)_{i=1}^{\infty} \in S^{\mathbb{N}}$ and positive integer n , we will denote by γ_n the composition $\gamma_n = \varphi_n \circ \cdots \circ \varphi_1$, and we will denote the degree of this composite map as

$$\deg(\gamma_n) = \prod_{i=1}^n \deg(\varphi_i).$$

As we will need to exclude these points from our equidistribution theorem, we define what it means for a point to be *exceptional* for the stochastic system S :

Definition 4.1. Let S be a countable set of rational maps defined over $\overline{\mathbb{Q}}$. We define the *exceptional set* of the system S to be the set E_S of all points $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$ such that the grand orbit

$$\mathcal{O}_S^{\pm}(\alpha) := \bigcup_{n=0}^{\infty} \bigcup_{\gamma_n \in S^n} (\gamma_n^{-1}(\alpha) \cup \{\gamma_n(\alpha)\})$$

is finite.

Remark 4.2.

- (1) While in classical dynamics the exceptional set is traditionally defined in terms of finiteness of the grand orbit, it is equivalent to require only that the backward orbit

$$\mathcal{O}_S^-(\alpha) := \bigcup_{n=0}^{\infty} \bigcup_{\gamma_n \in S^n} \gamma_n^{-1}(\alpha)$$

is finite. Indeed, any finite set which is backward invariant under a collection of surjective maps is necessarily forward invariant as well.

- (2) It is straightforward to see that the exceptional set for S must be contained in the exceptional set for every map $\varphi \in S$. In particular, this implies that $|E_S| \leq 2$.
- (3) On the other hand, a point α may be exceptional for every $\varphi \in S$ without being exceptional for S . For example, let $\varphi_1(z) = \frac{1}{z^2}$ and $\varphi_2(z) = z^2 + 1$, and take $S = \{\varphi_1, \varphi_2\}$. Then ∞ is exceptional for both maps φ_1 and φ_2 , but ∞ is not exceptional for S , since $\varphi_1^{-1}(\infty) = \{0\}$ and 0 is not exceptional for φ_2 .

Let $y \in Y$ be a place of $\overline{\mathbb{Q}}$. For each $\varphi \in S$ we can find a continuous function $g_{\varphi,y} : \mathbb{P}^1(\mathbb{C}_v) \rightarrow \mathbb{R}$ such that $g_{\varphi,y}$ is continuous and

$$(45) \quad \Delta g_{\varphi,y} = \frac{1}{\deg \varphi} \varphi^*(\lambda_y) - \lambda_y,$$

where λ_y denotes the equilibrium measure of the unit disc in \mathbf{A}_y^1 , which we remind the reader is the normalized Lebesgue measure of the unit circle in \mathbb{C} if y is archimedean, or the unit point mass at the Gauss point if y is non-archimedean. Such a $g_{\varphi,y}$ is not difficult to construct; for example, if we write $\varphi(z) = f(z)/g(z)$ where $f, g \in \overline{\mathbb{Q}}[z]$ are polynomials with no common factors, then

$$g_{\varphi,y}(z) = \frac{1}{\deg \varphi} \log \max\{|f(z)|_y, |g(z)|_y\} - \log^+ |z|_y$$

suffices, with the natural extension to the Berkovich line in the non-archimedean setting. Note that, depending on our choice of representation of $\varphi(z)$ as a quotient, our $g(z)$ functions may differ by a constant. We will choose to normalize our $g_{\varphi,y}(z)$ functions so that $g_{\varphi,y}(\infty) = 0$. In order to construct the canonical measure associated to φ at v , one ordinarily forms the telescoping series:

$$G_{\varphi,y}(z) = \sum_{n=0}^{\infty} \frac{1}{(\deg \varphi)^n} g_{\varphi,y}(\varphi^n(z)).$$

It follows that $G_{\varphi,y}(z)$ is also continuous on \mathbf{P}_y^1 and one can show (see, for example, [FRL06, Théorème 8]) that $\Delta G_{\varphi,y} = \mu_{\varphi,y} - \lambda_y$.

Each $g_{\varphi,y}(z)$ is continuous and bounded on \mathbf{P}_y^1 , but in order for our resulting system to be given by a generalized adelic measure, we will need an additional assumption on our family of maps. Define for each place $y \in Y$ and $\varphi \in S$

$$(46) \quad C_{\varphi}(y) = \sup_{z \in \mathbf{P}^1(\mathbb{C}_y)} |g_{\varphi,y}(z)|_y.$$

This exists, as each $g_{\varphi,y}$ is a continuous function on the compact space \mathbf{P}_y^1 . Notice that, as each map φ is defined over a single number field K and has good reduction at all but finitely many places of its base field, $C_{\varphi}(y)$ is a nonnegative simple function with compact support in Y . If we take the expected value over all maps φ

in our family according to our measure ν_1 , we obtain a function $C_S : Y \rightarrow \mathbb{R}$ given by

$$(47) \quad C_S(y) = \mathbb{E}_S C_\varphi(y) = \sum_{\varphi \in S} \nu_1(\varphi) C_\varphi(y).$$

Notice that, as an expectation of nonnegative compactly supported simple functions, $C_S(y)$ is a nonnegative measurable extended-real valued function. We are now ready to state our condition on the family of maps:

Definition 4.3. We say that our stochastic family of maps (S, ν_1) is L^1 height controlled if the associated function $C_S : Y \rightarrow \mathbb{R}$ defined above satisfies

$$(48) \quad \int_Y C_S(y) d\mu(y) < \infty.$$

Notice that if a family is L^1 height controlled, then it is also height controlled in the sense of Healey and Hindes [HH19]. Healey and Hindes [HH19, Thm. 1.2] defined, under this assumption, the *stochastic height* h_S associated to the system (S, ν_1) as

$$h_S(\alpha) = \lim_{n \rightarrow \infty} \mathbb{E}_{S^n} \frac{1}{\deg \gamma_n} h_{\gamma_n}(\alpha) \quad \text{for all } \alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}),$$

where we recall that for a rational map ψ we denote by h_ψ the canonical height associated to ψ . In the next theorem, we will show that h_S is really a height associated to a generalized adelic measure associated to (S, ν_1) :

Theorem 4.4. *Let S be a finite or countably infinite set of rational maps defined over an algebraic closure $\overline{\mathbb{Q}}$, with each map being of degree at least 2, and let ν_1 be a given probability measure on S . If the maps in S are L^1 height controlled, then there exists a unique generalized adelic measure ρ such that the stochastic height associated to (S, ν_1) is equal to h_ρ .*

Proof. We will construct ρ by defining ρ_y at almost all places $y \in Y$, and checking that the resulting measure meets the desired conditions. We start by constructing the Green's function associated to the stochastic height at y , continuing our construction from above. We let, for each $y \in Y$,

$$(49) \quad g_{1,y}(z) = \mathbb{E}_S g_{\varphi,y}(z) = \sum_{\varphi \in S} \nu_1(\varphi) g_{\varphi,y}(z),$$

where the expectation is taken over all φ in our probability space (S, ν_1) . Let $C(y)$ be the function given by (47) above. Notice that

$$(50) \quad \sup_{z \in \mathbb{P}_y^1} |g_{1,y}(z)| \leq \mathbb{E}_S C_\varphi(y) = C_S(y).$$

By our assumption that our family is L^1 height controlled, we know that the set of places y where $C_S(y) = \infty$ is of μ -measure 0. Let $y \in Y$ be a place where $C_S(y) < \infty$. Then from the condition $C_S(y) < \infty$, it follows from the Weierstrass M -test that $g_{1,y}(z)$, is a continuous function $\mathbb{P}^1(\mathbb{C}_y) \rightarrow \mathbb{R}$ as well. We now recursively define, for each $n \geq 1$, the functions

$$(51) \quad g_{n+1,y}(z) = \int_{S^n} \frac{1}{\deg(\gamma_n)} g_{1,y}(\gamma_n(z)) d\nu_n(\gamma_n) = \mathbb{E}_{S^n} \frac{g_{1,y}(\gamma_n(z))}{\deg(\gamma_n)}.$$

Let us recall (see for example the discussion in [FRL06, §6.1] or [BR10, Prop. 9.54]) that for any potential function $g(z)$,

$$(52) \quad \varphi^*(\Delta g) = \Delta(g \circ \varphi).$$

It follows that

$$(53) \quad \Delta g_{n+1,y} = \mathbb{E}_{S^n} \frac{\gamma_n^*(\Delta g_{1,y})}{\deg(\gamma_n)},$$

where the expectation is taken in (S^n, ν_n) and we denote the random variable by γ_n . Combining with (45) and using the linearity of the pullback, we see that:

$$(54) \quad \begin{aligned} \Delta g_{n+1,y} &= \mathbb{E}_{S^n} \mathbb{E}_S \frac{\gamma_n^*(\varphi^*(\lambda_y))}{\deg(\gamma_n) \deg(\varphi)} - \mathbb{E}_{S^n} \frac{\gamma_n^*(\lambda_y)}{\deg(\gamma_n)} \\ &= \mathbb{E}_{S^{n+1}} \frac{\gamma_{n+1}^*(\lambda_y)}{\deg(\gamma_{n+1})} - \mathbb{E}_{S^n} \frac{\gamma_n^*(\lambda_y)}{\deg(\gamma_n)} \\ &= \rho_{n+1,y} - \rho_{n,y}, \end{aligned}$$

where we set

$$(55) \quad \rho_{n,y} = \mathbb{E}_{S^n} \frac{\gamma_n^*(\lambda_y)}{\deg(\gamma_n)} \quad \text{for each } n \geq 1, \quad \text{and } \rho_{0,y} = \lambda_y.$$

Since the pullback of a positive Borel measure ρ under a rational map φ has mass $\varphi^*(\rho)(\mathbb{P}_y^1) = \deg(\varphi) \cdot \rho(\mathbb{P}_y^1)$, we have that each $\rho_{n,y}$ is a probability measure, and thus $\Delta g_{n+1,y}$ is the difference of two probability measures.

As our initial $g_{1,y}$ function is bounded, it follows that the $g_{n,y}$ functions are also bounded. Indeed,

$$|g_{2,y}(z)| \leq \mathbb{E}_S \frac{|g_{1,y}(\varphi(z))|}{\deg \varphi} \leq C_S(y) \mathbb{E}_S \frac{1}{\deg \varphi} = \frac{C_S(y)}{\delta_S},$$

where we define, following Healey and Hindes, the *stochastic degree* δ_S to be the harmonic mean of the degrees of our rational maps:

$$(56) \quad \delta_S = \left(\mathbb{E}_S \frac{1}{\deg \varphi} \right)^{-1} \geq 2,$$

where the last inequality follows from the fact that $\deg \varphi \geq 2$ for all $\varphi \in S$ by assumption. It follows by induction using (53) that

$$(57) \quad |g_{n,y}(z)| \leq \frac{C_S(y)}{\delta_S^{n-1}}$$

for all $n \geq 1$. Finally, we will define our stochastic telescoping series to be:

$$(58) \quad g_{S,y}(z) = \sum_{n=1}^{\infty} g_{n,y}(z).$$

By the Weierstrass M -test and (57), for all $y \in Y$ for which $C_S(y) < \infty$, this series converges uniformly to a continuous function $g_{S,y} : \mathbb{P}_y^1 \rightarrow \mathbb{R}$. Further, if we define

$$(59) \quad C(y) = \sup_{z \in \mathbb{P}_y^1} |g_{S,y}(z)|,$$

then by our telescoping construction,

$$C(y) \leq 2C_S(y)$$

for every $y \in Y$, and so since we assumed our family (S, ν_1) was L^1 height controlled, we have that $C_S \in L^1(Y)$, so $C \in L^1(Y)$, and $C(y) < \infty$ is true for μ -almost every $y \in Y$, which is what we desired.

We now proceed to prove that, when restricted to $y \in Y(\mathbb{Q}, p)$, the function $g_S(y, z) = g_{S,y}(z)$ defines a measurable function

$$g_S : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p) \rightarrow \mathbb{R}$$

for every rational prime $p \in M_{\mathbb{Q}}$. To see this, note that the functions $g_{\varphi}(y, z) = g_{\varphi,y}(z) : Y(\mathbb{Q}, p) \times \mathbb{P}^1(\mathbb{C}_p)$ are measurable (in fact, continuous), as any given φ is defined over a single number field, so they are naturally locally constant in the y variable over the sets $\{Y(K, v) : v \in M_K, v \mid p\}$, and their potential functions are well-known to be continuous. It follows that, as a limit of measurable functions, $g_{1,y}$ is measurable in both variables, and thus likewise so are $g_{n,y}$ and g_S .

We now define the measures ρ_y for each place $y \in Y$ by

$$(60) \quad \Delta g_{S,y}(z) = \rho_y - \lambda_y.$$

This is well-defined, as the negative part of $\Delta g_{S,y}(z)$ in the Jordan decomposition is indeed λ_y , since from the first construction of $g_{\varphi,y}$ in equation (45), the negative part of the measure was λ_y . As we formed the telescoping series, the negative part remained λ_y , and as we took the expectation over φ , it remained so as well. Since the associated potential functions g_S meet the criteria of Definition 2.1, it follows that $\rho = (\rho_y)_{y \in Y}$ is a generalized adelic measure. Further, it is clear that

$$(61) \quad \Delta \left(\sum_{n=1}^N g_{n,y} \right) = \rho_{N,y} - \lambda_y,$$

and since $\sum_{n=1}^N g_{n,y}(z) \rightarrow g_{S,y}(z)$ uniformly as $N \rightarrow \infty$, it follows (e.g., from [BR10, Prop. 5.32]) that $\rho_{N,y} \rightarrow \rho_y$ in the weak sense of convergence of measures.

Lastly, we will establish that $h_{\rho} = h_S$. To see this, recall that for a rational function φ defined over a number field K , if we pull back the standard adelic measure λ *once* by the map φ , we obtain an adelic measure (*not* yet the canonical measure associated to φ , in general) given by

$$\rho = \frac{\varphi^*(\lambda)}{\deg \varphi},$$

and we have

$$(62) \quad h_{\rho}(\alpha) = \frac{1}{\deg \varphi} h(\varphi(\alpha))$$

for all $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$. In other words, the effect of pulling back once by φ and normalizing by the degree of φ was to obtain the normalized standard height of $\varphi(\alpha)$ instead of α . If we took more pullbacks under φ , we would obtain in the limit the canonical φ -height of α . In our setting, it follows by taking a limit over the finite subsets of S^n ordered under inclusion that the resulting height is still defined over a number field, and our formula from (62) applies, that the generalized adelic measures $\rho_n = (\rho_{n,y})_{y \in Y}$ satisfy

$$h_{\rho_n}(\alpha) = \mathbb{E}_{S^n} \frac{1}{\deg \gamma_n} h_{\gamma_n}(\alpha) \quad \text{for all } \alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}),$$

and so we get that $h_{\rho} = h_S$, as claimed. Unicity follows easily from the equidistribution result proven below in Theorem 4.8. \square

4.2. An example with places without continuous potentials. A key fact that comes out of the proof of our equidistribution theorem for heights associated to generalized adelic measures is that it holds at the places where we have a continuous potential function $g_y : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ with $\Delta g_y = \rho_y - \lambda_y$. Our hypothesis that our stochastic family is L^1 height controlled guarantees that this holds at almost every place; however, it is possible to construct families where there is a nonempty set of places, albeit still of μ -measure zero, for which the potential functions are not bounded, and therefore not continuous, and therefore for which our equidistribution results will *not* apply. Note that this phenomenon cannot occur with adelic measures and quasi-adelic measures, as both of these constructions are defined over a single given number field, and there are no places of μ -measure zero in that context, as each place v of a number field K has μ -measure

$$\mu(Y(K, v)) = \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} > 0.$$

We now construct an example of a stochastic dynamical system where there is a nonempty, measure zero set of places for which this continuous potential function will not exist, and in particular, for which our equidistribution results will not apply.

Example 4.5. Let τ_n be a sequence of Salem numbers such that $\tau_n \searrow \theta_0$ as $n \rightarrow \infty$, where θ_0 is the real root of $x^3 - x - 1$. That such a sequence exists is consequence of the fact that θ_0 is a Pisot-Vijayaraghavan number and a result of Salem [Sal45, Theorem IV]. By Northcott's theorem, as the heights of the τ_n are bounded (recall that for a Salem number $\tau > 1$, $h(\tau) = (\log \tau)/[\mathbb{Q}(\tau) : \mathbb{Q}]$), we must have $[\mathbb{Q}(\tau_n) : \mathbb{Q}] \rightarrow \infty$. Therefore we can choose τ_n satisfying $[\mathbb{Q}(\tau_n) : \mathbb{Q}] \geq 3^n$ for every n , and we will assume that we have done so.

Let $K_n = \mathbb{Q}(\tau_1, \dots, \tau_n)$. By our choice of $\tau_n > 1$ we can assume that for each $n \in \mathbb{N}$, there is a place v_n of K_n such that $|\tau_i|_{v_n} > 1$ for each $1 \leq i \leq n$, and that $Y(K_{n+1}, v_{n+1}) \subset Y(K_n, v_n)$. Notice that, as an intersection of nested nonempty compact sets in a profinite space,

$$N := \bigcap_{n=1}^{\infty} Y(K_n, v_n) \neq \emptyset.$$

We will show that at all of the places of N , $\rho_{S,y}$ will fail to admit a continuous potential for the stochastic dynamical system which we will define. Notice that as there is just one real place of $\mathbb{Q}(\tau_n)$ where the absolute value is positive, we have

$$\mu(Y(K_n, v_n)) \leq \frac{1}{[\mathbb{Q}(\tau_n) : \mathbb{Q}]} \leq \frac{1}{3^n},$$

so in particular, it follows that $\mu(N) = 0$.

Define $d_n = [\mathbb{Q}(\tau_n) : \mathbb{Q}]$ and let $\alpha_n = \tau_n^{d_n}$. Notice that

$$h(\alpha_n) = \log \tau_n \searrow \log \theta_0 \quad \text{as } n \rightarrow \infty.$$

Define for $n \in \mathbb{N}$ the maps

$$\varphi_n(z) = \alpha_n^2 z^2.$$

Let $S = \{\varphi_n : n \in \mathbb{N}\}$, and endow S with the probability measure given by $\nu_1(\varphi_n) = 1/2^n$. We will show that the family (S, ν_1) is an L^1 height controlled family, and so there exists an adelic measure ρ_S such that for μ -almost every $y \in Y$, there exists a continuous function $g_y : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ such that $\Delta g_y = \rho_{S,y} - \lambda_y$.

As α_n is an algebraic unit, φ_n has good reduction at all finite places. Thus, $\rho_{S,y} = \lambda_y$ for all $y \nmid \infty$, and g_y is the identically 0 function at all finite y , and hence $C(y) = 0$ as well at these places. Now consider $y \in Y(\mathbb{Q}, \infty)$. Notice that if we define

$$p(\gamma_k, z) = \frac{1}{\deg(\gamma_k)} \log^+ |\gamma_k(z)|_y - \log^+ |z|_y,$$

then

$$\Delta p(\gamma_k, z) = \frac{\gamma_k^*(\lambda_y)}{\deg(\gamma_k)} - \lambda_y.$$

It follows that, up to a constant $c \in \mathbb{R}$,

$$g_y(z) = c + \lim_{k \rightarrow \infty} \mathbb{E}_{S^k} p(\gamma_k, z),$$

where the expectation is taken over $\gamma_k \in S^k$. We say *up to a constant* because, as the reader will recall, we assume that $g_y(z)$ is normalized at every place so that $g_y(z) \rightarrow 0$ as $z \rightarrow \infty$, but as we will observe, that requires here adjusting our choice of p function. To see this, note that if $\gamma_k = (\varphi_{n_1}, \dots, \varphi_{n_k})$, then

$$p(\gamma_k, z) = \log^+ |\alpha_{n_k}^{1/2^{k-1}} \cdots \alpha_{n_2}^{1/2} \alpha_{n_1} z|_y - \log^+ |z|_y,$$

so in order to assume that $p(\gamma_k, z) \rightarrow 0$ as $z \rightarrow \infty$, we define

$$p_0(\gamma_k, z) = p(\gamma_k, z) - \log |\alpha_{n_k}^{1/2^{k-1}} \cdots \alpha_{n_2}^{1/2} \alpha_{n_1}|_y,$$

and then we see that $g_y(z) = \lim_{k \rightarrow \infty} \mathbb{E}_{S^k} p_0(\gamma_k, z)$ and g_y is normalized as we desired. It easy to check that for $c > 0$, the function $|\log^+ |cz|_y - \log^+ |z|_y - \log c|$ has maximum $|\log c|$, and so it follows that

$$\sup_{z \in \mathbb{P}_y^1} |p_0(\gamma_k, z)| = |p_0(\gamma_k, 0)| = |\log |\alpha_{n_k}^{1/2^{k-1}} \cdots \alpha_{n_2}^{1/2} \alpha_{n_1}|_y|.$$

Now as each component of γ_k is chosen independently in S , it follows that

$$\mathbb{E}_{S^k} |p_0(\gamma_k, 0)| = \frac{2^k - 1}{2^k} \sum_{n=1}^{\infty} \nu_1(\varphi_n) \log |\alpha_n|_y = \frac{2^k - 1}{2^k} \sum_{n=1}^{\infty} \frac{1}{2^n} |\log |\alpha_n|_y|.$$

So in particular, we get that

$$(63) \quad C(y) = \sup_{z \in \mathbb{P}_y^1} |g_y(z)| = |g_y(0)| \leq \sum_{n=1}^{\infty} \frac{1}{2^n} |\log |\alpha_n|_y|.$$

Notice that for all $y \in N$, we have $|p_0(\gamma_k, 0)| = -p_0(\gamma_k, 0)$ for every γ_k , as $|\alpha_n|_y > 1$ for all $n \geq 1$ and $y \in N$, so in fact

$$-g_y(0) = \sum_{n=1}^{\infty} \frac{1}{2^n} \log |\alpha_n|_y \geq \sum_{n=1}^{\infty} \frac{1}{2^n} \log d_n \log \tau_n \geq \sum_{n=1}^{\infty} \frac{1}{2^n} \log 3^n \log \theta_0 = \infty,$$

so our potential functions cannot be bounded for the places $y \in N$.

Now it remains to show that $C(y)$ is still an integrable function on $Y(\mathbb{Q}, \infty)$ in order to prove that our family is still integrable. To see this, notice that by the product formula,

$$\int_{Y(\mathbb{Q}, \infty)} \log |\alpha_n|_y d\mu(y) = 0,$$

and thus, as our α_n is an algebraic unit,

$$\int_{Y(\mathbb{Q}, \infty)} |\log |\alpha_n|_y| d\mu(y) = 2h(\alpha_n)$$

(cf. [AV09, Equation 1.3]). It follows that

$$\begin{aligned} \int_Y C(y) d\mu(y) &\leq \sum_{n=1}^{\infty} \frac{1}{2^n} \int_{Y(\mathbb{Q}, \infty)} |\log |\alpha_n|_y| d\mu(y) \\ &= \sum_{n=1}^{\infty} \frac{2h(\alpha_n)}{2^n} \leq 2 \sup_n \log \tau_n, \end{aligned}$$

but this quantity is bounded as $\tau_n \searrow \theta_0$. It follows that (S, ν_1) is L^1 height controlled as claimed.

The reader may find it interesting to note as well that the Borel-Cantelli lemma implies that for μ -almost all $y \in Y(\mathbb{Q}, \infty)$, $|\alpha_n|_y = 1$ for all but finitely n . We note that our set N lies in the exceptional set where infinitely many $|\alpha_n|_y > 1$.

4.3. Pullback formula for stochastic heights. The stochastic generalized adelic measure shares some, but not all, of the properties of the usual dynamical measure associated to iteration of a single rational map. For example, if $\mu_{\varphi, y}$ denotes the canonical measure of φ at a place y , then it is well-known (cf. [BR10, Theorem 10.2]) that

$$\mu_{\varphi, y} = \frac{\varphi^*(\mu_{\varphi, y})}{\deg \varphi} \quad \text{and} \quad \varphi_*(\mu_{\varphi, y}) = \mu_{\varphi, y}.$$

We will prove a similar statement for the expected pullback of the stochastic canonical measure. However, the analogous stochastic pushforward formula does *not* hold in general for stochastic heights, and this is tied to the fact that the stochastic Fatou and Julia sets are not fully invariant under the maps $\varphi \in S$, which will be proved in the forthcoming paper [DFT].

Theorem 4.6 (Pullback formula for stochastic heights). *Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2, and let ρ denote the canonical generalized adelic measure associated to S . Then for μ -almost every place $y \in Y$,*

$$\rho_y = \mathbb{E}_S \frac{\varphi^*(\rho_y)}{\deg \varphi},$$

where the expectation is taken over all maps $\varphi \in S$ with respect to the measure ν_1 .

We note that, as usual, in the case that ρ is defined over a single number field K , the μ -almost every place condition becomes every place of K , as every place v of K gives rise to a set $Y(K, v)$ of positive μ measure.

Proof. Recall from above that we defined in (58) the Green's functions $g_{S, y}$, whose Laplacians satisfy $\Delta g_{S, y}(z) = \rho_y - \lambda_y$ at each place, provided that $C(y) < \infty$ at that place. As $C \in L^1(Y)$ by assumption that S is L^1 height controlled, this condition holds for μ -almost every place y . We will use the same notation as in the proof of Theorem 4.4, that is, the partial sum of $g_{S, y}(z)$ is denoted

$$g_{S, y, N}(z) = \sum_{n=1}^N g_{n, y}(z),$$

where for each $n \geq 1$, $g_{n,y}$ has Laplacian

$$\Delta g_{n,y} = \rho_{n,y} - \rho_{n-1,y},$$

where $\rho_{0,y} = \lambda_y$ and $\rho_{n,y}$ is defined in (55) as the expected pullback measure of λ_y under the maps $\gamma \in S^N$ weighted by $1/\deg(\gamma)$. It follows that the partial sum forms a telescoping series and has Laplacian

$$\Delta g_{S,y,N} = \rho_{N,y} - \lambda_y.$$

Notice that, as occurred above in the computation of (54),

$$\mathbb{E}_S \frac{\varphi^*(\rho_n)}{\deg \varphi} = \rho_{n+1,y},$$

so it follows that

$$\mathbb{E}_S \frac{\varphi^*(\Delta g_{S,y,N})}{\deg \varphi} = \rho_{N+1,y} - \rho_{1,y}.$$

Taking the limit as $N \rightarrow \infty$, we see that

$$(64) \quad \mathbb{E}_S \frac{\varphi^*(\Delta g_{S,y})}{\deg \varphi} = \rho_y - \rho_{1,y}.$$

But as $\Delta g_{S,y} = \rho_y - \lambda_y$, we have

$$(65) \quad \begin{aligned} \mathbb{E}_S \frac{\varphi^*(\Delta g_{S,y})}{\deg \varphi} &= \mathbb{E}_S \frac{\varphi^*(\rho_y)}{\deg \varphi} - \mathbb{E}_S \frac{\varphi^*(\lambda_y)}{\deg \varphi} \\ &= \mathbb{E}_S \frac{\varphi^*(\rho_y)}{\deg \varphi} - \rho_{1,y}. \end{aligned}$$

Combining equations (64) and (65), we see that

$$(66) \quad \mathbb{E}_S \frac{\varphi^*(\rho_y)}{\deg \varphi} = \rho_y,$$

which is what we wanted to show. \square

4.4. Stochastic dynamical equidistribution. We are now ready to state our main equidistribution theorem. We start by defining the random backwards orbit measure of a point:

Definition 4.7. Let $\alpha \in \overline{\mathbb{Q}}$. We define the *random backwards orbit measure* of α under the stochastic dynamical system (S, ν_1) recursively by $\Delta_{0,\alpha} = \delta_\alpha$, and

$$\Delta_{n+1,\alpha} = \mathbb{E}_S \left(\frac{\varphi^*(\Delta_n)}{\deg(\varphi)} \right) \quad \text{for } n \geq 0.$$

Here, δ_z denotes the Dirac point mass at z , and $\varphi^*(\Delta)$ denotes the pullback of the measure Δ under φ . Note that for a point mass δ_z we have

$$\varphi^*(\delta_z) = \sum_{w \in \varphi^{-1}(z)} \delta_w,$$

where the w in the sum are counted with multiplicity, so that we must divide by the degree of φ to obtain a new probability measure.

Notice that, so long as the measure ν_1 is strictly positive, the support of $\Delta_{n,\alpha}$ is

$$\text{supp}(\Delta_{n,\alpha}) = \bigcup_{\gamma_n \in S^n} \gamma_n^{-1}(\alpha),$$

and the weights come from the probability of the individual likelihood of each γ_n in S^n under the product measure, along with the multiplicities as preimages under γ_n .

Theorem 4.8 (Stochastic dynamical equidistribution). *Let S be a countable set of rational maps defined over an algebraic closure $\overline{\mathbb{Q}}$, with each map being of degree at least 2, and let ν_1 be a probability measure on S with respect to which the maps in S are L^1 height controlled. Suppose $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}) \setminus E_S$ where E_S denotes the exceptional set of the stochastic dynamical system. Then for μ -almost every place y of $\overline{\mathbb{Q}}$, the backwards orbit measures $\Delta_{n,\alpha}$ under (S, ν_1) converge weakly to the stochastic dynamical measure ρ_y as $n \rightarrow \infty$.*

The substance of the theorem lies in proving that the $\Delta_{n,\alpha}$ measures are well-distributed and have $h_S(\Delta_{n,\alpha}) \rightarrow 0$. The equidistribution result then follows from applying Theorem 3.17 above, the equidistribution theorem for generalized adelic measures.

Before beginning our proof of stochastic dynamical equidistribution, we need to prove a few results. We start with a result which generalizes the main conclusion of Theorem 1.2(b) from [HH19].

Proposition 4.9. *Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2. Then for any discrete probability measure Δ on $\mathbb{P}^1(\overline{\mathbb{Q}})$ of finite height,*

$$\mathbb{E}_S \frac{h_S(\varphi_*(\Delta))}{\deg \varphi} = h_S(\Delta),$$

where $\varphi_*(\Delta)$ denotes the pushforward measure.

We remind the reader that the pushforward measure of a discrete measure

$$\Delta = \sum_i t_i \delta_{\alpha_i}$$

where $t_i \geq 0$, $\sum t_i = 1$, and $\alpha_i \in \mathbb{P}^1(\overline{\mathbb{Q}})$, is given by

$$\varphi_*(\Delta) = \sum_i t_i \delta_{\varphi(\alpha_i)}.$$

Proof. By Proposition 3.4, it suffices to prove the result for $\Delta = \delta_\alpha$ for $\alpha \in \overline{\mathbb{Q}}$. (We note in passing that $h(\delta_\alpha)$ is indeed equal to the usual height of α , see Proposition 2.13 above.) Recall that

$$h_S(\alpha) = \lim_{n \rightarrow \infty} \int_{S^n} \frac{h(\gamma_n(\alpha))}{\deg \gamma_n} d\nu_n(\gamma_n).$$

For $n \geq 2$, each $\gamma_n \in S^n$ can be written as $\gamma_n = \gamma_{n-1} \circ \varphi$ where $\gamma_{n-1} \in S^{n-1}$ and $\varphi \in S$. By the absolute convergence of the series involved,

$$\begin{aligned} h_S(\alpha) &= \lim_{n \rightarrow \infty} \int_{S^{n-1}} \int_S \frac{h(\gamma_{n-1} \circ \varphi(\alpha))}{\deg(\gamma_{n-1}) \deg(\varphi)} d\nu_1(\varphi) d\nu_{n-1}(\gamma_{n-1}) \\ &= \int_S \frac{1}{\deg(\varphi)} \lim_{n \rightarrow \infty} \int_{S^{n-1}} \frac{h(\gamma_{n-1} \circ \varphi(\alpha))}{\deg(\gamma_{n-1})} d\nu_{n-1}(\gamma_{n-1}) d\nu_1(\varphi) \\ &= \int_S \frac{1}{\deg(\varphi)} h_S(\varphi(\alpha)) d\nu_1(\varphi) \\ &= \mathbb{E}_S \frac{h_S(\varphi_*(\alpha))}{\deg \varphi}. \quad \square \end{aligned}$$

Before we move on, we state for its independent interest a restatement of Proposition 4.9 in the case where $\Delta = \delta_\alpha$:

Corollary 4.10. *If (S, ν_1) is an L^1 height controlled family of dynamical rational maps, each of degree at least 2, and $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$, then the stochastic height $h_S(\alpha)$ satisfies*

$$(67) \quad h_S(\alpha) = \mathbb{E}_S \frac{h_S(\varphi(\alpha))}{\deg \varphi}.$$

We note that, in the case where the family S meets the conditions of the [HH19] paper, this corollary is essentially equivalent to the $k = 1$ case of their Theorem 1.2(b). We state it primarily because it is conceptually important and aesthetically pleasing as the stochastic analogue of the usual scaling property satisfied by the dynamical height, namely, that

$$h_\varphi(\alpha) = \frac{h_\varphi(\varphi(\alpha))}{\deg \varphi}.$$

Lemma 4.11. *Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2. Let Δ be a discrete measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$ with finite height. Then*

$$\Gamma = \mathbb{E}_S \left(\frac{\varphi^*(\Delta)}{\deg(\varphi)} \right)$$

is a discrete measure on $\mathbb{P}^1(\overline{\mathbb{Q}})$ with finite height as well, and there exists a constant $C \geq 0$ (independent of Δ) such that

$$h(\Gamma) \leq \frac{1}{2} h(\Delta) + C.$$

Proof. We begin by recalling the notation of Theorem 3.1 above. For each $\varphi \in S$ we have an generalized adelic measure ρ_φ defined over the field of definition of φ . In particular, for each $y \in Y$ there exists a continuous function $g_y : \mathbb{P}_y^1 \rightarrow \mathbb{R}$ such that $\Delta g_{\varphi, y} = \rho_{\varphi, y} - \lambda_y$, and further, the function

$$C_\varphi(y) = \sup_{z \in \mathbb{P}_y^1} |g_{\varphi, y}(z)|$$

is compactly supported and constant on the sets $Y(K, v)$, where K is a number field over which φ is defined and v ranges over the places of K . In particular,

$C_\varphi \in L^1(Y)$. In fact, we proved in equation (21) that for all discrete measures Δ on $\overline{\mathbb{Q}}$, we have

$$|h_\varphi(\Delta) - h(\Delta)| \leq 3 \int_Y C_\varphi(y) d\mu(y).$$

Further, the condition that S is L^1 height bounded states that the function

$$C(y) = \mathbb{E}_S C_\varphi(y)$$

is also in $L^1(Y)$. Now,

$$h_\varphi(\varphi^*(\Delta)) = h_\varphi(\Delta).$$

(Recall that while $h_\varphi(\beta) = (1/\deg(\varphi))h_\varphi(\alpha)$ for any $\beta \in \varphi^{-1}(\alpha)$, the pullback measure has total mass $\deg(\varphi)$ as it weights each point in the pullback equally with multiplicity.) It follows from Proposition 3.4 that

$$\begin{aligned} h(\Gamma) &= \mathbb{E}_S \left(\frac{h(\varphi^*(\Delta))}{\deg(\varphi)} \right) \leq \mathbb{E}_S \frac{1}{\deg(\varphi)} \left(h_\varphi(\varphi^*(\Delta)) + 3 \int_Y C_\varphi d\mu \right) \\ &= \mathbb{E}_S \frac{1}{\deg(\varphi)} \left(h_\varphi(\Delta) + 3 \int_Y C_\varphi d\mu \right) \\ &\leq \mathbb{E}_S \frac{1}{\deg(\varphi)} \left(h(\Delta) + 6 \int_Y C_\varphi d\mu \right) \\ &\leq \frac{1}{2} h(\Delta) + 3 \int_Y C(y) d\mu(y), \end{aligned}$$

where we have used the fact that $\deg(\varphi) \geq 2$ for every $\varphi \in S$, which completes the proof of the claim with constant $C = 3 \int_Y C(y) d\mu(y)$. \square

Proposition 4.12. *Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2. For any $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$, the backwards orbit measures $\Delta_{n,\alpha}$ have finite height in the sense of Definition 2.3 above.*

Proof. It suffices to prove finite height with respect to the classical height, as by Corollary 3.2, a discrete $\mathbb{P}^1(\overline{\mathbb{Q}})$ measure has finite height with respect to one generalized adelic measure ρ if and only if it has finite height for all such heights. Note that $\Delta_{0,\alpha} = \delta_\alpha$ has finite height $h(\Delta_{0,\alpha}) = h(\alpha)$. Suppose that $\Delta_{n,\alpha}$ has finite height. Then by applying Lemma 4.11, we see that

$$h(\Delta_{n+1,\alpha}) \leq \frac{1}{2} h(\Delta_n(\alpha)) + C$$

for C a constant which depends only on the family (S, ν_1) , but not on α . It follows that each $\Delta_{n+1,\alpha}$ has finite height, as claimed. \square

Indeed, it follows from the proof above that the limit supremum of the standard heights of a backwards orbit is bounded, but we will prove in the course of our equidistribution theorem that in fact $h_S(\Delta_{n,\alpha}) \rightarrow 0$ as $n \rightarrow \infty$.

Proposition 4.13. *Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2. For any $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}) \setminus E_S$ not in the exceptional set of S , the backwards orbit measures $\Delta_{n,\alpha}$ are well-distributed.*

We refer the reader to Definition 4.1 above for the definition of the exceptional set of S .

In order to prove Proposition 4.13, we require a few lemmas. For the remainder of this section, suppose that S is as in Proposition 4.13 and for convenience let $\Delta_n = \Delta_{n,\alpha}$. For a rational function ψ , we define $e_z(\psi)$ to be ramification index of ψ at z .

Lemma 4.14. *Let S, α and $\Delta_n = \Delta_{n,\alpha}$ be as in Proposition 4.13 above. Assume $n \geq 0$ is fixed. Then for each $0 \leq k \leq n$, we have*

$$\Delta_n(z) = \sum_{\gamma_k \in S^k} \nu_k(\gamma_k) \frac{e_z(\gamma_k)}{\deg \gamma_k} \Delta_{n-k}(\gamma_k(z)).$$

Proof. The proof is by induction on n . Note that the case $n = 0$ is trivial, as we identify S^0 as consisting of just the identity map and with ν_0 as the trivial probability measure. Moreover, for any $n \geq 0$ the case $k = 0$ is trivial for the same reason, so we will henceforth assume $k \geq 1$. Now let $n \geq 1$. We have

$$\begin{aligned} \Delta_n(z) &= \mathbb{E}_S \left(\frac{\varphi^*(\Delta_{n-1})(z)}{\deg \varphi} \right) \\ &= \sum_{\varphi \in S} \nu_1(\varphi) \frac{\varphi^*(\Delta_{n-1})(z)}{\deg \varphi} \\ &= \sum_{\varphi \in S} \frac{\nu_1(\varphi)}{\deg \varphi} \sum_{x \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{n-1}(x) \sum_{w \in \varphi^{-1}(x)} e_w(\varphi) \delta_w(z) \\ &= \sum_{\varphi \in S} \nu_1(\varphi) \frac{e_z(\varphi)}{\deg \varphi} \Delta_{n-1}(\varphi(z)). \end{aligned}$$

By induction, for all $w \in \mathbb{P}^1(\overline{\mathbb{Q}})$ we have for $1 \leq k \leq n$ that

$$\Delta_{n-1}(w) = \sum_{\gamma_{k-1} \in S^{k-1}} \nu_{k-1}(\gamma_{k-1}) \frac{e_w(\gamma_{k-1})}{\deg \gamma_{k-1}} \Delta_{n-k}(\gamma_{k-1}(w));$$

applying this to $w = \varphi(z)$ with $\varphi \in S$ yields

$$\begin{aligned} \Delta_n(z) &= \sum_{\varphi \in S} \nu_1(\varphi) \frac{e_z(\varphi)}{\deg \varphi} \sum_{\gamma_{k-1} \in S^{k-1}} \nu_{k-1}(\gamma_{k-1}) \frac{e_{\varphi(z)}(\gamma_{k-1})}{\deg \gamma_{k-1}} \Delta_{n-k}(\gamma_{k-1}(\varphi(z))) \\ &= \sum_{\varphi \in S} \sum_{\gamma_{k-1} \in S^{k-1}} \nu_1(\varphi) \nu_{k-1}(\gamma_{k-1}) \frac{e_z(\varphi)}{\deg \varphi} \frac{e_{\varphi(z)}(\gamma_{k-1})}{\deg \gamma_{k-1}} \Delta_{n-k}(\gamma_{k-1}(\varphi(z))), \end{aligned}$$

and by writing $\gamma_k = \gamma_{k-1} \circ \varphi \in S^k$ we get

$$\Delta_n(z) = \sum_{\gamma_k \in S^k} \nu_k(\gamma_k) \frac{e_z(\gamma_k)}{\deg \gamma_k} \Delta_{n-k}(\gamma_k(z)). \quad \square$$

Lemma 4.15. *Let S, α and $\Delta_n = \Delta_{n,\alpha}$ be as in Proposition 4.13 above. Suppose $e_z(\gamma_3) = \deg \gamma_3$ for all $\gamma_3 \in S^3$. Then $z \in E_S$.*

Proof. Define the set

$$\mathcal{R}_S = \{x \in \mathbb{P}^1(\overline{\mathbb{Q}}) : e_x(\varphi) = \deg \varphi \text{ for all } \varphi \in S\}.$$

By the Riemann-Hurwitz theorem, we have $|\mathcal{R}_S| \leq 2$.

Let $\varphi_1, \varphi_2, \varphi_3 \in S$ be arbitrary. Taking $\gamma_3 = \varphi_3 \circ \varphi_2 \circ \varphi_1$, the equality $e_z(\gamma_3) = \deg \gamma_3$ implies that

$$e_z(\varphi_1) = \deg \varphi_1, \quad e_{\varphi_1(z)}(\varphi_2) = \deg \varphi_2, \quad \text{and} \quad e_{\varphi_2(\varphi_1(z))}(\varphi_3) = \deg \varphi_3.$$

Since $\varphi_1, \varphi_2, \varphi_3$ were arbitrary, it follows that

- (1) $z \in \mathcal{R}_S$;
- (2) $\varphi(z) \in \mathcal{R}_S$ for all $\varphi \in S$; and
- (3) $\varphi_2(\varphi_1(z)) \in \mathcal{R}_S$ for all $\varphi_1, \varphi_2 \in S$.

Suppose that $\varphi(z) = z$ for all $\varphi \in S$. Then, since φ is totally ramified at z , we have $\varphi^{-1}(z) = \{z\}$ for all $\varphi \in S$, and therefore $z \in E_S$.

Now suppose there exists $\varphi_1 \in S$ such that $w := \varphi_1(z) \neq z$. By (b), we have $\varphi(z) \in \mathcal{R}_S$ for all $\varphi \in S$; in particular, we have $w \in \mathcal{R}_S$, so $\mathcal{R}_S = \{z, w\}$. By (c), we also have $\varphi(w) \in \mathcal{R}_S$ for all $\varphi \in S$. Thus, for all $\varphi \in S$, we have $\varphi(z) \in \{z, w\}$ and $\varphi(w) \in \{z, w\}$. Now fix $\varphi \in S$; since φ is totally ramified at both z and w , we have $\varphi(z) \neq \varphi(w)$, hence $\varphi(\mathcal{R}_S) = \mathcal{R}_S$, and therefore—once again using the fact that φ is totally ramified at z and w —we have $\varphi^{-1}(\mathcal{R}_S) = \mathcal{R}_S$. Therefore, z is exceptional. \square

Remark 4.16. The conclusion of Lemma 4.15 need not hold if we only assume that $e_z(\gamma_2) = \deg \gamma_2$ for all $\gamma_2 \in S^2$. For example, let $\phi_1(z) = \frac{1}{z^2}$ and $\phi_2(z) = z^2 + 1$, and take $S = \{\phi_1, \phi_2\}$. For all four second iterates $\gamma_2 \in S^2$, we have $e_\infty(\gamma_2) = 4 = \deg \gamma_2$. However, ∞ is not exceptional; see Remark 4.2.

Lemma 4.17. *Let S, α and $\Delta_n = \Delta_{n, \alpha}$ be as in Proposition 4.13 above. For $z \in \mathbb{P}^1(\overline{\mathbb{Q}})$, define*

$$\sigma(z) = \sum_{\gamma_3 \in S^3} \nu_3(\gamma_3) \frac{e_z(\gamma_3)}{\deg \gamma_3}.$$

There exists $M < 1$ such that $\sigma(z) < M$ for all $z \notin E_S$.

Proof. Fix $\gamma_3 \in S^3$, and define

$$\mathcal{A} := \{z \in \mathbb{P}^1(\overline{\mathbb{Q}}) : e_z(\gamma_3) = \deg \gamma_3\}.$$

By the Riemann-Hurwitz theorem, the set \mathcal{A} has at most two elements, so in particular it is finite.

If $z \notin \mathcal{A}$, we have $e_z(\gamma_3) < \deg \gamma_3$, hence

$$\sigma(z) \leq M_1 := 1 - \frac{\nu_3(\gamma_3)}{\deg \gamma_3} < 1.$$

If $z \in \mathcal{A} \setminus E_S$, there exists (by Lemma 4.15) an element $\gamma'_3 \in S^3$ such that $e_z(\gamma'_3) < \deg \gamma'_3$, and arguing as above yields $\sigma(z) < 1$. Since \mathcal{A} is finite, the quantity

$$M_2 := \max_{z \in \mathcal{A} \setminus E_S} \sigma(z)$$

exists and is strictly less than 1. The lemma follows by taking $M := \max\{M_1, M_2\}$. \square

We are now ready to prove Proposition 4.13.

Proof of Proposition 4.13. It suffices to show that

$$\lim_{n \rightarrow \infty} \sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z) = 0,$$

since we have

$$\sum_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z)^2 \leq \left(\sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z) \right) \cdot \sum_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z) = \sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z).$$

First, we observe that for all $z \in \mathbb{P}^1(\overline{\mathbb{Q}})$ we have

$$\begin{aligned} \Delta_n(z) &= \sum_{\varphi \in S} \nu_1(\varphi) \frac{e_z(\varphi)}{\deg \varphi} \Delta_{n-1}(\varphi(z)) \\ &\leq \left(\sup_{w \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{n-1}(w) \right) \sum_{\varphi \in S} \nu_1(\varphi) \\ &= \sup_{w \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{n-1}(w), \end{aligned}$$

so the sequence $\left(\sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_n(z) \right)_{n \in \mathbb{N}}$ is non-increasing.

Finally, we claim that the subsequence $\left(\sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3k}(z) \right)_{k \in \mathbb{N}}$ decreases at least geometrically, from which our result follows. Observe that since $\alpha \notin E_S$, Δ_n is not supported on E_S for any n ; that is, $\Delta_n(w) = 0$ for all $w \in E_S$.

Now fix $k \geq 1$ and $z \in \mathbb{P}^1(\overline{\mathbb{Q}}) \setminus E_S$. Let M be as in Lemma 4.17. Then we have

$$\begin{aligned} \Delta_{3k}(z) &= \sum_{\gamma_3 \in S^3} \nu_3(\gamma_3) \frac{e_z(\gamma_3)}{\deg \gamma_3} \Delta_{3(k-1)}(\gamma_3(z)) \\ &\leq \left(\sup_{w \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3(k-1)}(w) \right) \sum_{\gamma_3 \in S^3} \nu_3(\gamma_3) \frac{e_z(\gamma_3)}{\deg \gamma_3} \\ &= \left(\sup_{w \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3(k-1)}(w) \right) \cdot \sigma(z) \\ &\leq M \left(\sup_{w \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3(k-1)}(w) \right). \end{aligned}$$

Since z was arbitrary, we have

$$\sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3k}(z) \leq M \left(\sup_{z \in \mathbb{P}^1(\overline{\mathbb{Q}})} \Delta_{3(k-1)}(z) \right),$$

and since $M < 1$, we have the claimed geometric decay. \square

We are now ready to prove our stochastic equidistribution theorem:

Proof of Theorem 4.8. The theorem will follow primarily from demonstrating that

$$h_S(\Delta_{n,\alpha}) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

so we will start by showing this. Notice that, by construction of our random backwards orbit measures $\Delta_{n,\alpha}$, we have

$$\mathbb{E}_S \varphi_*(\Delta_{n+1,\alpha}) = \Delta_{n,\alpha} \quad \text{for all } n \geq 0.$$

Applying Proposition 4.9, we see that

$$h_S(\Delta_{n+1,\alpha}) = \mathbb{E}_S \frac{h_S(\varphi_*(\Delta_{n+1,\alpha}))}{\deg \varphi} \leq \frac{1}{2} \mathbb{E}_S h_S(\varphi_*(\Delta_{n+1,\alpha})),$$

as $\deg \varphi \geq 2$ for every $\varphi \in S$. Now, applying Lemma 3.4, we can say that

$$(68) \quad \mathbb{E}_S h_S(\varphi_*(\Delta_{n+1,\alpha})) = h_S(\mathbb{E}_S \varphi_*(\Delta_{n+1,\alpha})) = h_S(\Delta_{n,\alpha}).$$

It follows that

$$h_S(\Delta_{n+1,\alpha}) \leq \frac{1}{2} h_S(\Delta_{n,\alpha}),$$

so by induction,

$$h_S(\Delta_{n,\alpha}) \leq \frac{1}{2^n} h_S(\alpha) \quad \text{for all } n \geq 0.$$

It follows that $h_S(\Delta_{n,\alpha}) \rightarrow 0$ as claimed. The h_S height is given by a generalized adelic measure by 4.4, and by Proposition 4.13, the $\Delta_{n,\alpha}$ measures are well-distributed. Theorem 3.17 therefore applies and gives the desired result. \square

4.5. Stochastic Julia and Fatou sets. Let (S, ν_1) be an L^1 height controlled stochastic family of rational maps, each of degree at least 2, and suppose that ν_1 is strictly positive. Let ρ denote the canonical generalized adelic measure such that $h_S = h_\rho$, as constructed above. For any place $y \in Y$ for which $C(y) < \infty$, we define the *stochastic Julia set* at the place y to be

$$\mathcal{J}_{S,y} = \text{supp}(\rho_y),$$

and we define the *stochastic Fatou set* to be

$$\mathcal{F}_{S,y} = \mathbb{P}_y^1 \setminus \mathcal{J}_{S,y}.$$

Many of the usual results one expects for Julia sets and Fatou sets are not always true for stochastic systems. We conclude with the following simple example which illustrates some of the key differences.

Example 4.18. Let $S = \{\varphi, \psi\}$ where $\varphi(z) = z^2$, $\psi(z) = 2z^2$ with $\nu_1(\varphi) = \nu_1(\psi) = 1/2$. Notice that S is defined over \mathbb{Q} , and since S contains only finitely many maps, there are only finitely many bad places (namely, $p = 2, \infty$), and so the associated generalized adelic measure is also an adelic measure over \mathbb{Q} in the original sense of Favre and Rivera-Letelier.

We claim that, at the archimedean place, the Julia set is the annulus

$$\mathcal{J}_{S,\infty} = \{z \in \mathbb{C} : 1/2 \leq |z| \leq 1\}$$

and $\rho_{S,\infty}$ is the radial measure which satisfies

$$\int_{D(0,r)} d\rho_{S,\infty} = \begin{cases} 0 & \text{if } r \leq 1/2, \\ 1 + \log_2(r) & \text{if } 1/2 \leq r \leq 1, \\ 1 & \text{if } r \geq 1, \end{cases}$$

or equivalently,

$$d\rho_{S,\infty}(re^{i\theta}) = \begin{cases} \frac{dr}{r \log 2} \frac{d\theta}{2\pi} & \text{if } 1/2 \leq r \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

To see this, note that

$$g_{1,\infty}(z) = \frac{1}{2} \left(\frac{1}{2} \log^+ |2z^2| + \frac{1}{2} \log^+ |z^2| \right) - \log^+ |z|.$$

If we denote the n th approximation to the local height by

$$\hat{\lambda}_{n,\infty}(z) = \mathbb{E}_{S^n} \frac{1}{\deg \gamma_n} \log^+ |\gamma_n(z)|,$$

and let $\hat{\lambda}_{0,\infty}(z) = \log^+ |z|$ by convention, then our n th term in the telescoping series that determines the Green's function is

$$g_{n,\infty}(z) = \hat{\lambda}_{n,\infty}(z) - \hat{\lambda}_{n-1,\infty}(z) \quad \text{for all } n \geq 1.$$

It follows that

$$\rho_{n,\infty} = \Delta \hat{\lambda}_{n,\infty}|_{\mathbb{C}} \quad \text{and that} \quad \lim_{n \rightarrow \infty} \rho_{n,\infty} = \rho_{\infty}$$

in the sense of weak convergence of measures.³ Now, it is easy to check that

$$\hat{\lambda}_{n,\infty}(z) = \frac{1}{2^n} \sum_{i=0}^{2^n-1} \log^+ \left| 2^{i/2^n} z \right| \rightarrow \int_0^1 \log^+ |2^x z| dx \quad \text{as } n \rightarrow \infty.$$

Of course, for each $r > 0$, the function of the form $\log^+ |z/r|$ is subharmonic and has Laplacian the normalized arc length measure on the circle $\{z \in \mathbb{C} : |z| = r\}$. It follows that ρ_{∞} does not charge any individual circle, and so the Portmanteau theorem for weak convergence of measures tells us that

$$(69) \quad \begin{aligned} \rho_{\infty}(D(0,r)) &= \lim_{n \rightarrow \infty} \Delta \hat{\lambda}_{n,\infty}(D(0,r)) \\ &= \lim_{n \rightarrow \infty} \frac{\#\{0 \leq i < 2^n : 2^{-i/2^n} < r\}}{2^n}, \end{aligned}$$

which gives us the desired result.

At the prime $p = 2$, the result is similar. We claim that

$$\mathcal{J}_{S,2} = [\zeta_{0,1}, \zeta_{0,2}],$$

where $\zeta_{a,b}$ is the point on the Berkovich projective line $\mathbb{P}_2^1 = \mathbb{P}^1(\mathbb{C}_2)$ which, as a seminorm on $\mathbb{C}_2[x]$, corresponds to the sup norm on the disc $D(a,b)$ in $\mathbb{P}^1(\mathbb{C}_2)$, and $[\zeta, \eta]$ denotes the natural line segment on the Berkovich tree connecting two points such that $\zeta \prec \eta$ under the natural ordering on the Berkovich tree. Let dx denote the path length metric on $[\zeta_{0,0}, \zeta_{0,\infty}]$ (the 'backbone' of the Berkovich tree for \mathbb{P}_2^1). Then, extending the function $\log^+ |2^{i/2^n} z|_2$ to \mathbb{P}_2^1 naturally by continuity, one has the Laplacian

$$\Delta \log^+ |2^{i/2^n} z|_2 = \delta_{\zeta_{0,2^{i/2^n}}} - \delta_{\infty},$$

and so the same argument as above in the archimedean case establishes that

$$d\rho_{S,2}(x) = \begin{cases} \frac{dx}{x \log 2} & \text{if } 1 \leq x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Next, we observe that the exceptional set of this system is $E_S = \{0, \infty\}$, as both points are exceptional and fixed for each map in S , so they are exceptional for the whole system.

Finally, we observe that by Theorem 4.8, for any $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}}) \setminus \{0, \infty\}$, the random backwards orbit measures $\Delta_{\alpha,n}$ converge in the weak sense of measures to $\rho_{S,\infty}$ at the infinite place, $\rho_{S,2}$ at $p = 2$, and to the standard measure $\rho_{S,p} = \lambda_p$ for $p \neq 2, \infty$.

³We restrict to \mathbb{C} here out of convenience because the logarithmic singularity of $\hat{\lambda}_{n,\infty}$ at $z = \infty$ results in a term of the form $-\delta_{\infty}$ in the Laplacian, which is cancelled by the same term from the Laplacian of the $\log^+ |z|$ function.

In the sequel [DFT] to this paper the authors will prove various analogues of classical results about the Julia and Fatou sets for stochastic dynamics, including backwards invariance of the Julia set (and hence forward invariance of the Fatou set) and various topological characterizations of the Julia set which can be proven using the equidistribution results of this article. We note that there is a large existing literature of dynamics for semigroups of rational functions, originating in the work of [HM96]. Note that our stochastic dynamical systems are in fact semigroups of rational functions, but with the additional data of a probability measure on the generating set. However, the theory is closely related, and in fact, we will show that support of our canonical measure coincides with the definition of the Julia set of the semigroup; in particular, so long as the measure ν_1 remains strictly positive (i.e., no map is assigned probability 0, and thus effectively removed from the system) then the support of the Julia set is independent of the particular choice of probability ν_1 . The particular distribution on the Julia set given by the canonical measure, however, still depends on ν_1 .

REFERENCES

- [AV09] Daniel Allcock and Jeffrey D. Vaaler, *A Banach space determined by the Weil height*, Acta Arith. **136** (2009), no. 3, 279–298. MR MR2475695 (2009j:11115)
- [Bak06] Matthew Baker, *A lower bound for average values of dynamical Green's functions*, Math. Res. Lett. **13** (2006), no. 2-3, 245–257. MR 2231115 (2007c:11078)
- [Bak08] ———, *An introduction to Berkovich analytic spaces and non-Archimedean potential theory on curves*, *p*-adic geometry, Univ. Lecture Ser., vol. 45, Amer. Math. Soc., Providence, RI, 2008, pp. 123–174. MR 2482347 (2010g:14029)
- [BR06] Matthew H. Baker and Robert Rumely, *Equidistribution of small points, rational dynamics, and potential theory*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 3, 625–688. MR 2244226 (2007m:11082)
- [BR10] Matthew Baker and Robert Rumely, *Potential theory and dynamics on the Berkovich projective line*, Mathematical Surveys and Monographs, vol. 159, American Mathematical Society, Providence, RI, 2010. MR 2599526
- [CL06] Antoine Chambert-Loir, *Mesures et équidistribution sur les espaces de Berkovich*, J. Reine Angew. Math. **595** (2006), 215–235. MR 2244803 (2008b:14040)
- [CLT09] Antoine Chambert-Loir and Amaury Thuillier, *Mesures de Mahler et équidistribution logarithmique*, Ann. Inst. Fourier (Grenoble) **59** (2009), no. 3, 977–1014. MR 2543659 (2011a:37189)
- [DeM03] Laura DeMarco, *Dynamics of rational maps: Lyapunov exponents, bifurcations, and capacity*, Math. Ann. **326** (2003), no. 1, 43–73. MR 1981611
- [DFT] John R. Doyle, Paul Fili, and Bella Tobin, *Julia sets for stochastic dynamical systems*, in preparation.
- [Fil] Paul Fili, *A metric of mutual energy and unlikely intersections for dynamical systems*, Preprint.
- [FP19] Paul Fili and Lukas Pottmeyer, *Quantitative height bounds under splitting conditions*, Trans. Amer. Math. Soc. **372** (2019), no. 7, 4605–4626.
- [FRL06] Charles Favre and Juan Rivera-Letelier, *Équidistribution quantitative des points de petite hauteur sur la droite projective*, Math. Ann. **335** (2006), no. 2, 311–361. MR 2221116 (2007g:11074)
- [FRL07] ———, *Corrigendum to: “Quantitative uniform distribution of points of small height on the projective line” (French) [Math. Ann. **335** (2006), no. 2, 311–361; mr2221116]*, Math. Ann. **339** (2007), no. 4, 799–801. MR 2341901 (2008h:11069)
- [Gub97] Walter Gubler, *Heights of subvarieties over M-fields*, Arithmetic geometry (Cortona, 1994), Sympos. Math., XXXVII, Cambridge Univ. Press, Cambridge, 1997, pp. 190–227. MR 1472498
- [HH19] Vivian Olsiewski Healey and Wade Hindes, *Stochastic canonical heights*, J. Number Theory **201** (2019), 228–256. MR 3958050

- [HM96] A. Hinkkanen and G. J. Martin, *The dynamics of semigroups of rational functions. I*, Proc. London Math. Soc. (3) **73** (1996), no. 2, 358–384. MR 1397693
- [Kaw07] Shu Kawaguchi, *Canonical heights for random iterations in certain varieties*, Int. Math. Res. Not. IMRN (2007), no. 7, Art. ID rnm 023, 33. MR 2345343
- [Mah64] K. Mahler, *An inequality for the discriminant of a polynomial*, Michigan Math. J. **11** (1964), 257–262. MR 0166188 (29 #3465)
- [MY] Niki Myrto Mavraki and Hexi Ye, *Quasi-adelic measures and equidistribution on \mathbb{P}^1* , Preprint.
- [Sal45] R. Salem, *Power series with integral coefficients*, Duke Math. J. **12** (1945), 153–172. MR MR0011720 (6,206b)
- [Thu05] Amaury Thuillier, *Théorie du potentiel sur les courbes en géométrie analytique non archimédienne. applications à la théorie d’arakevov*, Thèse de l’Université de Rennes 1 (2005), viii + 184 p.
- [Ye15] Hexi Ye, *Rational functions with identical measure of maximal entropy*, Adv. Math. **268** (2015), 373–395. MR 3276598
- [Yua08] Xinyi Yuan, *Big line bundles over arithmetic varieties*, Invent. Math. **173** (2008), no. 3, 603–649. MR 2425137 (2010b:14049)

DEPARTMENT OF MATHEMATICS, OKLAHOMA STATE UNIVERSITY, STILLWATER, OK 74078
Email address: john.r.doyle@okstate.edu

DEPARTMENT OF MATHEMATICS, OKLAHOMA STATE UNIVERSITY, STILLWATER, OK 74078
Email address: paul.fili@okstate.edu

DEPARTMENT OF MATHEMATICS, OKLAHOMA STATE UNIVERSITY, STILLWATER, OK 74078
Email address: bella.tobin@okstate.edu