The Annals of Probability
1996, Vol. 24, No. 3, 1490–1506

ON THE CONVERGENCE OF SCALED RANDOM SAMPLES[1]

By Geoffrey Pritchard

University of Wisconsin—Madison

The scaled-sample problem asks the following question: given a distribution on a normed linear space E, when do there exist constants $\{\gamma_n\}$ such that $\{X^{(j)}/\gamma_n\}_{j=1}^n$ converges as $n \to \infty$ (in the Hausdorff metric given by the norm) to a fixed set K? (Here $\{X^{(j)}\}$ are i.i.d. with the given distribution.) The main result presented here relates the convergence of scaled samples to a large deviation principle for single observations, thereby achieving a dimension-free description of the problem.

1. Introduction. The *scaled-sample problem* asks the following question: given a probability measure μ on a normed linear space E, when do there exist constants $\gamma_n \uparrow \infty$ such that $\{X^{(j)}/\gamma_n\}_{j=1}^n \to K$ (a.s., or in probability), where $\{X^{(j)}\}_{j=1}^\infty$ are i.i.d. with law μ, K is a (necessarily deterministic and compact) closed subset of E and the convergence is in the Hausdorff metric on closed bounded sets induced by the norm? [The Hausdorff metric is that given by $h(K,L) = (\sup_{x\in K} \inf_{y\in L}\|x-y\|) \vee (\sup_{x\in L}\inf_{y\in K}\|x-y\|)$.]

The simplest examples occur with $E = \mathbb{R}$ and μ supported on $[0,\infty)$. In this case the scaled-sample problem is equivalent to the question of whether there exist γ_n such that $(\max_{m=1}^n X^{(j)})/\gamma_n \to 1$ (a.s., or in probability) (see [6]). This is sometimes referred to as *relative stability* (a.s., or in probability) of $\{\max_{j=1}^n X^{(j)}\}_{n=1}^\infty$. Equivalent conditions can be given in terms of the distribution function F of μ.

The almost sure convergence occurs iff [8]

$$\int_1^\infty \frac{dF(x)}{1-F(\varepsilon x)} < \infty \quad \text{for } 0 < \varepsilon < 1.$$

The convergence in probability occurs iff [5]

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = \begin{cases} \infty, & \text{if } x < 1, \\ 1, & \text{if } x = 1, \\ 0, & \text{if } x > 1. \end{cases}$$

Also, $\gamma_n \sim F^{\leftarrow}(1-1/n) = \inf\{t\colon F(t) \ge 1-1/n\}$ [8]. (We will everywhere use $a_n \sim b_n$ to mean $a_n/b_n \to 1$.) A simple sufficient condition for the a.s.

Received September 1994; revised April 1995.
[1] Supported by NSF grant DMS-90-24961.
AMS 1991 *subject classification*. Primary 60G70, 60B12; secondary 60B11, 60F15, 60D05.
Key words and phrases. Scaled sample, large deviations, regular variation.

convergence is the regular variation of $-\ln(1-F)$, that is, the condition that for some $\alpha > 0$ we have

$$\frac{-\ln(1-F(tx))}{-\ln(1-F(t))} \to x^{\alpha} \quad \text{as } t \to \infty, \forall x > 0.$$

This condition (or rather a multidimensional version of it) was considered in [3]. It is a stronger condition than scaled-sample convergence, and we shall consider shortly to what scaled-sample behaviour it is equivalent.

The situation when $E = \mathbb{R}^d$ was considered by Kinoshita and Resnick [6], who described the a.s. convergence of such scaled samples by a polar coordinate decomposition: $\{\max_{j=1}^{n} \|X^{(j)}\|\}$ must be a.s. relatively stable, and another condition on the relative extent of the distribution in different directions determines the shape of the limiting set K. This approach yields quite simple necessary and sufficient conditions for scaled-sample convergence, but it is unfortunately limited to finite-dimensional spaces. (An infinite-dimensional example which satisfies the hypotheses of Proposition 4.8 in [6], but whose scaled samples do not converge is as follows: let $\{e_i\}_{i=1}^{\infty}$ be orthogonal unit vectors in a Hilbert space and let Z be a random vector which takes the value e_i with probability 2^{-i}. Let Y be an exponentially distributed random variable of mean 1 which is independent of Z and let $X = YZ$.)

In infinite dimensions the best-known example of scaled-sample convergence occurs for Gaussian measures. The limiting set is then the unit ball of the reproducing kernel Hilbert space associated with the measure (see, for example, [1]).

Finally, Fisher [4] considered the scaled-sample problem for $E = \mathbb{R}^2$ and μ a product of two identical measures supported on $[0, \infty)$. He discovered that the limiting sets could only be balls (i.e., $\{(x, y): x \geq 0,\ y \geq 0 \text{ and } x^p + y^p \leq 1\}$ for some $p \in [0, \infty]$). His results contain the essential idea of *regular* scaled-sample convergence defined in the next section.

2. The main result. We use the following notation. For a subset K of a vector space E, $d_K: E \to [0, \infty]$ is given by $d_K(x) = (\sup\{t \geq 0: tx \in K\})^{-1}$ (or ∞ if the sup is 0). Also,

$$B_m(K, p) = \left\{ (x_1, \dots, x_m) \in E^m : \sum_{i=1}^{m} d_K(x_i)^p \leq 1 \right\} \quad \text{for } 0 < p < \infty,$$

$$B_m(K, 0) = \bigcap_{p>0} B_m(K, p), \qquad B_m(K, \infty) = K^m$$

and we write $B(K, p)$ for $B_2(K, p)$.

We shall say that a family $\{\mu_\varepsilon\}_{\varepsilon > 0}$ of measures on E indexed by the positive real numbers satisfies a large deviation principle (LDP) with good rate function $I: E \to [0, \infty]$ if the following hold: I is lower semicontinuous, the "level sets" $\{x: I(x) \le a\}$ are compact $\forall a > 0$ and for all measurable A,

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \mu_\varepsilon(A) \le -\inf_{\mathrm{cl}\, A} I,$$

$$\liminf_{\varepsilon\to 0} \varepsilon \ln \mu_\varepsilon(A) \ge -\inf_{A^\circ} I.$$

A family of random elements will be said to satisfy a LDP if the family of probability laws it generates does so.

THEOREM 1. *Let E be a normed linear space and use the Hausdorff metric on its closed subsets. Let μ be a probability measure on E and let X, $\{X^{(j)}\}_{j=1}^\infty$ and $\{X_k^{(j)}\}_{j,k=1}^\infty$ all be i.i.d. with law μ. Let $\{\gamma_n\}_{n=1}^\infty$ be a positive sequence increasing to ∞ and let $f: (0,\infty) \to (0,\infty)$ be decreasing with $f(1/\ln n) = \gamma_n$. Let $K \subset E$ be closed with $\{0\}$ a proper subset of K.*

(i) *Let $p \in (0,\infty]$. The following are equivalent:*

(a) *$\{X/f(\varepsilon)\}_{\varepsilon>0}$ satisfies a large deviation principle as $\varepsilon \to 0$, with good rate function*

$$I(x) = \begin{cases} d_K(x)^p, & \text{if } 0 < p < \infty,\\ 0, & \text{if } p = \infty \text{ and } x \in K,\\ \infty, & \text{if } p = \infty \text{ and } x \notin K.\end{cases}$$

(b) $$\left\{\frac{\left(X_1^{(j)}, \dots, X_m^{(j)}\right)}{\gamma_n}\right\}_{j=1}^n \to B_m(K,p) \quad a.s. \text{ for all } m \in \mathbb{N}.$$

(c) $$\left\{\frac{\left(X_1^{(j)}, X_2^{(j)}\right)}{\gamma_n}\right\}_{j=1}^n \to_p B(K,p)$$

(d) $$\left\{\frac{X^{(j)}}{\gamma_n}\right\}_{j=1}^n \to_p K \quad \text{and} \quad \frac{\gamma\lfloor n^r \rfloor}{\gamma_n} \to r^{1/p} \qquad \forall r > 0.$$

(*If $p = \infty$, the limit is* 1 *for all $r > 0$.*)

(ii) *If $\{(X_1^{(j)}, X_2^{(j)}/\gamma_n\}_{j=1}^n \to_p N$, then N has the form $B(K,p)$ for some $p \in [0,\infty]$.*

(iii) *Suppose $\{X/f(\varepsilon)\}_{\varepsilon>0}$ satisfies an LDP with a rate function I which is "good"* (*has compact level sets*) *and such that* (a) *if $I(x) = 1$, then $\inf_U I < 1$ for all neighborhoods U of X,* (b) *$\exists x \ne 0$ with $I(x) \le \frac{1}{2}$ and* (c) *$I(tx) \le I(x)$ for $x \in E$, $t \in [0,1]$. Then the equivalent conditions of* (i) *hold, for some $p \in (0,\infty]$.*

REMARKS. (ii) is a dimension-free generalization of the result in [4]. It is equally valid for products of any finite number of copies of μ. [Apply a projection onto any pair of components. If $p > 0$, we can use (i); if $p = 0$, note that a set in E^m whose every projective onto two coordinates is $B(K,0)$ must be $B_m(K,0)$.]

The requirement that I have compact level sets is a natural one, since a limiting set K of scaled samples must always be compact. (For any $\varepsilon > 0$, it will be covered by a family of balls of radius ε centered at the points of some sufficiently large scaled sample.)

The equivalent conditions of (i) define, for each $p \in (0, \infty]$, a class of distributions on E which will be described as exhibiting "regular scaled-sample convergence of index p."

The large deviation principle is a dimension-free generalization of the regular variation condition mentioned in the previous section. This is demonstrated (in one dimension) by the next result.

PROPOSITION 2.1. *With notation as in Theorem* 1, *suppose that* $E = \mathbb{R}$, μ *is supported on* $[0, \infty)$, $0 < p < \infty$, *and* $K = [0, 1]$. *Then the following statements are equivalent*:

(i) *The large deviation principle of* (i)(a) *above holds.*

(ii) ϕ *given by* $\phi(x) = -\ln \mu((x, \infty))$ *is regularly varying of index p at* ∞ *and* $\gamma_n \sim \phi^{\leftarrow}(\ln n)$.

PROOF. First note that (i) is equivalent to $\lim_{\varepsilon \to 0} \varepsilon \ln \mu((af(\varepsilon), \infty)) = -a^p$ for all $a > 0$ [i.e., the LDP just for sets of the form $[a, \infty)$]. This implies

$$\varepsilon\phi(f(\varepsilon)) \to 1 \quad \text{as } \varepsilon \to 0. \tag{2.1}$$

We now show that (2.1) is also implied by (ii). For $\phi(\gamma_n) \sim \phi(\phi^{\leftarrow}(\ln n)) \sim \ln n$ by regular variation of ϕ [note $\phi(\phi^{\leftarrow}(\ln n))$ differs from $\ln n$ only when ϕ is discontinuous at $\phi^{\leftarrow}(\ln n)$; the difference can be no more than the size of the discontinuity, which (again by regular variation) must be $o(\ln n)$]. We thus have $\varepsilon_n \phi(f(\varepsilon_n)) \to 1$ when $\varepsilon_n = (\ln n)^{-1}$, and hence (2.1) in general.

The equivalence now follows by writing

$$\varepsilon\phi(af(\varepsilon)) = \varepsilon\phi(f(\varepsilon)) \cdot \frac{\phi(af(\varepsilon))}{\phi(f(\varepsilon))}. \qquad \square$$

3. Examples.

EXAMPLE 1. Let $p \in (0, \infty)$ and $s \in \mathbb{R}$. Consider the distribution on $[2, \infty)$ given by

$$\Phi(x) = P(X > x) = \begin{cases} \exp(-x^p(p \ln n)^{-ps}, & \text{for } x \geq 2, \\ 1, & \text{for } x < 2. \end{cases}$$

Let $\gamma_n = \Phi^{-1}(1/n)$. Calculation shows that $\gamma_n \sim (\ln n)^{1/p}(\ln \ln n)^s$. Then by Proposition 2.1, we see that the LDP of Theorem 1 is satisfied (for any suitable f) with rate function $I(x) = x^p$. Hence this distribution exhibits regular scaled-sample convergence of index p. The limiting set is $[0, 1]$. The member of this family with $p = 2$, $s = 0$ is closely analogous to (and has the same scaled-sample behaviour as) the one-dimensional Gaussian distribution.

EXAMPLE 2. The distribution on $[0, \infty)$ with $\Phi(x) = P(X > x) = \exp(-e^x)$. Here $\gamma_n = \ln \ln n$ is the natural choice. The function $-\ln \Phi$ is regular varying of order ∞, so this distribution exhibits scaled-sample convergence of order ∞. The limiting set is $[0, 1]$.

EXAMPLE 3 (Example 5.2 from [6]). The distribution on $[0,\infty)^2$ given by

$$\mu([x,\infty) \times [y,\infty)) = (e^x + e^y - 1)^{-1} \quad \text{for } x, y > 0.$$

The LDP for single observations for this distribution has normalizer $f(\varepsilon) = \varepsilon^{-1}$ and rate function $I(x_1, x_2) = 2(x_1 \vee x_2) - (x_1 \wedge x_2)$ for $x_1 \geq 0$, $x_2 \geq 0$. To check this, it is enough to check it only for sets of the form (i) $[a,\infty) \times [b_1, b_2]$, where $0 \leq b_1 < b_2 \leq a$ (and images of such sets under interchanging the coordinates), and (ii) $[a,\infty)^2$, where $a > 0$. For the large deviation lower bound notice that any ball contains a difference of two sets of type (i) with different values of a, hence different large deviation rates. For the upper bound notice that $\{x: I(x) \geq c\}$ is always contained in a union U of finitely many sets of types (i) and (ii) with $\inf_U I$ arbitrarily close to c. For the type (i) sets,

$$\begin{aligned}
\lim_{\varepsilon \to 0} & \varepsilon \ln P(X/f(\varepsilon) \in [a,\infty) \times [b_1, b_2]) \\
&= \lim_{\varepsilon \to 0} \varepsilon \ln\bigl(1/(e^{a/\varepsilon} + e^{b_1/\varepsilon} - 1) - 1/(e^{a/\varepsilon} + e^{b_2/\varepsilon} - 1)\bigr) \\
&= b_2 - 2a \\
&= -I(a, b_2) \\
&= -\inf_{[a,\infty)\times[b_1,b_2]} I.
\end{aligned}$$

For the type (ii) sets,

$$\begin{aligned}
\lim_{\varepsilon \to 0} \varepsilon \ln P\bigl(X/f(\varepsilon) \in [a,\infty)^2\bigr) &= -a \\
&= -\inf_{[a,\infty)} I, \quad \text{similarly.}
\end{aligned}$$

This rate function has the form $I(x) = d_K(x)^p$ with $p = 1$ and K the convex hull of $\{(0,0), (\frac{1}{2}, 0), (0, \frac{1}{2}), (1,1)\}$. By Theorem 1, the distribution exhibits regular scaled-sample convergence of index 1 to K. The normalizers are $\gamma_n = f(1/\ln n) = \ln n$.

EXAMPLE 4. This example shows that convergence of scaled samples need not be regular of any order.

Let $\lambda: (0,\infty) \to (0,\infty)$ be the piecewise linear, continuous, increasing function with $\lambda(0) = 0$ and

$$\lambda'(t) = \begin{cases} s_1, & \text{on } (n, n+1), \text{ for all even } n, \\ s_2, & \text{on } (n, n+1), \text{ for all odd } n. \end{cases}$$

Here $0 < s_1 < s_2$. Define a distribution on $[1,\infty)$ by

$$\Phi(t) = P(X > t) = \exp(-\exp(\lambda(\ln t))) \quad \text{for } t \geq 1.$$

First, we show that the scaled samples from this distribution converge (a.s.) by checking the integral condition of [8]. We need that for every $\varepsilon \in (0,1)$,

$$-\int_1^\infty \frac{d\Phi(t)}{\Phi(\varepsilon t)} < \infty, \quad \text{that is,} \quad \int_1^\infty \frac{|\Phi'(t)|}{\Phi(\varepsilon t)}\,dt < \infty.$$

We have

$$\begin{aligned}|\Phi'(t)| &= |\Phi(t)\ln\Phi(t)\lambda'(\ln t)t^{-1}| \\ &\le \Phi(t)|\ln\Phi(t)|s_2 t^{-1}.\end{aligned}$$

Also (using the mean value theorem twice),

$$\begin{aligned}\frac{\Phi(t)}{\Phi(\varepsilon t)} &= \exp(\exp(\lambda(\ln\varepsilon + \ln t)) - \exp(\lambda(\ln t))) \\ &\le \exp(\exp(\lambda(\ln t))(\lambda(\ln\varepsilon + \ln t) - \lambda(\ln t))) \\ &\le \exp(\exp(\lambda(\ln t))s_1\ln\varepsilon).\end{aligned}$$

Since $\lambda(t) \le s_2 t$ we have

$$\begin{aligned}\frac{|\Phi'(t)|}{\Phi(\varepsilon t)} &\le \exp(\exp(\lambda(\ln t))s_1\ln\varepsilon + \lambda(\ln t))s_2 t^{-1} \\ &\le \exp(t^{s_2}s_1\ln\varepsilon + s_2\ln t)s_2 t^{-1},\end{aligned}$$

which is integrable on $[1,\infty)$ for $0 < \varepsilon < 1$. Hence the scaled samples convergence when they are normalized by, for example, constants $\gamma_n = \Phi^{-1}(1/n) = \exp(\lambda^{-1}\ln n)$).

Second, we show that the LDP does not hold; equivalently that $-\ln\Phi$ is not regularly varying. For $a, t > 0$,

$$\begin{aligned}\frac{-\ln\Phi(at)}{-\ln\Phi(t)} &= \exp(\lambda(\ln a + \ln t) - \lambda(\ln t)) \\ &= \exp((\ln a)\lambda'(c)),\end{aligned}$$

where c lies between $\ln t$ and $\ln t + \ln a$. If $|\ln a| \le 1$, then $\limsup_{t\to\infty}\lambda'(c) = s_2$ and $\liminf_{t\to\infty}\lambda'(c) = s_1$.

EXAMPLE 5. Let $\{e_i\}_{i=1}^\infty$ be the usual basis vectors of l_2; that is, e_i is an infinite vector with a 1 in the ith position and 0 in all other positions. Choose a positive sequence $r_k \to 0$. Let N be a $\mathbb{N}$-valued random variable with $P(N = i) > 0\ \forall i$ and let Y be an exponential random variable independent of N. Let μ be the law of $X = Yr_N e_N$. (Compare this example with the infinite-dimensional example of Section 1.)

The relevant LDP for this μ has $f(\varepsilon) = \varepsilon^{-1}$ and $I = d_K$, where

$$K = \bigcup_{i=1}^\infty \{te_i : 0 \le t \le r_i\}.$$

To show the large deviation upper bound, let A be measurable and let $t = \inf_{\mathrm{cl}\, A} I$. If $t = \infty$, then $P(X/f(\varepsilon) \in A) = 0\ \forall \varepsilon > 0$, so this case is trivial. Otherwise, the upper bound follows from

$$P(X/f(\varepsilon) \in A) \le P\big(d_K(X/f(\varepsilon)) \ge t\big) = P(Y \ge t/\varepsilon) = \exp(-t/\varepsilon).$$

To show the lower bound, let A be measurable and let $t = \inf_{A^\circ} I$. If $t = \infty$ there is nothing to show. Otherwise, for $\delta > 0$ arbitrary, we can find i, r, s such that $t \le r < s \le t + \delta$ and $ur_i e_i \subseteq A$ for $r \le u \le s$. The lower bound then follows from

$$\begin{aligned} P(X/f(\varepsilon) \in A) &\ge P(X/f(\varepsilon) = ur_i e_i \text{ for some } u \in [r, s]) \\ &= P(N = i) P(Y/f(\varepsilon) \in [r, s]). \end{aligned}$$

Samples from μ scaled by the normalizers $\gamma_n = \ln n$ thus have limiting set K.

4. Proof of Theorem 1. Our first result will eventually show part (iii) of Theorem 1.

PROPOSITION 4.1. *Suppose the conditions of Theorem* 1(iii) *hold. Then* $\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{n=1}^n \to N$ *a.s. for some* $N \subset E^2$. *Furthermore*, $N \ne B(K, 0)$ *for any* K.

LEMMA 4.2. *Let* $\{\mu_\varepsilon\}_{\varepsilon > 0}$ *be a family of probability measures on a topological space* S (*with a* σ-*algebra containing the Borel* σ-*algebra*), *satisfying a large deviation lower bound with rate* I, *that is*, $\liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(A) \ge -\inf_{A^\circ} I$ *for all measurable* A. *Then* $\{\mu_\varepsilon \times \mu_\varepsilon\}_{\varepsilon > 0}$ *satisfies a large deviation lower bound on* S^2 *with rate* $J(x, y) = I(x) + I(y)$, *which is good* (*i.e., has compact level sets*) *if* I *is.*

PROOF. This is the lower bound of Lemma 2.8 in [7]. Note that the space is not required to be Polish for this lower bound. □

REMARK. The large deviation upper bound for compact sets could be proved similarly, but the upper bound for closed sets cannot be proved without an additional exponential tightness assumption. We can prove Proposition 4.1 without this additional assumption, however.

PROOF OF PROPOSITION 4.1. We use the product norm $\|(x, y)\| = \|x\| \vee \|y\|$ in E^2. For any set S, S^ε denotes the set of points whose distance from S is less than ε, and S^c denotes the complement of S. Let $F_n = \{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n$. Let $J(x, y) = I(x) + I(y)$. We will show that the limiting set is $N = \{z: J(z) \le 1\}$. Fix $\varepsilon > 0$. We show that eventually $F_n \subset N^\varepsilon$ and $N \subset F_n^\varepsilon$.

$F_n \subset N^\varepsilon$ *eventually*. Let $A = (N^\varepsilon)^c$. The proof falls into two parts: first showing that

$$\begin{aligned} A &\subseteq \Big(\{(x, y): I(x) \le 1\}^{\varepsilon'}\Big)^c \cup \Big(\{(x, y): I(y) \le 1\}^{\varepsilon'}\Big)^c \\ &\quad \cup \bigcup_{i=1}^k \Big[\Big(\{(x, y): I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c \cap \Big(\{(x, y): I(y) \le \beta_i\}^{\varepsilon'}\Big)^c\Big] \end{aligned} \tag{4.1}$$

for some $\varepsilon' > 0$, where each $\alpha_i + \beta_i > 1$, and then using (4.1) to complete the proof.

To show (4.1), first we cover the compact set

$$\mathrm{cl}\Big((N^{\varepsilon/2})^c\Big) \cap \{(x, y): I(x) \le 1, I(y) \le 1\}$$

in a similar manner. Let $(a, b) \in \mathrm{cl}((N^{\varepsilon/2})^c)$. Then $J(a, b) > 1$. By lower semicontinuity of J, we have $\delta > 0$ and λ such that $\inf_{B(a,\delta)\times B(b,\delta)} J \ge \lambda > 1$. [$B(x, r)$ denotes a ball of center x, radius r.] Let $\alpha = \inf_{B(a,\delta)} I$ and $\beta = \inf_{B(b,\delta)} I$. Then $\alpha + \beta > 1$. Take α', β' such that $\alpha' < \alpha$, $\beta' < \beta$ and $\alpha' + \beta' > 1$. Then $(a, b) \in (\{(x, y)\colon I(x) \le \alpha'\}^{\delta})^c \cap (\{(x, y)\colon I(y) \le \beta'\}^{\delta})^c$. By compactness, then,

$$\begin{aligned} &\mathrm{cl}\Big((N^{\varepsilon/2})^c\Big) \cap \{(x, y): I(x) \le 1, I(y) \le 1\} \\ &\quad \subseteq \bigcup_{i=1}^{k} \Big[\Big(\{(x, y): I(x) \le \alpha_i\}^{\varepsilon_i}\Big)^c \cap \Big(\{(x, y): I(y) \le \beta_i\}^{\varepsilon_i}\Big)^c\Big], \end{aligned} \tag{4.2}$$

where each $\alpha_i + \beta_i > 1$ and $\varepsilon_i > 0$.

Now to show (4.1) itself, choose $\varepsilon' < \frac{1}{2}(\varepsilon \wedge \varepsilon_1 \wedge \cdots \wedge \varepsilon_k)$ and take $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k$ as above. Suppose $(a, b) \in A$, but $(a, b) \in \{(x, y)\colon I(x) \le 1\}^{\varepsilon'}$ and $(a, b) \in \{(x, y)\colon I(y) \le 1\}^{\varepsilon'}$. Then $\exists (a_1, b_1)$ with $I(a_1) \le 1$ and $\|(a, b) - (a_1, b_1)\| < \varepsilon'$, and also (a_2, b_2) with $I(b_2) \le 1$ and $\|(a, b) - (a_2, b_2)\| < \varepsilon'$. Put $(c, d) = (a_1, b_2)$. Then $\|(a, b) - (c, d)\| < \varepsilon'$. Since $(a, b) \in (N^{\varepsilon})^c$, we have $(c, d) \in (N^{\varepsilon/2})^c$ as well as $I(c) \le 1$ and $I(d) \le 1$. By (4.2) we have

$$(c, d) \in \Big(\{(x, y): I(x) \le \alpha_i\}^{\varepsilon_i}\Big)^c \cap \Big(\{(x, y): I(y) \le \beta_i\}^{\varepsilon_i}\Big)^c$$

for some i, and so

$$(a, b) \in \Big(\{(x, y): I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c \cap \Big(\{(x, y): I(y) \le \beta_i\}^{\varepsilon'}\Big)^c.$$

To use (4.1), let X, Y be i.i.d. with law μ. We use (4.1) to get an upper bound on $P((X, Y)/f(\delta) \in A)$. We have for each i,

$$\begin{aligned} &P\Big((X, Y)/f(\delta) \in \Big(\{(x, y): I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c \cap \Big(\{(x, y): I(y) \le \beta_i\}^{\varepsilon'}\Big)^c\Big) \\ &\quad = P\Big(X/f(\delta) \in \Big(\{x: I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c\Big) P\Big(X/f(\delta) \in \Big(\{x: I(x) \le \beta_i\}^{\varepsilon'}\Big)^c\Big). \end{aligned}$$

Let $r_i = \inf\{I(x)\colon x \in \mathrm{cl}((\{x\colon I(x) \le \alpha_i\}^{\varepsilon'})^c)\}$. Note that $r_i > \alpha_i$ since by the goodness of J, the infimum will be achieved. The large deviation principle then gives

$$\limsup_{\delta \to 0} \delta \ln P\Big(X/f(\delta) \in \Big(\{x: I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c\Big) \le -r_i < \alpha_i$$

and so

$$P\Big(X/f(\delta) \in \Big(\{x: I(x) \le \alpha_i\}^{\varepsilon'}\Big)^c\Big) \le \exp(-\alpha_i/\delta)$$

for all sufficiently small δ. Together with similar arguments for β_i, this produces

$$P\left(\frac{(X,Y)}{f(\delta)} \in \left(\{(x,y): I(x) \le \alpha_i\}^{\varepsilon'}\right)^c \cap \left(\{(x,y): I(y) \le \beta_i\}^{\varepsilon'}\right)^c\right)$$
$$\le \exp\left(-\frac{\alpha_i + \beta_i}{\delta}\right)$$

for all sufficiently small δ. Similarly,

$$P\left((X,Y)/f(\delta) \in \left(\{(x,y): I(x) \le 1\}^{\varepsilon'}\right)^c\right) \le \exp(-s/\delta)$$

for all small enough δ, where $s > 1$. Using (4.1), then, we have that for some $t > 1$,

$$P((X,Y)/f(\delta) \in A) \le (k+2)\exp(-t/\delta)$$

for all small enough δ. Put $\delta = (\ln n)^{-1}$. Then for all large enough n,

$$P((X,Y)/\gamma_n \in A) \le (k+2)n^{-t}$$

and so these probabilities are summable $n \to \infty$. By the Borel–Cantelli lemma, $(X^{(n)}, Y^{(n)})/\gamma_n \in A$ only finitely often, a.s. Thus $F_n \subseteq N^\varepsilon$ eventually, a.s.

$N \subseteq F_n^\varepsilon$ *eventually*. Cover N by balls $B(z_1, \varepsilon/2), \dots, B(z_k, \varepsilon/2)$ with each $z_i \in N$. Then

$$\begin{aligned} P(N \not\subseteq F_n^\varepsilon) &\le P\left(F_n \text{ does not intersect each ball } B\left(z_i, \frac{\varepsilon}{2}\right)\right) \\ &\le \sum_{i=1}^k P\left(F_n \text{ disjoint from } B\left(z_i, \frac{\varepsilon}{2}\right)\right) \\ &\le \sum_{i=1}^k P\left(\frac{(X_1, X_2)}{\gamma_n} \notin B\left(z_i, \frac{\varepsilon}{2}\right)\right)^n \\ &\le \sum_{i=1}^k \exp\left(-nP\left(\frac{(X_1, X_2)}{\gamma_n} \in B\left(z_i, \frac{\varepsilon}{2}\right)\right)\right). \end{aligned}$$

Now by our assumption on I, we have $\inf_{B(z_i, \varepsilon/2)} J < 1\ \forall i$, so choose r such that $\max_i \inf_{B(z_i, \varepsilon/2)} J < r < 1$. By Lemma 4.2,

$$\liminf_{\delta \to 0} \delta \ln P\left(\frac{(X_1, X_2)}{f(\delta)} \in B\left(z_i, \frac{\varepsilon}{2}\right)\right) \ge -\inf_{B(z_i, \varepsilon/2)} J > -r \qquad \forall i.$$

Hence,

$$P((X_1, X_2)/f(\delta) \in B(z_i, \varepsilon/2)) \ge \exp(-r/\delta) \qquad \forall i, \text{ for small enough } \delta,$$

and so

$$P((X_1, X_2)/\gamma_n \in B(z_i, \varepsilon/2)) \ge n^{-r} \qquad \forall i, \text{ for large enough } n.$$

We thus have

$$\sum_{n=1}^\infty P(N \not\subseteq F_n^\varepsilon) \le \sum_{n=1}^\infty k \exp(-n^{1-r}) < \infty$$

and so the result follows.

The assertion that $N \neq B(K,0)$ follows from the assumption that $\exists x \neq 0$ with $I(x) \leq \frac{1}{2}$ [since points in $B(K,0)$ can have at most one nonzero component]. □

The next major step in the proof of Theorem 1 is to show that the limiting set of scaled samples for a product of two identical measures can only be a "ball" $B(K, p)$. We first prove some lemmas

LEMMA 4.3. *If* $\{X^{(j)}/\gamma_n\}_{j=1}^n \to_p K$, *then* $\{\gamma_n\}$ *is slowly varying* (*i.e.*, $\gamma_{\lfloor rn \rfloor}/\gamma_n \to 1$ *for all* $r > 0$).

PROOF. It is enough to show the required limit for any positive integer r. Suppose $\exists n_k \to \infty$ with $\gamma_{rn_k}/\gamma_{n_k} \to c$, where $c \neq 1$; necessarily $c > 1$. Let $F_n = \{X^{(j)}/\gamma_n\}_{j=1}^n$ and $G_k = \{X^{(j)}/\gamma_{n_k}\}_{j=1}^{rn_k}$. We see that the Hausdorff distance $h(c^{-1}G_k, F_{rn_k}) \to 0$ as $k \to \infty$. Hence $G_k \to_p cK$ as $k \to \infty$. Choose $\varepsilon > 0$ such that $cK \not\subseteq K^{2\varepsilon}$. Then $P(G_k \subseteq K^{\varepsilon}) \to 0$, but since G_k consists of r independent copies of F_{n_k}, $P(G_k \subseteq K^{\varepsilon}) = P(F_{n_k} \subseteq K^{\varepsilon})^r \to 1$, a contradiction. □

REMARK. We used above the useful observation that if $F_n \to_p K$, then (i) $\forall x \in K$ and neighborhood A of x, we have $P(F_n \cap A \neq \varnothing) \to 1$ and (ii) $\forall x \notin K$ there is a neighborhood A of x such that $P(F_n \cap A \neq \varnothing) \to 0$. We will slightly strengthen this point in a moment.

LEMMA 4.4. *Let* $F_n \to_p K$, *as in Lemma* 4.3. *Then K is star-shaped.*

PROOF. Let $x \in K$ and $0 < t < 1$. For $\varepsilon > 0$ we have $P(F_n \cap B(x, \varepsilon/2) \neq \varnothing) \to 1$ as $n \to \infty$. However, for ε small enough,

$$\{F_n \cap B(x, \varepsilon/2) \neq \varnothing\} \subseteq \{F_{m(n)} \cap B(tx, \varepsilon) \neq \varnothing\},$$

where $m(n)$ is chosen so that $|\gamma_n/\gamma_{m(n)} - t| < (\varepsilon/3)\|x\|$ (this is possible by the slow variation of γ_n). Hence $P(F_{m(n)} \cap B(tx, \varepsilon) \neq \varnothing) \nrightarrow 0$ as $n \to \infty$. Since ε is arbitrary, $tx \in K$. □

DEFINITION. A *semicone* is a subset S of E such that $tS \subseteq S$ $\forall t \geq 1$.

REMARK. Now that we know K is star-shaped, we can strengthen the previous remark to: if $x \notin K$, there is a semicone neighborhood A of x such that $P(F_n \cap A \neq \varnothing) \to 0$.

LEMMA 4.5. (a) *Let* $\{X^{(j)}/\gamma_n\}_{j=1}^n \to_p K$ *as in Lemma* 4.3. *Then* (i) $\forall x \in K$ *and every neighborhood A of x*, $nP(X/\gamma_n \in A) \to \infty$ *and* (ii) $\forall x \notin K$ *and every semicone neighborhood A of x*, $nP(X/\gamma_n \in A) \to 0$.

(b) *Let* $\{(X_1^{(j)}, X_2^{(j)}/\gamma_n\}_{j=1}^n \to_p N$ *in* E^2. *Then* (iii) $\forall (x, y) \in N$ *and every pair of neighborhoods A, B of x, y, respectively*, $nP(X/\gamma_n \in A)P(X/\gamma_n \in B) \to \infty$ *and* (iv) $\forall (x, Y) \notin N$ *and every semicone neighborhoods A, B of x, y, respectively*, $nP(X/\gamma_n \in A)P(X/\gamma_n \in B) \to 0$.

REMARK. The quantities in the conclusions of (a) and (b) are the expected number of points from the scaled sample in A and $A \times B$, respectively.

PROOF OF LEMMA 4.5(i)–(iii). Use the remark following Lemma 4.4 together with the fact that if $0 \le x_n \le 1$, then

$$(1 - x_n)^n \to 0 \Rightarrow nx_n \to \infty$$

and

$$(1 - x_n)^n \to 1 \Rightarrow nx_n \to 0.$$

[For the first implication, note that if $nx_n \le M$ for infinitely many n, then $\limsup_n (1 - x_n)^n \ge \liminf_n (1 - M/n)^n = e^{-M}$. For the second, take logarithms.] We show (i) and (ii) by letting $x_n = P(X/\gamma_n \in A)$, and we show (iii) by letting $x_n = P(X/\gamma_n \in A)P(X/\gamma_n \in B)$. This would also show (iv) if A and B were not required to be semicones. □

LEMMA 4.6. *Let* $\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \to_p N$. *If* $(x, y) \in N$, *then* $(\alpha x, \beta y) \in N$ *for all* $\alpha, \beta \in [0, 1]$.

PROOF. We show the result for $\beta = 1$. This suffices since N is star-shaped and invariant under exchanging the two coordinates.

Let $\varepsilon > 0$. We have

$$n\mu(\gamma_n B(x, \varepsilon/2))\mu(\gamma_n B(y, \varepsilon)) \to \infty.$$

It will be enough to show

$$n\mu(\gamma_n B(\alpha x, \varepsilon))\mu(\gamma_n B(y, \varepsilon)) \nrightarrow 0,$$

which we will achieve by showing that $\mu(\gamma_n B(\alpha x, \varepsilon)) \ge \frac{1}{2}\mu(\gamma_n B(x, \varepsilon/2))$ for infinitely many n. Were this not the case, $2\mu(\gamma_n B(\alpha x, \varepsilon)) < \mu(\gamma_n B(x, \varepsilon/2))$ would hold for all large enough n. As in (4.4), for all large enough n there exists $m(n)$ with $|\gamma_n/\gamma_{m(n)} - \alpha| < \varepsilon/3\|x\|$, giving $\mu(\gamma_n B(x, \varepsilon/2)) \le \mu(\gamma_{m(n)} B(\alpha x, \varepsilon))$. These inequalities yield $2\mu(\gamma_n B(\alpha x, \varepsilon)) < \mu(\gamma_{m(n)} B(\alpha x, \varepsilon))$ for all large enough n, which quickly gives, for some n and all k,

$$\mu\Big(\gamma_{\underbrace{m(m(\cdots m(n)\cdots))}_{k}} B(\alpha x, \varepsilon)\Big) > 2^k \mu(\gamma_n B(\alpha x, \varepsilon)),$$

a contradiction since all probabilities are bounded by 1. □

PROOF OF LEMMA 4.5(iv). Let $(x, y) \notin N$. By Lemma 4.6, $C = \{(sx, ty): s, t \ge 1\}$ is disjoint from N. Define the semicone neighborhoods $A_\varepsilon, B_\varepsilon$ of x, y, respectively, by $A_\varepsilon = \bigcup_{s \ge 1} sB[x, \varepsilon]$ and $B_\varepsilon = \bigcup_{t \ge 1} tB[y, \varepsilon]$. Then $\bigcap_{\varepsilon > 0} A_\varepsilon \times B_\varepsilon = C$. By the finite intersection property of the compact set N, some $A_{2\varepsilon} \times B_{2\varepsilon}$ is disjoint from N, making $A_\varepsilon \times B_\varepsilon$ disjoint from N^ε. Hence $P(\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \cap A_\varepsilon \times B_\varepsilon \ne \emptyset) \to 0$ and so Lemma 4.5(iv) obtains by the same argument used for the other parts of Lemma 4.5. □

LEMMA 4.7. *Suppose* $\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \to_p N$.

(a) *If* $(x, x) \notin N$ *and* $(y, y) \notin N$, *then* $(x, y) \notin N$.
(b) *If* $(x, x) \in N$ *and* $(y, y) \in N$, *then* $(x, y) \in N$.

REMARK. This is a generalization of Lemma 5 in [4].

PROOF OF LEMMA 4.7. (a) By Lemma 4.5, there are semicone neighborhoods A, B of x, y, respectively, such that $nP(X/\gamma_n \in A)^2 \to 0$ and $nP(X/\gamma_n \in B)^2 \to 0$. Hence

$$nP(X/\gamma_n \in A)P(X/\gamma_n \in B) = \left(nP(X/\gamma_n \in A)^2 \cdot nP(X/\gamma_n \in B)^2\right)^{1/2} \to 0$$

and so $(x, y) \notin N$. (b) is similar. □

LEMMA 4.8. (a) *Let* $\gamma_n \uparrow \infty$ *and suppose there exists* $p \in (0, \infty)$ *be such that* $\liminf_n \gamma_{n^2}/\gamma_n > 2^{1/p}$. *Then there exists* $n_k \to \infty$ *such that for any* $r_0 \in (0,1)$ $\exists k_0 \in \mathbb{N}$,

$$\gamma_{\lfloor n_k^r \rfloor} \le r^{1/p}\gamma_{n_k} \quad \text{for } r \in [r_0, 1] \text{ and } k \ge k_0.$$

(b) *The same holds if the* lim inf *condition is replaced by* $\limsup_n \gamma_{n^2}/\gamma_n < 2^{1/p}$ *and the inequality in the conclusion is reversed.*

PROOF. We prove (a); (b) is similar. Let $f(x) = \gamma_{\lfloor \exp(x) \rfloor}$. Then f is nondecreasing and $\liminf_{x\to\infty} f(2x)/f(x) > 2^{1/p}$. Pick q so that $\liminf_{x\to\infty} f(2x)/f(x) > 2^{1/q} > 2^{1/p}$ and pick x_0 so that $f(2x) > 2^{1/q}f(x)$ when $x \ge x_0$. Let $g(x) = f(x)x^{-1/q}$, so $g(2x) \ge g(x)$ for $x \ge x_0$. For each $k \in \mathbb{N}$, choose x_k in the interval $[\ln k, 2\ln k)$ to maximize g within that interval. Since g decreases on each interval $[\ln j, \ln(j+1))$, the maximum will be attained and, in fact, $x_k = \ln n_k$ for some integer n_k. For $r_0 \in (0,1)$, take $k_0 \in \mathbb{N}$ such that $r_0 \ln k_0 \ge x_0$. Then if $r \in [r_0, 1]$ and $k \ge k_0$, we can find a nonnegative integer m such that $r2^m x_k \in [\ln k, 2\ln k)$. This gives $g(rx_k) \le g(r2^m x_k) \le g(x_k)$, which is $\gamma_{\lfloor n_k^r \rfloor} \le r^{1/q}\gamma_{n_k}$. □

PROOF OF THEOREM 1(ii). We are to show that if $\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \to_p N$, then $N = B(K, p)$ for some $K \subset E$ and $p \in [0, \infty]$. In fact, K must be the projection of N onto one copy of E.

By projecting it is clear that $\{X^{(j)}/\gamma_n\}_{j=1}^n \to_p K$. Also, $N \subset K^2$. Our first goal will be to show that

$$\lim_n \frac{\gamma_{n^2}}{\gamma_n} = \sup\{t : tx \in K, \text{for some } x \text{ with } (x, x) \notin N\}$$
$$= \inf\{t : tx \notin K, \text{for some } x \text{ with } (x, x) \in N\}.$$

Refer to the above sup and inf as I_1 and I_2, respectively. By inspection, $I_1 \ge I_2$. Suppose $(x, x) \notin N$ and $tx \in K$. By Lemma 4.5(iv) there is a semicone neighborhood A of x such that $nP(X/\gamma_n \in A)^2 \to 0$. Replacing n by n^2

yields $n^2P(X/\gamma_{n^2}\in A)^2\to 0$, and so $nP(X/\gamma_{n^2}\in A)\to 0$. However, by Lemma 4.5(i), $nP(X/\gamma_n\in tA)\to\infty$, that is, $nP(X\in t\gamma_n A)\to\infty$. Thus $P(X\in t\gamma_n A)\ge P(X\in\gamma_{n^2}A)$ eventually, and so $\gamma_{n^2}\ge t\gamma_n$ eventually (since A is a semicone). Hence $\liminf_n \gamma_{n^2}/\gamma_n\ge t\ge I_1$. Now suppose $(x,x)\in N$ and $tx\notin K$. Take A such that $nP(X/\gamma_n\in tA)\to 0$ [using Lemma 4.5(ii)]. We also have [by Lemma 4.5(iii)] that $nP(X/\gamma_n\in A)^2\to\infty$. Arguing as before we find that $\gamma_{n^2}\le t\gamma_n$ eventually. We can now write

$$\liminf_n \frac{\gamma_{n^2}}{\gamma_n}\ge I_1\ge I_2\ge \limsup_n \frac{\gamma_{n^2}}{\gamma_n},$$

showing the result.

If $\lim_n \gamma_{n^2}/\gamma_n=\infty$, then (considering the expression for I_2 as an inf) we have $(x,x)\in N\Rightarrow tx\notin K$ for any $t>0\Rightarrow x=0$ (since K is bounded). Making use of Lemma 4.7 gives that $N=B(K,0)$.

If $\lim_n \gamma_{n^2}/\gamma_n=1$, then (considering the expression for I_1 as a sup) we have $d_K(x)<1\Rightarrow(x,x)\in N$. Making use of Lemma 4.7 and noting that N is closed gives that

$$N\supseteq \mathrm{cl}\{(x,y)\colon d_K(x)<1 \text{ and } d_K(y)<1\}.$$

Hence $N=K^2$.

For the rest of the proof, then, we assume $\lim_n \gamma_{n^2}/\gamma_n\in(1,\infty)$. Let $p\in(0,\infty)$ be such that $\lim_n \gamma_{n^2}/\gamma_n=2^{1/p}$. To complete the proof we show that $N=\{(x,y)\colon d_K(x)^p+d_K(y)^p\le 1\}$. First, let (x,y) be such that $d_K(x)^p+d_K(y)^p>1$. Take $s,t\in(0,1)$ and $p'>p$ such that $s^{p'}+t^{p'}>1$, $x/s\notin K$ and $y/t\notin K$. Choose semicone neighborhoods A,B of x/s, y/t, respectively, with $n\mu(\gamma_n A)\to 0$ and $n\mu(\gamma_n B)\to 0$. By Lemma 4.8(a), take $n_k\to\infty$ such that $\gamma(\lfloor n_k^{s^{p'}}\rfloor)\le s\gamma_{n_k}$ and $\gamma(\lfloor n_k^{t^{p'}}\rfloor)\le t\gamma_{n_k}$ for all k. Then

$$\begin{aligned} n_k\,\mu(\gamma_{n_k}sA)(\gamma_{n_k}tB) &\le n_k\,\mu\big(\gamma\big(\lfloor n_k^{s^{p'}}\rfloor A\big)\big)\mu\big(\gamma\big(\lfloor n_k^{t^{p'}}\rfloor B\big)\big)\\ &\le \underbrace{\big(\lfloor n_k^{s^{p'}}\rfloor\,\mu\big(\gamma\big(\lfloor n_k^{s^{p'}}\rfloor A\big)\big)\big)}_{\to 0}\underbrace{\big(\lfloor n_k^{t^{p'}}\rfloor\,\mu\big(\gamma\big(\lfloor n_k^{t^{p'}}\rfloor B\big)\big)\big)}_{\to 0}\end{aligned}$$

and we conclude $(x,y)\notin N$.

Second, let (x,y) be such that $d_K(x)^p+d_K(y)^p<1$. Take $s,t\in(0,1)$ and $p'<p$ such that $x/s\in K$, $y/t\in K$ and $s^{p'}+t^{p'}<1$. Let A,B be any semicone neighborhoods of x/s, y/t, respectively. Then $n\mu(\gamma_n A)\to\infty$ and $n\mu(\gamma_n B)\to\infty$. By Lemma 4.8(b), take $n_k\to\infty$ with $\gamma(\lfloor n_k^{s^{p'}}\rfloor)\ge s\gamma_{n_k}$ and $\gamma(\lfloor n_k^{t^{p'}}\rfloor)\ge t\gamma_{n_k}$ for all k. Then

$$\begin{aligned} n_k\,\mu(\gamma_{n_k}sA)\mu(\gamma_{n_k}tB) &\ge n_k\,\mu\big(\gamma\big(\lfloor n_k^{s^{p'}}\rfloor A\big)\big)\mu\big(\gamma\big(\lfloor n_k^{t^{p'}}\rfloor B\big)\big)\\ &\ge \underbrace{\big(\lfloor n_k^{s^{p'}}\rfloor\,\mu\big(\gamma\big(\lfloor n_k^{s^{p'}}\rfloor A\big)\big)\big)}_{\to \infty}\underbrace{\big(\lfloor n_k^{t^{p'}}\rfloor\,\mu\big(\gamma\big(\lfloor n_k^{t^{p'}}\rfloor B\big)\big)\big)}_{\to \infty}\end{aligned}$$

and we conclude $(x,y)\in N$. Since N is closed, it follows that N contains all (x,y) with $d_K(x)^p+d_K(y)^p\le 1$. □

PROOF OF (c) ⇒ (d) IN THEOREM 1(i). Suppose $0 < p \le \infty$ and $\{(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \to_p B(K, p)$. Choose (appealing to the Hahn–Banach theorem) a bounded linear map ϕ: $E \to \mathbb{R}$ such that $\phi(K) \subseteq [-1,1]$ and $1 \in \phi(K)$. Then

$$\{\phi(X^{(j)})/\gamma_n\}_{j=1}^n \to_p \phi(K)$$

[since the map $C \mapsto \phi(C)$, taking the compact subsets of E to those of $\mathbb{R}$, is continuous] and similarly,

$$\{\Phi(X_1^{(j)}, X_2^{(j)})/\gamma_n\}_{j=1}^n \to_p \Phi(B(K, p)),$$

where $\Phi(x, y) = (\phi(x), \phi(y))$. Letting ϕ_+ denote $\phi \vee 0$, we then have

$$\{\phi_+(X^{(j)})/\gamma_n\}_{j=1}^n \to_p [0,1]$$

and

$$\{(\phi_+(X_1^{(j)}), \phi_+(X_2^{(j)}))/\gamma_n\}_{j=1}^n \to_p \Phi(B(K, p)) \cap [0,\infty)^2 = B([0,1], p).$$

By the result of [8] mentioned in the first section, we note that we must have $\gamma_n \sim L(n)$, where $L(n) = \inf\{y: P(\phi(X) \le y) \ge 1 - 1/n\}$. By the result of [4], we have $L(n^r)/L(n) \to r^{1/p}$ for all $r \in (0,1)$, and hence for all $r > 0$. Hence

$$\frac{\gamma_{\lfloor n^r \rfloor}}{\gamma_n} = \underbrace{\frac{\gamma_{\lfloor n^r \rfloor}}{L(\lfloor n^r \rfloor)}}_{\to 1} \cdot \underbrace{\frac{L(\lfloor n^r \rfloor)}{L(n^r)}}_{\to 1} \cdot \underbrace{\frac{L(n^r)}{L(n)}}_{\to r^{1/p}} \cdot \underbrace{\frac{L(n)}{\gamma_n}}_{\to 1}$$

[the second limit on the right-hand side holds by Lemma 4.3, since $\{L(n)\}$ is also a valid normalising sequence for scaled samples of $\phi_+(X)$]. □

PROOF OF (d) ⇒ (a) IN THEOREM 1(i). Suppose $\{X^{(j)}/\gamma_n\}_{j=1}^n \to_p K$ and that the normalizers $\{\gamma_n\}$ satisfy $\gamma_{\lfloor n^r \rfloor}/\gamma_n \to r^{1/p}$ $\forall r > 0$. (If $p = \infty$, the limit is 1 for all r.) We wish to show that the large deviation principle of (a) holds.

It is straightforward to see that I is lower semicontinuous and that it has compact level sets. (Recall that the limit set in a scaled-sample convergence is always compact.)

Large deviation upper bound:

$$\limsup_{\varepsilon \to 0} \varepsilon \ln P(X/f(\varepsilon) \in A) \le -\inf_{\mathrm{cl}\, A} I \quad \text{for measurable } A.$$

It is enough to show this for closed A. Also, we may assume $\inf_A I > 0$ (otherwise there is nothing to prove). Take t with $0 < t < \inf_A d_K$. If $p = \infty$, we further require that $t > 1$ (which is possible since A is disjoint from the star-shaped K and d_K has compact level sets, so $\inf_A d_K$ is attained). Then $t^{-1}A$ is closed and disjoint from K. By compactness of K we have $\delta > 0$ such that $t^{-1}A \subseteq (K^\delta)^c$. Now $nP(X \in \gamma_n(K^\delta)^c) \to 0$ [since $P(\{X^{(j)}/\gamma_n\}_{j=1}^n \cap (K^\delta)^c \ne \emptyset) \to 0$; see Lemma 4.5] and $(K^\delta)^c$ is a semicone. Hence $nP(X \in \gamma_{\lfloor n^r \rfloor} t(K^\delta)^c) \to 0$ for any $r > t^{-p}$ (or if $p = \infty$, we only need $r > 0$). We then see that

$$\exp\left(\frac{1}{r\varepsilon}\right) P\left(X \in f(\varepsilon) t (K^\delta)^c\right) \to 0 \quad \text{as } \varepsilon \to 0$$

[this is true along the sequence $\varepsilon_n = 1/\ln\lfloor n^r \rfloor$, and note $P(X \in f(\varepsilon)t(K^\delta)^c)$ is a monotone function of ε]. Thus

$$\limsup_{\varepsilon} \varepsilon \ln P(X \in f(\varepsilon)A) \le -\frac{1}{r}.$$

Hence the result since r and t were arbitrary.

Large deviation lower bound:

$$\liminf_{\varepsilon \to 0} \varepsilon \ln P(X/f(\varepsilon) \in A) \ge -\inf_{A^\circ} I \quad \text{for measurable } A.$$

It is enough to show this for A open and nonempty.

For $0 < p < \infty$, take x, δ such that $0 \ne B(x, 3\delta) \subseteq A$. Let $u = 1 - 2\delta/\|x\|$. Take v, s such that $0 < ud_K(x) \le v < s < d_K(x)$. Take A_1 to be the semicone generated by some ball centered at x with radius less than δ, such that $\inf_{\mathrm{cl}\, A_1} d_K > s$ (this is possible since d_K is lower semicontinuous). Then

$$\limsup_{\varepsilon \to 0} \varepsilon \ln P(X \in f(\varepsilon)A_1) \le -\inf_{\mathrm{cl}\, A_1} d_K^p$$

so that if $r_1 > s^{-p}$, then $P(X \in f(\varepsilon)A_1) \le \exp(-1/r_1\varepsilon)$ for all small enough ε. Note also that $uA_1 - A_1 \subseteq A$. In addition, $v^{-1}ux \in K$, so $nP(X \in \gamma_n v^{-1}uA_1) \to \infty$. This gives

$$nP(X \in \gamma_{\lfloor n^{r_2} \rfloor} uA_1) \to \infty \quad \text{whenever } r_2^{1/p} < v^{-1}.$$

For such r_2, then, $\exp(1/r_2\varepsilon)P(X \in f(\varepsilon)uA_1) \to \infty$ as $\varepsilon \to 0$ (see the upper bound argument), so $P(X \in f(\varepsilon)uA_1) \ge \exp(-1/r_2\varepsilon)$ for all small enough ε. Choose r_1, r_2 such that $s^{-1} < r_1^{1/p} < r_2^{1/p} < v^{-1}$. Then we have

$$\begin{aligned} P(X \in f(\varepsilon)A) &\ge P(X \in f(\varepsilon)(uA_1 - A_1)) \\ &\ge \exp\left(\frac{-1}{r_2\varepsilon}\right) - \exp\left(\frac{-1}{r_1\varepsilon}\right) \end{aligned}$$

for all small ε. From this we obtain

$$\liminf_{\varepsilon \to 0} \varepsilon \ln P(X \in f(\varepsilon)A) \ge \frac{-1}{r_2} \ge -s^p \ge -d_K(x)^p.$$

Since $x \in A$ was arbitrary, the result follows.

For $p = \infty$, the proof is similar. If $d_K(y) > 1\ \forall y \in A$, there is nothing to prove. If some $y \in A$ has $d_K(y) \le 1$, then note that the result for $A/d_K(y)$ implies that for A. [This is true at least if f is continuous and strictly decreasing, since then we can write $f(\varepsilon)/d_K(y) = f(e(\varepsilon))$ for some $e(\varepsilon) < \varepsilon$. However, to show the result for any f we have only to show it along the sequence $\varepsilon_n = (\ln n)^{-1}$.] Thus we can assume $\exists y \in A$ with $d_K(y) = 1$. Take y, δ such that $0 \notin B(y, \delta) \subseteq A$ and $d_K(y) = 1$. Let $x = y(1 + \delta/\|y\|)$ and $u = 1 - 2\delta/\|x\|$. Take v, s such that $0 < ud_K(x) \le v < 1 < s < d_K(x)$. Take A_1 to be a semicone generated by a ball with center x and radius less than δ. Then $\varepsilon \ln P(X \in f(\varepsilon)A_1) \to -\infty$ as $\varepsilon \to 0$. The rest is similar to the case $0 < p < \infty$. Note that when we choose r_1, r_2 the only requirement is $r_1 < r_2$.

□

PROOF OF THE REMAINING PARTS OF THEOREM 1. Part (i). (b) ⇒ (c) is obvious; (c) ⇒ (d) ⇒ (a) has already been shown. We thus have (a) ⇒ (b) left to prove.

If (a) holds, then we have the hypothesis of part (iii). By Proposition 4.1 we have (b) holding for $m = 2$, but now note that if the condition of (b) holds for any particular m, then (c) and (d) follow. Hence (d) holds with μ replaced by $\mu \times \mu$. We thus have (a) for this measure, and hence the hypothesis of (iii) for it. Using Proposition 4.1, we then obtain

$$\left\{ \frac{\left(X_1^{(j)}, X_2^{(j)}, X_3^{(j)}, X_4^{(j)} \right)}{\gamma_n} \right\}_{j=1}^{n} \to B(B(K,p),p) = B_4(K,p).$$

Iterating this argument gives us (b) when m is any power of 2; hence for all m (by projecting onto smaller product spaces).

Part (ii). Is already shown.

Part (iii). Apply Proposition 4.1. As we now know, the limiting set N must be a ball $B(K, p)$; as we noted in Proposition 4.1, it cannot be $B(K, 0)$. □

5. Some further remarks on $\{\gamma_n\}$. The asymptotic behaviour of the normalizers γ_n for a distribution showing regular scaled-sample convergence of index p can be fairly closely described.

LEMMA 5.1. *Let $0 < p \le \infty$ and let $\gamma_n \uparrow \infty$ be a positive sequence with the property that $\gamma_{\lfloor n^r \rfloor}/\gamma_n \to r^{1/p}$ for all $r > 0$.*

(i) *If $0 < p' < p$, then $\gamma_n = o((\ln n)^{1/p'})$.*
(ii) *If $p < p'' < \infty$, then $(\ln n)^{1/p''} = o(\gamma_n)$.*

PROOF. Let $f(x) = \gamma_{\lfloor \exp(x) \rfloor}$. Then f is regularly varying at ∞ of order p^{-1} (slowing varying if $p = \infty$) since, if $r \in (0, \infty)$, we have for any $s \in (r, \infty)$ and large x that $\lfloor x \rfloor^r \le x^r \le \lfloor x \rfloor^s$ and so, as $x \to \infty$,

$$r^{1/p} \leftarrow \frac{\gamma_{\lfloor \lfloor x \rfloor^r \rfloor}}{\gamma_{\lfloor x \rfloor}} = \frac{f(r \ln \lfloor x \rfloor)}{f(\ln x)} \le \frac{f(r \ln x)}{f(\ln x)} \le \frac{f(s \ln \lfloor x \rfloor)}{f(\ln x)} = \frac{\gamma_{\lfloor \lfloor x \rfloor^s \rfloor}}{\gamma_{\lfloor x \rfloor}} \to s^{1/p}.$$

The result then follows by Potter's theorem on regularly varying functions ([2], Theorem 1.5.6). □

REFERENCES

[1] DE ACOSTA, A. and KUELBS, J. (1983). Limit theorems for moving averages of independent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 67–123.

[2] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. 1987. *Regular Variation.* Cambridge Univ. Press.

[3] DAVIS, R., MULROW, E. and RESNICK, S. (1988). Almost sure limit sets of random samples in $\mathbb{R}^d$. *Adv. in Appl. Probab.* **20** 573–599.

[4] FISHER, L. (1969). Limiting sets and convex hulls of samples from product measures. *Ann. Math. Statist.* **40** 1824–1832.

[5] GNEDENKO, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44** 423–453.

[6] KINOSHITA, K. and RESNICK, S. (1991). Convergence of scaled random samples in $\mathbb{R}^d$. *Ann. Probab.* **19** 1640–1663.

[7] LYNCH, J. and SETHURAMAN, J. (1987). Large deviations for processes with independent increments. *Ann. Probab.* **15** 610–627.

[8] RESNICK, S. and TOMKINS, R. J. (1973). Almost sure stability of maxima. *J. Appl. Probab.* **10** 387–401.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN—MADISON
MADISON, WISCONSIN 53706
E-MAIL: pritchar@math.wisc.edu