

Computing the distribution of the sum of dependent random variables via overlapping hypercubes

Marcello Galeotti*

Department of Statistics, Informatics and Applications, University of Florence

Abstract

The original motivation of this work comes from a classic problem in finance and insurance: that of computing the value-at-risk (VaR) of a portfolio of dependent risky positions, i.e. the quantile at a certain level of confidence of the *loss* distribution. In fact, it is difficult to overestimate the importance of the concept of VaR in modern finance and insurance: it has been recommended, although with several warnings, as a measure of risk and the basis for capital requirement determination both by the guidelines of international committees (such as Basel 2 and 3, Solvency 2 etc.) and the internal models adopted by major banks and insurance companies. However the actual computation of the VaR of a portfolio constituted by several dependent risky assets is often a hard practical and theoretical task. To this purpose here we prove the convergence of a geometric algorithm (alternative to Monte Carlo and quasi Monte Carlo methods) for computing the value-at-risk of a portfolio of any dimension, i.e. the distribution of the sum of its components, which can exhibit any dependence structure. Moreover our result has a relevant measure-theoretical meaning. What we prove, in fact, is that the H -measure of a d -dimensional simplex (for any $d \geq 2$ and any absolutely continuous with respect to Lebesgue measure H) can be approximated by convergent algebraic sums of H -measures of hypercubes (obtained through a *self-similar* construction).

Keywords: finance; applied probability; algorithm convergence; measure theory
JEL: :C6 (Mathematical and Quantitative Methods)

1. Introduction and preliminaries

1.1. Statement of the problem

A portfolio \mathcal{P} of risky positions, in finance and insurance models, can be represented as a vector of random *losses*, i.e. $\mathcal{P} = (X_1, \dots, X_d)$ (to be precise, each X_i represents the random *profit-and-loss* result of an investment, within a given horizon, a negative value corresponding to a profit, a positive value to a loss).

Then the *total loss* of the portfolio is given by the random variable $X = X_1 + \dots + X_d$. As a consequence, the value-at-risk (VaR) of the portfolio at a confidence level $\alpha \in (0, 1)$ (e.g. $\alpha = 0.95, 0.99$ etc.) is defined as the α -quantile of X , i.e.

$$\text{VaR}_\alpha(X) = \inf \{l : \text{prob}(X > l) \leq 1 - \alpha\} = \inf \{l : \text{prob}(X \leq l) \geq \alpha\}$$

*Corresponding author: Department of Statistics, Informatics and Applications, University of Florence, Via delle Pandette 9, Ed. D6 2.15, 50127 Firenze, Italy.

Email address: marcello.galeotti@unifi.it (Marcello Galeotti)

Hence $VaR_\alpha(X)$ has been interpreted as the reserve (*regulatory*) capital that a bank or an insurance company must hold when it undertakes the risk represented by the portfolio \mathcal{P} in order to avoid with probability at least α a negative budget result at the end of the given horizon. Although failing, in some circumstances, to be *sub-additive* (thus, according to a most quoted definition Artzner, Delbaen, Eber and Heath, 1999, *coherent*) the value-at-risk has been recommended, although with several warnings, as a measure of risk and the basic criterion for capital adequacy calculation both by the guidelines of international committees (such as Basel 2 and 3, Solvency 2 etc.) and the internal models adopted by major banks and insurance companies. However the actual computation of the VaR of a portfolio constituted by several dependent risky assets is often a hard practical and theoretical task

To this purpose, in an article published in 2011 Arbenz, Embrechts and Puccetti have proposed a new algorithm, called AEP after the names of the authors, to compute numerically the distribution function of the sum of d dependent, non negative random variables with given absolutely continuous joint distribution. Briefly, given a joint distribution H , the algorithm approximates the H -measure of a simplex (hence the distribution of the sum of the random variables) by an algebraic sum of H -measures of hypercubes (which can be easily calculated). Besides providing the motivations for the algorithm (in particular as far as the calculation of value-at-risk, in finance and insurance, is concerned), the authors have underlined the novelties of the AEP algorithm, with respect to more usual Monte Carlo and quasi-Monte Carlo methods Glasserman (2004). Precisely such an algorithm, first, is deterministic (hence independent from sample choice), and, secondly, it is also independent from the specific distribution H , that is from the dependence structure (i.e. copula) of the random variables. Moreover, the AEP algorithm is beautifully *self-similar*, i.e. the same algorithm is applied to each newly generated simplex: a property which will be most exploited in the following.

In front of these advantages, two open problems were detected (see Arbenz, Embrechts and Puccetti, 2011):

1. The numerical complexity of the algorithm increases, at each step, exponentially, making it hardly manageable for dimension $d > 5$.
2. In the original article Arbenz, Embrechts and Puccetti (2011) the convergence of the algorithm was proven only for dimension $d \leq 8$ (under further differentiability assumptions on the function H and for a particular choice of a pivotal parameter α).

In the present paper we solve Problem 2, proving that the AEP algorithm converges for any $d \geq 2$ and any absolutely continuous distribution H (with bounded density in a neighborhood of the simplex *diagonal*), when the above mentioned parameter α varies in a specified interval. We do not exclude that such a result may be also preliminary to a partial solution of Problem 1, for example through some efficient extrapolation of the AEP Arbenz, Embrechts and Puccetti (2012).

1.2. Notations

In the article we adopt basically the notations of Arbenz, Embrechts and Puccetti (2011).

First of all, we denote vectors in boldface, i.e. $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$, $d \geq 2$. In particular we indicate by $\mathbf{i}_0, \dots, \mathbf{i}_N$, $N = 2^d - 1$, the 2^d vectors in $\{0, 1\}^d$ (e.g. $\mathbf{i}_0 = (0, \dots, 0)$, $\mathbf{i}_N = (1, \dots, 1)$) and by $\#\mathbf{i}$ the number of 1's in the vector \mathbf{i} (e.g. $\#\mathbf{i}_0 = 0$, $\#\mathbf{i}_N = d$). Moreover we set $\lambda(\mathbf{x}) = x_1 + \dots + x_d$. As in the following we will consider vectors $\mathbf{x} \in \mathbb{R}_+^d$ (the non-negative orthant of \mathbb{R}^d), $\lambda(\mathbf{x})$ can be seen as the l^1 norm in \mathbb{R}_+^d .

Then, fixed $h > 0$ and $\mathbf{b} = (b_1, \dots, b_d)$, we define the following simplexes:

$$S(\mathbf{b}, h) := \{\mathbf{x} \in \mathbb{R}^d : \lambda(\mathbf{b}) < \lambda(\mathbf{x}) \leq \lambda(\mathbf{b}) + h, x_k - b_k > 0, k = 1, \dots, d\}$$

and

$$S(\mathbf{b}, -h) := \{\mathbf{x} \in \mathbb{R}^d : \lambda(\mathbf{b}) - h < \lambda(\mathbf{x}) \leq \lambda(\mathbf{b}), x_k - b_k \leq 0, k = 1, \dots, d\}$$

Analogously we define the hypercubes

$$Q(\mathbf{b}, h) := \{\mathbf{x} \in \mathbb{R}^d : b_k < x_k \leq b_k + h, k = 1, \dots, d\}$$

and

$$Q(\mathbf{b}, -h) := \{\mathbf{x} \in \mathbb{R}^d : b_k - h < x_k \leq b_k, k = 1, \dots, d\}$$

Clearly h is the side length of the above hypercubes, while, as our arguments will be developed in \mathbb{R}_+^d , where $\lambda(\mathbf{x})$ represents the l^1 norm, h will be called the *radius* of $S(\mathbf{b}, h)$ (and $S(\mathbf{b}, -h)$).

Then, given an absolutely continuous joint distribution $H(x_1, \dots, x_d)$ (with support in \mathbb{R}_+^d), we will denote by v_H the relative probability measure, while vol will indicate the Lebesgue measure. Hence $vol(Q(\mathbf{b}, \pm h)) = h^d$ and $vol(S(\mathbf{b}, \pm h)) = \frac{h^d}{d!}$. For sake of completeness, we define also $S(\mathbf{b}, 0) = Q(\mathbf{b}, 0) = \emptyset$. Obviously $v_H(S(\mathbf{b}, 0)) = v_H(Q(\mathbf{b}, 0)) = 0$.

1.3. The AEP algorithm

The aim of the AEP algorithm is to approximate the H -measure of a d -dimensional simplex (where H is an absolutely continuous joint distribution in \mathbb{R}^d) by an algebraic sum of H -measures of hypercubes (overlapping when $d > 2$). The reason is that the H -measure of a hypercube is easily computed. In fact, by the notations of the previous paragraph, consider $Q(\mathbf{b}, l)$, $l \leq 0$. Then, as it is well-known,

$$v_H(Q(\mathbf{b}, l)) = \sum_{k=0}^N (-1)^{\frac{d(1+sgn(l))}{2} - \#\mathbf{i}_k} H(\mathbf{b} + l\mathbf{i}_k)$$

Hence let us sum up the scenario described in Arbenz, Embrechts and Puccetti (2011). X_1, \dots, X_d are non-negative (or, what is the same after a translation, bounded from below) random variables and $H(x_1, \dots, x_d)$ is their joint absolutely continuous distribution function. Hence, H being known, the aim is to compute, for a positive s , $Prob(X_1 + \dots + X_d \leq s) = v_H(S(\mathbf{0}, s))$. In the following, having fixed s , we will consider the rescaling $x_i \rightarrow \frac{x_i}{s}$, so that our problem will be the computation of $v_H(S(\mathbf{0}, 1))$. The first step of the AEP algorithm consists in replacing $S(\mathbf{0}, 1)$ by a hypercube $Q_1^1 = Q(\mathbf{0}, \alpha)$ with $\frac{1}{d} \leq \alpha < 1$. Then, among the vertices of the hypercube different from $\mathbf{0}$, i.e. $\alpha\mathbf{i}_k$, $k = 1, \dots, N = 2^d - 1$, there are those lying in $S(\mathbf{0}, 1)$, when $\#\mathbf{i}_k \leq \frac{1}{\alpha}$, and (possibly) those lying outside the simplex, when $\#\mathbf{i}_k > \frac{1}{\alpha}$. To each such vertex corresponds a simplex given, with the previous notations, by $S_k^2 := S(\alpha\mathbf{i}_k, 1 - \alpha(\#\mathbf{i}_k))$. Pose $l_k = 1 - \alpha(\#\mathbf{i}_k)$: hence $l_k \geq 0$. It is easily calculated that

$$v_H(S(\mathbf{0}, 1)) = v_H(Q_1^1) + \sum_{k=1}^N \sigma_k^2 v_H(S_k^2) \quad (1)$$

where $\sigma_k^2 = (-1)^{\mu_k}$, $\mu_k = \#\mathbf{i}_k + 1 - d \frac{1 - sgn(l_k)}{2}$, if $l_k \geq 0$, $\sigma_k^2 = 0$ if $l_k = 0$.

$P_1 := v_H(Q_1^1)$ is the first approximation of $v_H(S(\mathbf{0}, 1))$. Then the algorithm proceeds recursively, by replacing each simplex S_k^2 of radius $|l_k|$ with a corresponding hypercube $Q_k^2 := Q(\alpha\mathbf{i}_k, (1 - \alpha(\#\mathbf{i}_k))\alpha)$ of side length $\alpha|l_k|$. Therefore

$$P_2 := P_1 + \sum_{k=1}^N \sigma_k^2 v_H(Q_k^2) \quad (2)$$

and so on (Figure 1 illustrates the simplest case $d = 2$, when the new simplexes generated, at each step, by the algorithm do not overlap). In particular, denote by S_k^{n+1} , $k = 1, \dots, N^n$, the simplexes

received as input by the algorithm at the beginning of the $(n + 1)$ – *th* iteration. Then the following recursive formula is proven in Arbenz, Embrechts and Puccetti (2011) :

$$v_H(S(\mathbf{0}, 1)) = P_n + \sum_{k=1}^{N^n} \sigma_k^{n+1} v_H(S_k^{n+1}) \quad (3)$$

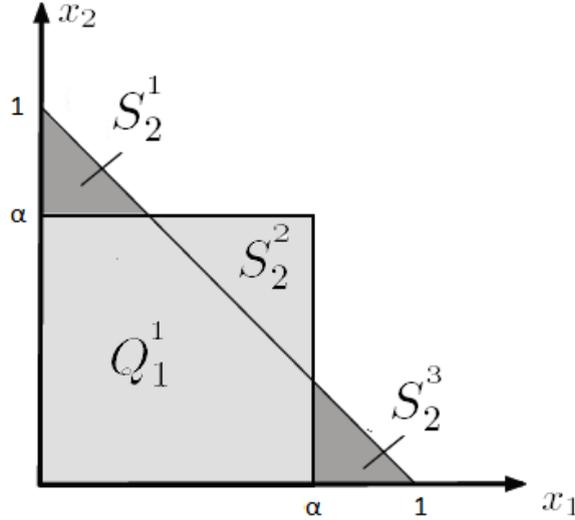


Figure 1: The AEP algorithm for $d = 2$.

1.4. Steps of the convergence proof

As we mentioned, in the original article Arbenz, Embrechts and Puccetti (2011) the convergence of the AEP was proven, when $\alpha = \frac{2}{d+1}$, for $d \leq 5$ and any absolutely continuous distribution H (with bounded density in a neighborhood of the simplex *diagonal*) and for $d \leq 8$ with further conditions (differentiability) on H . In this work, instead, we prove that the AEP algorithm converges, when $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, in any dimension d for any absolutely continuous distribution H with bounded density in a neighborhood of the simplex *diagonal*.

The proof is given through a Lemma and a Theorem. The Lemma proves that the algorithm converges for the Lebesgue measure when $d \geq 2$ and $\alpha \in \left[\frac{1}{d}, \frac{1}{\sqrt[d]{d!}}\right)$. Then the Theorem states that such a result holds for any absolutely continuous (with respect to Lebesgue one) measure as well, when $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ (observe that $\frac{2}{d+1} < \frac{1}{\sqrt[d]{d!}}$ when $d \geq 2$). The basic idea underlying the Theorem's proof is fairly simple. Suppose one can show that, at any step of the algorithm, a *corresponding* sub-simplex of S is *exactly filled up*, by summing *positive and negative* hypercubes, while in a suitably chosen strip outside the simplex *positive and negative* hypercubes exactly compensate. Then, if this way the simplex S is *geometrically* approximated, the convergence eventually follows from the assumed boundedness of the density in a neighborhood of the simplex *diagonal*. However such a proof cannot be so *direct* (e.g. merely combinatorial), due to the growing intricacy of hypercube *overlapping* when the dimension d increases. Therefore the Theorem's proof will be divided into five steps, which can

appear rather technical, since they are, precisely, designed to overcome technical difficulties, but, on the other hand, follow a *natural* path of argumentation. Below we illustrate them, in order to help the comprehension of the actual proof.

First step The scope of this step is to provide an algebraic construction which allows to directly add and subtract the hypercubes of the algorithm, rather than their *volumes* in some absolutely continuous measure. This way, grossly speaking, we can think of such hypercubes as sort of "bricks", which are "brought in" when their coefficient is +1 and "taken away" when their coefficient is -1. To this end we construct a \mathbb{Z} -module Ω , generated by the Lebesgue measurable subsets of \mathbb{R}^d . Precisely, $\Omega = \{a_1 A_1 + \dots + a_k A_k\}$, where $a_1, \dots, a_k \in \mathbb{Z}$ and A_1, \dots, A_k are Lebesgue measurable subsets of \mathbb{R}_+^d , defining in a suitable way the sum in Ω . At the n -th step of the AEP the algebraic sum of the

hypercubes Q_k^n is given by $\Pi_n = \sum_{k=1}^{\rho(n)} \sigma_k^n Q_k^n$, where $\rho(n) = \frac{N^n - 1}{N - 1}$, $N = 2^d - 1$, and $\sigma_k^n = \pm 1$ according to the algorithm rules. Hence $\Pi_n \in \Omega$. Moreover in Ω a partial ordering, denoted by the symbol \succeq , is defined. Without entering, for the moment, into details, we observe that, since any absolutely continuous measure v_H can be extended by linearity to Ω , $A \succ B$ implies $v_H(A) > v_H(B)$ and $A \simeq B$ implies $v_H(A) = v_H(B)$.

Second step We will consider, for any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, a sequence of sub-simplexes of $S(\mathbf{0}, 1)$ defined by

$$S_n = \{0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^n, x_1, \dots, x_d \geq 0, n \geq 1\}$$

Then we take $\alpha = \frac{1}{d}$ and prove that, for any $n \geq 1$, the following equivalence holds:

$$\sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap S_n) \simeq S_n \quad (4)$$

The meaning of the equivalence is, roughly speaking, that, for $\alpha = \frac{1}{d}$, the algebraic sum of the hypercubes at each n -th step of the algorithm produces an *exact filling* (with respect to any absolutely continuous measure) of the corresponding sub-simplex S_n .

The proof exploits an induction argument to show that for any $n \geq 1$

$$\sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap S_n) \succeq S_n$$

and, subsequently, the equivalence follows from the Lemma.

Third step We consider now $\alpha = \frac{1+\varepsilon}{d}$, where $0 < \varepsilon \ll \frac{1}{d-1}$, so that each hypercube generated *inside* $S(\mathbf{0}, 1)$ has only one vertex lying (*just*) outside the simplex. In fact, we take a sequence of ε_m (satisfying the above inequalities) tending to zero. Then, through combinatorial arguments and exploiting again the Lemma, an extension of the equivalence (4) is proven for any ε_m . As a matter of fact, in the end, we prove the following. Consider the strips

$$\begin{aligned} T_n &= \{0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^n\} \\ &\quad \text{and} \\ T'_n &= \left\{1 + (d\alpha - 1)(1 - \alpha)^{n-1} \leq x_1 + \dots + x_d \leq d\alpha\right\} \end{aligned}$$

Then

$$\sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap (T_n \cup T'_n)) \simeq S_n \quad (5)$$

where S_n is defined as above and $\alpha = \frac{1+\varepsilon_m}{d}$ (observe that, for $\alpha = \frac{1}{d}$, $d\alpha = 1$ and T'_n is reduced to the hyperplane $x_1 + \dots + x_d = 1$).

Fourth step Here we use the elementary property of one-variable analytic functions, which are identically zero if their zeroes have an accumulation point, in order to extend the measure equality derived from the above equivalence. Let $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ and consider (in \mathbb{R}_+^d) an analytic distribution H . Let us define, for any $n \geq 1$,

$$g_n^H(\alpha) = \sum_{k=1}^{\rho(n)} \sigma_k^n v_H(Q_k^n \cap (T_n \cup T'_n)) - v_H(S_n)$$

It is easily checked that $g_1^H(\alpha) = g_2^H(\alpha) \equiv 0$ as $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$. Moreover, when $n \geq 3$, $g_n^H(\alpha)$ has a sequence of zeroes $\alpha_m \rightarrow \frac{1}{d}$. Hence, due to the analyticity of H , $g_n^H(\alpha)$ is identically zero in an interval of analyticity $\left[\frac{1}{d}, \hat{\alpha}\right]$. A value where $g_n^H(\alpha)$ might loose analyticity corresponds to the case of some hypercubes Q_k^n crossing $T_n(\alpha)$ or $T'_n(\alpha)$, when α crosses $\hat{\alpha}$. However a rather technical argument allows to prove that $g_n^H(\alpha)$ is still zero in a right neighborhood of $\hat{\alpha}$. Hence

$$g_n^H(\alpha) \equiv 0 \tag{6}$$

for any $n \geq 1$, any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ and any analytic distribution H . But since an absolutely continuous function can be approximated as well as we want by analytic functions, (6) holds for any absolutely continuous distribution as well.

Fifth step Having proven(6), the conclusion of the Theorem appears quite close. In fact the last step consists, precisely, in proving that

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^{\rho(n)} \sigma_k^n v_H(Q_k^n) = v_H(S) \tag{7}$$

for any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ and any absolutely continuous distribution H whose density is bounded in a neighborhood of the simplex *diagonal*

$$D = \{x_1 + \dots + x_d = 1, x_1, \dots, x_d \geq 0\}$$

Such a final step is rather straightforward, even if a fairly *subtle* argument is still required.

1.5. The (almost) trivial case $d = 2$

In the case $d = 2$ we also proceed by the first step. However the squares considered at each step, when $\alpha = \frac{1}{2}$, do not overlap (i.e. their intersections have zero Lebesgue measure): in fact $2\alpha = 1$. Hence the equivalence (4) is immediately checked (see Figure 2). By self-similarity the equivalence

(5), i.e. $\sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap (T_n \cup T'_n)) \simeq S_n$, is also easily verified for any $n \geq 1$ and $\alpha \in \left[\frac{1}{2}, \frac{2}{3}\right]$ (see Figure 3).

Hence the equality $g_n^H(\alpha) \equiv 0$ follows for any absolutely continuous distribution H . The fifth step requires, finally, the general argument described in the Theorem's proof. Therefore the convergence proof when $d = 2$ still needs the assumption of density boundedness on the *triangle diagonal* but not the Lemma below on the convergence in the Lebesgue case.

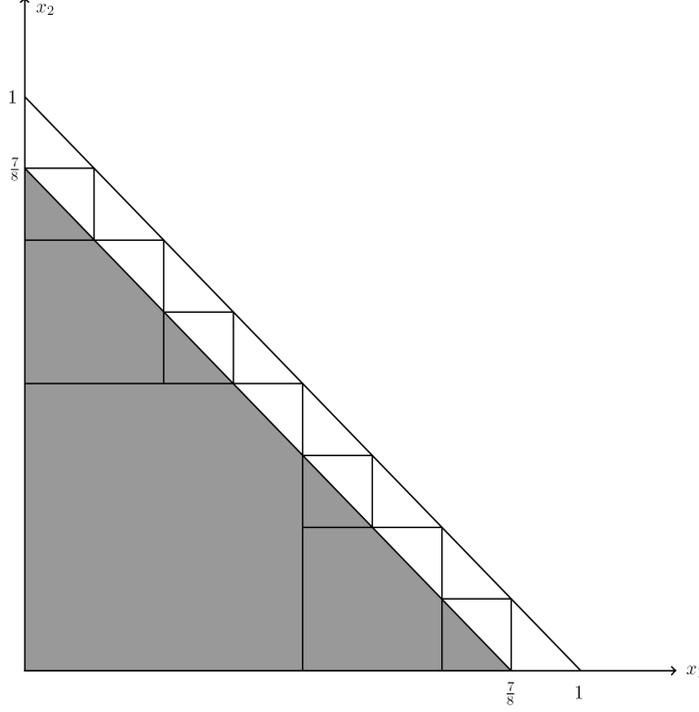


Figure 2: Convergence of the algorithm when $d=2$ and $\alpha = \frac{1}{2}$.

1.6. A useful Proposition

We end the Section by proving the following

Proposition 1. For any $d \geq 2$ and any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ the sub-simplex

$$S_n = \{0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^n, x_1, \dots, x_d \geq 0\}$$

is covered by the hypercubes of Π_n with sides $\alpha(1 - \alpha)^s$, $0 \leq s \leq n - 1$

Proof. Recall $\lambda(\mathbf{x}) = x_1 + \dots + x_d$. We prove the Proposition by induction. Clearly the property holds for $n = 1$. So, assume it holds for some $n \geq 1$ and consider $\mathbf{x} = (x_1, \dots, x_d) \in S_{n+1} - S_n$. Then $\lambda(\mathbf{x}) \leq 1 - (1 - \alpha)^{n+1}$. Take the $\max(x_1, \dots, x_d)$. To fix the ideas, assume it is x_1 . Hence there exists \bar{x}_1 , $0 < \bar{x}_1 < x_1$, such that $\lambda(\mathbf{x}) = 1 - (1 - \alpha)^n$ and, by the induction hypothesis, a point $\mathbf{y} = (y_1, \dots, y_d)$ with $y_1 + \dots + y_d = \alpha + \alpha(1 - \alpha) + \dots + \alpha(1 - \alpha)^t$, $0 \leq t < n - 1$, satisfying $y_1 \leq \bar{x}_1 \leq y_1 + \alpha(1 - \alpha)^{t+1}$, \dots , $y_d \leq x_d \leq y_d + \alpha(1 - \alpha)^{t+1}$.

Then, in case $x_1 > y_1 + \alpha(1 - \alpha)^{t+1}$, replace y_1 by $y_1 + \alpha(1 - \alpha)^{t+1}$ and proceed recursively, considering the $\max(x_1 - y_1, \dots, x_d - y_d)$. Therefore, since $\lambda(\mathbf{x}) \leq 1 - (1 - \alpha)^{n+1}$, we will eventually find $\mathbf{y} = (y_1, \dots, y_d)$ with $y_1 + \dots + y_d = \alpha + \alpha(1 - \alpha) + \dots + \alpha(1 - \alpha)^{t+r}$, $t + r \leq n - 1$, such that $y_k \leq x_k \leq y_k + \alpha(1 - \alpha)^{t+r+1}$ for $k = 1, \dots, d$ ■

Remark 2. The convergence problem arises precisely from the fact that the above hypercubes overlap (i.e. they have intersections of positive volume) for some $n > 2$ when $d > 2$ and $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right)$ (when $d > 3$ if $\alpha = \frac{2}{d+1}$).

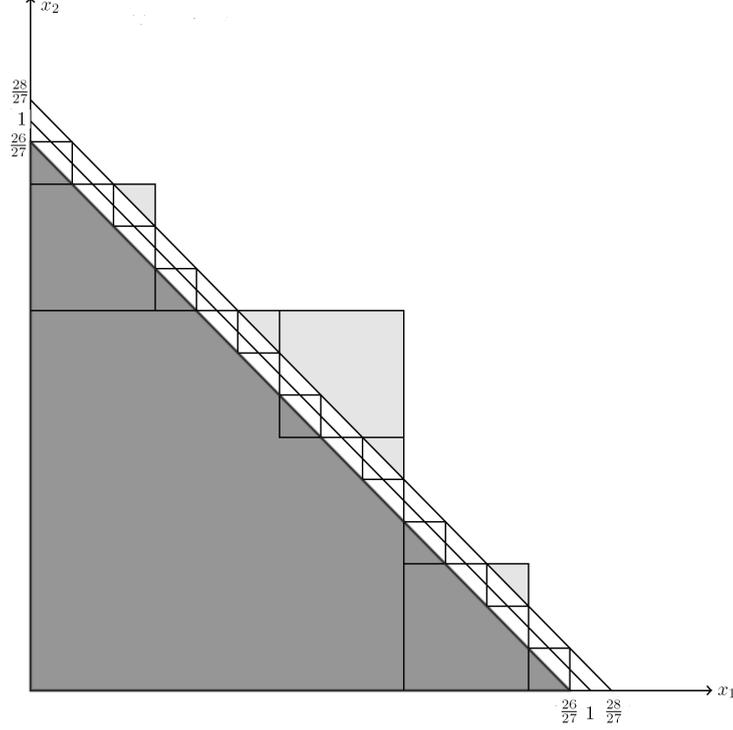


Figure 3: Convergence of the algorithm when $d=2$ and $\alpha = \frac{2}{3}$.

2. Convergence of the algorithm in any dimension

2.1. Convergence for the Lebesgue measure

Lemma 3. *The AEP algorithm converges for the Lebesgue measure when $d \geq 2$ and $\alpha \in \left[\frac{1}{d}, \frac{1}{\sqrt[d]{d!}}\right)$.*

Proof. Let us denote by $\text{vol}(A)$ the Lebesgue measure (volume) of a Lebesgue measurable set $A \subseteq \mathbb{R}^d$. Then, with the above choice of α ,

$$\text{vol}(S) - P_1 = \frac{1}{d!} - \alpha^d = \sum_k \sigma_k^2 \text{vol}(S_k^2) > 0$$

where the S_k^2 are simplexes and $\sigma_k^2 = \pm 1$.

Set $\sum_k \sigma_k^2 \text{vol}(S_k^2) = R_1$, so that

$$\text{vol}(S) = P_1 + R_1 \tag{8}$$

Add and subtract to the second member of (8) $\sum_k \sigma_k^2 \text{vol}(Q_k^2)$, where Q_k^2 are hypercubes and $\text{vol}(Q_k^2) = \text{vol}(S_k^2) \alpha^d d!$

Hence

$$\text{vol}(S) = P_2 + \sum_k \sigma_k^2 \text{vol}(S_k^2) d! \left(\frac{1}{d!} - \alpha^d \right)$$

where $P_2 = P_1 + \sum_k \sigma_k^2 \text{vol}(Q_k^2)$.

So

$$R_2 := \sum_k \sigma_k^2 \text{vol}(S_k^2) d! \left(\frac{1}{d!} - \alpha^d \right) = (1 - \alpha^d d!) \left(\frac{1}{d!} - \alpha^d \right) \quad (9)$$

and moreover $R_2 = \sum_h \sigma_h^3 \text{vol}(S_h^3)$, with $\sigma_h^3 = \pm 1$ and S_h^3 simplexes.

Hence, recursively,

$$0 < \text{vol}(S) - P_{n+1} = R_{n+1} = (1 - \alpha^d d!)^n \left(\frac{1}{d!} - \alpha^d \right) \rightarrow 0 \text{ as } n \rightarrow +\infty$$

■

Remark 4. *It is easily checked that for any $d \geq 2$ $\frac{2}{d+1} < \frac{1}{\sqrt[d]{d!}}$.*

2.2. The main Theorem

Theorem 5. *Let H be a probability distribution with support in \mathbb{R}_+^d , $d \geq 2$, absolutely continuous with bounded density in a neighborhood of the simplex diagonal. Then the AEP algorithm converges for any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1} \right]$.*

The proof will be divided in five steps

2.2.1. First step: construction of a \mathbb{Z} -module

First of all, we want to define an algebraic operation (called sum) among *extended sets*, where *positive sets* are generated by Lebesgue measurable subsets of \mathbb{R}^d and *negative sets* are generated by *subsets* of $-\mathbb{R}^d$, the *negative copy* of \mathbb{R}^d . To be precise, called \mathcal{M} the family of Lebesgue measurable subsets of \mathbb{R}^d , we define

$$\Omega = \{a_1 A_1 + \dots + a_k A_k, A_1, \dots, A_k \in \mathcal{M}, a_1, \dots, a_k \in \mathbb{Z}\}.$$

Then the elements of Ω are finite sequences of measurable subsets of \mathbb{R}^d , each one *multiplied* by an integer (positive, negative or zero). At the moment $+$ is just a punctuation sign. We also set $(-1)A = -A$ and $A + (-B) = A - B$.

Then we define a sum in Ω , still denoted by $+$, commutative and associative, by the following rules:

1) for any $A \in \Omega$ and $h \in \mathbb{Z}$, $hA = A + \dots + A$ h times if $h > 0$, $hA = -|h|A$ if $h < 0$, $0A = \emptyset$, $-\emptyset = \emptyset$, $A + \emptyset = A$;

2) for any $A, B \in \mathcal{M}$ $A + B = A \cup B + A \cap B$;

3) for any $A, B \in \mathcal{M}$ $A - B = A/B - B/A$, where $A/B = A \cap B^c$.

It follows, in particular, $A - A = \emptyset \forall A \in \Omega$. Hence Ω is a \mathbb{Z} -module.

This way we can extend the Lebesgue (and any equivalent) measure to Ω as a linear functional. In fact, denote by $\text{vol}(A)$ the Lebesgue measure of $A \in \mathcal{M}$. Then, for $a_1 A_1 + \dots + a_k A_k \in \Omega$, we define

$$\text{vol}(a_1 A_1 + \dots + a_k A_k) := a_1 \text{vol}(A_1) + \dots + a_k \text{vol}(A_k) \quad (10)$$

Moreover the sum operation induces a partial ordering in Ω as follows.

If $A \in \Omega$ we say that $A \succ \emptyset$ if $A = a_1 A_1 + \dots + a_h A_h + a_{h+1} A_{h+1} + \dots + a_s A_s$, with $a_1, \dots, a_h > 0$, $A_1, \dots, A_h, A_{h+1}, \dots, A_s \in \mathcal{M}$, $\text{vol}(A_1), \dots, \text{vol}(A_h) > 0$, $\text{vol}(A_{h+1}) = \dots = \text{vol}(A_s) = 0$. We say that $A \simeq \emptyset$ if in the above expression $h = 0$. Then, if $A, B \in \Omega$, we say that $A \succ B$, $A \simeq B$ if, respectively, $A - B \succ \emptyset$, $A - B \simeq \emptyset$. In particular, observe that, for $A, B \in \mathcal{M}$, $A \subseteq B \implies A \lesssim B$, but not vice-versa. Then $A \succ B \implies \text{vol}(A) > \text{vol}(B)$ and $A \simeq B \implies \text{vol}(A) = \text{vol}(B)$

By the above definitions, considering the AEP algorithm for $d \geq 2$ and $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, we can replace, at the n -th step, the sum P_n of volumes with a sum Π_n of elements of Ω , i.e.

$$\Pi_n = \sum_k \sigma_k^n Q_k^n$$

where $\sigma_k^n = \pm 1$ and Q_k^n are hypercubes, so that $P_n = \text{vol}(\Pi_n)$.

2.2.2. Second step: proof of an equivalence when $\alpha = \frac{1}{d}$

Now, take as in the Lemma $S = S(\mathbf{0}, 1)$. We start by considering the case $\alpha = \frac{1}{d}$, although in the following we will continue to use the symbol α , as most arguments apply to all $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$.

Then define

$$S_n = \left\{ 0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^n = 1 - \left(1 - \frac{1}{d}\right)^n, x_1, \dots, x_d \geq 0 \right\} \quad (11)$$

We recall that, given $\mathbf{x} = (x_1, \dots, x_d)$, $\lambda(\mathbf{x}) = x_1 + \dots + x_d$.

We want to prove that, at any step of the AEP algorithm,

$$S_n \simeq \sum_k \sigma_k^n (Q_k^n \cap S_n) \quad (12)$$

where \simeq is the above defined equivalence in Ω .

We will prove (12) by induction. In fact (12) holds for any n when $d = 2$ and for $n = 1, 2$ when $d > 2$, as it is easily checked. Therefore, fixed $d > 2$, assume (12) holds for some $n \geq 2$. Since we have seen (Proposition 1) that S_{n+1} is covered by the hypercubes of Π_{n+1} with sides $\alpha(1 - \alpha)^s$, $0 \leq s \leq n$, having a positive sign in Π_{n+1} , we have to consider the contribution of hypercubes with a negative sign in Π_{n+1} . Let us illustrate the situation by taking the second step of the algorithm and considering a hypercube $Q_r = Q(\mathbf{b}_r, \alpha(1 - r\alpha))$, where \mathbf{b}_r has r coordinates equal to α and $d - r$ equal to zero, $2 \leq r \leq d - 1$ (if $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, $2 \leq r < \frac{1}{\alpha}$). Thus we can find $n \geq 2$ such that

$$1 - (1 - \alpha)^n < r\alpha \leq 1 - (1 - \alpha)^{n+1}$$

Then, consider $Q_l = Q(\mathbf{b}_l, \alpha(1 - l\alpha))$, $1 \leq l < r$, where \mathbf{b}_l is obtained from \mathbf{b}_r by replacing $r - l$ α 's with 0 's: to fix the ideas,

$$\mathbf{b}_r = \left(\overbrace{\alpha, \dots, \alpha}^{r \text{ times}}, 0, \dots, 0 \right), \mathbf{b}_l = \left(\overbrace{\alpha, \dots, \alpha}^{l \text{ times}}, 0, \dots, 0 \right) \quad (13)$$

Next, consider the hypercube $Q_l^1 = Q(\mathbf{b}_l^1, \alpha(1 - l\alpha)(1 - \alpha)^{r-l})$, where

$$\mathbf{b}_l^1 = (\alpha, \dots, \alpha, \alpha(1 - l\alpha), \dots, \alpha(1 - l\alpha)(1 - \alpha)^{r-l-1}, 0, \dots, 0).$$

Hence Q_l^1 has, in the development of the AEP, the same sign as $Q(\mathbf{b}_l, \alpha(1-l\alpha))$. By adding $\alpha(1-l\alpha)(1-\alpha)^{r-l}$ and then, if necessary, $\alpha(1-l\alpha)(1-\alpha)^{r-l+1}$ and so on to the smallest coordinate between the $(l+1)$ -th and the r -th place of the vertex \mathbf{b}_l^j of Q_l^j , finally we get a hypercube $Q_l^t = Q(\mathbf{b}_l^t, \alpha(1-l\alpha)(1-\alpha)^m)$ such that $\mathbf{b}_r \in Q_l^t$ and $m < n$.

Thus $Q_r \cap S_{n+1}$ can be covered by *positive* hypercubes of the AEP approximation of a simplex $\widetilde{S}^l = S(\mathbf{b}_l, 1-l\alpha)$. Then, by applying the AEP to \widetilde{S}^l , we can find

$$\bar{m} = \inf \left\{ m : l\alpha + (1-l\alpha)(1-(1-\alpha)^m) \geq 1 - (1-\alpha)^{n+1} \geq r\alpha \right\}$$

Hence $\bar{m} \leq n$, i.e. \bar{m} corresponds to a step $< n+1$ of the algorithm. Therefore, by the induction hypothesis

$$\sum_h \sigma_h^{\bar{m}} \left(\widetilde{Q}_h^{\bar{m}} \cap Q_r \cap \widetilde{S}_{\bar{m}} \right) \simeq Q_r \cap \widetilde{S}_{\bar{m}}$$

and the same holds replacing $\widetilde{S}_{\bar{m}}$ by S_{n+1} , since $Q_r \cap \widetilde{S}_{\bar{m}} \succeq Q_r \cap S_{n+1}$.

However such a $Q_r \cap S_{n+1}$ must be accounted for, in the analogous expression relative to original AEP, with a sign $(-1)^{1+l}$ and clearly that must be repeated $\binom{r}{l}$ times (the number of ways by which $r-l$ α 's can be replaced by 0 's). As a consequence, in the AEP expression relative to the $(n+1)$ -th step, $Q_r \cap S_{n+1}$ is multiplied by an integer

$$z_r = (-1)^{1+r} + (-1)^{2+r} \binom{r}{1} + \dots + (-1)^{2r} \binom{r}{r-1} = 1 \quad (14)$$

since from $(1-1)^r = 0$ it follows

$$1 - \binom{r}{1} + \dots + (-1)^{r-1} \binom{r}{r-1} = (-1)^{r+1} \quad (15)$$

and (14) is obtained from (15) by multiplying both members of the equality by $(-1)^{r+1}$.

The above argument can be implemented recursively. In fact, consider, with the above notations, the simplex $\widetilde{S}^r = S(\mathbf{b}_r, 1-r\alpha)$, $2 \leq r < \frac{1}{\alpha}$. Then, for $1 \leq l < r$, take a simplex $\widetilde{S}^{l_t} = S(\mathbf{b}_l^t, (1-l\alpha)(1-\alpha)^{m^*})$, where \mathbf{b}_l^t corresponds, for example, to our previous construction and

$$m^* = \sup \{ m / (1-l\alpha)(1-\alpha)^m > 1-r\alpha \}$$

Hence it is easily checked that $\widetilde{S}^r \subseteq \widetilde{S}^{l_t}$. Thus take, by self-similarity, the application of the AEP to both the above simplexes. Consider, for any sufficient high n , the intersection $\widetilde{Q} \cap S_{n+1}$ of a hypercube \widetilde{Q} belonging, with a positive sign, to the AEP development relative to \widetilde{S}^r with

$$S_{n+1} = \left\{ 0 \leq x_1 + \dots + x_d \leq 1 - (1-\alpha)^{n+1}, x_1, \dots, x_d \geq 0 \right\}$$

Then, as $\widetilde{Q} \cap S_{n+1}$ corresponds to a step $p < n+1$ of the AEP algorithm applied to \widetilde{S}^r , by the induction hypothesis

$$\sum_h \sigma_h^p \left(\widetilde{Q}_h^p \cap \widetilde{Q} \cap S_{n+1} \right) \simeq \widetilde{Q} \cap S_{n+1}$$

As above, this implies that $\widetilde{Q} \cap S_{n+1}$ will be accounted for, in the analogous expression relative to the AEP approximation of $S = S(\mathbf{0}, 1)$, with a sign $(-1)^{1+r}$. Similarly, as $\widetilde{S}^r \subseteq \widetilde{S}^{l_t} \subseteq \widetilde{S}^l = S(\mathbf{b}_l, 1-l\alpha)$,

where \mathbf{b}_l is defined as in (13), $\tilde{Q} \cap S_{n+1}$ will be accounted again with a coefficient $(-1)^{1+l}$. Therefore, as above, recalling (14), we can conclude that in $\sum_k \sigma_k^{n+1} (Q_k^{n+1} \cap \tilde{Q} \cap S_{n+1})$ $\tilde{Q} \cap S_{n+1}$ is accounted for with coefficient 1.

Moreover the same argument holds if we replace the simplex $S(\mathbf{0}, 1)$ by some $S(\mathbf{b}, (1-\alpha)^q)$, where $\lambda(\mathbf{b}) = \alpha + \alpha(1-\alpha) + \dots + \alpha(1-\alpha)^{q-1} = 1 - (1-\alpha)^q$, $q \geq 1$.

Finally $\sum_k \sigma_k^{n+1} (Q_k^{n+1} \cap S_{n+1})$ includes the intersections with S_{n+1} of the hypercubes generated at

the $(n+1)$ -th step with vertices on the axes, say $Q(\mathbf{b}^h, \alpha(1-\alpha)^n)$, $\mathbf{b}^h = \left(0, \dots, \overbrace{1 - (1-\alpha)^n}^{h\text{-th place}}, \dots, 0 \right)$,

$h = 1, \dots, d$. Thus, clearly, $Q(\mathbf{b}^h, \alpha(1-\alpha)^n) \cap S_{n+1}$ is accounted for with a positive sign.

What we have proven, in fact, is that $S_p \simeq \sum_k \sigma_k^p (Q_k^p \cap S_p)$ for $p \leq n$ (and we know the equivalence holds for $p = 1, 2$) implies $S_{n+1} \preceq \sum_k \sigma_k^{n+1} (Q_k^{n+1} \cap S_{n+1})$. So assume by contradiction that there is a first $n^* > 2$ such that

$$S_{n^*} \prec \sum_k \sigma_k^{n^*} (Q_k^{n^*} \cap S_{n^*}) \quad (16)$$

Consequently there will be an *excess of volume* measured by $\sum_k \sigma_k^{n^*} \text{vol}(Q_k^{n^*} \cap S_{n^*}) - \text{vol}(S_{n^*})$. We want to show that such an *excess* does not decrease (in fact it increases) through the subsequent iterations of the algorithm.

To this end recall that, by Proposition 1, S_{n^*} is covered by hypercubes of sides $\alpha(1-\alpha)^s$, $0 \leq s < n^*$. Therefore we can detect one of them, say Q^* , such that

$$\sum_k \sigma_k^{n^*} (Q_k^{n^*} \cap Q^* \cap S_{n^*}) - Q^* \cap S_{n^*} \succeq A$$

where $A \in \mathcal{M}$ and $\text{vol}(A) > 0$.

Hence consider the AEP applied to \widetilde{S}^r , $2 \leq r \leq d-1$, defined as above. Then, after n^* steps (corresponding to $n^* + m_r$ ones, for a suitable m_r , of the original algorithm) there will be a hypercube \widetilde{Q}^* such that

$$\sum_k \sigma_k^{n^*} (\widetilde{Q}_k^{n^*} \cap \widetilde{Q}^* \cap \widetilde{S}_{n^*}^r) - \widetilde{Q}^* \cap \widetilde{S}_{n^*}^r \succeq \widetilde{A}$$

where $\text{vol}(\widetilde{A}) = (1-r\alpha)^d \text{vol}(A)$. Therefore, by the above arguments, recalling (14),

$$\sum_k \sigma_k^{n^*+m_r} (Q_k^{n^*+m_r} \cap \widetilde{Q}^* \cap (S_{n^*+m_r} - S_{n^*})) - Q^* \cap (S_{n^*+m_r} - S_{n^*}) \succeq \widetilde{A}$$

In other words, for any $m \geq n^*$,

$$\sum_k \sigma_k^m (Q_k^m \cap S_m) - S_m \succeq A \quad (17)$$

where $\text{vol}(A) > 0$.

Now, let n be sufficiently high and $p \geq 1$. We indicate by $\rho(m) = \frac{N^m - 1}{N - 1}$, $N = 2^d - 1$, the number of hypercubes produced by the AEP in the first m steps. Then, as it is easily checked,

$$\begin{aligned} \sum_{k=1}^{\rho(n+p)} \sigma_k^{n+p} (Q_k^{n+p} \cap S_{n+p}) &= \sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap S_n) + \sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap (S_{n+p} - S_n)) + \\ &+ \sum_{k=\rho(n)+1}^{\rho(n+p)} \sigma_k^{n+p} (Q_k^{n+p} \cap (S_{n+p} - S_n)) \end{aligned}$$

Choose $p = tn$, in such a way that $(1 - \alpha)^t \leq (\frac{1}{4})^d$. Then, fixed $q \geq 1$, for any $1 \leq k \leq \rho(n)$, $\text{vol}(Q_k^n \cap (S_{n+p+q} - S_{n+p})) \leq (\frac{1}{4})^{dn}$.

Hence $\left| \sum_{k=1}^{\rho(n)} \sigma_k^n \text{vol}(Q_k^n \cap (S_{n+p+q} - S_{n+p})) \right| < (\frac{1}{2})^{dn}$, as $\rho(n) \leq 2^{dn}$.

Thus

$$\begin{aligned} \sum_{k=\rho(n)+1}^{\rho(n+p)} \sigma_k^{n+p} \text{vol}(Q_k^{n+p} \cap (S_{n+p+q} - S_{n+p})) &= \sum_{h=1}^{\rho(n+p+q)} \sigma_k^{n+p+q} \text{vol}(Q_k^{n+p+q} \cap S_{n+p+q}) - \\ &- \sum_{k=1}^{\rho(n+p)} \sigma_k^{n+p} \text{vol}(Q_k^{n+p} \cap S_{n+p}) + 0 \left((\frac{1}{2})^{dn} \right) \end{aligned}$$

Since we have shown that $\sum_{k=1}^{\rho(m)} \sigma_k^m \text{vol}(Q_k^m \cap S_m)$ is increasing with m , it follows through straightforward steps that

$$\limsup \sum_{k=1}^{\rho(m)} \sigma_k^m \text{vol}(Q_k^m) \geq \lim_{m \rightarrow +\infty} \sum_{k=1}^{\rho(m)} \sigma_k^m \text{vol}(Q_k^m \cap S_m) \quad (18)$$

Therefore, because of (17)

$$\limsup_{m \rightarrow +\infty} \sum_{k=1}^{\rho(m)} \sigma_k^m \text{vol}(Q_k^m) - \text{vol}(S) \geq \text{vol}(A) > 0$$

which, being $\alpha = \frac{1}{d}$, contradicts the Lemma.

Hence we have proven that, when $\alpha = \frac{1}{d}$, for any n , (12) holds, i.e.

$$\sum_k \sigma_k^n (Q_k^n \cap S_n) \simeq S_n$$

2.2.3. Third step: extension of the above equivalence

Now we want to prove that for any $\alpha \in \left(\frac{1}{d}, \frac{2}{d+1} \right]$ an analogous equivalence holds, when, however, the totality of the hypercubes relative to the n -th step of the AEP is replaced by a selection (or *extrapolation*) consisting in those hypercubes, say \widehat{Q}_k^n , whose sides are of the type $\alpha(1 - l_1\alpha) \dots (1 - l_q\alpha)$, where $0 \leq l_1, \dots, l_q < \frac{1}{\alpha}$.

To this purpose we start by taking $\alpha = \frac{1+\varepsilon}{d}$, where $\varepsilon > 0$ is sufficiently small and in any case $\varepsilon \ll \frac{1}{d-1}$. As a matter of fact, for sake of simplification, we take ε satisfying

$$\left(1 - \frac{1+\varepsilon}{d} \right)^{\overline{m}} = \varepsilon \quad (19)$$

which can be done choosing a suitable $\varepsilon(\overline{m})$, when \overline{m} is large enough: in fact $(1 - \frac{1}{d})^{\overline{m}+1} < \varepsilon(\overline{m}) < (1 - \frac{1}{d})^{\overline{m}}$. Hence each hypercube with the *main* (i.e. the one with the smallest l^1 norm) vertex in the strip $\{0 \leq x_1 + \dots + x_d < 1\}$ has only one vertex in the strip $\{1 < x_1 + \dots + x_d \leq 1 + \varepsilon\}$.

Consider, then, the above mentioned extrapolation. When $n = 1, 2$, as we have often recalled, the following equivalence holds:

$$\sum_k \sigma_k^n (\widehat{Q}_k^n \cap S_n) \simeq S_n$$

and the above arguments show that for any n

$$\sum_k \sigma_k^n \left(\widehat{Q}_k^n \cap S_n \right) \succeq S_n$$

Suppose, then, there exists a first $n^* > 2$ for which

$$\sum_k \sigma_k^{n^*} \left(\widehat{Q}_k^{n^*} \cap S_{n^*} \right) \succ S_{n^*} \quad (20)$$

and let

$$\sum_k \sigma_k^{n^*} \left(\widehat{Q}_k^{n^*} \cap S_{n^*} \right) - S_{n^*} \simeq A, \text{ where } A = a_1 A_1 + \dots + a_h A_h, \text{ with } a_1, \dots, a_h > 0, \text{ vol}(A_1), \dots, \text{vol}(A_h) > 0.$$

The problem now is that, in order to be able to utilize the Lemma, we have to consider as well the *excesses* produced, via self-similarity, by the application of the algorithm to the *exterior* simplexes generated in the strip $\{1 < x_1 + \dots + x_d \leq 1 + \varepsilon\}$, which, in their turn, generate new simplexes lying inside $S(\mathbf{0}, 1)$ and so on. The following construction allows, precisely, to deal with this problem.

Take $m^* = \min(n^*, \bar{m})$, recalling (19).

Consider the sub-simplexes $S^p = S(\mathbf{b}^p, (1 - \alpha)^{m^*})$, where $\lambda(\mathbf{b}^p) = \alpha + \alpha(1 - \alpha) + \dots + \alpha(1 - \alpha)^{m^* - 1}$, so that, as it is easily seen, $1 \leq p \leq d^{m^*}$. Hence we can compare the *distinct* (i.e. not contained in previous ones) simplexes generated, by the extrapolated hypercubes, at the $(j + 1)$ -th step of the AEP, $j = 1, \dots, m^*$, in the strip $\{1 < x_1 + \dots + x_d \leq 1 + \varepsilon\}$, with the ones produced at the j -th extrapolation relative to the sub-simplex S^{jd} .

For example, consider the simplex $\widetilde{S}^{l\varepsilon} = S(\mathbf{c}, -\varepsilon(1 - l\alpha))$, where $\lambda(\mathbf{c}) = l\alpha + d\alpha(1 - l\alpha) = 1 + \varepsilon(1 - l\alpha)$. Then, applying the extrapolated AEP to $\widetilde{S}^{l\varepsilon}$, after n^* steps we get

$$\sum_k \sigma_k^{n^*} \text{vol} \left(\widehat{Q}_k^{n^*} \cap \widetilde{S}_{n^*} \right) - \text{vol} \left(\widetilde{S}_{n^*} \right) = \varepsilon(1 - l\alpha) \text{vol}(A)$$

which will be accounted for with a sign $(-1)^{2+l}$ in the overall algorithm. However this contribution will be compensated by one equal to $(-1)^{1+l}(1 - \alpha)^{m^*}(1 - l\alpha) \text{vol}(A)$ (as $\varepsilon \leq (1 - \alpha)^{m^*}$) provided by the AEP development of the corresponding simplex we have picked up (in this case a sub-simplex of S^{2d}).

Moreover, observe that, for $\alpha = \frac{1+\varepsilon}{d}$,

$$d(1 - \alpha) - \binom{d}{2}(1 - 2\alpha) + \dots + (-1)^d d(1 - (d - 1)\alpha) = 1 + (-1)^{d-1} \varepsilon \quad (21)$$

This implies that, if the extrapolated algorithm applied, say, to one of the above simplexes S^p produces after n^* steps an *excess* (in the sense of (20)) of volume $q > 0$, then after a sufficiently higher number of steps the *excess* of volume will be at least $q + (1 - \varepsilon)q$. For example, denote by h the *radius* of S^p and apply the extrapolated algorithm both to S^p and to the sub-simplexes of *radii* $h(1 - \alpha), h(1 - 2\alpha), \dots, h(1 - (d - 1)\alpha)$ (recall $\alpha = \frac{1+\varepsilon}{d}$ and $\varepsilon \ll \frac{1}{d-1}$). As we have seen in § 2.2.2, if, after a certain number of steps, an excess of volume is *accumulated*, then it does not decrease in the following steps. Hence we can *put aside* (save) such excesses and sum them up, algebraically, at the end (i.e. when the one relative to the simplex with *radius* $h(1 - (d - 1)\alpha)$ is, for the first time, produced). Clearly the process is recursive and we can combine it with the above construction, which

allows to *compensate* the *negative excesses* of volume produced by the application of the extrapolated AEP to *exterior* simplexes. To fix the ideas, consider, after a suitable partition, the first d sub-simplexes S^p , $1 \leq p \leq d$. Then, following formula (21), after a sufficiently high number of steps, they will contribute to the *excess* of volume of the overall extrapolated algorithm in measure, say, $\mu(1 - \varepsilon)$ (e.g. $\mu = (1 - \alpha)^{m^* - 1} \text{vol}(A)$). But, for our construction, we have to subtract a quantity, due to the corresponding *excess* of volume caused by the *first exterior* simplex, at most $\mu(1 - \alpha) = \mu(1 - \frac{1+\varepsilon}{d})$ (since $\varepsilon \leq (1 - \alpha)^{m^*}$). Hence, if $\varepsilon \ll \frac{1}{2d}$, the contribution will be in any case greater than $\mu\varepsilon$. The results holds, *a fortiori*, for the other *strings*, of length d , of sub-simplexes S^p , as in those cases the quantities to be subtracted are, for the same number of steps, smaller. As an example, consider the string $\{S^p, d + 1 \leq p \leq 2d\}$. After a convenient number of steps, recalling the above arguments, we can denote the *excess* of volume produced by the application of the extrapolated algorithm to any S^p as, say, $[\rho + \rho(1 \pm \varepsilon)](1 - \alpha)$ (e.g. $\rho = (1 - \alpha)^{m^* - 1} \text{vol}(A)$). Then, after a sufficiently high number of steps, we will have, due to (21), an *excess* of volume originated from the string $\{S^p, d + 1 \leq p \leq 2d\}$ which is (at least) $[\rho + \rho(1 \pm \varepsilon)](1 \pm \varepsilon)$. But now, applying our comparison, what we have to subtract is (at most) $\rho(1 \pm \varepsilon)(1 - \alpha)$.

Hence, recall the arguments of § 2.2.2 and consider the two strips

$$T_q = \left\{ 0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^{m^*} (1 - \alpha)^q \right\}$$

and

$$T'_q = \{1 + \varepsilon(1 - \alpha)^q \leq x_1 + \dots + x_d \leq 1 + \varepsilon\}$$

Then it follows through straightforward steps that, for a sufficiently high \tilde{q} ,

$$\sum_k \sigma_k^{\tilde{q}+m^*} \text{vol} \left(Q_k^{\tilde{q}+m^*} \cap (T_{\tilde{q}} \cup T'_{\tilde{q}}) \right) - \text{vol}(S_{\tilde{q}+m^*}) \geq \text{vol}(A) + c\varepsilon + 0(\varepsilon^2) \quad (22)$$

where $c > 0$, and we can go on recursively.

Eventually, by the same arguments we utilized above, it follows

$$\limsup_{m \rightarrow +\infty} \sum_{k=1}^{\rho(m)} \sigma_k^m \text{vol}(Q_k^m) - \text{vol}(S) > 0$$

contradicting, once again, the Lemma.

2.2.4. **Fourth step: analytic measures**

Then we have proven that for $\alpha = \frac{1+\varepsilon(\bar{m})}{d}$, with $\lim_{\bar{m} \rightarrow +\infty} \varepsilon(\bar{m}) = 0$,

$$\sum_k \sigma_k^n \left(\widehat{Q}_k^n \cap S_n \right) \simeq S_n \quad (23)$$

holds for any $n \geq 1$, where \widehat{Q}_k^n are hypercubes of the above defined extrapolation. Hence, if we consider an analytic distribution H in \mathbb{R}_+^d , the function

$$f_n^H(\alpha) = \sum_k \sigma_k^n v_H \left(\widehat{Q}_k^n \cap S_n \right) - v_H(S_n) \quad (24)$$

analytic in $[\frac{1}{d}, \hat{\alpha}_n]$, is zero on a sequence of values tending to $\frac{1}{d}$. Therefore $f_n^H(\alpha) \equiv 0$ in $[\frac{1}{d}, \hat{\alpha}_n]$. As it is easily seen, a value $\hat{\alpha}$ (we drop the pedex in order to simplify the notation) where $f_n^H(\alpha)$ may loose analyticity is such that there exist hypercubes $\hat{Q}_j^n(\hat{\alpha}) = Q(\mathbf{b}_j, \hat{\alpha}(1-l_1\hat{\alpha}) \dots (1-l_{q_j}\hat{\alpha}))$ with $\lambda(\mathbf{b}_j) = 1 - (1-l_1\hat{\alpha}) \dots (1-l_{q_j}\hat{\alpha}) = 1 - (1-\hat{\alpha})^n$, but $l_1 + \dots + l_{q_j} < n^1$. Therefore, assume this is the case and take a sufficiently small interval of $\hat{\alpha}$, say $[\hat{\alpha} - \delta, \hat{\alpha} + \delta]$ for a small $\delta > 0$. Choose a suitable analytic function $\varphi(\alpha)$ defined in $[\hat{\alpha} - \delta, \hat{\alpha} + \delta]$, $\varphi(\alpha)$ having the sign of $\alpha - \hat{\alpha}$, such that the simplex $S^*(\alpha) = \{0 \leq x_1 + \dots + x_d \leq 1 - (1-\alpha)^n + \varphi(\alpha), x_1, \dots, x_d \geq 0\}$, $\alpha \in [\hat{\alpha} - \delta, \hat{\alpha} + \delta]$, intersects each hypercube $\hat{Q}_j^n(\alpha)$ at most at one point. Now we want to show that, for $\alpha \in [\hat{\alpha} - \delta, \hat{\alpha}]$

$$\tilde{f}_n^H(\alpha) = \sum_k \sigma_k^n v_H(\hat{Q}_k^n \cap S^*(\alpha)) - v_H(S^*(\alpha)) = 0 \quad (25)$$

In fact, this is true when $\alpha = \hat{\alpha}$, since in this case $S^*(\hat{\alpha}) = S_n(\hat{\alpha})$, while for $\alpha \in [\hat{\alpha} - \delta, \hat{\alpha}]$

$$f_n^{H'}(\alpha) = \sum_k \sigma_k^n v_{H'}(\hat{Q}_k^n \cap S_n(\alpha)) - v_{H'}(S_n(\alpha)) = 0$$

for any analytic measure (equivalent to Lebesgue) H' .

Hence, posed $\varphi(\alpha) = -\beta$, $\beta > 0$, we can choose, for any $m > 1$, an analytic distribution H'_m and a number ρ , $0 < \rho \ll \beta$, such that, called $\delta_{H'_m}(\mathbf{x})$ and $\delta_H(\mathbf{x})$ the densities, respectively, of H'_m and H ,

$$\begin{aligned} |\delta_{H'_m}(\mathbf{x}) - \delta_H(\mathbf{x})| &\leq 2^{-dn} \rho^m \quad \text{when } \mathbf{x} \in S^*(\alpha) \\ &\text{and} \\ \delta_{H'_m}(\mathbf{x}) &< 2^{-dn} \rho^m \quad \text{when } 1 - (1-\alpha)^n \geq \lambda(\mathbf{x}) \geq 1 - (1-\alpha)^n - \beta + \rho^m, \mathbf{x} \geq 0 \end{aligned} \quad (26)$$

It follows that

$$0 = \sum_k \sigma_k^n v_{H'_m}(\hat{Q}_k^n \cap S_n(\alpha)) - v_{H'_m}(S_n(\alpha)) = \sum_k \sigma_k^n v_H(\hat{Q}_k^n \cap S^*(\alpha)) - v_H(S^*(\alpha)) + 0(\rho^m) \quad (27)$$

implying (25) when $m \rightarrow +\infty$.

Hence, because of analyticity, $\tilde{f}_n^H(\alpha) \equiv 0$ in $[\hat{\alpha} - \delta, \hat{\alpha} + \delta]$. But for any $\alpha \in (\hat{\alpha}, \hat{\alpha} + \delta]$ $\varphi(\alpha)$ can be chosen arbitrarily small, e.g. $\varphi(\alpha) < \rho^m$ for an arbitrary $0 < \rho \ll 1$. So, in fact, $f_n^H(\alpha) \equiv 0$ in $[\hat{\alpha} - \delta, \hat{\alpha} + \delta]$, which extends the analyticity of $f_n^H(\alpha)$.

As a consequence, for any $n \geq 1$, any $\alpha \in [\frac{1}{d}, \frac{2}{d+1}]$ and any analytic distribution H

$$f_n^H(\alpha) = 0. \quad (28)$$

Remark 6. The above arguments show in particular that, for a given $\alpha \in [\frac{1}{d}, \frac{2}{d+1}]$ and an analytic distribution H , if a sub-simplex $S^*(\alpha)$ satisfies $S_{n-1}(\alpha) \subset S^*(\alpha) \subset S_n(\alpha)$, then $\sum_k \sigma_k^n v_H(\hat{Q}_k^n \cap S^*(\alpha)) - v_H(S^*(\alpha)) = 0$, where \hat{Q}_k^n denote the hypercubes of the above described extrapolation.

Remark 7. Recall the definition (24) of $f_n^H(\alpha)$. Then we observed that, for an analytic H , $f_n^H(\alpha)$ is analytic in $[\frac{1}{d}, \hat{\alpha}_n]$, where $\hat{\alpha}_n$ denotes the first value at which $f_n^H(\alpha)$ might loose analyticity. In fact, it is easily checked that such $\hat{\alpha}_n$ constitute a non-increasing sequence (in particular $\hat{\alpha}_1 = \hat{\alpha}_2 = \frac{2}{d+1}$).

¹if, say, $r\hat{\alpha} = 1 - (1-\hat{\alpha})^n$, $2 \leq r < \frac{1}{\hat{\alpha}}$ and consequently $n > r$, then, set $\psi(\alpha) = r\alpha - 1 + (1-\alpha)^n$, it is easily checked that $\psi'(\hat{\alpha}) < 0$

Moreover a conclusion analogous to (28) holds if the extrapolation is replaced by the whole AEP development and we consider, instead of $f_n^H(\alpha)$,

$$g_n^H(\alpha) = \sum_k \sigma_k^n v_H(Q_k^n \cap (T_n \cup T_n')) - v_H(S_n)$$

where

$$T_n = \{0 \leq x_1 + \dots + x_d \leq 1 - (1 - \alpha)^n\}$$

and

$$T_n' = \left\{1 + (d\alpha - 1)(1 - \alpha)^{n-1} \leq x_1 + \dots + x_d \leq d\alpha\right\}$$

In fact we observe, first of all, that $g_1^H(\alpha) = g_2^H(\alpha) = 0$ for any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ (this follows from the fact that, when $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, $d\alpha - 1 \leq 1 - \alpha$). Then we can proceed by induction. Assume, for some $n \geq 2$, $g_n^H(\alpha) \equiv 0$ in $\left[\frac{1}{d}, \frac{2}{d+1}\right]$. Then, exploiting the self-similarity of the AEP algorithm and the above Remark 6, it is proven through straightforward arguments that $g_{n+1}^H(\alpha) \equiv 0$ in the same interval of analyticity $\left[\frac{1}{d}, \hat{\alpha}_{n+1}\right]$ of $f_{n+1}^H(\alpha)$. Consequently the analytic extension is proven exactly in the same way as above.

Hence for any $n \geq 1$, $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$ and analytic distribution H

$$g_n^H(\alpha) = \sum_k \sigma_k^n v_H(Q_k^n \cap (T_n \cup T_n')) - v_H(S_n) = 0 \quad (29)$$

Now recall that an absolutely continuous function H can be approximated as well as we want by some analytic function H' . Then fix an absolutely continuous distribution H , $n \geq 3$ and (an arbitrarily small) $\varepsilon > 0$. By straightforward arguments it follows that there exists an analytic distribution H' such that, for any $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$,

$$\left|g_n^H(\alpha) - g_n^{H'}(\alpha)\right| < \varepsilon \quad (30)$$

(in fact a detailed proof of this statement, such as the one given in Appendix, requires a result known as the *absolute continuity of the Lebesgue integral* (see, e.g., Kolmogorov and Fomin, 1977, p.294).

Hence(29) holds for any absolutely continuous distribution H .

2.2.5. Fifth step: concluding the convergence proof

The last step consists in computing the $\lim_{n \rightarrow +\infty} \sum_{k=1}^{\rho(n)} \sigma_k^n \text{vol}(Q_k^n)$ when $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$. To this end we utilize an argument we have already introduced and here we repeat in detail.

Let n be sufficiently high and $p \geq 1$. Indicate by $\rho(m) = \frac{N^m - 1}{N - 1}$, $N = 2^d - 1$, the number of hypercubes produced by the AEP in the first m steps. Then it is easily checked that

$$\begin{aligned} & \sum_{k=1}^{\rho(n+p)} \sigma_k^{n+p} (Q_k^{n+p} \cap (T_{n+p} \cup T_{n+p}')) = \sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap (T_n \cup T_n')) + \\ & \sum_{k=1}^{\rho(n)} \sigma_k^n (Q_k^n \cap [(T_{n+p} - T_n) \cup (T_{n+p}' - T_n')]) + \\ & + \sum_{k=\rho(n)+1}^{\rho(n+p)} \sigma_k^{n+p} (Q_k^{n+p} \cap [(T_{n+p} - T_n) \cup (T_{n+p}' - T_n')]) \end{aligned}$$

Choose $p = tn$ in such a way that $(1 - \alpha)^t \leq (\frac{1}{4})^d$. Then, fixed $q \geq 1$, recalling that the density of H is bounded in a neighborhood of the diagonal $\{x_1 + \dots + x_d = 1, x_1, \dots, x_d \geq 0\}$, for a sufficiently high n and any $1 \leq k \leq \rho(n)$ it follows $v_H(Q_k^n \cap [(T_{n+p} - T_n) \cup (T'_{n+p} - T'_n)]) \leq C (\frac{1}{4})^{dn}$, for some $C > 0$.

Hence, being $\rho(n) \leq 2^{dn}$,

$$\left| \sum_{k=1}^{\rho(n)} \sigma_k^n v_H(Q_k^n \cap [(T_{n+p+q} - T_{n+p}) \cup (T'_{n+p+q} - T'_{n+p})]) \right| < C \left(\frac{1}{2}\right)^{dn}$$

Thus

$$\begin{aligned} & \sum_{k=\rho(n)+1}^{\rho(n+p)} \sigma_k^{n+p} v_H(Q_k^{n+p} \cap [(T_{n+p+q} - T_{n+p}) \cup (T'_{n+p+q} - T'_{n+p})]) = \\ & \sum_{h=1}^{\rho(n+p+q)} \sigma_k^{n+p+q} v_H(Q_k^{n+p+q} \cap (T_{n+p+q} \cup T'_{n+p+q})) - \\ & - \sum_{k=1}^{\rho(n+p)} \sigma_k^{n+p} v_H(Q_k^{n+p} \cap (T_{n+p} \cup T'_{n+p})) + 0 \left(\left(\frac{1}{2}\right)^{dn}\right) \end{aligned}$$

In fact, posed $n + p = n(t + 1) = m$, then $n = \frac{m}{t+1}$. So, recalling (29), as

$$v_H(S_{m+q}) - v_H(S_m) \leq C^* (1 - \alpha)^m$$

for a suitable $C^* > 0$,

$$\sum_{k=1}^{\rho(m)} \sigma_k^m v_H(Q_k^m \cap [(T_{m+q} - T_m) \cup (T'_{m+q} - T'_m)]) = 0 \left(\left(\frac{1}{2}\right)^{\frac{dm}{t+1}} \right) \quad (31)$$

for any $q \geq 1$. Therefore, finally,

$$\lim_{m \rightarrow +\infty} \sum_{k=1}^{\rho(m)} \sigma_k^m v_H(Q_k^m) = v_H(S) \quad (32)$$

which concludes the proof of the Theorem.

Remark 8. *We believe that the restriction, in the proof of the Theorem, of the density boundedness in a neighborhood of the simplex diagonal could be removed.*

3. Conclusions

What has been proven, i.e. the convergence of the above *geometric* algorithm, is, in our perspective, only the first stage of a fairly ambitious program of research. The goal is, in fact, to develop a whole family of *deterministic* (possibly geometric) algorithms, which, besides presenting the advantages recalled in the Introduction, can also compete in velocity with Monte Carlo methods. Hence specific algorithms of such a type could be applied for computing not only the distribution of the sum of dependent random variables (equivalent, when the variables are non-negative, to the l^1 norm), but also other distributions (e.g., distributions relative to the norms l^2 , l^∞ , in general l^p with $p > 1$, extreme value distributions etc..) of relevant interest in finance and insurance.

4. Appendix

Proposition 9. *With the notations of § 2.2.4, for any absolutely continuous distribution H , $g_n^H(\alpha) = 0 \ \forall n \geq 1$ and $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$*

Proof. Thanks to self-similarity, it suffices to prove that, for any absolutely continuous distribution H ,

$$f_n^H(\alpha) = \sum_{k=1}^{\widehat{\rho}(n)} \sigma_k^n v_H(\widehat{Q}_k^n \cap S_n(\alpha)) - v_H(S_n(\alpha)) = 0$$

where \widehat{Q}_k^n are the hypercubes of the *extrapolation* defined in § 2.2.3 (in fact we can assume $n \geq 3$, as the cases $n = 1, 2$ are trivial).

To this end we start by considering, in the hypercube $[0, 1]^d$, the *trapezoid* T defined by

$$T = \left\{ \mathbf{x}' = (x_1, \dots, x_{d-1}) \in [0, 1]^{d-1}, 0 \leq x_d \leq 1 - \lambda(\mathbf{x}') \right\}$$

Then the Theorem on the absolute continuity of the Lebesgue integral (see Kolmogorov and Fomin, 1977), implies that we can find $\delta > 0$ such that:

- letting $T_\delta = \left\{ \mathbf{x}' = (x_1, \dots, x_{d-1}) \in [0, 1]^{d-1}, 0 \leq x_d \leq 1 - \lambda(\mathbf{x}') - \delta \right\}$, $v_H(T) - v_H(T_\delta) < \frac{\varepsilon}{16\widehat{\rho}(n)}$;
- we can tile $[0, 1]^{d-1}$ by hypercubes of side length $\frac{1}{m}$, m being sufficiently high, in such a way that for each *tile* Q_t , $1 \leq t \leq m^d$, we can consider a rectangular *hyperprism* R_t having basis Q_t and height $h_t = 1 - \delta - \min \lambda(\mathbf{x}) \leq 1 - \frac{\delta}{2} - \max \lambda(\mathbf{x})$: hence $v_H(T_\delta) \leq v_H\left(\bigcup_{t=1}^{m^d} R_t\right) \leq v_H(T)$.

Now, fixed n and $\alpha \in \left[\frac{1}{d}, \frac{2}{d+1}\right]$, by the mentioned Theorem the above construction can be reproduced (i.e. *rescaled*) for any *trapezoid* $\widehat{Q}_k^n \cap S_n$, replacing 1 by $1 - (1 - \alpha)^n$ and $1 - \delta$ by $(1 - (1 - \alpha)^n)(1 - \delta)$. Moreover, set $\widehat{Q}_k^n = Q(\mathbf{b}, l)$, $l > 0$, and define $Q' = Q(\mathbf{b}, l) \cap \{x_d = b_d\}$. Then

$$\widehat{Q}_k^n \cap S_n = \left\{ \mathbf{x}' = (x_1, \dots, x_{d-1}) \in Q', b_d \leq x_d \leq (1 - (1 - \alpha)^n) \left(1 - \frac{\delta}{2}\right) - \lambda(\mathbf{x}') \right\}$$

We can also choose δ so small that, if \widehat{Q}_k^n has an intersection of positive volume with S_n , then it has an intersection of positive volume also with $S_n^\delta = \{0 \leq \lambda(\mathbf{x}) \leq 1 - (1 - \alpha)^n - \delta, x_1, \dots, x_d \geq 0\}$.

Now we can consider an analytic distribution H' such that in

$$T_{\frac{\delta}{2}} = \left\{ \mathbf{x}' = (x_1, \dots, x_{d-1}) \in [0, 1]^{d-1}, 0 \leq x_d \leq (1 - (1 - \alpha)^n) \left(1 - \frac{\delta}{2}\right) - \lambda(\mathbf{x}') \right\}$$

$|H'(\mathbf{x}) - H(\mathbf{x})| < \frac{\varepsilon}{16\widehat{\rho}(n)(2m)^d}$, while in $T_0 - T_{\left(\frac{\delta}{2} - \sigma\right)}$, for any arbitrarily small $\sigma > 0$, the density of H' can be chosen as small as we want. Moreover, by Proposition 1 of § 1.4, S_n has a cover of hypercubes, from which we can extract a cover of $p \leq \widehat{\rho}(n)$ not overlapping rectangular *hyperprisms*, to which the above construction can be analogously applied.

Then, through straightforward steps, it follows that for any arbitrarily small $\varepsilon > 0$

$$|f_n^H(\alpha)| < \varepsilon$$

Hence $f_n^H(\alpha) = 0$ ■

References

- P. Artzner, F. Delbaen, J.M. Eber, D.Heath, Coherent measures of risk, *Mathematical Finance* 9 (1999), 203-228
- P. Arbenz, P. Embrechts, G. Puccetti, The AEP algorithm for the fast computation of the distribution of the sum of dependent random variables, *Bernoulli* 17(2) (2011), 562–591.
- P. Arbenz, P. Embrechts, G. Puccetti, The GAEP algorithm for the fast computation of the distribution of a function of dependent random variables, *Stochastics* 84 (2012), 569-597
- P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer, New York, NY, 2004
- A. N. Kolmogorov, S.V. Fomin, *Eléments de la théorie des fonctions et de l'analyse fonctionnelle*, Editions MIR, Moscou, 1977