RANDOM CONCAVE FUNCTIONS

BY PETER BAXENDALE[1,a] AND TING-KAM LEONARD WONG[2,b]

[1]*Department of Mathematics, University of Southern California,* [a]*baxendal@usc.edu*
[2]*Department of Statistical Sciences, University of Toronto,* [b]*tkl.wong@utoronto.ca*

Spaces of convex and concave functions appear naturally in theory and applications. For example, convex regression and log-concave density estimation are important topics in nonparametric statistics. In stochastic portfolio theory, concave functions on the unit simplex measure the concentration of capital, and their gradient maps define novel investment strategies. The gradient maps may also be regarded as optimal transport maps on the simplex. In this paper we construct and study probability measures supported on spaces of concave functions. These measures may serve as prior distributions in Bayesian statistics and Cover's universal portfolio, and induce distribution-valued random variables via optimal transport. The random concave functions are constructed on the unit simplex by taking a suitably scaled (mollified, or soft) minimum of random hyperplanes. Depending on the regime of the parameters, we show that as the number of hyperplanes tends to infinity there are several possible limiting behaviors. In particular, there is a transition from a deterministic almost sure limit to a nontrivial limiting distribution that can be characterized using convex duality and Poisson point processes.

1. Introduction.

1.1. *Motivations.* In this paper we study probability measures on spaces of concave functions. We first describe some applications that motivated our study. In the first two applications there is an infinite dimensional parameter space consisting of convex or concave functions, and the problem is to find mathematically tractable prior distributions on the space.

1.1.1. *Nonparametric Bayesian statistics.* Consider a nonlinear regression problem where data is drawn according to the model

$$(1.1) \qquad\qquad Y_i = f(X_i) + \epsilon_i.$$

In many applications the regression function f is known to satisfy certain shape constraints such as monotonicity or convexity/concavity. Without assuming further structures on f, this problem is nonparametric as f is an element of an infinite dimensional function space. See [15] for various shape constraints in economics and operations research. Although the shape-constrained estimation problem can be studied by various methods (see, e.g., [12, 13, 31] in the references therein), it is both important and interesting to consider the Bayesian approach. To do this we need suitable prior distributions for the convex function f. In [12] Hannah and Dunson proposed to generate random convex functions on $\mathbb{R}^n$ by taking the maximum of a (random) number of random hyperplanes, and established rates of convergence of the Bayes estimator. While in [12] the main concern is the support and concentration properties of the prior, we will establish concrete results about the limiting distributions as the number of hyperplanes tends to infinity.

Another important class of shape-constrained inference problems is density estimation. A classic example, studied in [4, 7] among many other papers, is log-concave density estimation. Here we observe data $X_1, \ldots, X_N$ with values in $\mathbb{R}^n$, where

$$X_i \overset{\text{i.i.d.}}{\sim} f$$

and f is a log-concave density, that is, $\log f$ is concave. For example, the densities of the normal and gamma distributions are log-concave. Again, to use the Bayesian approach we need to introduce suitable prior distributions on the space of concave functions. So far there is little work on this topic except the one-dimensional case [21]. For recent progress in log-concavity in general and density estimation we refer the reader to [28, 29].

In nonparametric Bayesian statistics a very useful class of prior distributions is the Dirichlet process introduced by Ferguson [9]. Realizations of the Dirichlet process are random discrete probability measures on a given state space. In one dimension, the Dirichlet process can be used to enforce shape constraints. For example, a convex function on an interval has a nondecreasing first derivative which can be identified with the distribution function of a measure. However, similar arguments do not extend immediately to multidimensions as the second derivative of a convex function, if exists, is matrix-valued.[1] See Section 1.1.3 for more discussion involving ideas from optimal transport.

1.1.2. *Stochastic portfolio theory and Cover's universal portfolio.*

Consider the open unit simplex in $\mathbb{R}^n$, $n \geq 2$, defined by

$$(1.2) \qquad \Delta_n := \{ p \in (0, 1)^n : p_1 + \cdots + p_n = 1 \}.$$

We let $\overline{\Delta}_n$ be its closure in $\mathbb{R}^n$. Let $e_1, \ldots, e_n$ be the standard Euclidean basis of $\mathbb{R}^n$ which represents the vertices of the simplex. We denote by $\overline{e} := (\frac{1}{n}, \ldots, \frac{1}{n})$ the barycenter of the simplex.

In stochastic portfolio theory (see [10, 17] for introductions) the open simplex Δ_n represents the state space of an equity market with n stocks. If $X_i(t) > 0$ denotes the market capitalization of stock i at time t, we call

$$(1.3) \qquad m_i(t) = \frac{X_i(t)}{X_1(t) + \cdots + X_n(t)}$$

the market weight of stock i. The vector $m(t) = (m_i(t))_{1 \leq i \leq n}$ then defines a process evolving in the simplex Δ_n. Let $\Phi : \Delta_n \to (0, \infty)$ be a positive concave function on Δ_n. In this context the function Φ plays two related roles. First, Φ can be regarded as a generalized measure of diversity (analogous to the Shannon entropy) which quantifies the concentration of capital in the market [10], Chapter 3. Second, the concave function Φ can be used to define an investment strategy, called functionally generated portfolio, with remarkable properties. Here is how the strategy is defined when Φ is differentiable. If the market weight is $m(t) = p \in \Delta_n$ at time t, invest $100\pi_i\%$ of the current capital in stock i, where

$$(1.4) \qquad \pi_i = p_i \big(1 + D_{e_i - p} \log \Phi(p) \big),$$

and $D_{e_i - p}$ is the directional derivative. We call the mapping $p \mapsto \boldsymbol{\pi}(p) = \pi \in \overline{\Delta}_n$ the portfolio map generated by Φ.[2] To give an example which will be relevant later, the geometric mean $\Phi(p) = p_1^{\pi_1} \cdots p_n^{\pi_n}$, where $\pi = (\pi_1, \ldots, \pi_n) \in \overline{\Delta}_n$ is fixed, generates the constant-weighted portfolio $\boldsymbol{\pi}(p) \equiv \pi$ [10], Example 3.1.6. As shown in [10, 17, 23, 38], the concavity of Φ

---

[1] The subgradient (as a set-valued mapping) satisfies a condition known as cyclical monotonicity; see [27], Section 24.

[2] We may also use the concave function Φ to define portfolios generated by ranked weights; see [10], Chapter 4.

allows the portfolio to diversify and capture volatility of the market through systematic re-balancing.

In the seminal paper [2] Cover constructed what is now called an online investment algorithm by forming a Bayesian average over the constant-weighted portfolios. The main idea is that strategies which have been performing well receive additional weights that are computed using an algorithm analogous to Bayes' theorem (where the portfolio value plays the role of the likelihood). To start the algorithm one needs an initial (i.e., prior) distribution on the space of portfolio strategies. In a nonprobabilistic framework it can be shown that Cover's universal portfolio tracks asymptotically the best strategy in the given (finite dimensional) family, in the sense that the average regret (measured by the growth rate) with respect to the best strategy tends to zero as the number of time steps tends to infinity. In [3, 35] the second author and his collaborators extended Cover's approach to the nonparametric family of functionally generated portfolios. Nevertheless, for practical applications and to obtain quantitative estimates we need tractable prior distributions for the generating function Φ. The distributions constructed in this paper may serve as a building block of the prior. Also see [14] for a study of minimums of hyperplanes in the context of robust portfolio optimization in stochastic portfolio theory.

1.1.3. *Optimal transport.*   Convex and concave functions are also interesting from the viewpoint of optimal transport (see [32, 33] for in-depth overviews). Given a cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ and probability measures P on $\mathcal{X}$ and Q on $\mathcal{Y}$, the Monge–Kantorovich problem is the minimization of the transport cost

$$\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \, dR(x, y)$$

over all couplings R of (P, Q). When $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$ and $c(x, y) = |x - y|^2$ is the squared Euclidean distance, Brenier's theorem [1] asserts that there is a deterministic optimal transport map of the form

$$(1.5) \qquad\qquad\qquad\qquad y = \nabla \phi(x),$$

where ϕ is a convex function (this holds, e.g., when P and Q have finite second moments and P is absolutely continuous). Conversely, given P fixed (e.g., standard normal) and a convex function ϕ, the transport map (1.5) is optimal with respect to P and the pushforward $Q = (\nabla \phi)_\# P$. Thus a probability distribution over Q (i.e., an element of $\mathcal{P}(\mathcal{P}(\mathbb{R}^n))$) can be used to define a distribution over the space of convex functions on $\mathbb{R}^n$.

In a series of papers [23–25] Pal and the second author studied a novel optimal transport problem, that we call the Dirichlet transport, on the unit simplex Δ_n.[3] The cost function is given by

$$c(p, q) = \log\left(\frac{1}{n} \sum_{i=1}^{n} \frac{q_i}{p_i}\right) - \sum_{i=1}^{n} \frac{1}{n} \log \frac{q_i}{p_i}, \quad p, q \in \Delta_n.$$

For this cost function we proved an analogue of Brenier's theorem in [25], Theorem 4: under natural conditions on P and Q, there exists a positive concave function Φ on Δ_n such that the optimal transport map is given by

$$(1.6) \qquad\qquad\qquad\qquad q = p \odot \pi\left(p^{-1}\right), p \in \Delta_n,$$

---

[3] As shown in [25, 37, 39], this transport problem also has remarkable properties from the information geometric point of view.
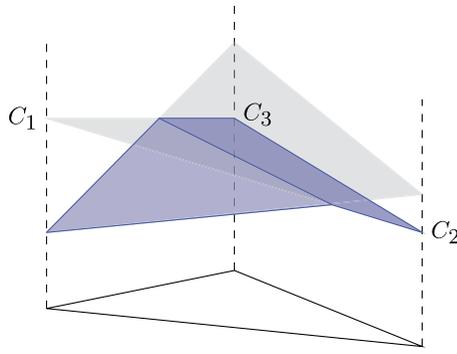
FIG. 1. *A random nonnegative concave function on $\overline{\Delta}_n$ given as the minimum of several hyperplanes. Here $C = (C_1, \ldots, C_n)$ is a random vector which determines the coefficients of the hyperplane.*

where π is the portfolio map generated by Φ in the sense of (1.4),

$$p^{-1} = \left(\frac{1/p_i}{\sum_j 1/p_j} \right)_{1 \le i \le n} \quad \text{and} \quad a \odot b = \left(\frac{a_i b_i}{\sum_j a_j b_j} \right)_{1 \le i \le n}.$$

When $P = Q$, the identity transport $q = p$ is realized by the geometric mean $\Phi(p) = (p_1 \cdots p_n)^{1/n}$. The weighted geometric mean $\Phi(p) = p_1^{\pi_1} \cdots p_n^{\pi_n}$ corresponds to a deterministic translation under an exponential coordinate system [24], Proposition 2.7(iii).

It follows that an element of $\mathcal{P}(\mathcal{P}(\Delta_n))$ induces a probability distribution over positive concave functions on Δ_n. Measures over spaces of probability distributions are important in optimal transport, information geometry and statistics. For example, in [34] von Renesse and Sturm constructed an entropic measure on the Wasserstein space and defined a Wasserstein diffusion. In principle, one can use Dirichlet processes on Δ_n to define random concave functions via the Dirichlet transport problem. Further properties of this construction are left for future research.

REMARK 1.1. Since a convex function can be identified with its epigraph, the results of this paper can be formulated in terms of random convex sets in $\mathbb{R}^n$ such that part of the boundary is fixed to be the unit simplex. While random convex sets (given, e.g., by the convex hulls of random points) have been studied extensively in the literature (see, e.g., [16, 22, 30] and the references therein), the motivations and questions studied in this paper are quite different.

1.2. *Summary of the paper.* Motivated by the applications described in Sections 1.1.2 and 1.1.3, in this paper we focus on random nonnegative concave functions on the unit simplex $\overline{\Delta}_n$. Thus we let

(1.7) $$\mathcal{C} := \{ \psi : \overline{\Delta}_n \to [0, \infty) \text{ continuous and concave} \},$$

and our aim is to construct and study probability measures on $\mathcal{C}$. We equip $\mathcal{C}$ with the topology of uniform convergence on compact subsets of Δ_n and the associated Borel σ-algebra. Properties of $\mathcal{C}$ are given in Section 2.1.

We study a natural probabilistic model for generating random concave functions. Namely, they are given by suitably scaled minimums of i.i.d. random hyperplanes (see Figure 1). More generally, we also consider a soft minimum m_λ where $\lambda \in (0, \infty]$ is an inverse smoothness parameter and $\lim_{\lambda \to \infty} m_\lambda = m_\infty = \min$ (see Definition 2.4). Thus, given a positive integer K, the number of hyperplanes, we consider the random concave function given by

(1.8) $$\Psi_K = a_K m_{\lambda_K}(\ell_1, \ldots, \ell_K),$$

where $a_K > 0$ is a scaling constant, $\lambda_K \in (0, \infty]$, and $\ell_1, \ldots, \ell_K$ are i.i.d. random hyperplanes. This model is rigorously defined in Section 2. Our main objective in this paper is to study the limiting behavior of the distribution of Ψ_K as the number of hyperplanes tends to infinity.

In Section 2.3 we consider the softmin case $\Psi_K = \mathsf{m}_\lambda(\ell_1, \ldots, \ell_K)$, where $\lambda \in (0, \infty)$ is finite and fixed. We show that there exists a deterministic concave function Ψ_∞, given in terms of the distribution of ℓ_k, such that $\Psi_K \to \Psi_\infty$ almost surely.

Section 3 studies the case of hardmin, that is, $\lambda_K \equiv \infty$ for all K. Under suitable conditions on the distribution of the hyperplanes ℓ_k, we show that the distribution of Φ_K converges weakly to a nontrivial limit μ as $K \to \infty$. Moreover, for some constant $\alpha > 0$, if $\Psi \sim \mu$ then the distribution of $\Psi(p)^{n+\alpha}$ for each fixed $p \in \Delta_n$ is exponential (see Proposition 3.8). This can be regarded as a functional limit theorem in the context of extreme value theory. Pointwise extremas of i.i.d. stochastic processes have been studied extensively in the literature (see, e.g., [5, 19, 26] and the references therein). Here, the novelty of our results lies in the shape constraint and an elegant convex duality for concave functions on the unit simplex which allows us to study the limiting distribution in detail. See Remarks 3.6 and 3.15 for more discussions about connections with extreme value theory.

Various properties of this limiting distribution μ are established, under additional conditions, in Section 4. In particular, we show that the geometric mean, which plays a special role in stochastic portfolio theory and the Dirichlet transport, arises as the expected value of the limiting random concave function. Using differential geometric methods, we also give an interesting explicit formula for the tail probability $\mathbb{P}(\Psi \geq \psi)$ for a given $\psi \in \mathcal{C}$.

Finally, in Section 5, we consider the mathematically more challenging case where the smoothness parameter λ_K depends on K. We identify regimes which give different limiting behaviors. Our analysis involves studying laws of large numbers for soft minimums of i.i.d. random variables, again related to Poisson point processes, which may be of independent interest.

In this paper we studied some probabilistic properties of random concave functions defined by the model (1.8). To address the applications described in Section 1.1 we need to develop efficient computational methods; the model (1.8) may also need to be modified to suit the specific needs. We plan to study these questions in future research.

2. Concave functions on the simplex.

2.1. *Preliminaries.*   As noted in Section 1 we will focus on the space $\mathcal{C}$, defined by (1.7), consisting of nonnegative continuous concave functions on $\overline{\Delta}_n$. We also let

$$\mathcal{C}_+ := \{\psi \in \mathcal{C} : \psi > 0 \text{ on } \Delta_n\}$$

be those functions in $\mathcal{C}$ that are strictly positive in the (relative) interior.

Our choice of using the simplex as the domain has the following mathematical advantages apart from the motivations described above. First, the simplex $\overline{\Delta}_n$ is a symmetric polyhedron, and in this case the duality of concave function takes a special form which is useful for our analysis. Second, if we specify a finite number of points $p^{(i)} \in \Delta_n$ and constants $r^{(i)} > 0$, the smallest function $\psi \in \mathcal{C}$ such that $\psi(p^{(i)}) \geq r^{(i)}$ for all i is polyhedral, that is, it is the minimum of a finite collection of hyperplanes. This is not the case if the boundary is smooth. Last but not least, the duality allows us to connect the limiting distributions of our model with Poisson point processes on the positive quadrant. While it may be possible to extend some results to general convex domains, we believe the unit simplex is of special interest.

Functions in $\mathcal{C}$ enjoy strong analytical properties (we refer the reader to [27] for standard results in convex analysis). For example, if $\psi \in \mathcal{C}$, then ψ is locally Lipschitz on Δ_n. Moreover, the superdifferential

$$\partial \psi(p) := \{\xi \in \mathbb{R}^n : \xi_1 + \cdots + \xi_n = 0, \psi(p) + \langle \xi, q - p \rangle \geq \psi(q) \text{ for all } q \in \overline{\Delta}_n\}$$

is nonempty, convex and compact for every $p \in \Delta_n$; moreover ψ is differentiable (i.e., the superdifferential $\partial \psi(p)$ reduces to a singleton) Lebesgue almost everywhere on Δ_n. By Aleksandrov's theorem (see, e.g., [8], Theorem 6.9) even the Hessian can be defined almost everywhere, but this result is not needed in this paper.

We equip the space $\mathcal{C}$ with the topology of local uniform convergence. By definition, a sequence $\{\psi_k\}$ converges to ψ in $\mathcal{C}$ if and only if for any compact subset Ω of Δ_n we have $\psi_k \to \psi$ uniformly on Ω. A metric which is compatible with this topology is

$$(2.1) \qquad d(\varphi, \psi) := \sum_{k=1}^{\infty} 2^{-k} \Big(\sup_{p \in \Delta_{n,k}} |\varphi(p) - \psi(p)| \wedge 1 \Big),$$

where $\Delta_{n,k} = \{p \in \Delta_n : p_i \geq 1/k, i = 1, \ldots, n\}$ is compact in Δ_n. Note that by [27], Theorem 10.3, any nonnegative concave function on Δ_n has a unique continuous extension to $\overline{\Delta}_n$. This implies that if $\varphi, \psi \in \mathcal{C}$ and $d(\varphi, \psi) = 0$ then $\varphi \equiv \psi$ on $\overline{\Delta}_n$; thus the metric is well defined on $\mathcal{C}$ even though the boundary is not explicitly included in (2.1). It is easy to verify that $(\mathcal{C}, d)$ is complete and separable. The following lemma is standard and a proof (which uses [27], Theorem 10.6) can be found in [36].

LEMMA 2.1. *For any $M \geq 0$ the set $\{\psi \in \mathcal{C} : \psi \leq M\}$ is compact in $\mathcal{C}$.*

Let $\mathcal{B} = \mathcal{B}(\mathcal{C})$ be the Borel σ-field generated by this topology. In this paper we are interested in probabilistic models for generating random elements of $\mathcal{C}$, or equivalently probability measures on $(\mathcal{C}, \mathcal{B})$. It is easy to see that $\mathcal{B}$ is generated by the collection of finite dimensional cylinder sets. This implies the following lemma.

LEMMA 2.2. *Let ν and $\tilde{\nu}$ be probability measures on $\mathcal{C}$. If they have the same finite dimensional distributions, then $\nu = \tilde{\nu}$.*

REMARK 2.3. Apart from the topology of uniform convergence over compact subsets of Δ_n as in (2.1), one may consider, for example, the topology of uniform convergence on $\overline{\Delta}_n$. We argue that our choice is more natural, and the main reason is that convergence theorems in convex analysis (such as [27], Theorem 10.9) are usually formulated in the topology of local uniform convergence. To give a concrete example, consider on $[0, 1]$ the sequence $\{\psi_k\}_{k \geq 2}$ of concave functions given by

$$\psi_k(x) = \begin{cases} kx & \text{for } 0 \leq x \leq \dfrac{1}{k}; \\ 1 & \text{for } \dfrac{1}{k} \leq x \leq 1 - \dfrac{1}{k}; \\ k(1 - x) & \text{for } \leq 1 - \dfrac{1}{k} \leq x \leq 1. \end{cases}$$

Then ψ_k converges with respect to metric d, but not uniformly, to the constant function $\psi(x) \equiv 1$.

Let $\mathcal{A}_+$ denote the set

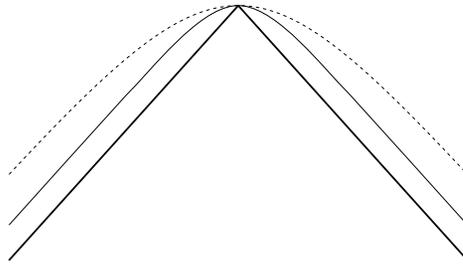$$\mathcal{A}_+ := \{\ell : \overline{\Delta}_n \to (0, \infty) \text{ affine}\}$$

FIG. 2. *Graphs of* $m_\lambda(x, 2 - x)$ *on* $[0, 2]$ *for* $\lambda = 2$ *(dashed),* $\lambda = 5$ *(thin solid) and* $\lambda = \infty$ *(thick solid).*

consisting of (strictly) positive affine functions on $\overline{\Delta}_n$. Clearly $\mathcal{A}_+ \subset \mathcal{C}_+ \subset \mathcal{C}$. Note that every element of $\mathcal{A}_+$ can be written in the form

$$(2.2) \qquad \ell(p) = \sum_{i=1}^{n} p_i x_i =: \langle p, x \rangle,$$

for some positive constants $x_1, \ldots, x_n > 0$, where $x_i = \ell(e_i)$ is the value of ℓ at the vertex e_i. Thus we may identify $\mathcal{A}_+$ with the positive quadrant $\mathbb{R}^n_+ := (0, \infty)^n$. By concavity, for any $\psi \in \mathcal{C}$ we have

$$(2.3) \qquad \psi = \inf\{\ell \in \mathcal{A}_+ : \ell \geq \psi\}.$$

Since every element of $\mathcal{C}$ can be written as the infimum of a collection of hyperplanes, to generate a random concave function in $\mathcal{C}$ it suffices to generate a random collection of hyperplanes in $\mathcal{A}_+$.

While every concave function in $\mathcal{C}$ can be generated in the form (2.3), in applications (e.g., in stochastic portfolio theory) it may be desirable to use a smooth approximation of the minimum operation, so that each realization is itself smooth (when the number of planes is finite). For this reason we introduce the softmin which is often used in convex optimization and machine learning. The smoothness parameter also adds an extra dimension to the mathematical analysis.

DEFINITION 2.4 (Softmin). Let $\lambda > 0$. For $K \geq 1$ and $x_1, \ldots, x_K \in \mathbb{R}$ we define the softmin (with parameter λ) by

$$(2.4) \qquad m_\lambda(x_1, \ldots, x_K) := \frac{-1}{\lambda} \log\left(\frac{1}{K} \sum_{k=1}^{K} e^{-\lambda x_k}\right).$$

By continuity, we define

$$m_\infty(x_1, \ldots, x_K) := \min\{x_1, \ldots, x_K\}$$

and call this the hardmin (see Figure 2 for an illustration). We also write $m_\lambda(x_1, \ldots, x_K) = m_\lambda\{x_k\}$ when the context is clear.

LEMMA 2.5 (Properties of softmin). *Let* $\lambda > 0$ *be fixed.*

(i) *For* $x_1, \ldots, x_K \in \mathbb{R}$ *we have*

$$(2.5) \qquad \min\{x_1, \ldots, x_K\} \leq m_\lambda(x_1, \ldots, x_K) \leq \min\{x_1, \ldots, x_K\} + \frac{1}{\lambda} \log K.$$

Also, for any $x \in \mathbb{R}$ *we have* $m_\lambda(x, \ldots, x) = x$ *and*

$$m_\lambda(x_1 + c, \ldots, x_K + c) = m_\lambda(x_1, \ldots, x_K) + c, \quad c \in \mathbb{R}.$$

(ii) *For $K \geq 1$ fixed, the softmin m_λ is a smooth and symmetric concave function of $x_1, \ldots, x_K$.*

(iii) *If $\Phi^{(1)}, \ldots, \Phi^{(K)}$ are finite concave functions defined on a convex set, then so is*

$$(2.6) \qquad \Phi = \mathsf{m}_\lambda(\Phi^{(1)}, \ldots, \Phi^{(K)}).$$

PROOF. All statements can be proved by elementary means, and for completeness we give the proof of (iii). First we observe that if $y_k \geq x_k$ for all k (possibly after a permutation of the elements), then

$$\mathsf{m}_\lambda\{y_k\} \geq \mathsf{m}_\lambda\{x_k\}.$$

Let p, q be elements of the domain of the $\Phi^{(k)}$, and let $0 < \alpha < 1$. By the previous remark as well as (ii) and the concavity of the $\Phi^{(k)}$, we have

$$\begin{aligned}
\Phi((1-\alpha)p + \alpha q) &= \mathsf{m}_\lambda\{\Phi^{(k)}((1-\alpha)p + \alpha q)\} \\
&\geq \mathsf{m}_\lambda\{(1-\alpha)\Phi^{(k)}(p) + \alpha\Phi^{(k)}(q)\} \\
&\geq (1-\alpha)\mathsf{m}_\lambda\{\Phi^{(k)}(p)\} + \alpha\mathsf{m}_\lambda\{\Phi^{(k)}(q)\} \\
&= (1-\alpha)\Phi(p) + \alpha\Phi(q).
\end{aligned}$$

This proves that the softmin Φ is concave as well. $\square$

By Lemma 2.5, if $\Phi^{(1)}, \ldots, \Phi^{(K)} \in \mathcal{C}$ then so is $\Phi = \mathsf{m}_\lambda(\Phi^{(1)}, \ldots, \Phi^{(K)})$. Moreover, by the differentiability of the softmin, if each $\Phi^{(k)}$ is differentiable, then so is Φ.

REMARK 2.6. As explained in Section 1.1.2, every element of $\mathcal{C}$ can be regarded as a portfolio generating function. Suppose $\Phi^{(1)}, \ldots, \Phi^{(K)} \in \mathcal{C}$ are differentiable, and let $\Phi = \mathsf{m}_\lambda(\Phi^{(1)}, \ldots, \Phi^{(K)})$ be their softmin. Also let $\boldsymbol{\pi}^{(k)}$ be the portfolio map generated by $\Phi^{(k)}$ in the sense of (1.4). Then it can be shown by a straightforward computation that the portfolio map generated by Φ is given by

$$(2.7) \qquad \boldsymbol{\pi}(p) = \left(1 - \sum_{k=1}^{K} a_k(p)\right)p + \sum_{k=1}^{K} a_k(p)\boldsymbol{\pi}^{(k)}(p), \quad p \in \Delta_n,$$

where

$$(2.8) \qquad a_k(p) = \frac{e^{-\lambda\Phi_k(p)}}{\sum_{\ell=1}^{K} e^{-\lambda\Phi_\ell(p)}} \cdot \frac{\Phi_k(p)}{\Phi(p)} = \frac{\Phi_k(p)e^{-\lambda\Phi_k(p)}}{K\Phi(p)e^{-\lambda\Phi(p)}}, \quad k = 1, \ldots, K.$$

Thus $\boldsymbol{\pi}$ is a linear combination of the market portfolio and the portfolios generated by $\{\Phi^{(k)}\}$. In the limiting case $\lambda \to \infty$ (i.e., the hardmin), (2.7) gives

$$\boldsymbol{\pi}(p) = \boldsymbol{\pi}^{(k)}(p), \quad \text{if } \Phi^{(k)}(p) < \Phi^{(\ell)}(p) \text{ for } \ell \neq k.$$

To conclude this subsection, we state some estimates of concave functions which are useful in several results below. The proof is given in the Appendix.

LEMMA 2.7.

(i) *For each $q \in \Delta_n$ there exists an explicit constant $M_q > 0$ such that*

$$(2.9) \qquad \psi(p) \leq M_q\psi(q) \quad \text{for all } p \in \overline{\Delta}_n \text{ and all } \psi \in \mathcal{C}.$$

(ii) *For $1 \leq j \leq n$ and $0 < \epsilon \leq \frac{1}{n}$ we define $c^{(j,\epsilon)} \in \Delta_n$ by*

(2.10)
$$c_i^{(j,\epsilon)} = \begin{cases} \epsilon, & i = j, \\ \dfrac{1-\epsilon}{n-1}, & i \neq j. \end{cases}$$

Thus $c^{(j,\epsilon)}$ is the center of the slice $p_j = \epsilon$ through the simplex Δ_n. Suppose $p \in \overline{\Delta}_n$ satisfies $0 \leq p_j \leq \epsilon \leq 1/n$ for some j. Then

(2.11)
$$\psi(p) \leq n\psi(c^{(j,\epsilon)}) \quad \text{for all } \psi \in \mathcal{C}.$$

In particular we have

$$\psi(p) \leq n\psi(\overline{e}) \quad \text{for all } p \in \overline{\Delta}_n \text{ and } \psi \in \mathcal{C}.$$

2.2. The probabilistic model.

In this paper we study a natural implementation of the representation (2.3). Namely, we consider random concave functions given as (soft) minimums of i.i.d. random hyperplanes.

Let $C = (C_1, \ldots, C_n)$ be a random vector with values in the quadrant $\mathbb{R}_+^n$. Note that the components of C may be dependent. Throughout the paper we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which the required random elements are defined. Given C, we define a random element ℓ of $\mathcal{A}_+$ given by

(2.12)
$$\ell(p) = \langle C, p \rangle = \sum_{i=1}^{n} p_i C_i, \quad p \in \overline{\Delta}_n.$$

For $K = 1, 2, \ldots$, let $\ell_1, \ell_2, \ldots, \ell_K$ be independent copies of ℓ. Define a random concave function Ψ_K by

(2.13)
$$\Psi_K = a_K \mathsf{m}_\lambda(\ell_1, \ldots, \ell_K),$$

where $\lambda = \lambda_K \in (0, \infty]$ possibly depends on K and $a_K > 0$ is a scaling constant to be chosen. The law of Ψ_K defines a probability measure ν_K on $\mathcal{C}$ which depends on K, λ_K, a_K and the distribution of C. We are interested in the limiting behavior of ν_K as $K \to \infty$.

2.3. Deterministic limit for softmin.

To give quickly a concrete result, in this subsection we consider the model (2.13) where $0 < \lambda < \infty$ is fixed and independent of K, and there is no scaling, that is, $a_K \equiv 1$. Using the strong law of large numbers, we show that there is a deterministic almost sure limit as $K \to \infty$.

THEOREM 2.8. *Fix $\lambda > 0$. Let*

(2.14)
$$\psi(t_1, \ldots, t_n) := \log \mathbb{E} e^{t_1 C_1 + \cdots + t_n C_n}$$

be the cumulant generating function of $C = (C_1, \ldots, C_n)$ which is finite and convex on $(-\infty, 0]^n$. Let $\{\ell_k\}_{k=1}^{\infty}$ be a sequence of independent copies of ℓ as in (2.12), and, for each $K \geq 1$, let Ψ_K be the random concave function defined by $\Psi_K = \mathsf{m}_\lambda(\ell_1, \ldots, \ell_K)$. Then

$$\lim_{K \to \infty} \Psi_K = \Psi_\infty \quad \mathbb{P}\text{-a.s.},$$

where $\Psi_\infty \in \mathcal{C}_+$ is the deterministic concave function given by

(2.15)
$$\Psi_\infty(p) = \frac{-1}{\lambda} \psi(-\lambda p), \quad p \in \Delta_n.$$

Here the convergence means that $d(\Psi_K, \Psi_\infty) \to 0$ a.s., where d is the metric on $\mathcal{C}$ defined by (2.1).
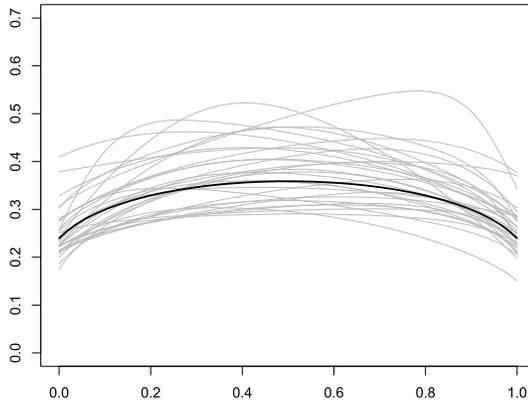
FIG. 3. *Samples (in grey) of Ψ_K in Example 2.9. Here $n = 2$ (so that the simplex can be identified with the unit interval), $\alpha = 1$, $\lambda = 10$ and $K = 30$. The limiting function Ψ_∞ is shown by the thick black curve.*

PROOF. Let $p \in \Delta_n$ be fixed. By definition, we have

$$\Psi_K(p) = \frac{-1}{\lambda} \log\left(\frac{1}{K} \sum_{k=1}^{K} e^{-\lambda \ell_k(p)} \right).$$

By the strong law of large numbers, we have the almost sure limit

$$\lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} e^{-\lambda \ell_k(p)} = \mathbb{E} e^{-\lambda \sum_{i=1}^{n} C_i p_i} = e^{\psi(-\lambda p)}.$$

Taking logarithm and dividing by $-\lambda$ shows that $\Psi_K(p) \to \Psi_\infty(p)$ almost surely. The pointwise convergence then holds, with probability 1, over a countable dense subset of Δ_n. By [27], Theorem 10.8, we have uniform convergence over compact subsets which implies convergence in the metric d. $\square$

EXAMPLE 2.9. In the context of Theorem 2.8, suppose $C_1, \ldots, C_n$ are i.i.d. exponential random variables with rate $\alpha > 0$. The cumulant generating function is given by

$$\psi(t) = \sum_{i=1}^{n} \log \frac{\alpha}{\alpha - t_i}.$$

It follows that the limiting function (2.15) is given by

$$(2.16) \qquad \Psi_\infty(p) = \sum_{i=1}^{n} \frac{1}{\lambda} \log\left(1 + \frac{\lambda}{\alpha} p_i \right).$$

Some samples from this model are given in Figure 3.

An interesting question is what happens when $\lambda \to \infty$. From (1.4), Ψ and $c\Psi$ generate the same portfolio map for any $c > 0$. Thus we consider instead the limit of $\Psi_\infty(p) / \Psi_\infty(\bar{e})$ as $\lambda \to \infty$ (recall that $\bar{e}$ is the barycenter of Δ_n). It turns out that

$$(2.17) \qquad \lim_{\lambda \to \infty} \frac{\Psi_\infty(p)}{\Psi_\infty(\bar{e})} = \lim_{\lambda \to \infty} \frac{\sum_{i=1}^{n} \log(1 + \frac{\lambda}{\alpha} p_i)}{n \log(1 + \frac{\lambda}{\alpha} \frac{1}{n})} = 1, \quad p \in \Delta_n.$$

Accordingly, as $\lambda \to \infty$ the corresponding portfolio converges to the market portfolio $\pi(p) \equiv p$. This result suggests that the limits $\lim_{\lambda \to \infty} \lim_{K \to \infty}$ and $\lim_{K \to \infty} \lim_{\lambda \to \infty}$ are different in our model; the difference will become clear in the next section.

3. Weak limit for hardmin. Now we consider the case of hardmin $m_\infty = \min$ so that $\lambda_K = \infty$ for all K. We show that a suitable scaling gives a nontrivial limiting distribution. This distribution can be characterized in terms of its tail probability, that is, $\mathbb{P}(\Psi \geq \psi)$ for $\psi \in \mathcal{C}$, and can be realized by a Poisson point process via a novel duality.

3.1. *Main results.* In this section we impose the following conditions on the random vector C which defines the random hyperplane $\ell(\cdot) = \langle C, \cdot \rangle$.

ASSUMPTION 3.1. The random vector $C = (C_1, \ldots, C_n)$ has a joint density ρ on $\mathbb{R}^n_+$ which is asymptotically homogeneous of order α near the origin. More precisely, there exist $\alpha \in \mathbb{R}$ and a nonnegative measurable function h on $\mathbb{R}^n_+$ such that $\int_{\mathbb{R}^n_+} h(x)\, dx \in (0, \infty]$ and

$$(3.1) \qquad \lim_{\kappa \to 0^+} \frac{1}{\kappa^\alpha} \rho(\kappa x) = h(x), \quad x \in \mathbb{R}^n_+,$$

uniformly for $x \in \Delta_n$.

REMARK 3.2. Since we use the hardmin, as K grows the (unnormalized) minimum $\min\{\ell_1, \ldots, \ell_K\}$ becomes smaller and smaller. Consequently, the weak limit of the scaled minimum, if it exists, only depends on the distribution of C in a neighborhood of the origin. This consideration motivates Assumption 3.1.

Assumption 3.1 imposes rather strong conditions on the function h. In the next lemma we gather some properties that are used subsequently. The proof is given in the Appendix.

LEMMA 3.3. *Under Assumption 3.1 the exponent α is uniquely determined and $\alpha > -n$. Also $h(\kappa x) = \kappa^\alpha h(x)$ for all $\kappa > 0$ and $x \in \mathbb{R}^n_+$, so that h is homogeneous of order α. Moreover, we have*

$$(3.2) \qquad \lim_{\kappa \to 0^+} \frac{1}{\kappa^\alpha} \int_A \rho(\kappa x)\, dx = \int_A h(x)\, dx < \infty$$

for every bounded Borel set $A \subset \mathbb{R}^n_+$. Thus h is locally integrable.

Here we give some examples of random vectors that satisfy Assumption 3.1.

EXAMPLE 3.4.

(i) Suppose C has a continuous density ρ on $\mathbb{R}^n_+$ with $\lim_{x \to 0} \rho(x) = \gamma > 0$. Then (3.1) holds with $\alpha = 0$ and $h(x) \equiv \gamma$.

(ii) Suppose $C_1, \ldots, C_n$ are independent and C_i has the gamma distribution with shape parameter $1 + \alpha_i > 0$ and scale parameter β_i. The density of C is given by

$$\rho(x) = \prod_{i=1}^n \frac{\beta_i^{1+\alpha_i}}{\Gamma(1+\alpha_i)} x_i^{\alpha_i} e^{-\beta_i x_i}.$$

Then (3.1) holds with $\alpha = \sum_{i=1}^n \alpha_i$ and

$$h(x) = \prod_{i=1}^n \frac{\beta_i^{1+\alpha_i}}{\Gamma(1+\alpha_i)} x_i^{\alpha_i}.$$

Our first main result is that a weak limit exists under a suitable scaling.

THEOREM 3.5. *Under Assumption* 3.1, *as* $K \to \infty$ *the distribution* ν_K *of the random function* $\Psi_K := K^{\frac{1}{n+\alpha}} \min\{\ell_1, \ldots, \ell_K\}$ *converges weakly to a probability measure* μ *supported on* $\mathcal{C}_+ \subset \mathcal{C}$.

Further properties of the limiting distribution μ will be studied in this and the next sections. To help the reader navigate through the hardmin case, here are pointers to the main results:

- In Proposition 3.8 we prove weak convergence of the finite dimensional distributions. In particular, if $\Psi \sim \mu$, then for $p \in \Delta_n$ fixed the random variable $\Psi(p)^{n+\alpha}$ is exponentially distributed. We compute $\mathbb{E}[\Psi]$ in Corollary 4.1.
- In Theorem 3.12 we characterize μ in terms of the tail probability functional $\mathsf{T}_\mu(\psi) = \mu\{\omega \in \mathcal{C} : \omega \geq \psi\}$, $\psi \in \mathcal{C}$. Explicit formulas are given for an important special case in Theorems 4.7 and 4.8.
- In Theorem 3.14 we provide an explicit representation of the limit distribution via a Poisson point process on the positive quadrant $\mathbb{R}_+^n$. Consequently, if $\Psi \sim \mu$ then Ψ is locally piecewise affine.
- Further properties, including boundary behaviors and connections to portfolio maps, are given in Theorem 4.6 and Proposition 4.14.

REMARK 3.6 (Connections with extreme value theory). Theorem 3.5 is a limit theorem for the minimum of (properly rescaled) i.i.d. hyerplanes over the simplex. As shown in the classic references [5, 19, 26], extrema of i.i.d. random variables, random vectors and stochastic processes have been studied extensively in extreme value theory. It is customary in this theory to consider maxima instead of minima. Expressing our results in terms of maximums (by considering instead negative convex functions), we have that the marginal $-\Psi(p)$, for $p \in \Delta_n$, has a Weibull-type distribution with extreme value index -1 (see, e.g., [5], Theorem 1.1.3). To the best of our knowledge, the constraint of concavity and the unit simplex have not been considered in the literature. In particular, by using an elegant duality for positive concave functions on Δ_n, to be described in the next subsection, we are able to obtain deep properties of the limiting distribution. These results suggest that concave functions on the unit simplex are natural and fundamental objects to study. Also see Remark 3.15 for more discussions about the Poisson representation (Theorem 3.14). Finally, we remark that our general framework (2.13) involves the softmin which is rarely, if at all, considered in extreme value theory.

3.2. *Duality for nonnegative concave functions on* $\overline{\Delta}_n$. Duality plays a major role in the proof of Theorem 3.5 and several other results of this paper. For nonnegative concave functions on the simplex the duality has an elegant form which is useful for our analysis. The reason is that each positive affine function on $\overline{\Delta}_n$ is specified by its values over the vertices $e_1, \ldots, e_n$ (see (2.2) and Figure 1), so we do not need to specify the constant term separately.

For $p \in \Delta_n$ and $a > 0$, we denote by $R(p, a)$ the region

$$(3.3) \qquad R(p, a) := \{x \in \mathbb{R}_+^n : \langle p, x \rangle < a\},$$

which is an open convex set in $\mathbb{R}_+^n$. Since the hyperplane $\langle p, x \rangle = a$ intersects the ith coordinate axis at $x_i = a/p_i$, the Euclidean volume of $R(p, a)$ is

$$(3.4) \qquad \mathrm{vol}(R(p, a)) = \frac{a^n}{n! p_1 p_2 \cdots p_n}.$$

If $\mathbf{p} = (p^{(1)}, p^{(2)}, \ldots, p^{(r)})$ is a collection of r distinct points in Δ_n and $\mathbf{a} = (a_1, \ldots, a_r)$ is a collection of positive real numbers, we define

$$(3.5) \qquad \widehat{R}(\mathbf{p}, \mathbf{a}) := R(p^{(1)}, a_1) \cup \cdots \cup R(p^{(r)}, a_r).$$
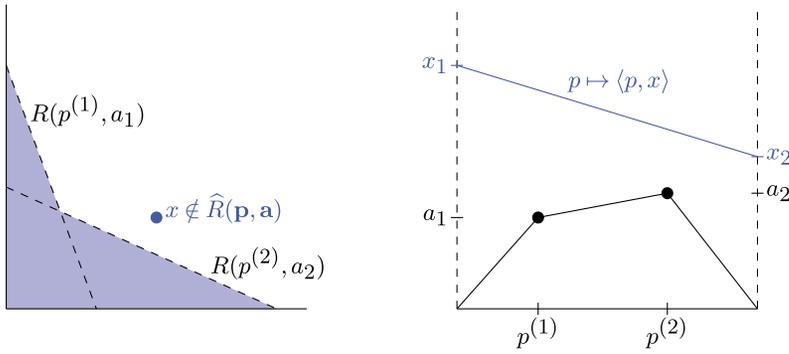
FIG. 4. *Left*: *The region* $\widehat{R}(\mathbf{p}, \mathbf{a})$ *when* $r = 2$, *and a point* $x \notin \widehat{R}(\mathbf{p}, \mathbf{a})$. *Right*: $x \notin \widehat{R}(\mathbf{p}, \mathbf{a})$ *if and only if* $\langle p, x \rangle \geq \psi_{\mathbf{p}, \mathbf{a}}(p)$.

Identifying $x \in \mathbb{R}_+^n$ with the positive affine function $p \mapsto \langle p, x \rangle$ on Δ_n, we see that $x \notin \widehat{R}(\mathbf{p}, \mathbf{a})$ if and only if $\langle p^{(i)}, x \rangle \geq a_i$ for all i. This implies that $\langle p, x \rangle \geq \psi_{\mathbf{p}, \mathbf{a}}(p)$ on $\overline{\Delta}_n$, where

$$(3.6) \qquad \psi_{\mathbf{p}, \mathbf{a}} := \inf\{\psi \in \mathcal{C} : \psi(p^{(i)}) \geq a_i, i = 1, \ldots, r\}$$

is the smallest nonnegative concave function generated by the given data. See Figure 4 for an illustration where $n = 2$ and we identify $\overline{\Delta}_2$ with the interval from e_1 to e_2. Since the domain is the unit simplex the function $\psi_{\mathbf{p}, \mathbf{a}}$ is polyhedral, that is, it is the minimum of a finite collection of hyperplanes.

More generally, for $\psi \in \mathcal{C}_+$ we define

$$(3.7) \qquad \widehat{R}(\psi) := \bigcup_{p \in \Delta_n} R(p, \psi(p))$$

and

$$(3.8) \qquad \widehat{S}(\psi) := \mathbb{R}_+^n \setminus \widehat{R}(\psi) = \{x \in \mathbb{R}_+^n : \langle p, x \rangle \geq \psi(p) \text{ for all } p \in \Delta_n\}.$$

Note that $\widehat{S}(\psi)$ is a convex set, and by the identification $x \leftrightarrow (p \mapsto \langle p, x \rangle)$ we see that $\widehat{S}(\psi)$ is equivalent to the set $\{\ell \in \mathcal{A}_+ : \ell \geq \psi\}$, so the operation $\psi \in \mathcal{C}_+ \mapsto \widehat{S}(\psi)$ is one-to-one. Thus the set $\widehat{S}(\psi)$ plays the role of the conjugate. This duality can be formalized by adapting the concept of support function from convex analysis (see [27], Section 13).

PROPOSITION 3.7. *For any* $\psi \in \mathcal{C}_+$ *we have*

$$\psi(p) = \inf_{x \in \widehat{S}(\psi)} \langle p, x \rangle \quad \text{for } p \in \overline{\Delta}_n.$$

PROOF. Let $x \in \widehat{S}(\psi)$. By definition we have $\langle p, x \rangle \geq \psi(p)$ for $p \in \Delta_n$, and by continuity the inequality extends to $\overline{\Delta}_n$. Taking the infimum over x we have

$$(3.9) \qquad \inf_{x \in \widehat{S}(\psi)} \langle p, x \rangle \geq \psi(p) \quad \text{for } p \in \overline{\Delta}_n.$$

For the other direction, let $p \in \Delta_n$ be an interior point so that $\partial \psi(p)$ is nonempty. So there exists $x \in [0, \infty)^n$ such that $\langle q, x \rangle \geq \psi(q)$ for all $q \in \overline{\Delta}_n$ and $\langle p, x \rangle = \psi(p)$. If $x \in \mathbb{R}_+^n$, then $x \in \widehat{S}(\psi)$ and we have

$$\inf_{y \in \widehat{S}(\psi)} \langle p, y \rangle \leq \langle p, x \rangle = \psi(p).$$

If not, then $x^{(r)} = x + (\frac{1}{r}, \ldots, \frac{1}{r}) \in \widehat{S}(\psi)$ for all $r \geq 1$, and so

$$\inf_{y \in \widehat{S}(\psi)} \langle p, y \rangle \leq \liminf_{r \to \infty} \langle p, x^{(r)} \rangle = \langle p, x \rangle = \psi(p).$$

Being the infimum of a collection of hyperplanes, $\inf_{x \in \widehat{S}(\psi)} \langle p, x \rangle$ is concave whose hypograph is closed. By [27], Theorem 10.2, it is continuous on $\overline{\Delta}_n$. Since $\psi \in \mathcal{C}_+$ is also continuous, the equality extends to the boundary. $\square$

3.3. *Proof of Theorem 3.5.* Recall that $\Psi_K = K^{\frac{1}{n+\alpha}} \min_{1 \leq k \leq K} \ell_k$ is the normalized minimum of K i.i.d. hyperplanes. First we show convergence of the finite dimensional distributions.

PROPOSITION 3.8. *Let $p^{(1)}, \ldots, p^{(r)}$ be a collection of distinct points in Δ_n, and let $a_1, \ldots, a_r > 0$. Denote the data by $(\mathbf{p}, \mathbf{a})$. Then*

$$(3.10) \qquad \lim_{K \to \infty} \mathbb{P}\big(\Psi_K(p^{(i)}) \geq a_i, i = 1, \ldots, r\big) = \exp\Big(- \int_{\widehat{R}(\mathbf{p},\mathbf{a})} h(x)\, dx\Big),$$

where h is given by (3.1). Thus, the joint distribution of the random vector $(\Psi_K(p^{(i)}))_{1 \leq i \leq r}$ converges weakly to the distribution defined by the right hand side of (3.10).

In particular, for $p \in \Delta_n$ fixed, let $\Psi(p)$ (which is a real-valued random variable) be distributed as the weak limit of $\Psi_K(p)$ as $K \to \infty$. Then the random variable $\Psi(p)^{n+\alpha}$ is exponentially distributed with rate $\int_{R(p,1)} h(y)\, dy$.

PROOF. We have

$$\mathbb{P}\big(\Psi_K(p^{(i)}) \geq a_i, i = 1, \ldots, r\big) = \mathbb{P}\Big(\min_{1 \leq k \leq K} \ell_k(p^{(i)}) \geq a_i K^{\frac{-1}{n+\alpha}} \forall i\Big)$$

$$= \mathbb{P}\big(\ell(p^{(i)}) \geq a_i K^{\frac{-1}{n+\alpha}} \forall i\big)^K.$$

Note that the last inequality holds since the random vector C (and hence $\ell(p) = \langle p, C \rangle$) has a density. We write

$$(3.11) \qquad \mathbb{P}\big(\ell(p^{(i)}) \geq a_i K^{\frac{-1}{n+\alpha}} \forall i\big)^K = \big(1 - \mathbb{P}(C \in K^{\frac{-1}{n+\alpha}} \widehat{R}(\mathbf{p}, \mathbf{a}))\big)^K.$$

We claim that

$$(3.12) \qquad \lim_{K \to \infty} K \mathbb{P}\big(C \in K^{\frac{-1}{n+\alpha}} \widehat{R}(\mathbf{p}, \mathbf{a})\big) = \int_{\widehat{R}(\mathbf{p},\mathbf{a})} h(x)\, dx.$$

Assuming (3.12), we may take limit in (3.11) to get

$$\lim_{K \to \infty} \mathbb{P}\big(\Psi_K(p^{(i)}) \geq a_i, i = 1, \ldots, r\big) = \lim_{K \to \infty} \Big(1 - \frac{1}{K} \int_{\widehat{R}(\mathbf{p},\mathbf{a})} h(x)\, dx\Big)^K$$

$$= \exp\Big(- \int_{\widehat{R}(\mathbf{p},\mathbf{a})} h(x)\, dx\Big),$$

which is the desired limit.

To prove (3.12), note that

$$K \mathbb{P}\big(C \in K^{\frac{-1}{n+\alpha}} \widehat{R}(\mathbf{p}, \mathbf{a})\big) = K \int_{K^{\frac{-1}{n+\alpha}} \widehat{R}(\mathbf{p},\mathbf{a})} \rho(x)\, dx$$

$$= K \int_{\widehat{R}(\mathbf{p},\mathbf{a})} \rho\big(K^{\frac{-1}{n+\alpha}} y\big) K^{\frac{-n}{n+\alpha}}\, dy$$

$$= \int_{\widehat{R}(\mathbf{p},\mathbf{a})} \frac{\rho\big(K^{\frac{-1}{n+\alpha}} y\big)}{K^{\frac{-\alpha}{n+\alpha}}}\, dy.$$

By Lemma 3.3, this converges to the integral of h over $\widehat{R}(\mathbf{p}, \mathbf{a})$.

Now let $p \in \Delta_n$ be given. Applying the above with $r = 1$, for $x > 0$ we have

$$\mathbb{P}\big(\Psi(p) \geq x\big) = \exp\bigg(-\int_{R(p,x)} h(y)\,dy\bigg).$$

Since $R(p, x) = x R(p, 1)$ and h is homogeneous with order α, we have

$$\int_{R(p,x)} h(y)\,dy = x^{n+\alpha} \int_{R(p,1)} h(y)\,dy.$$

Thus

$$\mathbb{P}\big(\Psi(p)^{n+\alpha} \geq x\big) = \mathbb{P}\big(\Psi(p) \geq x^{\frac{1}{n+\alpha}}\big) = \exp\bigg(-x \int_{R(p,1)} h(y)\,dy\bigg),$$

that is, $\Psi(p)^{n+\alpha}$ is exponentially distributed with rate $\int_{R(p,1)} h(y)\,dy$. $\quad\square$

Let ν_K be the law of Ψ_K regarded as random elements of the metric space $(\mathcal{C}, d)$.

LEMMA 3.9.   *The sequence $\{\nu_K\}_{K \geq 1}$ is tight.*

PROOF.   By Lemma 2.7, we have the bound

(3.13) $\psi(p) \leq n\psi(\overline{e}),$

which holds for any $\psi \in \mathcal{C}$.

Let $\epsilon > 0$ be given. By (3.10), the family of univariate distributions corresponding to $\{\Psi_K(\overline{e})\}_{K \geq 1}$ is tight. Thus, there exists $M > 0$ such that

$$\mathbb{P}\big(\Psi_K(\overline{e}) \leq M\big) \geq 1 - \epsilon \quad \forall K \geq 1.$$

By Lemma 2.1, the set $\mathcal{K} = \{\psi \in \mathcal{C} : \psi \leq nM\}$ is compact in $\mathcal{C}$. Using the uniform estimate (3.13), for any $K \geq 1$ we have

$$\nu_K(\mathcal{K}) \geq \mathbb{P}\big(\Psi_K(\overline{e}) \leq M\big) \geq 1 - \epsilon.$$

This establishes the tightness of $\{\nu_K\}_{K \geq 1}$. $\quad\square$

Now we are ready to prove Theorem 3.5.

PROOF OF THEOREM 3.5.   By Lemma 3.9 and Prokhorov's theorem, the sequence $\{\nu_K\}$ is relatively compact in the topology of weak convergence. This means that for any subsequence $\{\nu_{K'}\}$ of $\{\nu_K\}$, there exists a further subsequence $\{\nu_{K''}\}$ that converges weakly to some probability measure, say ν^*, on $\mathcal{C}$. However, by Proposition 3.8 the finite dimensional distributions of ν^* are given by the right hand side of (3.10) which does not depend on the subsequence chosen. Thus by Lemma 2.2 there is a unique weak limit point. Consequently, the original sequence $\{\nu_K\}$ converges weakly to $\mu = \nu^*$ whose finite dimensional marginals are given by the right hand side of (3.10).

It remains to verify that μ is supported on $\mathcal{C}_+$, the subset of functions in $\mathcal{C}$ that are strictly positive on Δ_n. By Lemma 2.7, if $\psi \in \mathcal{C}$ is positive at some $p \in \Delta_n$, then $\psi(q) > 0$ for all $q \in \Delta_n$. From (3.10) it is clear that if $\Psi \sim \mu$ then $\Psi(p) > 0$ with probability 1 for any $p \in \Delta_n$. This implies that $\mu(\mathcal{C}_+) = 1$ and the theorem is proved. $\quad\square$

3.4. *Tail probability.* Consider the limiting distribution μ given in Theorem 3.5. Proposition 3.8 characterizes the finite dimensional distributions of μ. Now we extend this result to the tail probability defined as follows.

DEFINITION 3.10. Given a Borel probability measure ν on $\mathcal{C}$, we define its tail probability as the functional $\mathsf{T}_\nu : \mathcal{C} \to [0, 1]$ defined by

$$(3.14) \qquad \mathsf{T}_\nu(\psi) := \nu\{\omega \in \mathcal{C} : \omega \geq \psi\}, \quad \psi \in \mathcal{C}.$$

We show that the tail probability T_ν characterizes ν. Recall from (3.6) that for a collection of points $\mathbf{p} = (p^{(1)}, \ldots, p^{(r)})$ in Δ_n and $\mathbf{a} = (a_1, \ldots, a_r)$ in $(0, \infty)$, $\psi_{\mathbf{p},\mathbf{a}}$ is the smallest concave function such that $\psi(p^{(i)}) \geq a_i$ for $i = 1, \ldots, r$.

LEMMA 3.11. *Let ν be a probability measure on $\mathcal{C}$. Given $\mathbf{p}$ and $\mathbf{a}$ as above, we have*

$$\nu\{\omega : \omega(p^{(i)}) \geq a_i\} = \nu\{\omega : \omega \geq \psi_{\mathbf{p},\mathbf{a}}\} = \mathsf{T}_\nu(\psi_{\mathbf{p},\mathbf{a}}).$$

Consequently the tail probability T_ν fully characterizes the measure ν.

PROOF. By definition, if $\omega \geq \psi_{\mathbf{p},\mathbf{a}}$ then clearly $\omega(p^{(i)}) \geq a_i$ for all i, and so

$$\nu\{\omega : \omega(p^{(i)}) \geq a_i\} \geq \mathsf{T}_\nu(\psi_{\mathbf{p},\mathbf{a}}).$$

On the other hand, if $\omega \in \mathcal{C}$ is such that $\omega(p^{(i)}) \geq a_i$ for all i, then by concavity of ω we have $\omega \geq \psi_{\mathbf{p},\mathbf{a}}$. This gives the reverse inequality. The last assertion follows from Lemma 2.2. □

Now we characterize the tail probability functional of the limiting measure μ. Recall that for $\psi \in \mathcal{C}$ we define $\widehat{R}(\psi) = \bigcup_{p \in \Delta_n} R(p, \psi(p)) \subset \mathbb{R}^n_+$.

THEOREM 3.12. *Consider the limiting measure μ as in Theorem 3.5. We have*

$$(3.15) \qquad \mathsf{T}_\mu(\psi) = \exp\left(-\int_{\widehat{R}(\psi)} h(x)\, dx\right) \quad \text{for } \psi \in \mathcal{C}.$$

PROOF. By Proposition 3.8 and Lemma 3.11, we have that (3.15) holds whenever $\psi = \psi_{\mathbf{p},\mathbf{a}}$ for some $\mathbf{p}$ and $\mathbf{a}$. We now extend this identity to an arbitrary $\psi \in \mathcal{C}$.

Let $D = \{p^{(1)}, p^{(2)}, \ldots\}$ be a countable dense set of Δ_n. By continuity of $\omega \in \mathcal{C}$, we have

$$\{\omega : \omega \geq \psi\} = \bigcap_{r=1}^\infty \{\omega : \omega(p^{(i)}) \geq \psi(p^{(i)}) \text{ for } 1 \leq i \leq r\}.$$

Since the events on the right hand side are decreasing in r, we have

$$(3.16) \qquad \mathbb{P}\{\omega : \omega \geq \psi\} = \lim_{r \to \infty} \mathbb{P}\{\omega : \omega(p^{(i)}) \geq \psi(p^{(i)}) \text{ for } 1 \leq i \leq r\}$$

$$(3.17) \qquad = \lim_{r \to \infty} \exp\left(-\int_{\widehat{R}_r} h(x)\, dx\right),$$

where $\widehat{R}_r$ is the $\widehat{R}(\mathbf{p}, \mathbf{a})$ generated by $p^{(1)}, \ldots, p^{(r)}$ and the values of ψ at these points.

It remains to show that $\lim_{r \to \infty} \int_{\widehat{R}_r} h(x)\, dx = \int_{\widehat{R}(\psi)} h(x)\, dx$. Indeed, we have

$$(3.18) \qquad \widehat{R}(\psi) = \bigcup_{i=1}^\infty R(p^{(i)}, \psi(p^{(i)})).$$

To see this, suppose by way of contradiction that there exists an element x of $\widehat{R}(\psi) \setminus \bigcup_{i=1}^\infty R(p^{(i)}, \psi(p^{(i)}))$. Then there exists $p \in \Delta_n$ such that $\langle p, x \rangle < \psi(p)$ but $\langle p', x \rangle \geq \psi(p')$ for all $p' \in D$. This is clearly a contradiction since we can approach p by a sequence of points p' in D. Now (3.18) and the monotone convergence theorem imply that $\lim_{r \to \infty} \int_{\widehat{R}_r} h(x)\, dx = \int_{\widehat{R}(\psi)} h(x)\, dx$. □

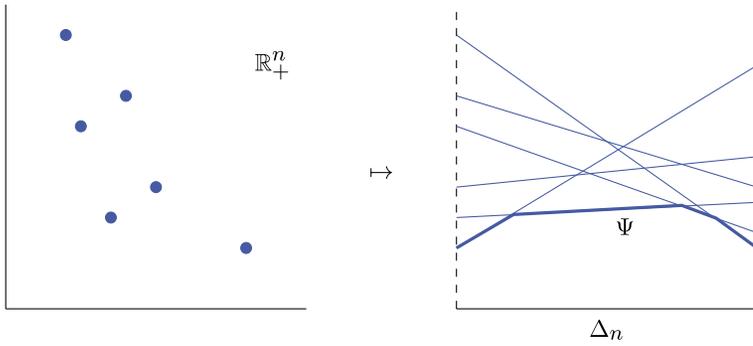FIG. 5.   *A realization of the point process N defines a nonnegative concave function via the correspondence between $\mathbb{R}_+^n$ and $\mathcal{A}_+$.*

3.5. *Realization of the weak limit by Poisson point processes.*   In Theorem 3.12 we may interpret the Borel set function

$$(3.19) \qquad\qquad A \mapsto \int_A h(x)\, dx$$

as a Radon measure m on $\mathbb{R}^n$, so that the integral in (3.15) is equal to $m(\widehat{R}(\psi))$. More generally, given an arbitrary Radon measure m on $\mathbb{R}_+^n$, can we construct a probability measure ν on $\mathcal{C}$ whose tail probability is given by

$$(3.20) \qquad\qquad \mathsf{T}_\nu(\psi) = \exp(-m(\widehat{R}(\psi)))?$$

We show that this can be achieved by a Poisson point process. This gives a direct probabilistic construction of the limiting distribution μ in Theorem 3.5 without going through the limiting process, and suggests an algorithm for simulating samples from μ. We leave practical implementation (possibly tailored to the financial applications) as a problem for future research.

We begin by recalling the defining property of a Poisson point process (for details see [18, 26]). Let m be a Radon measure on $\mathbb{R}_+^n$. A Poisson point process with intensity measure m is a random closed set N such that for any bounded Borel set $A \subset \mathbb{R}_+^n$, the random variable $|N \cap A|$ (here $|\cdot|$ denotes the cardinality of a set) follows the Poisson distribution with rate $m(A)$, and if $A_1, \ldots, A_m$ are disjoint, then $|N \cap A_1|, \ldots, |N \cap A_m|$ are independent.

Recall from Section 2.1 that any point of $\mathbb{R}_+^n$ can be identified with a positive affine function on Δ_n. Given the random set N, we can define a random concave function Ψ by

$$(3.21) \qquad\qquad \Psi(p) = \inf_{x \in N} \langle p, x \rangle, \quad p \in \Delta_n.$$

See Figure 5 for an illustration of this construction. It is easy to see that, as long as N is locally finite, the random concave function Ψ is locally piecewise affine on Δ_n. Note that this property may not hold in the diagonal softmin case studied in Section 5; see, in particular, Figure 8 where the realizations are almost surely differentiable on Δ_n.

REMARK 3.13.   It is clear that (3.21) makes sense for any point process (not necessarily Poisson). However, at this generality it is difficult to obtain concrete results about the resulting distribution apart from the tail probability.

THEOREM 3.14.   *Let ν be the distribution of the random concave function Ψ defined in (3.21), where N is a Poisson point process on $\mathbb{R}_+^n$ with intensity measure m. Then its tail probability functional is given by (3.20). In particular, the limiting distribution μ in Theorem 3.5 is the special case where the intensity measure m has density h with respect to Lebesgue measure.*

PROOF. Let $\psi \in \mathcal{C}$ be given and consider the event $\{\Psi \geq \psi\}$. Note that its complement $\{\Psi \geq \psi\}^c$ occurs if and only if there exists $p \in \Delta_n$ such that

$$\Psi(p) = \inf_{x \in N} \langle p, x \rangle < \psi(p),$$

that is, $N \cap \widehat{R}(\psi) \neq \varnothing$. It follows that

$$\mathbb{P}(\Psi \geq \psi) = \mathbb{P}(N \cap \widehat{R}(\psi) = \varnothing) = e^{-m(\widehat{R}(\psi))}$$

and the theorem is proved. $\square$

REMARK 3.15. Poisson point processes arise naturally in characterizations of extreme value distributions. In [11], Theorem 2.4, a representation of max-infinitely divisible and sample continuous processes (which arise as limits of pointwise maxima of i.i.d. processes) on a compact metric space is given in terms of Poisson point processes. In this general result, the point process lives on the function space itself, and the Poisson representation is not easy to apply directly. Using the duality in Section 3.2, we are able to obtain a clean and explicit probabilistic representation of the limiting distribution μ.

4. Further properties of the limiting distribution. Consider the limiting distribution μ in Theorem 3.5. Equivalently, we may consider the distribution constructed in Theorem 3.14 where the intensity measure has the form (3.19) and h satisfies the conditions of Lemma 3.3. In this section we develop further properties of μ, sometimes under additional conditions. Throughout this section we let Ψ be a random element in $\mathcal{C}$ with distribution μ.

4.1. *Exponential distribution and the geometric mean.* In Proposition 3.8 we have shown that $\Phi^{n+\alpha}(p)$ is exponentially distributed for $p \in \Delta_n$ fixed. A direct calculation using the exponential distribution gives a formula for the expected value of $\Psi(\cdot)$.

COROLLARY 4.1. *For any $p \in \Delta_n$, we have*

$$(4.1) \qquad \mathbb{E}[\Psi(p)] = \Gamma\left(\frac{1}{n+\alpha} + 1\right)\left(\int_{R(p,1)} h(y) \, dy\right)^{-\frac{1}{n+\alpha}},$$

where $\Gamma(\cdot)$ is the gamma function.

REMARK 4.2. As a by-product we have the following result: for any h on $\mathbb{R}_+^n$ which is homogeneous of order α and locally integrable, the function defined by the right hand side of (4.1) is concave. In Proposition 4.14 below we compute the derivative of $\mathbb{E}[\Psi(p)]$ and the corresponding portfolio map.

Under the following condition we can derive an explicit formula for the expected value of $\Psi(\cdot)$.

ASSUMPTION 4.3. In Assumption 3.1 we assume that h has the form

$$(4.2) \qquad h(x) = \gamma \prod_{i=1}^{n} x_i^{\alpha_i}$$

for some $\gamma > 0$ and exponents $\alpha_i > -1$ (see Example 3.4).

Note that the homogeneity of h implies that $\alpha = \sum_{i=1}^{n} \alpha_i$, and we have $n + \alpha > 0$.
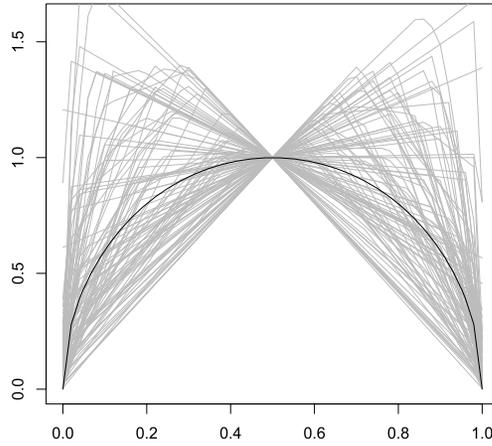
FIG. 6. *Samples of $\Psi_K(\cdot)/\Psi_K(\overline{e})$ where $n = 2$, $K = 300$ and the components of C are independent $U(0, 1)$ random variables (thus $\alpha_1 = \alpha_2 = 0$). The x-axis is $p_1 = 1 - p_2$. The solid black curve is the (limiting) expected value given by the geometric mean $2\sqrt{p_1 p_2}$.*

PROPOSITION 4.4. *Under Assumption 4.3, let $\pi_i = \frac{1+\alpha_i}{n+\alpha}$ so that $\pi = (\pi_1, \ldots, \pi_n) \in \Delta_n$. Also let*

(4.3) $$M = \gamma \int_{R(\overline{e}, 1/n)} \prod_{i=1}^n y_i^{\alpha_i} \, dy = \int_{R(\overline{e}, 1/n)} h(y) \, dy > 0.$$

Then for $p \in \overline{\Delta}_n$ we have

(4.4) $$\mathbb{E}\big[\Psi(p)\big] = L p_1^{\pi_1} \cdots p_n^{\pi_n},$$

where $L = \Gamma(\frac{1}{n+\alpha} + 1)/M^{1/(n+\alpha)}$.

PROOF. For $p \in \Delta_n$, letting $u_i = p_i y_i$, we have

(4.5) $$\int_{R(p,1)} h(y) \, dy = \frac{\gamma}{p_1^{1+\alpha_1} \cdots p_n^{1+\alpha_n}} \int_{R(\overline{e}, 1/n)} \prod_{i=1}^n u_i^{\alpha_i} \, du = \frac{M}{p_1^{1+\alpha_1} \cdots p_n^{1+\alpha_n}},$$

and (4.4) follows by Corollary 4.1. By the dominated convergence theorem we may extend (4.4) to the boundary. □

From Proposition 4.4, the geometric mean emerges as a limiting average shape of the random concave functions $\Psi_K = K^{\frac{1}{n+\alpha}} \min\{\ell_1, \ldots, \ell_K\}$. See Figure 6 for an illustration where we plot instead the normalized value $\Psi_K(\cdot)/\Psi_K(\overline{e})$ (so that the value is 1 at the barycenter). This result is interesting because the geometric mean (4.4) plays a fundamental role in stochastic portfolio theory and the Dirichlet transport. Indeed, as a portfolio generating function it generates, in the sense of (1.4), the constant-weighted portfolio $\pi(p) \equiv \pi$. In Section 4.4 we give further results about the random portfolio map generated by Ψ.

4.2. *Boundary behavior.* Next we study the behavior of the random function Ψ near the boundary $\partial \Delta_n$ of the simplex. In this subsection we work under Assumption 4.3. First we show that Ψ vanishes on the boundary almost surely.

PROPOSITION 4.5. *Under Assumption 4.3 we have $\mathbb{P}(\Psi|_{\partial \Delta_n} \equiv 0) = 1$.*

PROOF. Let $p \in \partial \Delta_n$ and let $\{p^{(r)}\}_{r=1}^{\infty}$ be a sequence in Δ_n converging to p. By continuity of Ψ, we have the almost sure limit

$$(4.6) \qquad \Psi(p) = \lim_{r \to \infty} \Psi(p^{(r)}).$$

Using the convergence (4.6) (which implies weak convergence), the Portmanteau lemma and Proposition 3.8, we have

$$\mathbb{P}(\Psi(p) > x) \leq \limsup_{r \to \infty} \exp\left(-x^{n+\alpha} \int_{R(p^{(r)}, 1)} h(y) \, dy\right)$$

for all $x > 0$. By (4.5), we have $\int_{R(p^{(r)}, 1)} h(y) \, dy \to \infty$ as $r \to \infty$. Thus $\Psi(p) = 0$ almost surely for any fixed $p \in \partial \Delta_n$. The previous argument then implies that $\mathbb{P}(\Psi|_D \equiv 0) = 1$, where D is a countable dense set of $\partial \Delta_n$. By the continuity of Ψ we have $\mathbb{P}(\Psi|_{\partial \Delta_n} \equiv 0) = 1$. $\square$

Indeed, under the same assumptions we can derive a Hölder estimate for Ψ near the boundary.

THEOREM 4.6. *Suppose Assumption 4.3 holds. Let $\pi_i = \frac{1+\alpha_i}{n+\alpha}$, $\kappa > 0$ and $\beta > 0$. Fix $j \in \{1, 2, \ldots, n\}$. Then, almost surely, there exists $\delta = \delta(\omega) > 0$ depending on $\omega \in \Omega$ such that*

$$(4.7) \qquad \Psi(p) \leq \kappa p_j^{(1-\beta)\pi_j}, \quad \text{for } p_j \leq \delta.$$

PROOF. We consider the behavior of Ψ near the face $p_j = 0$. Recall the notation $c^{(j,\epsilon)}$ defined in (2.10). Define the events

$$B(a, \epsilon) = \{\Psi(c^{(j,\epsilon)}) > a\}.$$

By Proposition 3.8 and (4.5), we have

$$\mathbb{P}(B(a, \epsilon)) = \exp\left\{-\frac{Ma^{n+\alpha}}{\epsilon^{1+\alpha_j}} \left(\frac{n-1}{1-\epsilon}\right)^{n-1+\alpha-\alpha_j}\right\},$$

where M is the constant defined by (4.3). Now let $B_k = B(a_k, \epsilon_k)$ where $\epsilon_k = \frac{1}{k}$ and

$$a_k = \frac{\kappa}{n}\left(\frac{1}{k+1}\right)^{\frac{1+\alpha_j}{n+\alpha}(1-\beta)} = \frac{\kappa}{n}\left(\frac{1}{k+1}\right)^{(1-\beta)\pi_j}.$$

With this choice of ϵ_k and a_k we have $\sum_{k=1}^{\infty} \mathbb{P}(B_k) < \infty$. By the Borel–Cantelli lemma, we have $\mathbb{P}(B_k \text{ i.o.}) = 0$, so that almost surely, there exists $k_0 = k_0(\omega)$ such that

$$\Psi(c^{(j,1/k)}) \leq \frac{\kappa}{n}\left(\frac{1}{k+1}\right)^{(1-\beta)\pi_j} \qquad \text{for } k \geq k_0.$$

Now define $\delta = \min\{1/k_0, 1/n\}$. If $p \in \Delta_n$ satisfies $p_j \leq \delta$, let $k \geq 1/\delta \geq k_0$ such that $\frac{1}{k+1} \leq p_j \leq \frac{1}{k}$. Then, from (2.11) in Lemma 2.7, we have

$$\Psi(p) \leq n\Psi(c^{(j,1/k)}) \leq \kappa p_j^{(1-\beta)\pi_j}. \qquad \square$$

4.3. *Explicit formula of the tail probability.* In this subsection we specialize to the case where $h(x) \equiv \gamma > 0$ on $\mathbb{R}_+^n$, that is, the intensity measure is a constant multiple of the Lebesgue measure. By Theorem 3.12, the tail probability of μ is given by

$$\mathsf{T}_\mu(\psi) = \exp(-\gamma \operatorname{vol}(\widehat{R}(\psi))), \quad \psi \in \mathcal{C},$$

where the volume $\operatorname{vol}(\widehat{R}(\psi))$ means the Lebesgue measure of $\widehat{R}(\psi)$. To understand it further it is desirable to have an explicit formula of the volume of the region $\widehat{R}(\psi)$. This will be computed under the assumption that ψ is C^2 (twice continuously differentiable). The computation, which is differential geometric, reveals an interesting geometric structure of the duality $\psi \leftrightarrow \widehat{R}(\psi)$ which may be of independent interest. We believe that the resulting formula can be extended to the general nonsmooth case by using the Monge–Ampère measure (see [6], Section 2.1), but this will not be pursued further in this paper.

First we introduce some notation that will be used in this subsection. We let $(q_1, \ldots, q_{n-1})$ be the coordinate system on Δ_n obtained by dropping the last component of $p = (p_1, \ldots, p_n)$. The domain of q is

$$D_{n-1} := \left\{ (q_1, \ldots, q_{n-1}) \in (0, 1) : \sum_{i=1}^{n-1} q_i < 1 \right\}.$$

Given $\psi \in \mathcal{C}$, by an abuse of notation we also regard it as a function of q on D_{n-1}:

$$\psi(q_1, \ldots, q_{n-1}) = \psi\left(q_1, \ldots, q_{n-1}, 1 - \sum_{i=1}^{n-1} q_i\right).$$

Also, by $D^2\psi(q)$ we mean the Hessian matrix of ψ as a function of the $(n-1)$-dimensional variable q.

To illustrate the technique we first assume that $\widehat{R}(\psi)$ is bounded. Note that by Proposition 4.5, if $\Psi \sim \mu$ then Ψ vanishes on the boundary. So in the tail probability it suffices to consider only functions that vanish on $\partial\Delta_n$.

THEOREM 4.7. *Suppose $\psi \in \mathcal{C}$ is C^2 on Δ_n, $\psi|_{\partial\Delta_n} \equiv 0$ and $\widehat{R}(\psi)$ is bounded. Then*

$$(4.8) \qquad \operatorname{vol}(\widehat{R}(\psi)) = \frac{1}{n} \int_{D_{n-1}} \psi(q) \det(-D^2\psi(q)) \, dq.$$

In particular, when $h(x) \equiv \gamma > 0$ we have

$$(4.9) \qquad \mathsf{T}_\mu(\psi) = \exp\left(\frac{-\gamma}{n} \int_{D_{n-1}} \psi(q) \det(-D^2\psi(q)) \, dq\right).$$

PROOF. Under the stated assumptions, the closure of $\widehat{R}(\psi)$ is an n-dimensional orientable compact manifold, denoted by M and embedded in $\mathbb{R}^n$, with piecewise C^1 boundary. The Euclidean coordinates $(x_1, \ldots, x_n)$ of $\mathbb{R}^n$ is a global coordinate system of M.

We will apply Stokes' theorem (see, e.g., [20])

$$(4.10) \qquad \int_{\partial M} \omega = \int_M d\omega$$

with the differential $(n-1)$-form given by

$$(4.11) \qquad \omega = \sum_{k=1}^n (-1)^{k+1} x_k \, dx_1 \wedge \cdots \wedge \widehat{dx_k} \wedge \cdots \wedge dx_n.$$

Here the notation $\widehat{dx_k}$ means that the term dx_k is omitted in the wedge product. It is immediate to check that the exterior derivative of ω is given by

$$d\omega = n\,dx_1 \wedge \cdots \wedge dx_n,$$

so that the right hand side of (4.10) is

$$\int_M d\omega = n\,\mathrm{vol}\big(\widehat{R}(\psi)\big).$$

It remains to compute $\int_{\partial M} \omega$. The boundary ∂M of M consists of parts of the coordinate hyperplanes as well as the (curved) part $S \subset \partial M$ parameterized by

$$(4.12) \qquad X_i(p) = \psi(p) + D\psi(p) \cdot (e_i - p), \quad 1 \le i \le n,\ p \in \Delta_n.$$

First we reparameterize N in terms of $q \in D_{n-1}$. Denote the partial derivatives of ψ (as a function of q) by

$$(4.13) \qquad \psi_i(q) = \frac{\partial \psi}{\partial q_i}(q) = D\psi(p) \cdot (e_i - e_n).$$

Since

$$e_i - p = e_i - e_n - \sum_{r=1}^{n-1} p_r(e_r - e_n),$$

we have

$$D\psi(p) \cdot (e_i - p) = \begin{cases} \psi_i(q) - \displaystyle\sum_{r=1}^{n-1} q_r \psi_r(q) & \text{if } 1 \le i \le n-1, \\[2mm] -\displaystyle\sum_{r=1}^{n-1} q_r \psi_r(q) & \text{if } i = n. \end{cases}$$

Writing X as a function of q, we can rewrite (4.12) as

$$X_i(q) = \begin{cases} \psi(q) + \psi_i(q) - \displaystyle\sum_{r=1}^{n-1} q_r \psi_r(q) & \text{if } 1 \le i \le n-1, \\[2mm] \psi(q) - \displaystyle\sum_{r=1}^{n-1} q_r \psi_r(q) & \text{if } i = n. \end{cases}$$

On each of the coordinate planes $x_i = 0$ we have, from (4.11),

$$\int_{\partial M \cap \{x_i = 0\}} \omega = 0.$$

On the curved part S where $x = X(q)$, consider the pullback

$$dx_i = \sum_{j=1}^{n-1} \frac{\partial X_i}{\partial q_j}\, dq_j.$$

Plugging this into (4.11) and computing the pullback of ω, we see that

$$\int_N \omega = \int_{D_{n-1}} \big|\det(A(q))\big|\, dq_1 \cdots dq_{n-1},$$

where $A(q)$ is the $n \times n$ matrix with entries $\frac{\partial}{\partial q_j} X_i(q)$ for $1 \le i \le n$, $1 \le j \le n-1$, and nth column $X(q)$. Now for $1 \le j \le n-1$ we have

$$\frac{\partial X_i}{\partial q_j}(q) = \begin{cases} \psi_{ij}(q) - \sum_{r=1}^{n-1} q_r \psi_{rj}(q) & \text{if } 1 \le i \le n-1, \\ -\sum_{r=1}^{n-1} q_r \psi_{rj}(q) & \text{if } i = n. \end{cases}$$

The matrix $A(q)$ can be written in block $((n-1)+1) \times ((n-1)+1)$ form as

$$A(q) = \begin{bmatrix} \psi_{ij}(q) - \sum_{r=1}^{n-1} q_r \psi_{rj}(q) & \psi(q) + \psi_i(q) - \sum_{r=1}^{n-1} q_r \psi_r(q) \\ -\sum_{r=1}^{n-1} q_r \psi_{rj}(q) & \psi(q) - \sum_{r=1}^{n-1} q_r \psi_r(q) \end{bmatrix}.$$

Subtracting the bottom row from all the other rows, we get $\det(A(q)) = \det(B(q))$ where

$$B(q) = \begin{bmatrix} \psi_{ij}(q) & \psi_i(q) \\ -\sum_{r=1}^{n-1} q_r \psi_{rj}(q) & \psi(q) - \sum_{r=1}^{n-1} q_r \psi_r(q) \end{bmatrix}.$$

In the matrix $B(q)$ the term $\psi(q)$ appears as an additive term in (n, n) entry, and nowhere else, so the calculation of $\det(B(q))$ involves some terms which do not involve $\psi(q)$, and all the other terms contain a simple factor $\psi(q)$. We get

$$\det(B(q)) = \det \begin{bmatrix} \psi_{ij}(q) & \psi_i(q) \\ -\sum_{r=1}^{n-1} q_r \psi_{rj}(q) & -\sum_{r=1}^{n-1} q_r \psi_r(q) \end{bmatrix} + \psi(q) \det[\psi_{ij}(q)]$$

$$=: \det(C(q)) + \psi(q) \det[\psi_{ij}(q)],$$

say. Now we can write

$$C(q) = \begin{bmatrix} \psi_{ij}(q) & \psi_i(q) \\ -\sum_{r=1}^{n-1} q_r \psi_{rj}(q) & -\sum_{r=1}^{n-1} q_r \psi_r(q) \end{bmatrix}$$

$$= \begin{bmatrix} \delta_{ij} \\ -q_j \end{bmatrix} [\psi_{ij}(q) \quad \psi_i(q)]$$

as the product of an $n \times (n-1)$ and an $(n-1) \times n$ matrix. This implies that the rank of $C(q)$ is at most $n-1$ and so $\det(C(q)) = 0$. Therefore

$$\det(A(q)) = \det(B(q)) = \psi(q) \det[\psi_{ij}(q)]$$

and the proof is complete. $\square$

Now we relax the boundedness assumption.

THEOREM 4.8.   *The volume formula* (4.8) *holds for any* $\psi \in \mathcal{C}$ *which is* C^2 *and satisfies* $\psi|_{\partial \Delta_n} \equiv 0$.

Proof. Here we apply Stokes' theorem with the same $(n-1)$-form ω on the region $\widehat{R}(\psi) \cap$
$(0, K)^n$. The boundary now consists of $S_K = S \cap [0, K]^n$ together with parts of the coordinate
planes $x_i = 0$ and parts of the planes $x_i = K$. For given K define sets

$$A_1 = \{(x_2, x_3, \ldots, x_n) \in (0, K)^{n-1} : (K, x_2, x_3, \ldots, x_n) \in \widehat{R}(\psi)\},$$

and

$$B_1 = (0, K) \times A_1 = \{x \in (0, K)^n : (K, x_2, x_3, \ldots, x_n) \in \widehat{R}(\psi)\}.$$

Note that $\{K\} \times A_1$ is the part of the boundary of $\widehat{R}(\psi) \cap (0, K)^n$ which lies in the plane
$x_1 = K$. Since $x_1 = K$ and $dx_1 = 0$ on the submanifold $\{K\} \times A_1$ we get

$$\omega = \sum_{k=1}^n (-1)^{k+1} x_k \, dx_1 \wedge \cdots \wedge \widehat{dx_k} \wedge \cdots \wedge dx_n = K \, dx_2 \wedge \cdots \wedge dx_n.$$

Since $dx_2 \wedge \cdots \wedge dx_n$ is the volume form on $\{K\} \times A_1$ given as the boundary of the oriented
domain $\widehat{R}(\psi) \cap (0, K)^n$, we obtain

$$\int_{\{K\} \times A_1} \omega = K \operatorname{vol}(A_1) = \operatorname{vol}(B_1).$$

Similar calculations are valid for the other coordinates, involving the sets

$$B_i = \{x \in (0, K)^n : (x_1, \ldots, x_{i-1}, K, x_{i+1}, \ldots, x_n) \in \widehat{R}(\psi)\}.$$

Now apply Stokes' theorem:

$$\int_{\widehat{R}(\psi) \cap (0, K)^n} d\omega = \int_{\partial(\widehat{R}(\psi) \cap (0, K)^n)} \omega.$$

Using the calculations above, we get

$$n \operatorname{vol}\left(\widehat{R} \cap (0, K)^n\right) = \sum_{i=1}^n \operatorname{vol}(B_i) + \int_{S_K} \omega.$$

Since $B_i \subset \widehat{R}(\psi) \cap (0, K)^n$ we can rewrite this as

$$\sum_{i=1}^n \operatorname{vol}\left((\widehat{R}(\psi) \cap (0, K)^n) \setminus B_i\right) = \int_{S_M} \omega.$$

Now we let $K \to \infty$. On the right side we have monotone behavior:

$$\int_{S_M} \omega = \frac{1}{n} \int_{X^{-1}(S_K)} \psi(q) \det\left(-D^2 \psi(q)\right) dp_1 \cdots dq_{n-1}$$

$$\nearrow \frac{1}{n} \int_{D_{n-1}} \psi(q) \det\left(-D^2 \psi(q)\right) dp_1 \cdots dq_{n-1}.$$

LEMMA 4.9. *For $K > 0$ write*

$$B_i^{(K)} = \{x \in (0, K)^n : (x_1, \ldots, x_{i-1}, K, x_{i+1}, \ldots, x_n) \in \widehat{R}(\psi)\}.$$

Then

$$\left(\widehat{R}(\psi) \cap (0, K)^n\right) \setminus B_i^{(K)} \nearrow \widehat{R}(\psi) \quad \text{as } K \nearrow \infty.$$

PROOF. It suffices to prove in the case $i = 1$. Suppose first $K < L$. If $x \in (\widehat{R}(\psi) \cap (0, K)^n) \setminus B_i^{(K)}$ then $x \in \widehat{R}(\psi) \cap (0, K)^n$ and $(K, x_2, \ldots, x_n) \notin \widehat{R}(\psi)$. Then $x \in \widehat{R}(\psi) \cap (0, L)^n$ and $(L, x_2, \ldots, x_n) \notin \widehat{R}(\psi)$, so that $x \in (\widehat{R}(\psi) \cap (0, L)^n) \setminus B_i^L$. This proves monotonicity.

To complete the proof it suffices to show that if $x \in \widehat{R}(\psi)$ then there exists K such that $x \in (0, K)^n$ and $(K, x_2, \ldots, x_n) \notin \widehat{R}(\psi)$. To show $(K_2, \ldots, x_n) \notin \widehat{R}(\psi)$ we need to show

$$K p_1 + \sum_{i=2}^{n} x_i p_i \geq \psi(p)$$

for all $p \in \Delta_n$. Let $\epsilon = \min(x_2, x_3, \ldots, x_n) > 0$. Since ψ has zero boundary values there is $\delta \in (0, 1/2)$ such that $p_1 < \delta$ implies $\psi(p) < \epsilon/2$. Finally we may choose $K > \max(\|\psi\|/\delta, x_1, x_2, \ldots, x_n)$.

If $p_1 \geq \delta$ then $K p_1 + \sum_{i=2}^{n} x_i p_i \geq K\delta > \|\psi\| \geq \psi(p)$. On the other hand if $p_1 < \delta$ then $\psi(p) < \epsilon$ and $p_2 + \cdots + p_n = 1 - p_1 \geq 1/2$, so that $K p_1 + \sum_{i=2}^{n} x_i p_i \geq \epsilon/2 > \psi(p)$. Together we have $(K, x_2, \ldots, x_n) \notin \widehat{R}(\psi)$. The condition $x \in (0, K)^n$ is trivially checked, and the proof is complete. $\square$

PROOF OF THE THEOREM, CONTINUED. Taking volumes of the increasing sequences of sets in the lemma, we get

$$\mathrm{vol}\big((\widehat{R}(\psi) \cap (0, K)^n) \setminus B_i^{(K)}\big) \nearrow \mathrm{vol}(\widehat{R}(\psi)$$

as $K \to \infty$ for each $1 \leq i \leq n$, and the proof is complete. $\square$

In the case $n = 2$ (so that D_{n-1} is one-dimensional) the formula (4.8) has an interesting alternative expression. To simplify the notation we write $q = q_1$.

COROLLARY 4.10. *Suppose $n = 2$ and let $\psi \in C$ be C^2 and (as a function of q) $\psi(0) = \psi(1) = 0$. Assume also*

$$\lim_{q \to 0} \psi(q)\psi'(q) = \lim_{q \to 1} \psi(q)\psi'(q) = 0.$$

Then

$$\mathrm{vol}\big(\widehat{R}(\psi)\big) = \frac{1}{2} \int_0^1 (\psi'(q))^2 \, dq, \qquad \mathsf{T}_\mu(\psi) = \exp\left(\frac{-\gamma}{2} \int_0^1 (\psi'(q))^2 \, dq\right).$$

PROOF. By Theorem 4.7 we have

$$\mathrm{vol}\big(\widehat{R}(\psi)\big) = \frac{1}{2} \int_0^1 \psi(q)(-\psi''(q)) \, dq.$$

Using integration by parts, this is equal to

$$-\frac{1}{2} \psi'(q)\psi(q)\big|_0^1 + \frac{1}{2} \int_0^1 (\psi'(q))^2 \, dq = \frac{1}{2} \int_0^1 (\psi'(q))^2 \, dq. \qquad \square$$

EXAMPLE 4.11. Let $n = 2$ and consider the geometric mean $\psi(q) = a\sqrt{q(1-q)}$ where $a > 0$. See Figure 7 for an illustration.

Since

$$\psi(q)\psi'(q) = a^2 \sqrt{q(1-q)} \frac{1 - 2q}{2\sqrt{q(1-q)}} = a^2 \frac{1 - 2q}{2}$$
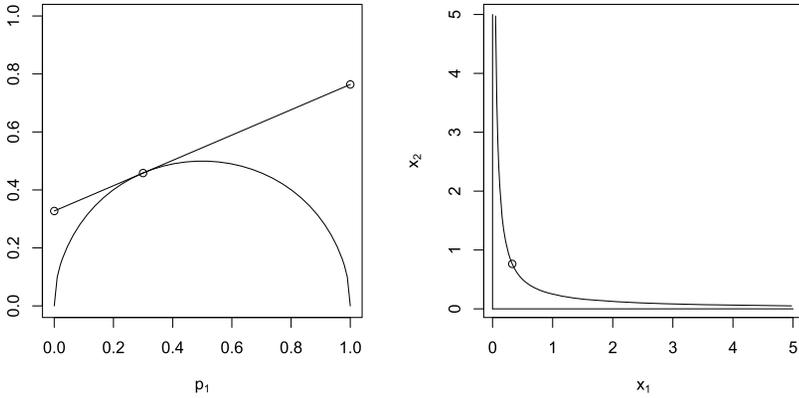
FIG. 7. *Left*: *Graph of* $\psi(p) = \sqrt{p_1 p_2}$ *(here* $\gamma = 1$*) as well as the tangent line at* $p_1 = 0.3$. *It defines a point* (x_1, x_2) *on the boundary of* $\widehat{R}(\psi)$. *Right*: *The region* $\widehat{R}(\psi)$ *and a point on its boundary. The area of* $\widehat{R}(\psi)$ *is infinite.*

does not vanish at 0 and 1, we cannot apply Corollary 4.10. Nevertheless, since

$$\int_0^1 \psi(q)(-\psi''(q)) \, dq = \int_0^1 \frac{a^2}{4x(1-x)} \, dx = \infty,$$

by Theorem 4.8 we have $\mathrm{vol}(\widehat{R}(\psi)) = \infty$. This proves that $\mathsf{T}_\mu(\psi) = 0$. This corresponds to the critical exponent in Theorem 4.6.

4.4. *Distribution of the portfolio weight.* Consider the random function Ψ with distribution μ. It generates, in the sense of (1.4), a portfolio map $\boldsymbol{\pi}$. The portfolio weight $\boldsymbol{\pi}(p)$ is uniquely defined whenever Ψ is differentiable at p (see [23], Section 2.3, for details). As an application of the representation of μ in terms of the Poisson point process, we derive the distribution of $\boldsymbol{\pi}(p)$ for each fixed p.

Let N denote the Poisson point process on $\mathbb{R}_+^n$ whose rate measure has density h. Thus we may write

(4.14) $$\Psi(p) = \inf_{x \in N} \langle p, x \rangle, \quad p \in \Delta_n.$$

Fix $p \in \Delta_n$ and consider the Poisson point process $N(p)$ on $\mathbb{R}_+$ given by $N(p) = \{\langle p, x \rangle : x \in N\}$. Since the rate measure for N has a density, then so does the rate measure for $N(p)$. It follows that $N(p)$ has no double points. In particular with probability 1 there is a unique $Z \in N$ such that $\langle p, Z \rangle$ is the minimum point of $N(p)$, and there exists a (random) $\delta > 0$ such that $\langle p, x \rangle \geq \langle p, Z \rangle + \delta$ for all $x \in N \setminus \{Z\}$.

From this observation and the representation (4.14), there exists a (random) neighborhood U of p such that $\Psi(q) = \langle q, Z \rangle$ for $q \in U$. Thus we have the following

LEMMA 4.12. *For any fixed* $p \in \Delta_n$, *the random function* Ψ *is* μ-*almost surely differentiable at* p. *Thus the portfolio weight* $\boldsymbol{\pi}(p)$ *is a.s.-defined.*

Again let $Z \in N$ be the point described above. Recall from (1.4) that the portfolio weight $\boldsymbol{\pi}(p)$ generated by Ψ is given by

$$\boldsymbol{\pi}_i(p) = p_i\big(1 + D_{e_i - p} \log \Psi(p)\big), \quad 1 \leq i \leq n.$$

Since $\Psi(q) = \langle q, Z \rangle$ for q near p, evaluating the derivative gives

$$\boldsymbol{\pi}(p) = \left(\frac{p_1 Z_1}{\langle p, Z \rangle}, \dots, \frac{p_n Z_n}{\langle p, Z \rangle} \right).$$

PROPOSITION 4.13. *Suppose $h(x)$ satisfies Assumption* 3.1. *Under μ, for any $p \in \Delta_n$ fixed, the portfolio weight $\boldsymbol{\pi}(p) = (Y_1, \ldots, Y_n)$ has density $ch(y_1/p_1, \ldots, y_n/p_n)$ with respect to the uniform distribution on Δ_n, where c is a normalizing constant. Equivalently, $(Y_1, \ldots, Y_{n-1})$ has density*

$$ch\left(\frac{y_1}{p_1}, \ldots, \frac{y_{n-1}}{p_{n-1}}, \frac{1 - \sum_{i=1}^{n-1} y_i}{p_n}\right)$$

on the set $D_{n-1} = \{y \in \mathbb{R}_+^{n-1} : \sum_{i=1}^{n-1} y_i < 1\}$.

PROOF. For fixed $p \in \Delta_n$ we consider the distribution of Z conditional on the value of $a = \langle p, Z \rangle$. Constants $c_1, c_2, \ldots$ will denote a sequence of normalizing constants. Conditioned on the value of $a = \langle p, Z \rangle$, the conditional distribution of $Z = (Z_1, \ldots, Z_n)$ has density $c_1 h(z_1, \ldots, z_n)$ with respect to uniform measure on the simplex $\{z \in \mathbb{R}_+^n : \sum_{i=1}^n p_i z_i = a\}$. More precisely, the conditional distribution of $(Z_1, \ldots, Z_{n-1})$ has density

$$c_1 h\left(z_1, \ldots, z_{n-1}, \left(a - \sum_{i=1}^{n-1} p_i z_i\right)/p_n\right)$$

with respect to Lebesgue measure on the set $\{z \in \mathbb{R}_+^{n-1} : \sum_{i=1}^{n-1} p_i z_i < a\}$. Now define $Y = (Y_1, \ldots, Y_n) \in \mathbb{R}_+^n$ by $Y_i = p_i Z_i/\langle p, Z \rangle$, so that $\boldsymbol{\pi}(p) = Y$. Conditioned on $\langle p, Z \rangle = a$, we have $Y_i = p_i Z_i/a$ and so the conditional distribution of $(Y_1, \ldots, Y_{n-1})$ has density

$$c_2 h\left(\frac{ay_1}{p_1}, \ldots, \frac{ay_{n-1}}{p_{n-1}}, \frac{a - \sum_{i=1}^{n-1} ay_i}{p_n}\right) = c_3 h\left(\frac{y_1}{p_1}, \ldots, \frac{y_{n-1}}{p_{n-1}}, \frac{1 - \sum_{i=1}^{n-1} y_i}{p_n}\right)$$

with respect to Lebesgue measure on the set D_{n-1}. On the right side of this expression the points where h is evaluated do not depend on a, and the set D_{n-1} does not depend on a, so the normalizing constant c_3 does not depend on a. Since this conditional density does not depend on the value of a, it is the unconditional density of $(Y_1, \ldots, Y_{n-1})$, and the proof is complete. □

The next proposition computes the expected value of $\boldsymbol{\pi}(p)$ at a fixed p. In words, it states that $\mathbb{E}[\boldsymbol{\pi}(p)]$ is equal to the portfolio $\overline{\boldsymbol{\pi}}(p)$ generated by the expected value $\overline{\Psi} = \mathbb{E}[\Psi(\cdot)]$. It is an interesting problem to study the properties of the random portfolio map $\boldsymbol{\pi}(\cdot) : \Delta_n \to \overline{\Delta}_n$ and their implications in optimal transport and Cover's universal portfolio.

PROPOSITION 4.14. *Under Assumption* 3.1, *the expected value of the portfolio weight $\boldsymbol{\pi}(p)$ generated by Ψ is the same as the portfolio weight $\overline{\boldsymbol{\pi}}(p)$ generated by the expectation $\overline{\Psi}(\cdot) = \mathbb{E}\Psi(\cdot)$.*

PROOF. For $p \in \mathbb{R}_+^n$ define $R(p) = \{x \in \mathbb{R}_+^n : \sum_{i=1}^n p_i x_i < 1\}$ and

$$F(p) = \int_{R(p)} h(x) \, dx.$$

Recall the set $D_{n-1} = \{y \in R_+^{n-1} : \sum_{i=1}^{n-1} y_i < 1\}$ and the notational convenience $y_n = 1 - \sum_{i=1}^{n-1} y_i$. Then (using elementary multivariate analysis and the scaling property of h) we have

$$D_{e_i} F(p) = -\frac{1}{p_1 \cdots p_n} \int_{D_{n-1}} \frac{y_i}{p_i} h\left(\frac{y_1}{p_1}, \ldots, \frac{y_n}{p_n}\right) dy_1 \cdots dy_{n-1}$$

for $1 \leq i \leq n$, and

$$F(p) = \frac{1}{(n+\alpha)p_1 \cdots p_n} \int_{D_{n-1}} h\left(\frac{y_1}{p_1}, \ldots, \frac{y_{n-1}}{p_{n-1}}, \frac{y_n}{p_n}\right) dy_1 \cdots dy_{n-1}.$$

From Corollary 4.1 we have

$$\overline{\Psi}(p) = \Gamma\left(\frac{1}{n+\alpha} + 1\right)(F(p))^{-\frac{1}{n+\alpha}}$$

and then

$$D_{e_i} \log \overline{\Psi}(p) = -\frac{D_{e_i} F(p)}{(n+\alpha)F(p)}$$

$$= \frac{\int_{D_{n-1}} \frac{y_i}{p_i} h(\frac{y_1}{p_1}, \ldots, \frac{y_n}{p_n}) dy_1 \cdots dy_{n-1}}{\int_{D_{n-1}} h(\frac{y_1}{p_1}, \ldots, \frac{y_n}{p_n}) dy_1 \cdots dy_{n-1}}$$

$$= \frac{1}{p_i} \mathbb{E}\boldsymbol{\pi}_i(p).$$

Therefore, using (1.4), the portfolio weight generated by $\overline{\Psi}$ is

$$p_i(1 + D_{e_i - p} \log \overline{\Psi}(p)) = p_i\left(1 + \frac{1}{p_i}\mathbb{E}\boldsymbol{\pi}(p)_i - \sum_{j=1}^{n} p_j \frac{1}{p_j}\mathbb{E}\boldsymbol{\pi}(p)_j\right) = \mathbb{E}\boldsymbol{\pi}_i(p),$$

and the proof is complete. $\square$

EXAMPLE 4.15 (Geometric mean and the constant-weighted portfolio). Suppose Assumption 4.3 holds. By Proposition 4.4, the expected value $\overline{\Psi}(\cdot)$ is a multiple of the geometric mean $p_1^{\pi_1} \cdots p_n^{\pi_n}$ which generates the constant-weighted portfolio $\overline{\boldsymbol{\pi}}(p) \equiv \pi_i$. By Proposition 4.14, if $\boldsymbol{\pi}$ is the (random) portfolio map generated by Ψ, then $\mathbb{E}\boldsymbol{\pi}(p) = \overline{\boldsymbol{\pi}}(p) \equiv \pi$.

5. Diagonal limits.

In this final section we study the model (2.13) where the parameter λ_K of the softmin depends on K, the number of hyperplanes. We have seen that when $\lambda_K \equiv \lambda < \infty$ is fixed there is a deterministic almost sure limit, and when $\lambda_K \equiv \infty$ a suitable scaling gives a nontrivial weak limit which can be described by a Poisson point process. Here we want to find explicit rates for λ_K which give possibly different limiting behaviors.

5.1. *Main results.* Before stating the main results we first set up some notation. Let $X_1, X_2, \ldots$ be i.i.d. copies of C, where C is the random vector in (2.12). Thus we may write the kth hyperplane as $\ell_k(p) = \langle p, X_k \rangle$. For $\lambda > 0$ and $K \geq 1$, let $\Phi_K(p) = \min_{1 \leq k \leq K} \langle p, X_k \rangle$ and

$$(5.1) \qquad \Phi_{\lambda, K}(p) := \mathsf{m}_\lambda\{\langle p, X_k \rangle : 1 \leq k \leq K\} = \frac{-1}{\lambda} \log\left(\frac{1}{K}\sum_{k=1}^{K} e^{-\lambda\langle p, X_k\rangle}\right).$$

Throughout Section 5 we work under Assumption 3.1. We let Ψ be a random concave function under the limiting distribution μ in Theorem 3.5. Note that it is the weak limit of the scaled hardmin $K^{\frac{1}{n+\alpha}} \Phi_K$.

The following are the main results of this section. First we give the case where the weak limit can be related to the hardmin limit Ψ.

THEOREM 5.1. *Let Assumption* 3.1 *be in force. Suppose* $\frac{\lambda_K}{K^{1/(n+\alpha)}} \to \infty$ *as* $K \to \infty$. *Then* $K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K} - \frac{K^{1/(n+\alpha)} \log K}{\lambda_K}$ *converges weakly to* Ψ. *In particular*:

(i) *if* $\frac{\lambda_K}{K^{1/(n+\alpha)} \log K} \to \infty$ *as* $K \to \infty$, *then* $K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}$ *converges weakly to* Ψ; *and*

(ii) *if* $\frac{\lambda_K}{K^{1/(n+\alpha)} \log K} \to c$ *as* $K \to \infty$, *where* $c > 0$ *is a fixed constant, then* $K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}$ *converges weakly to* $\Psi + c$ *which is a deterministic shift of* Ψ.

Next we consider a case where λ_K grows less quickly than in Theorem 5.1 above. Here the effect of the additive normalization is much stronger, and the weak limit may not be supported on $\mathcal{C}$, the space of nonnegative (continuous) concave functions on $\overline{\Delta}_n$. Indeed we will show by an example that the limit may become negative.

THEOREM 5.2. *Let Assumption* 3.1 *be in force. Suppose* $\frac{\lambda_K}{K^{1/(n+\alpha)}} \to c$ *as* $K \to \infty$, *where* $c > 0$ *is a fixed constant. Then the sequence*

$$K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K} - \frac{K^{1/(n+\alpha)} \log K}{\lambda_K}$$

converges weakly to a random concave function $\widetilde{\Psi}_c$ *on* Δ_n.

See Section 5.4 for more discussion including a probabilistic representation of the limit $\widetilde{\Psi}_c$ in terms of a Poisson point process (see (5.16)). Although Theorem 5.2 does not fit directly under the framework of Section 2.1 and thus cannot be used directly in the applications described in Section 1.1.2, it is mathematically interesting as it gives another limit which is genuinely different from that of the hardmin case.

REMARK 5.3. In the above results the functions may become negative and so may lie outside $\mathcal{C}$. To be precise, here we are using the topology of local uniform convergence on the space $\widetilde{\mathcal{C}}$ of real-valued continuous concave functions on $\overline{\Delta}_n$. We may continue to use the metric given by (2.1).

5.2. *Poisson convergence.* In this subsection we relate $X = \{X_k\}_{k \geq 1}$, regarded as a point process on $\mathbb{R}^n_+$, to a Poisson point process which is the main probabilistic tool of this section. For $K \geq 1$, let

(5.2) $$Y^{(K)} := \{Y_k^{(K)} : 1 \leq k \leq K\} := \{K^{\frac{1}{n+\alpha}} X_k : 1 \leq k \leq K\}.$$

So $Y^{(K)}$ is also a point process on $\mathbb{R}^n_+$. The following result shows that $Y^{(K)}$ converges in distribution to a Poisson point process. This gives an alternative method to prove Theorem 3.5 but we will not elaborate on this. For the precise meaning of the convergence of point processes we refer the reader to [26], Section 3.4.

PROPOSITION 5.4. *Suppose* $c_K \to 1$. *As* $K \to \infty$, *the point process*

$$c_K Y^{(K)} = \{c_K Y_k^{(K)} : 1 \leq k \leq K\}$$

converges in distribution to the Poisson process N *on* $\mathbb{R}^n_+$ *with intensity measure* $dm = h(x) dx$, *where* h *is the function specified in Assumption* 3.1.[4]

PROOF. By [26], Proposition 3.19, it suffices to prove that the Laplace transform of $c_K Y^{(K)}$ converges to that of N. More precisely, we will show that

(5.3) $$\lim_{K \to \infty} \mathbb{E}\big[e^{-\sum_{k=1}^K g(c_K Y_k^{(K)})}\big] = \exp\bigg[-\int_{\mathbb{R}^n_+} (1 - e^{-g(y)}) h(y) \, dy\bigg],$$

---

[4]We will first apply this result with $c_K \equiv 1$. The general case will be used in Section 5.4.

for any continuous function $g : \mathbb{R}^n \to \mathbb{R}$ with compact support. First we note that

$$\mathbb{E}\big[e^{-\sum_{k=1}^{K} g(c_K Y_k^{(K)})}\big] = \bigg[\int_{\mathbb{R}_+^n} e^{-g(c_K K^{1/(n+\alpha)} x)} \rho(x)\, dx\bigg]^K$$

(5.4)

$$= \bigg[1 - \int_{\mathbb{R}_+^n} \big(1 - e^{-g(c_K K^{1/(n+\alpha)} x)}\big) \rho(x)\, dx\bigg]^K,$$

where ρ is the common density of the X_k (see Assumption 3.1).

Since g is compactly supported, so is $1 - e^{-g}$. By an standard extension of the limit in Lemma 3.3 with $\lambda = c_K^{-1} K^{-1/(n+\alpha)}$, we have

$$K \int_{\mathbb{R}_+^n} \big(1 - e^{-g(c_K K^{1/(n+\alpha)} x)}\big) \rho(x)\, dx$$

$$= \frac{K^{\alpha/(n+\alpha)}}{c_K^n} \int_{\mathbb{R}_+^n} \big(1 - e^{-g(y)}\big) \rho\big(c_K^{-1} K^{-1/(n+\alpha)} y\big)\, dy$$

$$\to \int_{\mathbb{R}_+^n} \big(1 - e^{-g(y)}\big) h(y)\, dy.$$

Letting $K \to \infty$ in (5.4) and using the above limit, we obtain (5.3). $\quad\square$

By the continuous mapping theorem, we immediately obtain the following corollary.

COROLLARY 5.5. *Suppose $c_K \to 1$. For any $p \in \Delta_n$, the one-dimensional point process*

(5.5) $$c_K Y^{(K)}(p) := \langle p, c_K Y^{(K)} \rangle := \{\langle p, c_K Y_k^{(K)} \rangle : 1 \leq k \leq K\}$$

converges weakly to the Poisson point process $N(p) := \langle p, N \rangle := \{\langle p, x \rangle : x \in N\}$ on $\mathbb{R}_+$ whose intensity measure $\widetilde{m}$ is the pushforward of $dm(x) = h(x)\, dx$ under the mapping $x \mapsto \langle p, x \rangle$. Explicitly, we have

(5.6) $$\widetilde{m}(0, t] = \int_{R(p,t)} h(x)\, dx = t^{n+\alpha} \int_{R(p,1)} h(y)\, dy.$$

For $p \in \Delta_n$ fixed, we will often denote $\widetilde{Y}^{(K)} = Y^{(K)}(p)$ and $\widetilde{N} = N(p)$.

5.3. *Proof of Theorem 5.1.* First we quickly settle the relatively trivial case of Theorem 5.1(i). By Lemma 2.5(i), we have

$$K^{\frac{1}{n+\alpha}} \Phi_K(p) \leq K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}(p) \leq K^{\frac{1}{n+\alpha}} \Phi_K(p) + \frac{1}{\lambda_K} K^{\frac{1}{n+\alpha}} \log K.$$

Suppose $\frac{\lambda_K}{K^{1/(n+\alpha)} \log K} \to \infty$. Then uniformly in p we have

$$\big|K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}(p) - K^{\frac{1}{n+\alpha}} \Phi_K(p)\big| \to 0, \quad K \to \infty.$$

From this and Theorem 3.5 we see easily that $K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}$ converges weakly to Ψ.

For the general case we need a more delicate analysis. For $K \geq 1$ and $p \in \Delta_n$ fixed, relabel the points $\{X_k\}_{1 \leq k \leq K}$ so that

$$\langle p, X_{(1)} \rangle \leq \langle p, X_{(2)} \rangle \leq \cdots \leq \langle p, X_{(K)} \rangle.$$

Note that $\langle p, X_{(1)} \rangle = \min(\langle p, X_1 \rangle, \ldots, \langle p, X_K \rangle) = \Phi_K(p)$. Write

(5.7) $$\Phi_{\lambda, K}(p) = -\frac{1}{\lambda} \log\bigg(\frac{1}{K} \sum_{k=1}^{K} e^{-\lambda \langle p, X_k \rangle}\bigg) = \frac{\log K}{\lambda} - \frac{1}{\lambda} \log\bigg(\sum_{k=1}^{K} e^{-\lambda \langle p, X_{(k)} \rangle}\bigg).$$

Then

$$
K^{\frac{1}{n+\alpha}} \Phi_{\lambda,K}(p)
$$

(5.8)
$$
= \frac{K^{\frac{1}{n+\alpha}} \log K}{\lambda} + K^{\frac{1}{n+\alpha}} \Phi_K(p) - \frac{K^{\frac{1}{n+\alpha}}}{\lambda} \log\left(1 + \sum_{k=2}^{K} e^{-\lambda(\langle p, X_{(k)} \rangle - \langle p, X_{(1)} \rangle)}\right)
$$

$$
=: \frac{K^{\frac{1}{n+\alpha}} \log K}{\lambda} + K^{\frac{1}{n+\alpha}} \Phi_K(p) - \frac{K^{\frac{1}{n+\alpha}}}{\lambda} \Theta_{\lambda,K}(p),
$$

where the term $\Theta_{\lambda,K}(p)$ is nonnegative and satisfies

(5.9) $$
\Theta_{\lambda,K}(p) \leq \sum_{k=2}^{K} e^{-\lambda(\langle p, X_{(k)} \rangle - \langle p, X_{(1)} \rangle)} = \sum_{k=2}^{K} e^{-(\lambda/K^{1/(n+\alpha)})(\langle p, Y_{(k)}^{(K)} \rangle - \langle p, Y_{(1)}^{(K)} \rangle)},
$$

and the $\langle p, Y_{(k)}^{(K)} \rangle$ are the points in $Y^{(K)}(p) = \{\langle p, Y_k^{(K)} \rangle : 1 \leq k \leq K\}$ arranged in ascending order.

The following is the main technical result needed in the proof of Theorem 5.1.

THEOREM 5.6.   *Fix $p \in \Delta_n$ and write $\widetilde{Y}_{(k)}^{(K)} = \langle p, Y_{(k)}^{(K)} \rangle$. Then*

(5.10) $$
\sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} \to 0
$$

in probability as $K \to \infty$ and $\epsilon \to 0^+$.

To prove Theorem 5.6 we need some lemmas. Let $\widetilde{N} = N(p)$ be the one-dimensional Poisson point process given in Corollary 5.5. Since m and hence $\widetilde{m}$ are Radon measures, we may order the points in $\widetilde{N}$ and write

$$
\widetilde{N} = \{\widetilde{N}_{(1)} \leq \widetilde{N}_{(2)} \leq \cdots\}.
$$

In view of the convergence $\widetilde{Y}^{(K)} \to \widetilde{N}$, we expect that

(5.11) $$
\sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} \approx \sum_{k=2}^{\infty} e^{-(\widetilde{N}_{(k)} - \widetilde{N}_{(1)})/\epsilon},
$$

and the right side of (5.11), since it no longer depends on K, is easy to analyze as $\epsilon \to 0^+$. In what follows, we do not attempt to justify (5.11) rigorously, but instead use the convergence $\widetilde{Y}^{(K)} \to \widetilde{N}$ to convert simple estimates on $\widetilde{N}$ into corresponding ones on $\widetilde{Y}^{(K)}$ for large K.

If Z is a point process (with no double points) we let $Z(B)$ be the cardinality of $Z \cap B$.

LEMMA 5.7.   *Given $\delta_0 > 0$ there exist positive constants L, M and c such that for K sufficiently large we have*

$$
\mathbb{P}(\widetilde{Y}_{(1)}^{(K)} \geq L) \leq \delta_0,
$$

$$
\mathbb{P}(\widetilde{Y}^{(K)}(0, 2L) > M) \leq \delta_0,
$$

$$
\mathbb{P}(\widetilde{Y}_{(2)}^{(K)} - \widetilde{Y}_{(1)}^{(K)} < c) \leq \delta_0.
$$

PROOF. By (5.6), the intensity $\widetilde{m}$ for $\widetilde{N}$ satisfies $\widetilde{m}(0, L) \to \infty$ as $L \to \infty$ and $\widetilde{m}(0, 2L) < \infty$ for all L. Thus we have

$$\mathbb{P}(\widetilde{N}_{(1)} \geq L) = \mathbb{P}(\widetilde{N}(0, L) = 0) \to 0 \quad \text{as } L \to \infty,$$

and then for any given L we have

$$\mathbb{P}(\widetilde{N}(0, 2L) > M) \to 0 \quad \text{as } M \to \infty.$$

Also since $\widetilde{m}$ has no atoms then $\mathbb{P}(\widetilde{N}_{(2)} > \widetilde{N}_{(1)}) = 1$. So given $\delta_0 > 0$, there exist positive constants L, M and c such that

$$\mathbb{P}(\widetilde{N}_{(1)} \geq L) \leq \delta_0/2,$$

$$\mathbb{P}(\widetilde{N}(0, 2L) > M) \leq \delta_0/2,$$

$$\mathbb{P}(\widetilde{N}_{(2)} - \widetilde{N}_{(1)} < c) \leq \delta_0/2.$$

The corresponding estimates for $\widetilde{Y}^{(K)}$ now follow directly from the convergence of $\widetilde{Y}^{(K)}$ to $\widetilde{N}$. In particular, for the convergence of $(\widetilde{Y}_{(1)}^{(K)}, \widetilde{Y}_{(2)}^{(K)})$ to $(\widetilde{N}_{(1)}, \widetilde{N}_{(2)})$ see [26], Proposition 3.13. $\square$

We will also need an a priori estimate on the $\widetilde{Y}^{(K)}$. The proof of the following lemma will be given in the Appendix. Note that Lemma 5.8(iii) and Corollary 5.9(ii) below will not be used until Section 5.4.

LEMMA 5.8. *Fix $p \in \Delta_n$.*

(i) *The random variable $\langle p, X_1 \rangle$ has density $\widetilde{\rho}(t)$ which satisfies*

$$\lim_{t \to 0^+} \frac{1}{t^{n+\alpha-1}} \widetilde{\rho}(t) = (n + \alpha) \int_{R(p,1)} h(x)\, dx.$$

(ii) *There exists $B < \infty$ such that*

$$\gamma^{n+\alpha} \mathbb{E}[e^{-\gamma \langle p, X_1 \rangle}] \leq B \quad \text{for } \gamma > 0.$$

(iii)

$$\lim_{\gamma, L \to \infty} \gamma^{n+\alpha} \mathbb{E}[e^{-\gamma \langle p, X_1 \rangle} 1_{\gamma \langle p, X_1 \rangle \geq L}] = 0.$$

COROLLARY 5.9.

(i) *For all $\epsilon > 0$ and $K \geq 1$*

$$\mathbb{E}\left[\sum_{k=1}^{K} e^{-\langle p, Y_k^{(K)} \rangle / \epsilon} \right] \leq \epsilon^{n+\alpha} B.$$

(ii) *For each fixed $c > 0$*

$$\lim_{K, L \to \infty} \mathbb{E}\left[\sum_{k=1}^{K} e^{-c \langle p, Y_k^{(K)} \rangle} 1_{\langle p, Y_k^{(K)} \rangle \geq L} \right] = 0.$$

PROOF. We have

$$\mathbb{E}\left[\sum_{k=1}^{K} e^{-\langle p, Y_k^{(K)} \rangle / \epsilon} \right] = K \mathbb{E}[e^{-K^{1/(n+\alpha)} \langle p, X_1 \rangle / \epsilon}]$$

and the first result follows by taking $\gamma = K^{1/(n+\alpha)}/\epsilon$ in Lemma 5.8(ii) above. The proof of the second result is similar, using Lemma 5.8(iii). □

PROOF OF THEOREM 5.6.   For all $\delta_0 > 0$ and $\delta_1 > 0$ we will prove that there exists K_0 and $\epsilon_0 > 0$ such that

$$\mathbb{P}\left(\sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} > \delta_1\right) < 4\delta_0,$$

whenever $K \geq K_0$ and $\epsilon < \epsilon_0$.

Let L, M, c and the constants in Lemma 5.7, and K_0 be a positive integer such that the conclusion of the lemma holds for $K \geq K_0$. Applying Markov's inequality to Corollary 5.9, we get

$$\mathbb{P}\left(\sum_{k=1}^{K} e^{-\langle p, Y_k^{(K)}\rangle/\epsilon} \geq \delta_1/2\right) \leq \frac{2\epsilon^{n+\alpha} B}{\delta_1} < \delta_0$$

as long as $\epsilon < \epsilon_1 = (\delta_0 \delta_1/(2B))^{1/(n+\alpha)}$. Define the event

$$\Omega_{K,\epsilon} = \{\widetilde{Y}_{(1)}^{(K)} \geq L\} \cup \{\widetilde{Y}^{(K)}(0, 2L) > M\}$$

$$\cup \{\widetilde{Y}_{(2)}^{(K)} - \widetilde{Y}_{(1)}^{(K)} < c\} \cup \left\{\sum_{k=1}^{K} e^{-\widetilde{Y}_{(k)}^{(K)}/\epsilon} \geq \delta_1/2\right\}.$$

Then $\mathbb{P}(\Omega_{K,\epsilon}) < 4\delta_0$ if $K \geq K_0$ and $\epsilon < \epsilon_1$. Now

$$\sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} = \sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} 1_{\widetilde{Y}_{(k)}^{(K)} < 2L}$$

$$+ \sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} 1_{\widetilde{Y}_{(k)}^{(K)} \geq 2L}.$$

For $\widetilde{Y}^{(K)} \notin \Omega_{K,\epsilon}$ the first sum on the right has at most M terms and each term is at most $e^{-c/\epsilon}$. Also, for $\widetilde{Y}^{(K)} \notin \Omega_{K,\epsilon}$ each term in the second sum has $\widetilde{Y}_{(1)}^{(K)} < L \leq \widetilde{Y}_{(k)}^{(K)}/2$, so that $\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)} \geq \widetilde{Y}_{(k)}^{(K)}/2$ and the second term is at most

$$\sum_{k=2}^{K} e^{-\widetilde{Y}_{(k)}^{(K)}/(2\epsilon)} 1_{\widetilde{Y}_{(k)}^{(K)} \geq 2L} \leq \sum_{k=1}^{K} e^{-\widetilde{Y}_{(k)}^{(K)}/(2\epsilon)} = \sum_{k=1}^{K} e^{-\langle p, Y_k^{(K)}\rangle/(2\epsilon)} < \delta_1/2$$

so long as $\epsilon < \epsilon_1/2$. Together, if $K \geq K_0$ and $\epsilon < \epsilon_1/2$ we have

$$\mathbb{P}\left(\sum_{k=2}^{K} e^{-(\widetilde{Y}_{(k)}^{(K)} - \widetilde{Y}_{(1)}^{(K)})/\epsilon} \geq Me^{-c/\epsilon} + \delta_1/2\right) \leq \mathbb{P}(\Omega_{K,\epsilon}) < 4\delta_0.$$

Finally it suffices to choose ϵ_2 so that $Me^{-c/\epsilon_2} \leq \delta_1/2$ and take $\epsilon_0 = \min(\epsilon_1/2, \epsilon_2)$. □

Now it is easy to complete the proof of Theorem 5.1.

PROOF OF THEOREM 5.1.   Let $A \subset \Delta_n$ be a finite set. By assumption $\lambda_K/K^{\frac{1}{n+\alpha}} \to \infty$, so using (5.8), (5.9) and Theorem 5.6, it is easy to see that

$$\left(K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}(p) - \frac{K^{\frac{1}{n+\alpha}} \log K}{\lambda_K}\right)_{p \in A} \to (\Psi(p))_{p \in A}$$

weakly as $K \to \infty$. By the argument in the proof of Theorem 3.5, we have that the random concave function $K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K} - \frac{K^{\frac{1}{n+\alpha}} \log K}{\lambda_K}$ converges weakly to Ψ as $K \to \infty$. $\quad\square$

5.4. *Proof of Theorem* 5.2. Again we start from the identity (5.7). Writing $\lambda_K / K^{\frac{1}{n+\alpha}} = c_K$, we have

$$(5.12) \qquad K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K}(p) - \frac{K^{\frac{1}{n+\alpha}} \log K}{\lambda_K} = -\frac{1}{c_K} \log \left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} \right),$$

where we recall $Y_k^{(K)} = K^{\frac{1}{n+\alpha}} X_k$ for $1 \le k \le K$. Motivated by Corollary 5.5, since $c_K \to c > 0$ we expect that

$$(5.13) \qquad -\frac{1}{c_K} \log \left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} \right) \to -\frac{1}{c} \log \left(\sum_{x \in N} e^{-c \langle p, x \rangle} \right) =: \widetilde{\Psi}_c(p),$$

where N is the Poisson point process with intensity measure $dm(x) = h(x)\,dx$. In the following proposition we verify this fact which is the main ingredient of the proof of Theorem 5.2.

PROPOSITION 5.10. *For any finite subset $A \subset \Delta_n$ we have*

$$\left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} \right)_{p \in A} \to \left(\sum_{x \in N} e^{-c \langle p, x \rangle} \right)_{p \in A}$$

in distribution as $K \to \infty$.

We first verify that the right hand side is finite almost surely.

LEMMA 5.11. *For any $p \in \Delta_n$ we have $\mathbb{E}[\sum_{x \in N} e^{-c \langle p, x \rangle}] < \infty$ and hence*

$$\sum_{x \in N} e^{-c \langle p, x \rangle} < \infty$$

almost surely.

PROOF. By Corollary 5.5 and the property of Poisson point process, we have

$$\mathbb{E}\left[\sum_{x \in N} e^{-c \langle p, x \rangle} \right] = \int_0^{\infty} e^{-ct}\, d\widetilde{m}(t)$$

$$(5.14) \qquad\qquad = (n + \alpha) \left(\int_{R(p,1)} h(y)\,dy \right) \int_0^{\infty} t^{n+\alpha-1} e^{-ct}\, dt$$

$$= \left(\int_{R(p,1)} h(y)\,dy \right) \frac{\Gamma(n + \alpha + 1)}{c^{n+\alpha}} < \infty. \qquad\square$$

PROOF OF PROPOSITION 5.10. Write $\widetilde{c}_K = c_K / c$, so that $\widetilde{c}_K \to 1$ as $K \to \infty$. Since the point process $\widetilde{c}_K Y^{(K)} = \{ \widetilde{c}_K K^{\frac{1}{n+\alpha}} X_k : 1 \le k \le K \}$ on $\mathbb{R}_+^n$ converges weakly to the Poisson point process N (see Proposition 5.4), by Skorohod's theorem we may construct these processes on the same probability space in such a way that $\widetilde{c}_K Y^{(K)} \to N$ almost surely. It suffices to show that

$$\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} \to \sum_{x \in N} e^{-c \langle p, x \rangle}$$

in probability, for each fixed $p \in A$, as $K \to \infty$. The convergence then extends easily to the joint vector. In particular we will show that for all $\delta_0 > 0$ and $\delta_1 > 0$ there exists K_0, depending on p, such that

$$\mathbb{P}\left(\left|\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} - \sum_{x \in N} e^{-c \langle p, x \rangle}\right| > \delta_1\right) < \delta_0$$

for all $K \geq K_0$.

Applying the dominated convergence theorem to Lemma 5.11 gives

$$\mathbb{E}\left[\sum_{x \in N} e^{-c \langle p, x \rangle} 1_{\langle p, x \rangle \geq L}\right] \to 0$$

as $L \to \infty$, and then Markov's inequality gives L_0 such that

$$\mathbb{P}\left(\sum_{x \in N} e^{-c \langle p, x \rangle} 1_{\langle p, x \rangle \geq L} > \delta_1/3\right) \leq \delta_0/3$$

for all $L \geq L_0$.

Since $c_K \to c$, there exists K_1 such that $\frac{c}{2} \leq c_K \leq 2c$, and hence $\tilde{c}_K \leq 2$, for $K \geq K_1$. Let $K \geq K_1$. Using Corollary 5.9(ii), we get

$$\mathbb{P}\left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} 1_{\tilde{c}_K \langle p, Y_k^{(K)} \rangle \geq L} > \delta_1/3\right)$$

$$\leq \mathbb{P}\left(\sum_{k=1}^{K} e^{-(c/2) \langle p, Y_k^{(K)} \rangle} 1_{\langle p, Y_k^{(K)} \rangle \geq L/2} > \delta_1/3\right)$$

$$\leq \frac{3}{\delta_1} \mathbb{E}\left[\sum_{k=1}^{K} e^{-(c/2) \langle p, Y_k^{(K)} \rangle} 1_{\langle p, Y_k^{(K)} \rangle \geq L/2}\right]$$

$$\to 0$$

as $K, L \to \infty$. Therefore there exist $L \geq L_0$ and $K_2 \geq K_1$ such that

$$\mathbb{P}\left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} 1_{\tilde{c}_K \langle p, Y_k^{(K)} \rangle \geq L} > \delta_1/3\right) < \delta_0/3$$

whenever $K \geq K_2$. Now let $f : \mathbb{R} \to \mathbb{R}$ be continuous with compact support such that $f(z) = e^{-cz}$ for $0 < z \leq L$, and $0 \leq f(z) \leq e^{-cz}$ for all $z \geq L$. Since f has compact support we have

$$\sum_{k=1}^{K} f\left(\tilde{c}_K \langle p, Y_k^{(K)} \rangle\right) \to \sum_{x \in N} f(\langle p, x \rangle)$$

almost surely. In particular there exists $K_0 \geq K_2$ such that

$$\mathbb{P}\left(\left|\sum_{k=1}^{K} f\left(\tilde{c}_K \langle p, Y_k^{(K)} \rangle\right) - \sum_{x \in N} f(\langle p, x \rangle)\right| > \delta_1/3\right) < \delta_0/3$$

for $K \geq K_0$. Finally, since

$$\left| \sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} - \sum_{x \in N} e^{-c \langle p, x \rangle} \right| \leq \left| \sum_{k=1}^{K} f\left(\tilde{c}_K \langle p, Y_k^{(K)} \rangle \right) - \sum_{x \in N} f\left(\langle p, x \rangle \right) \right|$$

$$+ \left(\sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} 1_{\tilde{c}_K \langle p, Y_k^{(K)} \rangle \geq L} \right)$$

$$+ \left(\sum_{x \in N} e^{-c \langle p, x \rangle} 1_{\langle p, x \rangle \geq L} \right),$$

combining the above estimates we have

$$\mathbb{P}\left(\left| \sum_{k=1}^{K} e^{-c_K \langle p, Y_k^{(K)} \rangle} - \sum_{x \in N} e^{-c \langle p, x \rangle} \right| > \delta_1 \right) < \delta_0/3 + \delta_0/3 + \delta_0/3 = \delta_0$$

whenever $K \geq K_0$ and the proof is complete. $\square$

Now we are ready to prove Theorem 5.2.

PROOF OF THEOREM 5.2. Consider the sequence

$$\widehat{\Psi}_K = K^{\frac{1}{n+\alpha}} \Phi_{\lambda_K, K} - K^{\frac{1}{n+\alpha}} \log K / \lambda_K$$

given by (5.12). By Proposition 5.10, we have weak convergence of the finite dimensional distributions. Similar to the proof of Theorem 3.5, it remains to show that the sequence, as random elements with values in $\widetilde{C}$, is tight.

For $r \geq 1$ define $e_i^{(r)} = (1 - 1/r)e_i + (1/r)\overline{e}$, where we recall $\overline{e}$ is the barycenter $(1/n, \ldots, 1/n)$ of Δ_n, and define $\Omega_r = \text{conv}\{e_1^{(r)}, \ldots, e_n^{(r)}\}$. Then Ω_r is compact and $\bigcup_{r \geq 1} \Omega_r = \Delta_n$. By [27], Theorem 10.9, the set

$$\left\{ \psi \in \widetilde{C} : \sup_{p \in \Omega_r} |\psi(p)| \leq M_r \text{ for all } r \geq 1 \right\}$$

is compact in $\widetilde{C}$ for any sequence $\{M_r\}$ with $M_r > 0$.

LEMMA 5.12. *Suppose $\psi \in \widetilde{C}$ satisfies $\psi(e_i^{(r)}) \leq -L_1$ and $\psi(\overline{e}) \leq L_2$. Then*

$$-L_1 \leq \psi(p) \leq (n-1)L_1 + nL_2$$

for all $p \in \Omega_r$.

PROOF. The convexity of ψ implies that $\psi(p) \geq -L_1$ for all $p \in \Omega_r$. Then the method of proof of Lemma 2.7 applied to the nonnegative function $\psi + L_1$ on Ω_r implies that $\psi(p) + L_1 \leq n(\psi(\overline{e}) + L_1) \leq n(L_1 + L_2)$ for all $p \in \Omega_r$. $\square$

Write $A_r = \{e_i^{(r)} : 1 \leq i \leq n\} \cup \{\overline{e}\}$. Let $\epsilon > 0$ and $r \geq 1$. Since

$$\left(\widehat{\Psi}_K(p) \right)_{p \in A_r} \to \left(-\frac{1}{c} \log \sum_{x \in N} e^{-c \langle p, x \rangle} \right)_{p \in A_r}$$

in distribution, there exists $J_r > 0$ such that

(5.15) $$\mathbb{P}\left(|\widehat{\Psi}_K(p)| \leq J_r \text{ for all } p \in A_r \right) \geq 1 - \frac{\epsilon}{2^r}$$

for all $K \geq 1$. By Lemma 5.12 there exists M_r such that

$$\mathbb{P}\Big(\sup_{p \in \Omega_r} \big| \widehat{\Psi}_K(p) \big| \leq M_r \Big) \geq 1 - \frac{\epsilon}{2^r}$$

and so

$$\mathbb{P}\Big(\sup_{p \in \Omega_r} \big| \widehat{\Psi}_K(p) \big| \leq M_r \text{ for all } r \geq 1 \Big) \geq 1 - \epsilon$$

for all $K \geq 1$. This establishes the tightness of the sequence, and completes the proof of Theorem 5.2. $\square$

To finish this paper we point out that the limit $\widetilde{\Psi}_c$ in Theorem 5.2 is drastically different from the limit Ψ in the hardmin case. By Proposition 5.10, we may realize $\widetilde{\Psi}_c$ by

$$(5.16) \qquad \widetilde{\Psi}_c(p) = -\frac{1}{c} \log\Big(\sum_{x \in N} e^{-c \langle p, x \rangle} \Big),$$

where N is the Poisson point process on $\mathbb{R}_+^n$ with intensity $dm(x) = h(x)\,dx$. From (5.16), it is not difficult to verify that $\widetilde{\Psi}_c$ is differentiable on Δ_n. In contrast, in Theorem 3.5 and Theorem 5.1 the random concave function Ψ is locally piecewise affine on Δ_n. Moreover the following result shows very different boundary behavior of $\widetilde{\Psi}_c$ compared with that of Ψ (see Proposition 4.5). Intuitively, here we see a nonvanishing effect of the softmin as $K \to \infty$ when λ_K is of order $K^{\frac{1}{n+\alpha}}$.

PROPOSITION 5.13. *Under Assumption 4.3, for all $c > 0$ we have*

$$\mathbb{P}\big(\widetilde{\Psi}_c(p) \to -\infty \quad as \ p \to \partial \Delta_n \big) = 1.$$

PROOF. Using (5.16), it suffices to show that

$$\mathbb{P}\big(N\big(R(p, 1) \big) \to \infty \quad as \ p \to \partial \Delta_n \big) = 1,$$

where $N(U)$ is the cardinality of $N \cap U$ where $U \subset \mathbb{R}_+^n$ is Borel. Since $p = (p_1, \ldots, p_n) \to \partial \Delta_n$ implies at least one coordinate p_i tends to 0, it suffices to show that

$$\mathbb{P}\big(N\big(R(p, 1) \big) \to \infty \text{ as } p_n \to 0 \big) = 1.$$

Given $\epsilon > 0$, for $p_n < \epsilon$ we have

$$R(p, 1) = \Big\{ x \in \mathbb{R}_+^n : \sum_{i=1}^n p_i x_i < 1 \Big\}$$

$$\supseteq \Big\{ x \in \mathbb{R}_+^n : \sum_{i=1}^{n-1} p_i x_i < 1/2 \text{ and } p_n x_n < 1/2 \Big\}$$

$$\supseteq \Big\{ x \in \mathbb{R}_+^n : \sum_{i=1}^{n-1} x_i < 1/2 \text{ and } x_n < 1/(2\epsilon) \Big\} := R_{n,\epsilon}$$

and $R_{n,\epsilon} \nearrow \{x \in \mathbb{R}_+^n : \sum_{i=1}^{n-1} x_i < 1/2\} := R_n$ as $\epsilon \to 0^+$. Under Assumption 4.3 we have $m(R_n) = \infty$, and so $N(R_n) = \infty$ almost surely. Therefore $N(R(p, 1)) \geq N(R_{n,\epsilon}) \to \infty$ as $\epsilon \to 0^+$, and the proof is complete. $\square$
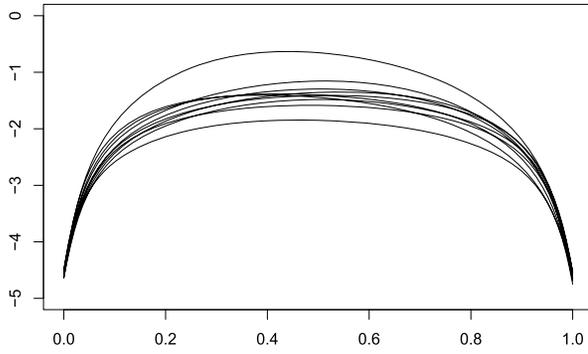
FIG. 8. *Ten approximate samples of $\widetilde{\Psi}_1$ when $n = 2$ and m is the Lebesgue measure on $\mathbb{R}^n_+$ (i.e., $\gamma = 1$ in Example 5.14). Note that the functions are differentiable in p. Here the samples are approximate because the Poisson point process is restricted to $[0, M]^n$ where $M > 0$ is a constant. When $M \to \infty$ the functions explode to $-\infty$ on the boundary of the simplex.*

EXAMPLE 5.14. Suppose that $h(x) \equiv \gamma > 0$ is constant, so that the intensity measure of N is proportional to the Lebesgue measure on $\mathbb{R}^n_+$. Some approximate examples of $\widetilde{\Psi}_1$ are shown in Figure 8. Specializing (5.14) to this case, we have

$$\mathbb{E}\left[\sum_{x \in N} e^{-\langle p, x \rangle} \right] = \gamma \int_{\mathbb{R}^n_+} e^{-\langle p, x \rangle} \, dx = \frac{\gamma}{p_1 p_2 \cdots p_n} \to \infty \quad \text{as } p \to \partial \Delta_n.$$

The proposition above shows that this blow-up also occurs in an almost-sure sense.

APPENDIX: MISCELLANEOUS PROOFS

PROOF OF LEMMA 2.7. (i) Let $q \in \Delta_n$ be given and we will prove (2.9) for $p \neq q$. For $p \in \overline{\Delta}_n$ with $p \neq q$ the half-ray $\{q + t(q - p) : t > 0\}$ intersects the boundary $\partial \Delta_n$ at a unique point p', say. Suppose $p' = q + \lambda(q - p)$, so that $q = \frac{\lambda}{1+\lambda} p + \frac{1}{1+\lambda} p'$. For $\psi \in \mathcal{C}$ the concavity of ψ along the line segment $[p, p']$ gives $\psi(q) \geq \frac{\lambda}{1+\lambda} \psi(p) + \frac{1}{1+\lambda} \psi(p')$. Since $\psi(p') \geq 0$ we get $\psi(q) \geq \frac{\lambda}{1+\lambda} \psi(p)$, so that $\psi(p) \leq \frac{1+\lambda}{\lambda} \psi(q)$.

Now $p' - q = \lambda(q - p)$ so that $\lambda = \|p' - q\| / \|q - p\|$ and then

$$\frac{1 + \lambda}{\lambda} = \frac{\|p' - p\|}{\|p' - q\|} \leq \frac{\text{diam}(\overline{\Delta}_n)}{\text{dist}(q, \partial \Delta_n)} =: M_q.$$

(ii) Now suppose $q = c^{(j,\epsilon)}$ and $0 \leq p_j \leq \epsilon$ with $\epsilon \leq 1/n$. The value $t = (1 - \epsilon)/(n - 2 + \epsilon)$ gives a point $\widetilde{p} = c^{(j,\epsilon)} + t(c^{(j,\epsilon)} - p)$ and it can be checked easily that $\widetilde{p}_i \geq 0$ for all i. Thus $\widetilde{p} \in \overline{\Delta}_n$ and so $\lambda \geq t$. Then

$$\frac{1 + \lambda}{\lambda} \leq \frac{1 + t}{t} = \frac{n - 1}{1 - \epsilon} \leq n.$$

The final statement now follows by taking $\epsilon = 1/n$ and $j = \arg\min\{p_i : 1 \leq i \leq n\}$, and noting that $\min\{p_i : 1 \leq i \leq n\} \leq 1/n$. $\square$

PROOF OF LEMMA 3.3. Replacing x by cx in (3.1) gives $h(cx) = c^\alpha h(x)$ so that h is homogeneous of order α. Since $\int_{\mathbb{R}^n_+} h(x) \, dx > 0$, h is not identically zero and it is easy to see that α is unique. If A is a bounded subset of $\mathbb{R}^n_+$ then there exists $\sigma > 0$ such that $\sum_{i=1}^n y_i \leq \sigma$ for all $y \in A$. Given $\epsilon > 0$ there is κ_0 such that

$$\left| \frac{1}{\kappa^\alpha} \rho(\kappa x) - h(x) \right| < \epsilon$$

whenever $x \in \Delta_n$ and $0 < \kappa < \kappa_0$. Let $y \in A$ then $y = cx$ for some $0 < c \le \sigma$ and $x \in \Delta_n$. For $\kappa < \kappa_0/\sigma$, so that $c\kappa < \kappa_0$, we have (using the homogeneity of h)

$$\left| \frac{1}{\kappa^\alpha} \rho(\kappa y) - h(y) \right| < c^\alpha \epsilon \le \sigma^\alpha \epsilon.$$

Thus $\frac{1}{\kappa^\alpha} \rho(\kappa x) \to h(x)$ uniformly on all bounded subsets of $\mathbb{R}_+^n$, and the limit (3.2) follows immediately.

Taking A to be the set $R = \{x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i \le 1\}$, we get

$$(A.1) \quad \mathbb{P}(C \in \kappa R) = \int_{\kappa R} \rho(x)\, dx = \kappa^n \int_R \rho(\kappa y)\, dy \ge \kappa^{n+\alpha} \left(\int_R h(y)\, dy - \epsilon \mathrm{vol}(R) \right)$$

for $\kappa < \kappa_0$. Since the left hand side is bounded above by 1, we have $\int_R h(y)\, dy < \infty$. Since h is homogeneous, a simple scaling argument gives $\int_A h(y)\, dy < \infty$ for all bounded sets $A \subset \mathbb{R}_+^n$, completing the proof of (3.2).

A similar scaling argument shows that $\int_R h(y)\, dy > 0$, since otherwise $\int_{\mathbb{R}_+^n} h(y)\, dy = 0$ and Assumption 3.1 is violated. So $\epsilon > 0$ can be chosen so that $\int_R h(y)\, dy - \epsilon \mathrm{vol}(R) > 0$. Since the left hand side of (A.1) converges to 0 as $\kappa \to 0^+$, we see that $\kappa^{n+\alpha} \to 0$ so that $n + \alpha > 0$ as desired. $\square$

PROOF OF LEMMA 5.8. Recall that the random vector X_1 has density ρ on $\mathbb{R}_+^n$. Given $p \in \Delta_n$, we define

$$D(p) = \left\{ (u_1, \ldots, u_{n-1}) \in \mathbb{R}_+^{n-1} : \sum_{j=1}^{n-1} p_j u_j < 1 \right\}.$$

By a straightforward computation, it can be verified that the random variable $\langle p, X_1 \rangle$ has density

$$\widetilde{\rho}(t) = \frac{t^{n-1}}{p_n} \int_{D(p)} \rho\left(t u_1, t u_2, \ldots, t u_{n-1}, t\left(1 - \sum_{j=1}^{n-1} p_j u_j\right)/p_n \right) du_1 \cdots du_{n-1}$$

for $t > 0$. By Lemma 3.3, as $t \to 0^+$ we have

$$\widetilde{\rho}(t) \sim \frac{t^{n+\alpha-1}}{p_n} \int_{D(p)} h\left(u_1, \ldots, u_{n-1}, \left(1 - \sum_{j=1}^{n-1} p_j u_j\right)/p_n \right) du_1 \cdots du_{n-1}$$

$$= (n + \alpha) t^{n+\alpha-1} \int_{R(p,1)} h(x)\, dx.$$

This proves (i).

Next we consider (ii) and (iii). By (i) there exist $\delta > 0$ and $A > 0$ such that $\widetilde{\rho}(t) \le A t^{n+\alpha-1}$ for $0 < t \le \delta$. Then

$$\gamma^{n+\alpha} \mathbb{E}\big[e^{-\gamma \langle p, X_1 \rangle} 1_{\gamma \langle p, X_1 \rangle \ge L} \big] = \gamma^{n+\alpha} \int_{L/\gamma}^\infty e^{-\gamma t} \widetilde{\rho}(t)\, dt$$

$$= \gamma^{n+\alpha} \int_{L/\gamma}^{\max(\delta, L/\gamma)} e^{-\gamma t} \widetilde{\rho}(t)\, dt$$

$$+ \gamma^{n+\alpha} \int_{\max(\delta, L/\gamma)}^\infty e^{-\gamma t} \widetilde{\rho}(t)\, dt$$

$$=: I_1 + I_2,$$

say. If $L/\gamma \geq \delta$ then $I_1 = 0$. Otherwise

$$
\begin{aligned}
I_1 = \gamma^{n+\alpha} \int_{L/\gamma}^{\delta} e^{-\gamma t} \widetilde{\rho}(t)\, dt &\leq A\gamma^{n+\alpha} \int_{L/\gamma}^{\delta} e^{-\gamma t} t^{n+\alpha-1}\, dt \\
&= A \int_{L}^{\gamma\delta} e^{-u} u^{n+\alpha-1}\, du \\
&\leq A \int_{L}^{\infty} e^{-u} u^{n+\alpha-1}\, du.
\end{aligned}
$$

Also $I_2 \leq \gamma^{n+\alpha} e^{-\gamma\delta}$. Together we get

$$
\gamma^{n+\alpha} \mathbb{E}\big[e^{-\gamma\langle p, X_1\rangle} 1_{\gamma\langle p, X_1\rangle \geq L}\big] \leq A \int_{L}^{\infty} e^{-u} u^{n+\alpha-1}\, du + \gamma^{n+\alpha} e^{-\gamma\delta}.
$$

Thus (iii) follows immediately, and taking $L = 0$ we obtain (ii) with $B = A\Gamma(n + \alpha) + \delta^{-(n+\alpha)} \sup\{v^{n+\alpha} e^{-v} : v > 0\}$. $\square$

REFERENCES

[1] BRENIER, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402

[2] COVER, T. M. (1991). Universal portfolios. *Math. Finance* **1** 1–29. MR1113417 https://doi.org/10.1111/j.1467-9965.1991.tb00002.x

[3] CUCHIERO, C., SCHACHERMAYER, W. and WONG, T.-K. L. (2019). Cover's universal portfolio, stochastic portfolio theory, and the numéraire portfolio. *Math. Finance* **29** 773–803. MR3974282 https://doi.org/10.1111/mafi.12201

[4] CULE, M., SAMWORTH, R. and STEWART, M. (2010). Maximum likelihood estimation of a multi-dimensional log-concave density. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 545–607. MR2758237 https://doi.org/10.1111/j.1467-9868.2010.00753.x

[5] DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering.* Springer, New York. MR2234156 https://doi.org/10.1007/0-387-34471-3

[6] DE PHILIPPIS, G. and FIGALLI, A. (2014). The Monge–Ampère equation and its link to optimal transportation. *Bull. Amer. Math. Soc.* (*N.S.*) **51** 527–580. MR3237759 https://doi.org/10.1090/S0273-0979-2014-01459-4

[7] DÜMBGEN, L. and RUFIBACH, K. (2009). Maximum likelihood estimation of a log-concave density and its distribution function: Basic properties and uniform consistency. *Bernoulli* **15** 40–68. MR2546798 https://doi.org/10.3150/08-BEJ141

[8] EVANS, L. C. and GARIEPY, R. F. (2015). *Measure Theory and Fine Properties of Functions*, Revised ed. *Textbooks in Mathematics.* CRC Press, Boca Raton, FL. MR3409135

[9] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

[10] FERNHOLZ, E. R. (2002). *Stochastic Portfolio Theory. Applications of Mathematics* (*New York*) **48**. Springer, New York. MR1894767 https://doi.org/10.1007/978-1-4757-3699-1

[11] GINÉ, E., HAHN, M. G. and VATAN, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165. MR1080487 https://doi.org/10.1007/BF01198427

[12] HANNAH, L. A. and DUNSON, D. B. (2011). Bayesian nonparametric multivariate convex regression. Preprint. Available at arXiv:1109.0322.

[13] HANNAH, L. A. and DUNSON, D. B. (2013). Multivariate convex regression with adaptive partitioning. *J. Mach. Learn. Res.* **14** 3261–3294. MR3144462

[14] ITKIN, D. and LARSSON, M. (2020). Robust asymptotic growth in stochastic portfolio theory under long-only constraints. Preprint. Available at arXiv:2009.08533.

[15] JOHNSON, A. L. and JIANG, D. R. (2018). Shape constraints in economics and operations research. *Statist. Sci.* **33** 527–546. MR3881207 https://doi.org/10.1214/18-STS672

[16] KABLUCHKO, Z., MARYNYCH, A., TEMESVARI, D. and THÄLE, C. (2019). Cones generated by random points on half-spheres and convex hulls of Poisson point processes. *Probab. Theory Related Fields* **175** 1021–1061. MR4026612 https://doi.org/10.1007/s00440-019-00907-3

[17] KARATZAS, I. and FERNHOLZ, R. (2009). Stochastic portfolio theory: An overview. In *Handbook of Numerical Analysis*, **15** 89–167. Elsevier.

[18] KINGMAN, J. F. C. (1993). *Poisson Processes. Oxford Studies in Probability* **3**. The Clarendon Press, New York. MR1207584

[19] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes. Springer Series in Statistics*. Springer, New York. MR0691492

[20] LEE, J. M. (2013). *Introduction to Smooth Manifolds*, 2nd ed. *Graduate Texts in Mathematics* **218**. Springer, New York. MR2954043

[21] MARIUCCI, E., RAY, K. and SZABÓ, B. (2020). A Bayesian nonparametric approach to log-concave density estimation. *Bernoulli* **26** 1070–1097. MR4058361 https://doi.org/10.3150/19-BEJ1139

[22] MOLCHANOV, I. (2005). *Theory of Random Sets. Probability and Its Applications* (*New York*). Springer London, Ltd., London. MR2132405

[23] PAL, S. and WONG, T.-K. L. (2016). The geometry of relative arbitrage. *Math. Financ. Econ.* **10** 263–293. MR3500452 https://doi.org/10.1007/s11579-015-0159-z

[24] PAL, S. and WONG, T.-K. L. (2018). Exponentially concave functions and a new information geometry. *Ann. Probab.* **46** 1070–1113. MR3773381 https://doi.org/10.1214/17-AOP1201

[25] PAL, S. and WONG, T.-K. L. (2020). Multiplicative Schrödinger problem and the Dirichlet transport. *Probab. Theory Related Fields* **178** 613–654. MR4146546 https://doi.org/10.1007/s00440-020-00987-6

[26] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes. Applied Probability. A Series of the Applied Probability Trust* **4**. Springer, New York. MR0900810 https://doi.org/10.1007/978-0-387-75953-1

[27] ROCKAFELLAR, R. T. (1997). *Convex Analysis. Princeton Landmarks in Mathematics*. Princeton Univ. Press, Princeton, NJ. MR1451876

[28] SAMWORTH, R. J. (2018). Recent progress in log-concave density estimation. *Statist. Sci.* **33** 493–509. MR3881205 https://doi.org/10.1214/18-STS666

[29] SAUMARD, A. and WELLNER, J. A. (2014). Log-concavity and strong log-concavity: A review. *Stat. Surv.* **8** 45–114. MR3290441 https://doi.org/10.1214/14-SS107

[30] SCHNEIDER, R. and WEIL, W. (2008). *Stochastic and Integral Geometry. Probability and Its Applications* (*New York*). Springer, Berlin. MR2455326 https://doi.org/10.1007/978-3-540-78859-1

[31] SEIJO, E. and SEN, B. (2011). Nonparametric least squares estimation of a multivariate convex regression function. *Ann. Statist.* **39** 1633–1657. MR2850215 https://doi.org/10.1214/10-AOS852

[32] VILLANI, C. (2003). *Topics in Optimal Transportation. Graduate Studies in Mathematics* **58**. Amer. Math. Soc., Providence, RI. MR1964483 https://doi.org/10.1090/gsm/058

[33] VILLANI, C. (2009). *Optimal Transport: Old and new. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Springer, Berlin. MR2459454 https://doi.org/10.1007/978-3-540-71050-9

[34] VON RENESSE, M.-K. and STURM, K.-T. (2009). Entropic measure and Wasserstein diffusion. *Ann. Probab.* **37** 1114–1191. MR2537551 https://doi.org/10.1214/08-AOP430

[35] WONG, T.-K. L. (2015). Universal portfolios in stochastic portfolio theory. Preprint. Available at arXiv:1510.02808.

[36] WONG, T.-K. L. (2015). Optimization of relative arbitrage. *Ann. Finance* **11** 345–382. MR3426111 https://doi.org/10.1007/s10436-015-0261-5

[37] WONG, T.-K. L. (2018). Logarithmic divergences from optimal transport and Rényi geometry. *Inf. Geom.* **1** 39–78. MR4010746 https://doi.org/10.1007/s41884-018-0012-6

[38] WONG, T.-K. L. (2019). Information geometry in portfolio theory. In *Geometric Structures of Information. Signals Commun. Technol.* 105–136. Springer, Cham. MR3887875

[39] WONG, T.-K. L. and YANG, J. (2019). Pseudo-Riemannian geometry embeds information geometry in optimal transport. Preprint. Available at arXiv:1906.00030.