Latin hypercube sampling with dependence and applications in finance

Natalie Packham and Wolfgang M. Schmidt

June 10, 2009

The authors thank Stefanie Müller from University of Kaiserslautern and members of the Financial Mathematics group at Fraunhofer Institut für Techno-und Wirtschaftsmathematik in Kaiserslautern for helpful comments.

Abstract: In Monte Carlo simulation, Latin hypercube sampling (LHS) [McKay *et al.* (1979)] is a well-known variance reduction technique for vectors of independent random variables. The method presented here, Latin hypercube sampling with dependence (LHSD), extends LHS to vectors of dependent random variables. The resulting estimator is shown to be consistent and asymptotically unbiased. For the bivariate case and under some conditions on the joint distribution, a Central Limit Theorem together with a closed formula for the limit variance are derived. It is shown that for a class of estimators satisfying some monotonicity condition, the LHSD limit variance is never greater than the corresponding Monte Carlo limit variance. In some valuation examples of financial payoffs, when compared to standard Monte Carlo simulation, a variance reduction of factors up to 200 is achieved. We illustrate that LHSD is suited for problems with rare events and for high-dimensional problems, and that it may be combined with Quasi-Monte Carlo methods.

Keywords: Monte Carlo simulation, variance reduction, Latin hypercube sampling, stratified sampling

JEL classification: C15, C63, G12

Contents

1. Introduction

Consider the problem of reducing the variance of a Monte Carlo estimator targeted at a vector of dependent random variables. Many existing variance reduction techniques are powerful, but exploit particular properties of the problem at hand; see [Glasserman (2004), Section 4.7] for a comparison of variance reduction techniques taking into account their complexity and effectiveness. The method proposed here, *Latin hypercube sampling with dependence (LHSD)*, is generally applicable, it is particularly simple, and it achieves an effective variance reduction for many estimation problems, including problems with rare events and high-dimensional problems. It is often effective even for low sample sizes, and it may easily be combined with other variance reduction techniques.

LHSD is a generalisation of a multivariate variance reduction technique known as *Latin hypercube sampling (LHS)*, introduced by [McKay *et al.* (1979)] and further studied by [Stein (1987)] and [Owen (1992)], amongst others. LHS relies on independence of the components of the random vector involved. Essentially, LHSD extends LHS to random vectors with dependent components. The method is mentioned by [Stein (1987)], but, to the best of our knowledge, it has not been analysed in detail and no results about its effectiveness have been derived yet.

On a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, let $(U^1, \dots, U^d)$ be a random vector with uniform marginals and with copula[a] C. Suppose the goal is to estimate $\mathbb{E}g(U^1, \dots, U^d)$ with $g : [0,1]^d \to \mathbb{R}$ Borel-measurable and C-integrable.

The usual Monte Carlo estimator based on n independent samples $(U_i^1, \dots, U_i^d)$, $i = 1, \dots, n$, is $1/n \sum_{i=1}^n g(U_i^1, \dots, U_i^d)$. It is a strongly consistent estimator, i.e., $1/n \sum_{i=1}^n g(U_i^1, \dots, U_i^d) \xrightarrow{\mathbf{P}\text{–a.s.}} \mathbb{E}g(U^1, \dots U^d)$ as $n \to \infty$. The Central Limit Theorem for sums of independent random variables states that the scaled estimator converges in distribution to a Normal distribution, i.e., $1/\sqrt{n} \sum_{i=1}^n \left[g(U_i^1, \dots, U_i^d) - \mathbb{E}g(U^1, \dots, U^d)\right] \xrightarrow{\mathcal{L}} \mathrm{N}(0, \sigma^2)$, with $\sigma^2 = \mathrm{Var}(g(U^1, \dots, U^d))$. The Central Limit Theorem serves as an indicator of the speed of convergence via the approximation $1/n \sum_{i=1}^n \left[g(U_i^1, \dots, U_i^d) - \mathbb{E}g(U^1, \dots U^d)\right] \approx X$, for some $X \sim \mathrm{N}(0, \sigma^2/n)$, from which we may derive confidence intervals and other statistics. In general, the variance of an estimator is a key figure for assessing the quality of an estimation.

LHSD transforms n independent samples $(U_i^1, \dots, U_i^d)$, $i = 1, \dots, n$, in such a way that for each dimension j, the marginals U_i^j, $i = 1, \dots, n$, are uniformly spread over $[0,1]$. At the same time, the transformation aims to preserve the copula. We show that the LHSD estimator of $\mathbb{E}g(U^1, \dots, U^d)$ is strongly consistent for bounded and C-a.e. continuous g. In the bivariate case, under some moderate conditions on the copula C of the underlying random vector, we derive a Central Limit Theorem, which states that the LHSD estimator converges to a Normal distribution. The Central Limit Theorem is derived by applying a result from [Fermanian *et al.* (2004)]. We show that, under some monotonicity conditions on g, the limit variance of the LHSD estimator is never greater than the respective Monte Carlo limit variance.

Monte Carlo simulation is widely used for the valuation of financial claims. The general approach to value a financial claim is to generate sample paths of the underlying financial securities. The discounted expectation of the claim's payoff under a risk-neutral measure is then an estimator of the claim's fair value. For a comprehensive overview of Monte Carlo simulation in financial applications, we refer to [Glasserman (2004)].

We consider two examples of financial claims that depend on the joint distribution of several underlying assets. A first-to-default credit basket is valued based on random numbers and Sobol sequences, both with and without LHSD. The variance (resp. mean square error) of the LHSD estimators is between 2.25 and 4 times smaller compared to the corresponding estimators without LHSD. Even though defaults are rare events, LHSD guarantees that for each credit in the credit basket a fixed number of default events is sampled; this is a consequence of the marginals being uniformly spread on $[0,1]$. The second example is concerned with the valuation of an Asian basket

[a] A copula C is the distribution function of a random vector with uniform marginals, see e.g. [Joe (1997)] and [Nelsen (1999)]. We also associate with C the measure induced by the copula C.

option, which may be formulated as a high-dimensional estimation problem (dimension 2500 in the example). The variance reduction achieved depends on the strike of the option and lies between factors of 6 and 200.

The outline of the paper is as follows: In Section 2 we introduce stratified sampling, a univariate variance reduction technique, and its multivariate extension, Latin hypercube sampling. We present the LHSD method in Section 3. Section 4 contains statements about the consistency and unbiasedness of the LHSD estimator. In Section 5, restricting ourselves to the bivariate case and under some conditions on the copula, we provide a Central Limit Theorem and we analyse the rate of convergence of the LHSD estimator. In Section 6 we show that the LHSD estimator for random vectors with uniform marginals extends naturally to random vectors with nonuniform marginals. As example applications we consider the valuation of first-to-default credit baskets and Asian basket options in Section 7.

2. Preliminaries

2.1. *Stratified sampling*

Stratified sampling is a variance reduction technique in a univariate setting that constrains the fraction of samples drawn from specific subsets, so-called *strata*. For a detailed exposition we refer to [Glasserman (2004), Chapter 4.3].

Suppose the goal is to estimate $\mathbb{E}g(U)$ with $U \sim U(0,1)$ (i.e., a uniform random variable on $[0,1]$), and with $g : [0,1] \to \mathbb{R}$ a Borel-measurable and integrable function. Let $A_1, \ldots, A_n$ be a partition of $[0,1]$. Then,

$$\mathbb{E}g(U) = \sum_{i=1}^{n} \mathbb{E}(g(U)|U \in A_i)\mathbf{P}(U \in A_i),$$

and a corresponding estimator of $\mathbb{E}g(U)$ is derived from sampling U conditional on $\{U \in A_i\}$, $i = 1, \ldots, n$. In the simplest case, the strata are chosen to be the equiprobable intervals $A_i = ((i-1)/n, i/n]$, $i = 1, \ldots, n$, and one sample is drawn from each stratum. This is achieved for example by drawing independent $U(0,1)$ samples, $U_1, \ldots, U_n$, and setting

$$V_i := \frac{i-1}{n} + \frac{U_i}{n}, \quad i = 1, \ldots, n. \tag{1}$$

The resulting estimator of $\mathbb{E}g(U)$, given by $1/n \sum_{i=1}^{n} g(V_i)$, is consistent, and by a Central Limit Theorem for the stratified estimator it follows that the limit variance is smaller than the Monte Carlo variance, cf. [Glasserman (2004), Section 4.3.1].

2.2. *Latin hypercube sampling*

Simply extending stratified sampling to d-dimensional random vectors by stratifying each dimension with n samples is unfeasible even for moderately small dimensions, since to have one sample in each stratum requires at least n^d samples. Latin hypercube sampling (LHS) efficiently extends stratified sampling to random vectors $(U^1, \ldots, U^d)$ whose components are independent (i.e., they are linked by the independence copula). It was introduced in [McKay *et al.* (1979)] and further developed by [Stein (1987)] and [Owen (1992)]. For an in-depth treatment of LHS see [Glasserman (2004), Section 4.4].

Assume that the goal is to estimate $\mathbb{E}g(U^1, \ldots, U^d)$ with $g : [0,1]^d \to \mathbb{R}$ Borel-measurable and integrable. Fixing a sample size n, generate n independent samples $(U_i^1, \ldots, U_i^d)$, $i = 1, \ldots, n$, and generate d independent permutations $\pi^1, \ldots, \pi^d$ of $\{1, \ldots, n\}$ drawn from the distribution that makes all permutations equally probable. Denoting by π_i^j the value to which i is mapped by the j-th permutation, a Latin hypercube sample is given by

$$V_i^j := \frac{\pi_i^j - 1}{n} + \frac{U_i^j}{n}, \quad j = 1, \ldots, d, \quad i = 1, \ldots, n.$$

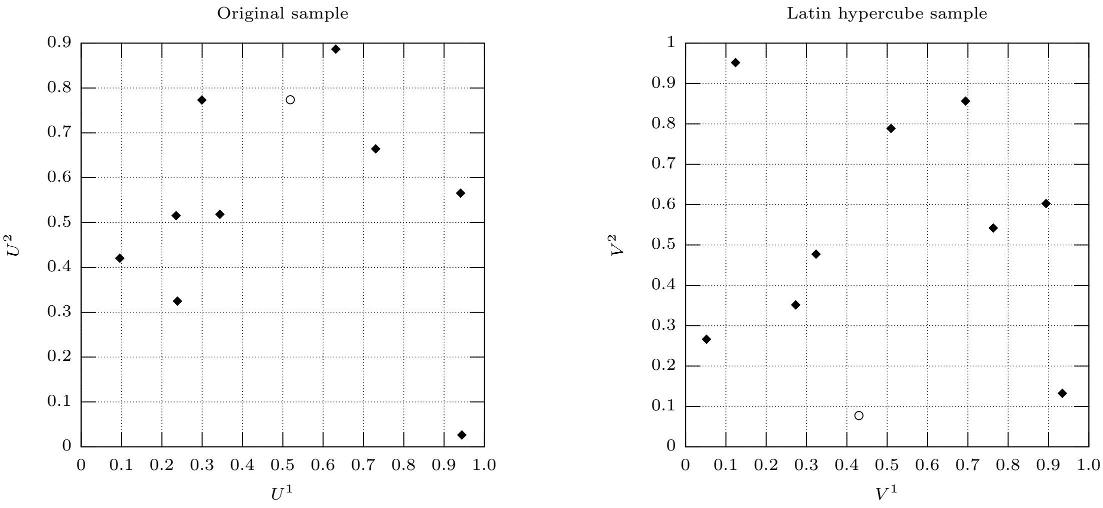

Fig. 1. Left: Original sample $(U_1^1, U_1^2), \ldots, (U_{10}^1, U_{10}^2)$, with (U_1^1, U_1^2) marked by a circle. Right: Corresponding Latin hypercube sample, with (V_1^1, V_1^2) marked by a circle. The permutations are $\pi^1 = \{5, 9, 7, 8, 1, 10, 4, 2, 3, 6\}$ and $\pi^2 = \{1, 7, 9, 6, 3, 2, 5, 10, 4, 8\}$.

An example of a Latin hypercube sample is shown in Figure 1. Observe that in each dimension j, $(V_1^j, \ldots, V_n^j)$ is a stratified sample. Furthermore, each point $(V_i^1, \ldots, V_i^d)$, is uniformly distributed on $[0,1]^d$, $1 \le i \le n$. The LHS estimator $1/n \sum_{i=1}^n g(V_i^1, \ldots, V_i^d)$ is consistent. [Stein (1987)] shows that, for functions g with finite second moment, the variance of the LHS estimator is smaller compared to the standard Monte Carlo estimator as long as the number of samples is sufficiently large. For bounded g, [Owen (1992)] derives a Central Limit Theorem for the LHS estimator.

Requiring independence of the components of the random vector is fundamental: Applying LHS to a sample of a random vector whose components are dependent destroys the dependence by application of random and independent permutations in each dimension. Conversely, applying first LHS to a sample of a random vector with independent components, and then applying a transform to introduce dependence breaks, in general, the stratification of the marginals, thereby possibly losing much of the appeal of LHS; see also [Glasserman (2004), p. 238]. Nonetheless, in our examples we compare LHSD with this naive method of applying LHS.

3. Latin hypercube sampling with dependence

We now describe an extension of LHS for random vectors with dependence. The general idea is to generate a Latin hypercube sample, albeit with the following modification: Instead of choosing a random permutation in each dimension, a particular permutation that depends on the samples of that dimension is chosen. For this we need the notion of a rank statistic.

Definition 1 (Rank statistic). *Let $X_1, \ldots, X_n$ be i.i.d. random variables with continuous distribution function. Reorder them such that $X^{(1)} < \ldots < X^{(n)}$ $\boldsymbol{P}$–a.s.. The index of X_i within $X^{(1)}, \ldots, X^{(n)}$ is the* i-th rank statistic*, given by*

$$r_{i,n}(X_1, \ldots, X_n) := \sum_{k=1}^n \mathbf{1}_{\{X_k \le X_i\}}. \tag{2}$$

That such an ordering exists $\mathbf{P}$–a.s. follows from the continuity of the distribution function. For ease of notation, we write just $r_{i,n}$ instead of $r_{i,n}(X_1, \ldots, X_n)$.

Consider a random vector $(U^1, \ldots, U^d)$, $U^j \sim U(0,1)$, $j = 1, \ldots, d$, whose components are linked by an arbitrary copula C, and let $(U_i^1, \ldots, U_i^d)$, $i = 1, \ldots, n$, be n independent samples of $(U^1, \ldots, U^d)$. For $1 \le i \le n$ and $1 \le j \le d$ denote by $r_{i,n}^j$ the i-th rank statistic of $(U_1^j, \ldots, U_n^j)$. A

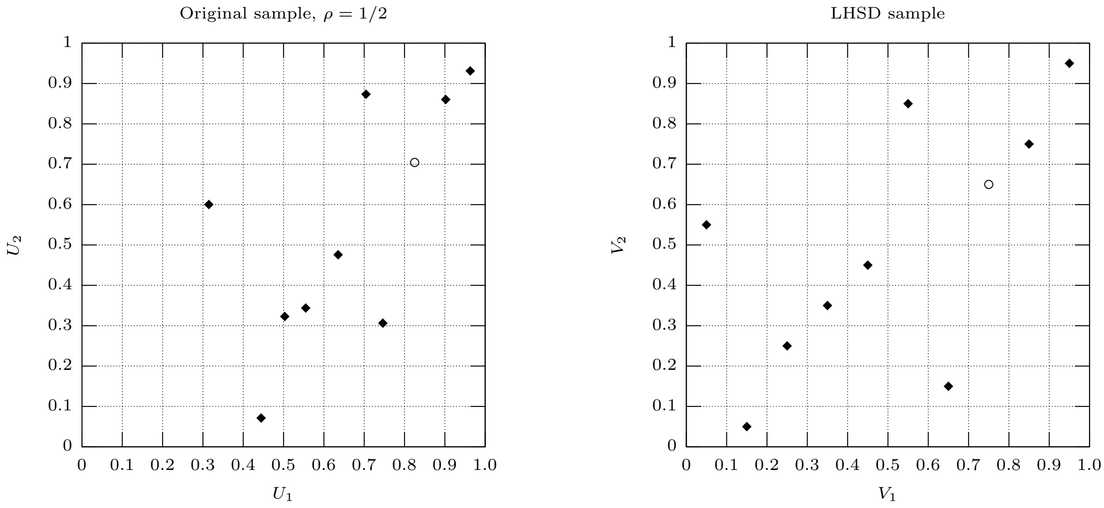

Fig. 2. Left: Original sample $(U_1^1, U_1^2), \ldots, (U_{10}^1, U_{10}^2)$ linked with a Gaussian copula with correlation $\rho = 1/2$; (U_1^1, U_1^2) is marked by a circle. Right: Corresponding LHSD sample, with $(V_{1,10}^1, V_{1,10}^2)$ marked by a circle. The rank statistics are $r^1 = \{8, 6, 1, 4, 3, 7, 5, 2, 9, 10\}$ and $r^2 = \{7, 9, 6, 4, 3, 2, 5, 1, 8, 10\}$, and $\eta_{i,10}^j := 1/2$, $j = 1, 2$, $i = 1, \ldots, 10$.

Latin hypercube sample with dependence is given by

$$V_{i,n}^j := \frac{r_{i,n}^j - 1}{n} + \frac{\eta_{i,n}^j}{n}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, d, \tag{3}$$

where $\eta_{i,n}^j$ are random variables taking values in $[0, 1]$, for which we specify some examples below. Figure 2 shows an example with 10 samples drawn from a bivariate Gaussian copula with correlation 1/2 and the corresponding LHSD samples.

Just as in regular LHS, $(V_1^j, \ldots, V_n^j)$ is a stratified sample in each dimension j. Recall that each sample from the stratified sample of Equation (1) is uniformly distributed within its stratum. If $\eta_{i,n}^j := U_i^j$ this property is lost by application of the rank statistic: in each dimension, the smallest sample is allocated to the first stratum, the second smallest to the second stratum, and so on. Conditional on $\{r_{i,n}^j = k\}$, U_i^j follows a beta distribution with parameters k and n, i.e., $\mathbf{P}(U_i^j \leq x | r_{i,n}^j = k) = B_k^n(x)$, which is the distribution of the k-th order statistic of n independent uniform random variables, see e.g. [Feller (1971), Ch. I.7]. The following choices produce a LHSD sample with uniform marginals:

(i) $\eta_{i,n}^j := B_{r_{i,n}^j}^n(U_i^j)$, $i = 1, \ldots, n$, $j = 1, \ldots, d$,

(ii) $(\eta_{i,n}^j)_{i=1,\ldots,n; j=1,\ldots,d}$ is a sample of independent $U(0, 1)$ random variables independent of $(U_i^j)_{i=1,\ldots,n; j=1,\ldots d}$.

If the primary goal is to capture the joint distribution, the following choices are computationally more efficient:

(iii) $\eta_{i,n}^j := 1/2$, which places each sample in the middle of its stratum,

(iv) $\eta_{i,n}^j := 1$, in which case $V_{i,n}^j$ is just the empirical distribution function of $(U_1^j, \ldots, U_n^j)$ at U_i^j,

$i = 1, \ldots, n$, $j = 1, \ldots, d$.

Remark 2. LHS is a special case of LHSD: Let $(U^1, \ldots, U^d)$ be independent, and let $(\eta_i^j)_{i=1,\ldots n; j=1,\ldots d}$ be chosen according to choice (ii). Then $(U_i^j)_{i=1,\ldots,n; j=1,\ldots,d}$ determine independent and equiprobable permutations that allocate samples to strata, and $(\eta_i^j)_{i=1,\ldots n; j=1,\ldots d}$ determine independently the position, uniformly distributed, of each sample in its stratum.

Assume that the quantity to estimate is $\mathbb{E}g(U^1,\dots,U^d)$ with $g:[0,1]^d\to\mathbb{R}$ Borel-measurable and integrable and $(U^1,\dots,U^d)$ a random vector with uniform marginals and copula C. The LHSD estimator is given by

$$\frac{1}{n}\sum_{k=1}^{n} g(V_{i,n}^1,\dots,V_{i,n}^d), \tag{4}$$

with $V_{i,n}^j$, $i=1,\dots,n$, $j=1,\dots,d$, obtained from the transformation of Equation (3).

Before we analyse the estimator formally, let us reflect why it would reduce the variance: Variance reduction over the usual Monte Carlo estimator is achieved by drawing "favourable" samples and avoiding "unfavourable" samples (i.e., samples with a large contribution to the variance of the estimator). For each dimension $1\le j\le d$, LHSD ensures that the samples $V_{1,n}^j,\dots,V_{n,n}^j$ are uniformly spread over the unit interval, thereby deleting inter-stratum variance and leaving only intra-stratum variance. As a consequence however, in general, the original dependence structure of the samples is broken, i.e., for fixed n, the copula of $(V_{i,n}^1,\dots,V_{i,n}^d)$, $i=1,\dots,n$, differs from the copula of $(U^1,\dots,U^d)$. On the other hand, as $n\to\infty$, each sample $V_{i,n}^j$ converges to U_i^j, since the fraction of samples $V_{k,n}^j$, $k=1,\dots,n$, such that $V_{k,n}^j\le V_{i,n}^j$, tends to U_i^j. This notion is captured by the rank statistic. We shall see below in Lemma 5 that the empirical distribution function of the LHSD samples tends to the original copula C. Summarising, an LHSD sample has marginals that are uniformly spread over the unit interval and, provided n is large enough, we can expect the error between the original copula and the copula of the LHSD samples to be small.

4. Consistency of the LHSD estimator

We establish consistency of the LHSD estimator, provided that g is bounded and fulfills some continuity conditions.

Observe that the usual laws of large numbers for sums of independent random variables do not apply, for the following reasons:

- In each dimension, by application of the rank statistic, the samples fail to be independent.
- For any i,j, $V_{i,n}^j\neq V_{i,n+1}^j$, hence, when progressing from n to $n+1$, we are not just adding an $(n+1)$-th term to the existing sum (4), but all terms of the sum change.

Proposition 3. *Let $g:[0,1]^d\to\mathbb{R}$ be bounded and continuous C-a.e.. Then the LHSD estimator (4) is strongly consistent, i.e.,*

$$\frac{1}{n}\sum_{i=1}^{n} g(V_{i,n}^1,\dots,V_{i,n}^d)\overset{\boldsymbol{P}\text{–a.s.}}{\longrightarrow}\mathbb{E}g(U^1,\dots,U^d),\qquad \text{as } n\to\infty.$$

It follows immediately by Dominated Convergence that the estimator is asymptotically unbiased:

Corollary 4. *Let $g:[0,1]^d\to\mathbb{R}$ be bounded and continuous C-a.e.. Then the LHSD estimator (4) is asymptotically unbiased, i.e.,*

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} g(V_{i,n}^1,\dots,V_{i,n}^d)\right)\longrightarrow\mathbb{E}g(U^1,\dots,U^d),\qquad \text{as } n\to\infty.$$

We require a preliminary result for the proof of Proposition 3.

Lemma 5. *For $0\le u^1,\dots,u^d\le 1$, define $C_n:[0,1]^d\to[0,1]$ by*

$$C_n(u^1,\dots,u^d):=\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\{V_{k,n}^1\le u^1,\dots,V_{k,n}^d\le u^d\}}.$$

Then C_n is a distribution function and

$$\sup_{(u^1,\dots,u^d)\in[0,1]^d}\left|C_n(u^1,\dots,u^d)-C(u^1,\dots,u^d)\right|\overset{\boldsymbol{P}\text{–a.s.}}{\longrightarrow}0,\qquad \text{as } n\to\infty.$$

Proof. It is straightforward to verify that C_n is a distribution function on $[0,1]^d$, $n \in \mathbb{N}$. For the second statement, let

$$F_n^j(u) = \frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{U_k^j \le u\}}, \quad u \in [0,1],$$

be the empirical distribution function based on $U_1^j, \dots, U_n^j$, $j = 1, \dots, n$. Define $\tilde{C}_n : [0,1]^d \to [0,1]$ as

$$\tilde{C}_n(u^1, \dots, u^d) := \frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{F_n^1(U_k^1) \le u^1, \dots, F_n^d(U_k^d) \le u^d\}}. \tag{5}$$

It is a consequence of [Deheuvels (1979), Théorème 3.1] (or [Deheuvels (1981), Lemmas 6 and 7]) that

$$\sup_{(u^1,\dots,u^d)\in[0,1]^d} \left|\tilde{C}_n(u^1, \dots, u^d) - C(u^1, \dots, u^d)\right| \quad \overset{\mathbf{P}\text{–a.s.}}{\longrightarrow} \quad 0, \qquad \text{as } n \to \infty.$$

Using the fact that $F_n^j(U_k^j) = r_{k,n}^j/n$, the claim follows from

$$|C_n(u^1, \dots, u^d) - \tilde{C}_n(u^1, \dots, u^d)| \le \frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\left\{u^1 \in \left[\frac{r_{k,n}^1 - 1}{n}, \frac{r_{k,n}^1}{n}\right), \dots, u^d \in \left[\frac{r_{k,n}^d - 1}{n}, \frac{r_{k,n}^d}{n}\right)\right\}} \le \frac{1}{n},$$

for any $(u^1, \dots, u^d) \in [0,1]^d$. □

Proof of Proposition 3. Assume first that g is continuous and observe that

$$\int_{[0,1]^d} g \, \mathrm{d}C_n = \frac{1}{n}\sum_{k=1}^n g(V_{k,n}^1, \dots, V_{k,n}^d),$$

which is just the LHSD estimator. It follows from Lemma 5 that C_n converges weakly to C for $\mathbf{P}$-almost all $\omega \in \Omega$, which is equivalent to

$$\int_{[0,1]^d} g \, \mathrm{d}C_n \quad \longrightarrow \quad \int_{[0,1]^d} g \, \mathrm{d}C = \mathbb{E}g, \quad \text{for } \mathbf{P} - \text{a.a. } \omega, \tag{6}$$

for every bounded, continuous function $g : [0,1]^d \to \mathbb{R}$. That Equation (8) holds for functions g that are C-a.e. continuous is a consequence of [Billingsley (1968), Theorem 5.2]. □

Remark 6. To relax the condition of boundedness on g, one would have to show that $g(V_{i,n}^1, \dots, V_{i,n}^d)$, $n \in \mathbb{N}$, are uniformly integrable. However, it seems that we have no means of establishing uniform integrability other than requiring boundedness, as in general the distribution of $(V_{i,n}^1, \dots, V_{i,n}^d)$ is not known. On the other hand, boundedness is an acceptable limitation when doing Monte Carlo simulation, since in most practical applications the function g can be sufficiently approximated by a bounded function, and ultimately, the output produced by a computer is bounded.

5. Central Limit Theorem for LHSD and variance reduction

It is natural to investigate the speed of convergence of the LHSD estimator and compare this to the rate of convergence of the standard Monte Carlo estimator. Assuming the bivariate case and posing some conditions on the copula, we state a Central Limit Theorem for the LHSD estimator and we establish that the limit distribution is Normal. We derive a closed-form expression for the LHSD estimator's limit variance, and we compare it to the corresponding Monte Carlo limit variance. Finally, we show that if the copula fulfills a certain positive dependence property and if the function to be estimated is nondecreasing in each argument, then the LHSD limit variance is always less or equal to the corresponding MC limit variance.

The empirical distribution function of the LHSD samples bears close resemblance to the empirical copula of the original sample, and it turns out the LHSD estimator is a special case of some multivariate rank-order statistics. For the study of empirical processes and empirical copulas, see e.g. [Deheuvels (1979)], [Deheuvels (1981)], [Gaenssler and Stute (1987)] and [Vaart and Wellner (1996)], [Fermanian *et al.* (2004)]. For results on multivariate rank-order statistics we refer to [Ruymgaart *et al.* (1972)], [Rüschendorf (1976)], [Genest *et al.* (1995)] and [Fermanian *et al.* (2004)]. The Central Limit Theorem stated below is derived from Theorem 6 of [Fermanian *et al.* (2004)]. Although the following analysis is restricted to the bivariate case, we presume that it can be extended to the multivariate case.

Definition 7. *A function $g : [0,1]^2 \to \mathbb{R}$ is of* bounded variation (in the sense of Hardy-Krause)*, if there exists a constant K such that*

(i) for every bounded rectangle $[a,b] \times [c,d] \subseteq [0,1]^2$, for all m, n and points $a = x_0 < x_1 < \cdots x_m = b$, $c = y_0 < y_1 < \cdots < y_n = d$,

$$\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |g(x_i, y_j) + g(x_{i+1}, y_{j+1}) - g(x_i, y_{j+1}) - g(x_{i+1}, y_j)| \leq K,$$

(ii) for every $u \in [0,1]$, $v \mapsto g(u,v)$ is a function whose variation is bounded by K,
(iii) for every $v \in [0,1]$, $u \mapsto g(u,v)$ is a function whose variation is bounded by K.

Note that there are different definitions of bounded variation in the bivariate case, see [Clarkson and Adams (1933)]. We use the term "bounded variation" as a synonym of "bounded variation in the sense of Hardy-Krause". For illustration we list some properties of bounded variation functions. It is a consequence of [Hobson (1921), §308] that if $g : [0,1]^2 \to \mathbb{R}$ is of bounded variation, then $\lim_{n\to\infty} g(u_n^1, u_n^2)$ exists for any sequence $(u_n^1, u_n^2)_{n\geq 1}$, with $(u_n^j)_{n\geq 1}$ monotone, $j = 1, 2$. By [Adams and Clarkson (1934), Corollary to Theorem 13], the discontinuities of a function of bounded variation are located on a denumerable number of parallels to the axes. Finally, note that a function of bounded variation is bounded [Clarkson and Adams (1933), p. 827].

Definition 8. *A function $g : [0,1]^2 \to \mathbb{R}$ is* right-continuous *if for any sequence $(u_n^1, u_n^2)_{n\geq 1}$, with $u_n^j \downarrow u^j$, $j = 1, 2$, $\lim_{n\to\infty} g(u_n^1, u_n^2) = g(u^1, u^2)$.*

See [Kallenberg (2001), Theorem 4.28] or [Jacod and Protter (2003), Theorem 18.8] for the following Lemma:

Lemma 9. *Let $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ be sequences of $\mathbb{R}$-valued random variables, with $X_n \xrightarrow{\mathcal{L}} X$ and $|X_n - Y_n| \xrightarrow{\mathbf{P}} 0$. Then $Y_n \xrightarrow{\mathcal{L}} X$.*

In the following, all integrals are Lebesgue-Stieltjes integrals and integrals are over $(0,1]$ if not stated otherwise. Throughout U, V are $U(0,1)$-distributed random variables.

Theorem 10 (Central Limit Theorem for LHSD). *Let the copula C of (U,V) have continuous partial derivatives and let $g : [0,1]^2 \to \mathbb{R}$ be of bounded variation and right-continuous. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(g(V_{i,n}^1, V_{i,n}^2) - \mathbb{E}g(U^1, U^2) \right) \quad \xrightarrow{\mathcal{L}} \quad \mathrm{N}(0, \sigma_{\mathrm{LHSD}}^2),$$

where, setting $\partial_1 C(u,v) = \partial C(u,v)/\partial u$ *and* $\partial_2 C(u,v) = \partial C(u,v)/\partial v$,

$$
\begin{aligned}
\sigma^2_{\mathrm{LHSD}} = {} & \iiiint C(u\wedge u', v\wedge v')\,\mathrm{d}g(u,v)\,\mathrm{d}g(u',v') - \left(\iint C(u,v)\,\mathrm{d}g(u,v)\right)^2 \\
& + \iiiint \Big\{ \partial_1 C(u',v')(C(u,v)u' - C(u\wedge u', v)) + \partial_1 C(u,v)(C(u',v')u - C(u\wedge u', v')) \\
& \qquad + \partial_2 C(u',v')(C(u,v)v' - C(u, v\wedge v')) + \partial_2 C(u,v)(C(u',v')v - C(u', v\wedge v')) \\
& \qquad + \partial_1 C(u,v)\,\partial_1 C(u',v')(u\wedge u' - uu') + \partial_2 C(u,v)\,\partial_2 C(u',v')(v\wedge v' - vv') \qquad (7) \\
& \qquad + \partial_1 C(u,v)\,\partial_2 C(u',v')(C(u,v') - uv') + \partial_1 C(u',v')\,\partial_2 C(u,v)(C(u',v) - u'v) \Big\} \\
& \qquad \mathrm{d}g(u,v)\,\mathrm{d}g(u',v').
\end{aligned}
$$

Remark 11. Observe that the conditions of bounded variation and right-continuity are required for the right-hand side of Equation (7) to be defined.

Proof. Theorem 6 of [Fermanian *et al.* (2004)] states that, under the above conditions on g and C,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(g(F_n^1(U_i^1), F_n^2(U_i^2)) - \mathbb{E} g(U^1, U^2) \right) \quad \xrightarrow{\mathcal{L}} \quad \int_{[0,1]^2} \mathbb{G}_C(u,v)\,\mathrm{d}g(u,v),$$

where F_n^j is the empirical distribution function based on the sample $U_1^j, \ldots, U_n^j$, $j = 1,2$, and

$$\mathbb{G}_C(u,v) = \{ B_C(u,v) - \partial_1 C(u,v) B_C(u,1) - \partial_2 C(u,v) B_C(1,v) \},$$

with B_C a Brownian bridge on $[0,1]^2$, i.e., a Gaussian family $(B_C(u,v))_{(u,v)\in[0,1]^2}$, with mean zero and covariance function

$$\mathbb{E}(B_C(u,v) B_C(u',v')) = C(u\wedge u', v\wedge v') - C(u,v)C(u',v'), \quad 0 \le u, u', v, v' \le 1.$$

In particular, the limit distribution is Gaussian.

Recall that $V_{i,n}^j = (r_{i,n}^j - 1 + \eta_{i,n}^j)/n$ and $F_n^j(U_i^j) = r_{i,n}^j/n$, $j = 1,2$. Fix n, and, for notational convenience, set $v_i^j := V_{i,n}^j$ and $u_i^j := F_n^j(U_i^j)$, $j = 1,2$. Assume that the variation of g is bounded by K. Then,

$$
\begin{aligned}
& \left| \sum_{i=1}^n \left[g(V_{i,n}^1, V_{i,n}^2) - g(F_n^1(U_i^1), F_n^2(U_i^2)) \right] \right| = \left| \sum_{i=1}^n \left[g(v_i^1, v_i^2) - g(u_i^1, u_i^2) \right] \right| \\
&= \Bigg| \sum_{i=1}^n \Big[g(v_i^1, v_i^2) + g(u_i^1, u_i^2) - g(v_i^1, u_i^2) - g(u_i^1, v_i^2) - 2g(u_i^1, u_i^2) + g(v_i^1, u_i^2) + g(u_i^1, v_i^2) \\
&\qquad - g(v_i^1, 0) + g(u_i^1, 0) + g(0, u_i^2) - g(0, v_i^2) + g(v_i^1, 0) - g(u_i^1, 0) - g(0, u_i^2) + g(0, v_i^2) \Big] \Bigg| \\
&\le \sum_{i=1}^n \left| g(v_i^1, v_i^2) + g(u_i^1, u_i^2) - g(v_i^1, u_i^2) - g(u_i^1, v_i^2) \right| + \sum_{i=1}^n \left| g(v_i^1, u_i^2) + g(u_i^1, 0) - g(u_i^1, u_i^2) - g(v_i^1, 0) \right| \\
&\quad + \sum_{i=1}^n \left| g(u_i^1, v_i^2) + g(0, u_i^2) - g(u_i^1, u_i^2) - g(0, v_i^2) \right| \\
&\quad + \sum_{i=1}^n \left| g(v_i^1, 0) - g(u_i^1, 0) \right| + \sum_{i=1}^n \left| g(0, v_i^2) - g(0, u_i^2) \right| \quad \le 4K,
\end{aligned}
$$

since each sum consists of terms that refer to non-overlapping intervals. Hence,

$$\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left[g(V_{i,n}^1, V_{i,n}^2) - g(F_n^1(U_i^1), F_n^2(U_i^2)) \right] \right| \quad \longrightarrow \quad 0, \quad \text{as } n \to \infty,$$

and the first statement follows by Lemma 9.

The expression for σ^2_{LHSD} is obtained by taking the second moment of the limit distribution, $\mathbb{E}\left(\iint \mathbb{G}_C(u,v)\,\mathrm{d}g(u,v)\right)^2$, and applying Fubini's Theorem, which is justified as follows: g as a function of bounded variation is the difference of two quasi-monotone functions (see e.g. [Adams and Clarkson (1934), Theorem 5]) and may be written as the difference of two integrals with respect to positive measures. Since g is bounded, the conditions for Fubini's Theorem are satisfied by observing that $\mathbb{E}|XY| < \infty$ for two jointly Normal random variables X and Y. □

We now examine the relationship between σ^2_{LHSD} and the limit variance of the standard Monte Carlo estimator, denoted by σ^2_{MC}. By the usual Central Limit Theorem for sums of i.i.d. random variables,

$$\sigma^2_{\mathrm{MC}} = \mathrm{Var}(g(U,V)) = \iint g(u,v)^2\,\mathrm{d}C(u,v) - \left(\iint g(u,v)\,\mathrm{d}C(u,v)\right)^2.$$

We first derive an expression for σ^2_{LHSD} when C is the independence copula. Recall that LHSD is a generalisation of Latin hypercube sampling (cf. Remark 2), so that σ^2_{LHSD} is a different way of writing the LHS limit variance derived in [Stein (1987)] and [Owen (1992)], where by a different argument, the LHS variance is derived as the "residual from additivity" of g.

We need the following Lemma:

Lemma 12. *Let C be a copula and let $h : [0,1]^4 \to \mathbb{R}$ be bounded. Then*

$$\iiiint h(u,v,u',v')\,\mathrm{d}C(u\wedge u', v\wedge v') = \iint h(u,v,u,v)\,\mathrm{d}C(u,v).$$

Proof. Observe that $C(u\wedge u', v\wedge v')$ is a copula, since by

$$C(u\wedge u', v\wedge v') = \mathbf{P}(U \le u\wedge u', V \le v\wedge v') = \mathbf{P}(U\le u, U\le u', V\le v, V\le v'), \tag{8}$$

it is a joint probability distribution with uniform marginals. By Equation (8),

$$\mathbb{E}h(U,V,U,V) = \iiiint h(u,v,u',v')\,\mathrm{d}C(u\wedge u', v\wedge v'),$$

and the statement follows. □

Proposition 13. *Let $g : [0,1]^d \to \mathbb{R}$ be of bounded variation and right-continuous, and let C be the independence copula, i.e., $C(u,v) = uv$, $u,v \in [0,1]$. Then for independent and $U(0,1)$-distributed U^1, U^2, U^3,*

$$\sigma^2_{\mathrm{LHSD}} = \sigma^2_{\mathrm{MC}} + 2\left(\mathbb{E}g(U^1,U^2)\right)^2 - \mathbb{E}(g(U^1,U^2)g(U^1,U^3)) - \mathbb{E}(g(U^1,U^3)g(U^2,U^3)) \le \sigma^2_{\mathrm{MC}}.$$

Proof. For the first statement, by Equation (7), after some computations,

$$\sigma^2_{\mathrm{LHSD}} = \iiiint \left\{(u\wedge u')(v\wedge v') + uvu'v' - (u\wedge u')vv' - uu'(v\wedge v')\right\}\mathrm{d}g(u,v)\,\mathrm{d}g(u',v').$$

By integration by parts (see Appendix A) and Lemma 12, after some calculations,

$$\begin{aligned}
\sigma^2_{\mathrm{LHSD}} &= \iint \left(g(1,1) + g(u,v) - g(u,1) - g(1,v)\right)^2\,\mathrm{d}u\,\mathrm{d}v \\
&+ \left(\iint \left(g(1,1) + g(u,v) - g(u,1) - g(1,v)\right)\mathrm{d}u\,\mathrm{d}v\right)^2 \\
&- \iiint \left(g(1,1) + g(u,v) - g(u,1) - g(1,v)\right)\left(g(1,1) + g(u,v') - g(u,1) - g(1,v')\right)\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}v' \\
&- \iiint \left(g(1,1) + g(u,v) - g(u,1) - g(1,v)\right)\left(g(1,1) + g(u',v) - g(u',1) - g(1,v)\right)\mathrm{d}u\,\mathrm{d}u'\,\mathrm{d}v \\
&= \iiiint \left(g(1,1) + g(u,v) - g(u,1) - g(1,v)\right)\underbrace{\left(g(u,v) + g(u',v') - g(u,v') - g(u',v)\right)}_{(\star)}\mathrm{d}u\,\mathrm{d}u'\,\mathrm{d}v\,\mathrm{d}v'
\end{aligned}$$

Observe that $\iiiint (g(1,1) - g(u,1) - g(1,v))(\star)\,\mathrm{d}u\mathrm{d}u'\mathrm{d}v\mathrm{d}v' = 0$, so that

$$\begin{aligned}\sigma^2_{\mathrm{LHSD}} &= \iiiint g(u,v)\,(g(u,v) + g(u',v') - g(u,v') - g(u',v))\,\mathrm{d}u\,\mathrm{d}u'\,\mathrm{d}v\,\mathrm{d}v' \\ &= \iint g(u,v)^2\,\mathrm{d}u\,\mathrm{d}v + \left(\iint g(u,v)\,\mathrm{d}u\,\mathrm{d}v\right)^2 \\ &\quad - \iiint g(u,v)g(u,v')\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}v' - \iiint g(u,v)g(u',v)\,\mathrm{d}u\,\mathrm{d}u'\,\mathrm{d}v,\end{aligned}$$

which establishes the first statement.

For the second statement, we show that $\mathbb{E}(g(U^1,U^2)g(U^1,U^3)) \geq \mathbb{E}g(U^1,U^2)\mathbb{E}g(U^1,U^3)$. For the left-hand side we obtain by the tower law for conditional expectations and conditional independence of U^2 and U^3 given U^1,

$$\begin{aligned}\mathbb{E}(g(U^1,U^2)g(U^1,U^3)) &= \mathbb{E}\left(\mathbb{E}(g(U^1,U^2)g(U^1,U^3)|U^1\right) \\ &= \mathbb{E}\left(\mathbb{E}(g(U^1,U^2)|U^1)\,\mathbb{E}(g(U^1,U^3)|U^1)\right) \\ &= \mathbb{E}(\left(h(U^1)^2\right),\end{aligned}$$

with $h(u) = \mathbb{E}g(u,U)$, $U \sim U(0,1)$. By Jensen's inequality

$$\begin{aligned}\mathbb{E}\left(h(U^1)^2\right) &\geq \left(\mathbb{E}h(U^1)\right)^2 \\ &= \mathbb{E}\left(\mathbb{E}(g(U^1,U^2)|U^1)\right)\,\mathbb{E}\left(\mathbb{E}(g(U^1,U^3)|U^1)\right) \\ &= \mathbb{E}g(U^1,U^2)\,\mathbb{E}g(U^1,U^3).\end{aligned}$$

By establishing $\mathbb{E}(g(U^1,U^3)g(U^2,U^3)) \geq \mathbb{E}g(U^1,U^3)\mathbb{E}g(U^2,U^3)$ in the same way, the second statement follows. □

The following Proposition gives us a means of comparing σ^2_{LHSD} and σ^2_{MC}.

Proposition 14. *Let the copula C of (U,V) have continuous partial derivatives and let $g : [0,1]^2 \to \mathbb{R}$ be of bounded variation and right-continuous. Then,*

$$\begin{aligned}\sigma^2_{\mathrm{LHSD}} &= \sigma^2_{\mathrm{MC}} - 2\mathrm{Cov}(g(U,V), g(U,0)) - 2\mathrm{Cov}(g(U,V), g(0,V)) + \mathrm{Var}(g(U,0) + g(0,V)) - C_g \\ &= \mathrm{Var}(g(U,V) - g(U,0) - g(0,V)) - C_g,\end{aligned}$$

where

$$\begin{aligned}C_g = \iiiint \Big\{&(1 - \partial_1 C(u',v'))(C(u,v)u' - C(u \wedge u', v)) \;+ (1 - \partial_1 C(u,v))(C(u',v')u - C(u \wedge u', v')) \\ &+ (1 - \partial_2 C(u',v'))(C(u,v)v' - C(u, v \wedge v')) + (1 - \partial_2 C(u,v))(C(u',v')v - C(u', v \wedge v')) \\ &+ (1 - \partial_1 C(u,v)\partial_1 C(u',v'))(u \wedge u' - uu') + (1 - \partial_2 C(u,v)\partial_2 C(u',v'))(v \wedge v' - vv') \\ &+ (1 - \partial_1 C(u,v)\partial_2 C(u',v'))(C(u,v') - uv') + (1 - \partial_1 C(u',v')\partial_2 C(u,v))(C(u',v) - u'v)\Big\} \\ &\mathrm{d}g(u,v)\,\mathrm{d}g(u',v'). \end{aligned} \tag{9}$$

Proof. By Lemma 12,

$$\begin{aligned}\sigma^2_{\mathrm{MC}} = \mathrm{Var}(g(U,V)) &= \iint g(u,v)^2\,\mathrm{d}C(u,v) - \iiiint g(u,v)g(u',v')\,\mathrm{d}C(u,v)\,\mathrm{d}C(u',v') \\ &= \iiiint g(u,v)g(u',v')\,\mathrm{d}C(u \wedge u', v \wedge v') - \iiiint g(u,v)g(u',v')\,\mathrm{d}C(u,v)\,\mathrm{d}C(u',v').\end{aligned}$$

Observe that the conditions required for integration by parts (see Appendix A) are satisfied; in particular every copula is continuous [Nelsen (1999), Theorem 2.2.4]. Integration by parts yields

$$\begin{aligned}\sigma^2_{\mathrm{MC}} = &\iiiint C(u\wedge u', v\wedge v')\,\mathrm{d}g(u,v)\,\mathrm{d}g(u',v') - \Big(\iint C(u,v)\,\mathrm{d}g(u,v)\Big)^2\\
&+ \iiiint \Big\{(C(u,v)u' - C(u\wedge u', v)) + (C(u',v')u - C(u\wedge u', v'))\\
&\qquad + (C(u,v)v' - C(u, v\wedge v')) + (C(u',v')v - C(u', v\wedge v'))\\
&\qquad + (u\wedge u' - uu') + (v\wedge v' - vv')\\
&\qquad + (C(u,v') - uv') + (C(u',v) - u'v)\Big\}\,\mathrm{d}g(u,v)\,\mathrm{d}g(u',v')\\
&+ 2\mathrm{Cov}(g(U,V), g(U,0)) + 2\mathrm{Cov}(g(U,V), g(0,V)) - \mathrm{Var}(g(U,0) + g(0,V)).\end{aligned}$$

The first statement follows by combination with Equation (7). The second statement follows from

$$\begin{aligned}&2\mathrm{Cov}(g(U,V), g(U,0)) + 2\mathrm{Cov}(g(U,V), g(0,V)) - \mathrm{Var}(g(U,0) + g(0,V))\\
&\qquad = \mathrm{Var}(g(U,V)) - \mathrm{Var}(g(U,V) - g(U,0) - g(0,V)). \quad \square\end{aligned}$$

For copulas with a specific dependence property and assuming that g is nondecreasing in each argument, σ^2_{LHSD} is never greater than σ^2_{MC} as we now show. For a comprehensive treatment of dependence properties of copulas, see [Nelsen (1999), Section 5.2] and [Joe (1997), Section 2.1].

Let X and Y be two random variables. We say that Y *is right-tail increasing in* X if, for all y, $x \mapsto \mathbf{P}(Y > y|X > x)$ is nondecreasing. If X and Y are continuous random variables whose copula C has continuous partial derivatives, then Y is right-tail increasing in X if and only if

$$\partial_1 C(u,v) \geq \frac{v - C(u,v)}{1-u}, \quad u, v \in [0,1],$$

cf. [Nelsen (1999), Corollary 5.2.6]. We say that C is *RTI* if X is right-tail increasing in Y and Y is right-tail increasing in X. An example of a copula that is RTI and that has continuous partial derivatives is the bivariate Normal copula with parameter $\rho \in (0,1)$; see [Joe (1997), Secion 5.1] for a comprehensive list of one- and two-parameter copulas that are RTI.

Proposition 15. *Let the copula C be RTI and have continuous partial derivatives and let $g : [0,1]^2 \to \mathbb{R}$ be right-continuous, of bounded variation and monotone nondecreasing in each argument. Then $\sigma^2_{\mathrm{LHSD}} \leq \sigma^2_{\mathrm{MC}}$.*

Proof. First note that if C is RTI then $C(u,v) \geq uv$, for all $u, v \in [0,1]$ (this property is called *positive quadrant dependence*).

Under the conditions stated, $\mathrm{Var}(g(U,V)) \geq \mathrm{Var}(g(U,V) - g(U,0) - g(0,V))$, which can be verified for example by integration by parts. It remains to be established that C_g given by Equation (9) is nonnegative. Consider first the case $u \leq u'$ and the first, second, fifth and seventh term of the integral of Equation (9):

$$
\begin{aligned}
&(1-\partial_1 C(u',v'))(C(u,v)u' - C(u,v)) + (1-\partial_1 C(u,v))(C(u',v')u - C(u,v')) \\
&\quad + (1-\partial_1 C(u,v)\partial_1 C(u',v'))(u-uu') + (1-\partial_1 C(u,v)\partial_2 C(u',v'))(C(u,v')-uv') \\
&= (1-\partial_1 C(u',v'))(1-u')(u-C(u,v)) - (1-\partial_1 C(u,v))u(v'-C(u',v')) \\
&\quad + \partial_1 C(u',v')(1-\partial_1 C(u,v))u(1-u') + \partial_1 C(u,v)(1-\partial_2 C(u',v'))(C(u,v')-uv') \\
&= (1-\partial_1 C(u',v'))(1-u')(u-C(u,v)) - (1-\partial_1 C(u,v))u\frac{v'-C(u',v')}{1-u'}(1-u') \\
&\quad + \partial_1 C(u',v')(1-\partial_1 C(u,v))u(1-u') + \partial_1 C(u,v)(1-\partial_2 C(u',v'))(C(u,v')-uv') \\
&\overset{\text{RTI}}{\geq} (1-\partial_1 C(u',v'))(1-u')(u-C(u,v)) - (1-\partial_1 C(u,v))u\partial_1 C(u',v')(1-u') \\
&\quad + \partial_1 C(u',v')(1-\partial_1 C(u,v))u(1-u') + \partial_1 C(u,v)(1-\partial_2 C(u',v'))(C(u,v')-uv') \\
&= (1-\partial_1 C(u',v'))(1-u')(u-C(u,v)) + \partial_1 C(u,v)(1-\partial_2 C(u',v'))(C(u,v')-uv') \\
&\geq 0,
\end{aligned}
$$

since all partial derivatives are in $[0,1]$, $u \geq C(u,v)$ and $C(u,v') \geq uv'$. In the case $v \leq v'$, the same computation may be applied for the remaining terms of the integral of Equation (9). In the same way nonnegativity for the case $u' \leq u, v' \leq v$ is obtained. Finally, consider the cases $u \leq u', v' \leq v$ and $u' \leq u, v \leq v'$. Observe that we may regroup the integrand of Equation (9), taking into account that $g(u,v)$ and $g(u',v')$ may be exchanged appropriately. In the case $u \leq u', v' \leq v$, write the last two terms of the integrand of Equation (9) as

$$\iiiint 2(1-\partial_1 C(u,v)\partial_2 C(u',v'))(C(u,v')-uv')\,\mathrm{d}g(u,v)\,\mathrm{d}g(u',v')$$

and in the case $u' \leq u, v \leq v'$ as

$$\iiiint 2(1-\partial_1 C(u',v')\partial_2 C(u,v))(C(u',v)-u'v)\,\mathrm{d}g(u,v)\,\mathrm{d}g(u',v'),$$

and repeat the computation above accordingly. □

Example 16. Let $g(u,v) = \ln(\ln(uv+1)+1)$ and let (U^1,U^2) be a random vector with uniform marginals and Normal copula with parameter $\rho = 0.5$. Numerical integration yields $\sigma^2_{\mathrm{MC}} = 0.022756$ and $\sigma^2_{\mathrm{LHSD}} = 0.001101$. We estimated σ^2_{MC} and σ^2_{LHSD} by running 1000 batches of n independent simulations of the respective estimators, for $n \in \{200, 400, 600, 800, 1000\}$. The deviations to the numbers from numerical integration are within 0.003 for MC and $4 \cdot 10^{-5}$ for LHSD.

We also estimated the variance of performing standard LHS on a sample of independent uniform random variables, which were then correlated with a Normal copula. Running 1000 batches of 1000 independent simulations we obtained a variance estimate of 0.012152 (scaled as if this were a limit variance). In the example, this yields a variance reduction over the standard Monte Carlo technique by the factor 2, but LHSD outperforms this method by a factor of 10.

Numerical examples indicate that the classes of functions and copulas for which the LHSD limit variance is bounded from above by the respective MC limit variance are much larger than the ones stated in Proposition 15.

We have focused on the bivariate case, but many multivariate financial applications are of higher than bivariate dimension, so that the Central Limit Theorem and other results of this section do not apply. However, since no conditions are posed that require the bivariate and exclude the multivariate case, we presume that an extension to the multivariate case is possible, albeit tedious, cf. Equation (7).

6. LHSD on random vectors with nonuniform marginals

So far, we have restricted our analysis to vectors of uniform random variables on $[0,1]$. We now provide the link to random vectors with nonuniform marginals. It is always possible to generate a random variable of arbitrary distribution from a uniform random variable on $[0,1]$ by applying the so-called inverse transform method. The association of a joint distribution function with a copula (a distribution function with uniform marginals on $[0,1]$) leads to methods for constructing random vectors $(X^1,\dots,X^d)$ with arbitrary marginals from random vectors $(U^1,\dots,U^d)$, where $U^j\sim U(0,1)$, $j=1,\dots,d$. We discuss this in more detail.

The inverse transform method is explained for example in [Glasserman (2004), Section 2.2.1] and [Nelsen (1999), Sections 2.3, 2.9]. Let X be a random variable with distribution function F. We shall assume F to be continuous, which implies $\mathbf{P}(X=x)=0$, $x\in\mathbb{R}$. The right-inverse of F is defined as the function $F^{(-1)}:[0,1]\to\mathbb{R}\cup\{\pm\infty\}$ with

$$F^{(-1)}(u):=\inf\{x:F(x)>u\},\quad u\in[0,1].$$

The right-inverse is right-continuous, strictly increasing and has at most countably many discontinuities. If F is strictly increasing, then $F^{(-1)}$ is just the inverse of F. From the monotonicity of distribution functions, $F^{(-1)}(u)<x$ if and only if $u<F(x)$. It follows that if $U\sim U(0,1)$, then $X\overset{\mathcal{L}}{=}F^{(-1)}(U)$, since

$$\mathbf{P}(X<x)=F(x)=\mathbf{P}(U<F(x))=\mathbf{P}(F^{(-1)}(U)<x).$$

Accordingly, for a Borel-measurable function $h:\mathbb{R}\to\mathbb{R}$, $h(X)\overset{\mathcal{L}}{=}g(U)$, with $g:=h\circ F^{(-1)}$.

Now consider the multivariate case. Recall that a copula is a multivariate distribution function whose margins are $U(0,1)$ distributions. By Sklar's Theorem [Nelsen (1999), Theorem 2.10.9], the copula associated with a d-dimensional distribution function F and univariate marginal distribution functions $F_1,\dots,F_d$ is the distribution function $C:[0,1]^d\to[0,1]$ that satisfies $F(x_1,\dots,x_d)=C(F_1(x_1),\dots,F_d(x_d))$. Conversely, for any $(u_1,\dots,u_d)\in[0,1]^d$,

$$C(u_1,\dots,u_d)=F(F_1^{(-1)}(u_1),\dots,F_d^{(-1)}(u_d)),$$

cf. [Nelsen (1999), Corollary 2.10.10]. If F is continuous, then C is unique, otherwise C is unique on $\mathrm{Ran}F_1\times\cdots\times\mathrm{Ran}F_d$, where $\mathrm{Ran}F_j\subseteq[0,1]$ denotes the range of F_j, $j=1,\dots,d$. The copula provides the link between the marginal distributions and the joint distribution of a random vector.

Now consider a random vector $(X^1,\dots,X^d)$ with marginal distribution functions $F_1,\dots,F_d$ and joint distribution function F. Then, for a Borel-measurable function $h:\mathbb{R}^d\to\mathbb{R}$,

$$h(X^1,\dots,X^d)\overset{\mathcal{L}}{=}h(F_1^{(-1)}(U^1),\dots,F_d^{(-1)}(U^d))=:g(U^1,\dots,U^d),\tag{10}$$

where the joint distribution of $(U^1,\dots,U^d)$ is determined by the copula corresponding to F and $F_1,\dots,F_d$. The following properties are immediate:

(i) If F_j, $j=1,\dots,d$ are continuous, and if h is F-a.e. continuous, then g is C-a.e. continuous.

(ii) If h is right-continuous, and $F_j^{(-1)}$, $j=1,\dots,d$, are the right-inverses of F_j, $j=1,\dots,d$, then g is right-continuous. Moreover, if h is of bounded variation, then so is g; this follows from the the strict monotonicity of the right-inverses.

Now, assuming that h is F-integrable, the LHSD estimator of $\mathbb{E}h(X^1,\dots,X^d)$ is given by

$$\frac{1}{n}\sum_{i=1}^{n}h(F_1^{(-1)}(V_{i,n}^1),\dots,F_d^{(-1)}(V_{i,n}^d)),\tag{11}$$

with $V_{i,n}^j$, $i=1,\dots,n$, $j=1,\dots,d$ as in Equation (3).

By the following Lemma, the ranks may be computed without first transforming the marginals $X^1,\dots,X^d$ into uniforms.

Lemma 17. *Let $X_1, \ldots, X_n$ be i.i.d. random variables whose distribution F is continuous. Then, for any $i = 1, \ldots, n$,*

$$r_{i,n}(X_1, \ldots, X_n) = r_{i,n}(F(X_1), \ldots, F(X_n)) \quad \boldsymbol{P}\text{–}a.s..$$

Proof. If F is strictly increasing, the statement is clear by Equation (2). By Equation (2) it suffices to show that $\mathbf{P}$–a.s. $X_i \leq X_j$ if and only if $F(X_i) \leq F(X_j)$, for any $i, j = 1, \ldots, n$. By monotonicity of F, $X_i \leq X_j$ implies $F(X_i) \leq F(X_j)$. For the reverse statement consider

$$\begin{aligned} \mathbf{P}(F(X_i) = F(X_j), X_i > X_j) &= \mathbf{P}(X_i \in (X_j, F^{(-1)}(F(X_j))]) \\ &= \int \mathbf{P}(X_i \in (y, F^{(-1)}(F(y))])\, F(\mathrm{d}y) = \int \Big[F(F^{(-1)}(F(y))) - F(y)\Big]\, F(\mathrm{d}y) = 0, \end{aligned}$$

where the last equality follows from $F(F^{(-1)}(z)) = z$ because of the continuity of F. □

7. Applications in finance

We demonstrate the effectiveness of LHSD with two examples. First, we value a first-to-default credit basket (FTD) - a contract that insures the loss incurred by the first default event in a basket of underlying securities. The value of an FTD depends crucially on the joint default probability distribution of the basket components. The example demonstrates that LHSD is an effective technique when sampling rare events; in fact, LHSD guarantees that, in each dimension, a certain number of rare events is sampled. We also combine Quasi-Monte Carlo (QMC) and LHSD by feeding our algorithm with Sobol sequences instead of random numbers. The combination of these techniques leads to a further pickup in efficiency, which is a surprising observation as Sobol sequences are already stratified, albeit before application of the copula transformation.

In the second example we value an Asian basket call option. Here, a call option is written on the weighted sum of a basket of securities monitored at several time points. The example is taken from [Imai and Tan (2007)], where a basket of 10 assets is monitored at 250 time points. [Imai and Tan (2007)] show that each simulation entails generating a correlated random vector of size 2500. This example demonstrates that LHSD can be used for high-dimensional problems. It has been observed that low discrepancy sequences lose their effectiveness in high dimensions [Owen (1998)], hence we do not test the combination of QMC and LHSD. [Imai and Tan (2007)] benchmark their results against a technique for QMC in a high-dimensional setting called Latin supercube sampling, see [Owen (1998)]; we compare our findings to these results. Combinations of such techniques with LHSD may again improve results.

7.1. *Example: Valuing a first-to-default credit basket*

An FTD is a contract between two counterparties, a protection buyer and a protection seller, that insures the protection buyer against the loss incurred by the first default event in a portfolio of some underlying risky entities over a fixed time horizon. The protection buyer regularly pays a constant premium s, called the *spread*, as a fraction of the notional until the first default event in the underlying portfolio takes place or until maturity of the FTD, whichever occurs first. This stream of payments is termed the *premium leg* of the FTD. In turn, the protection seller compensates the protection buyer for the loss incurred by the first default event at the time of default. This side of the contract is called the *default leg*.

For the valuation of an FTD we follow [Schmidt and Ward (2002)]. With each credit $j = 1, \ldots, d$ of the underlying portfolio we associate the random default time τ_j and the recovery rate R_j. We assume R_j to be constant and known. Furthermore, we assume the default distributions $\mathbf{P}(\tau_j \leq t)$, $t \geq 0$, $j = 1, \ldots, d$, to be given. These can be derived from the credit default swap (CDS) market; as an approximation, assuming a constant CDS spread s_j for credit j, we determine the default intensity λ_j, of credit j from the so-called credit triangle, $\lambda_j := s_j/(1 - R_j)$, and we set

$$F_j(t) := \mathbf{P}(\tau_j \leq t) = 1 - \mathbf{e}^{-\lambda_j t}, \quad t \geq 0. \tag{12}$$

Table 1. Parameters of FTD example; the fair spread of the FTD is 417.88bp.

Parameter	Value
Maturity	$T = 5$ (years)
spread payment dates (frequency)	$(t_k)_{k=1,\dots,K}$ (quarterly)
Default-free zero bond prices	$B_t = \mathbf{e}^{-.05t}$, $t \geq 0$
Number of underlying credits	$d = 5$
5yr.-CDS spread of each credit	$s_j = 1\%$, $j = 1, \dots, d$
Recovery rate of each credit	$R_j = 0.3$, $j = 1, \dots, d$
Correlation between any two credits	$\rho = 30\%$

For $t \geq 0$, denote by B_t today's default-free zero bond price with maturity t. Let $t_0 = 0$ and let $t_1 < t_2 < \dots < t_K = T$ be the spread payment dates of an FTD with maturity T, and set $\Delta_{t_k} := t_k - t_{k-1}$, $k = 1, \dots, K$. Denote the time of the FTD's default event by $\tau := \min(\tau_1, \dots, \tau_d)$. The discounted payoffs of the default leg and the premium leg are given by

$$h_d(\tau_1, \dots, \tau_d) = \sum_{j=1}^{d} (1 - R_j) B(\tau) \mathbf{1}_{\{(0,T]\}}(\tau)\, \mathbf{1}_{\{\tau = \tau_j\}} \tag{13}$$

$$h_p(\tau_1, \dots, \tau_d) = s \sum_{k=1}^{K} \Delta_{t_k} B(t_k) \mathbf{1}_{\{\tau > t_k\}}. \tag{14}$$

The fair spread s of the FTD is then obtained by equating the expected value (under the risk-neutral measure) of the premium and the default leg,

$$s \sum_{k=1}^{K} \Delta_{t_k} B_{t_k} \mathbf{P}(\tau > t_k) = \sum_{j=1}^{d} (1 - R_j) \int_0^T B_u \mathbf{P}(\tau \in \mathrm{d}u, \tau = \tau_j). \tag{15}$$

From this equation and from $\mathbf{P}(\tau \leq t) = 1 - \mathbf{P}(\tau_1 > t, \tau_2 > t, \dots, \tau_d > t)$ it is clear that the value of the FTD depends on the joint distribution of $\tau_1, \dots, \tau_d$. Setting $s = 1$, the left-hand side of Equation (15) can be interpreted as the present value of a risky basis point.

In our example we assume that the joint distribution of the default times $\tau_1, \dots, \tau_d$ is driven by a Normal copula (Gaussian copula),

$$\mathbf{P}(\tau_1 \leq t, \dots, \tau_j \leq t) = \mathrm{N}_\Sigma \left(\mathrm{N}^{(-1)}(F_1(t)), \dots, \mathrm{N}^{(-1)}(F_j(t)) \right),$$

with N_Σ the multivariate standard normal distribution function with correlation matrix Σ and $\mathrm{N}^{(-1)}$ the inverse of the univariate standard normal distribution function.

The valuation algorithm for the fair FTD spread is given by Algorithm 1. The input parameters for an example involving 5 homogeneous credits are given in Table 1. The fair FTD spread was first computed numerically using a factor-model approach, see e.g. [Hull and White (2004)]. It was then computed from simulations using random numbers and using low discrepancy sequences, both "as is" and adding a LHSD step. This leads to the following four simulation cases:

(i) Standard Monte Carlo simulation,
(ii) LHSD based on random numbers,
(iii) Simulation with low discrepancy sequence,
(iv) LHSD based on low discrepancy sequence.[b]

[b]Note that we have not proved convergence for the two simulation cases involving low discrepancy sequences, that is, Proposition 3 does not apply for these cases. They were included to illustrate that LHSD may be combined with techniques other than Monte Carlo simulation.

1: // n: number of simulations, d : number of credits
2: **for** $j = 1$ to d **do**
3: $\lambda_j \leftarrow s_j/(1 - R_j)$ // default intensities; credit triangle
4: **end for**
5: Compute A such that $AA^T = \Sigma$ // e.g. Cholesky factorisation
6: **for** $i = 1$ to n **do**
7: **for** $j = 1$ to d **do**
8: generate $X_i^j \sim \mathrm{N}(0, 1)$ // independent of X_k^m, $k = 1, \ldots i - 1$, $m = 1, \ldots, j - 1$
9: **end for**
10: $(Z_i^1, \ldots, Z_i^d)^T \leftarrow A \cdot (X_i^1, \ldots, X_i^d)^T$ // vector of correlated standard normal samples
11: **end for**
12: **for** $j = 1$ to d **do**
13: compute $r_{1,n}^j, \ldots, r_{n,n}^j$ // ranks from $(Z_1^j, \ldots, Z_n^j)$, cf. Lemma 17
14: **for** $i = 1$ to n **do**
15: $V_{i,n}^j \leftarrow (r_{i,n}^j - 1/2)/n$ // Equation (3)
16: $\tau_i^j \leftarrow F_j^{(-1)}(V_{i,n}^j)$ // default times; $F_j^{(-1)}(t) := -\ln(1 - t)/\lambda_j$, Equation (12)
17: **end for**
18: **end for**
19: $s \leftarrow 1$
20: **for** $i = 1$ to n **do**
21: $L_i \leftarrow h_d(\tau_i^1, \ldots, \tau_i^d)$ // discounted default leg, Equation (13)
22: $P_i \leftarrow h_p(\tau_i^1, \ldots, \tau_i^d)$ // Equation (14)
23: **end for**
24: $\bar{L} \leftarrow (L_1 + \cdots + L_n)/n$ // present value of expected loss, RHS of Equation (15)
25: $\bar{P} \leftarrow (P_1 + \cdots + P_n)/n$ // PV of a risky basis point, left-hand side of Equation (15)
26: **return** $s \leftarrow \bar{L}/\bar{P}$ // fair spread of FTD

Algorithm 1: FTD valuation

For the LHSD transformation, Equation (3), we have chosen $\eta_{i,n}^j = 1/2$, for all i, j. The implementation was done in `C++` with the `QuantLib` library [QuantLib (2008)] using the Mersenne twister algorithm for random number generation and Sobol sequences for low discrepancy sequences. Root mean square error estimates were obtained by simulating each estimator 100 times. The RMSE estimates and RMSE ratios for various samples sizes are given in Table 2. The ratios of CPU time consumed for generating samples with and without LHSD is also shown for various sample sizes. The CPU time ratios do not include the CPU time required for computing the FTD payoff; consequently the efficiency of LHSD increases with the CPU time required for computing the payoff function. The LHSD step involves sorting a sequence of random numbers (see e.g. [Press *et al.* (1992)], Chapter 8.4 for sorting algorithms), hence the computational overhead of the LHSD step is of the complexity of the sorting algorithm. On the other hand, by Lemma 17, the rank statistics can be computed from samples of any distribution (cf. Line 13 in Algorithm 1), whereas in a typical Monte Carlo simulation, the generated samples may additionally need to be transformed to uniforms. Observe that over all simulations, LHSD samples a fixed number of default events of the individual credits, but the occurence of joint defaults is random.

To interpret the performance pickup of LHSD over standard Monte Carlo involving the CPU time required, one has to take into account that the variance in Monte Carlo simulation decreases linearly with the number of simulations. The variance of LHSD over Monte Carlo simulation is reduced by a factor of approximately 4, while the simulation time increases by a factor of less than 1.82.

RMSE estimates using standard LHS on independent random numbers and then correlating them with a Gaussian copula were similar to the LHSD estimates. The CPU ratios of this method compared to standard Monte Carlo simulation are approximately 1.3, that is, slightly smaller than the CPU ratios for LHSD.

Table 2. Root mean square error of estimation in basis points and CPU time ratios for various sample sizes (100 simulations of estimator). Comparable ratios were obtained for smaller simulation sizes. The fair FTD spread is 417.88bp.

No. of sim. ($\times 10^3$)	200	400	600	800	1000
MC	2.02	1.47	1.10	0.89	0.80
MC + LHSD	1.00	0.61	0.53	0.45	0.39
Sobol	0.30	0.20	0.16	0.14	0.11
Sobol + LHSD	0.21	0.12	0.11	0.09	0.08
MC/(MC + LHSD)	2.02	2.41	2.08	1.98	2.05
Sobol/(Sobol + LHSD)	1.43	1.67	1.45	1.56	1.38
CPU time (MC + LHSD)/MC	1.66	1.71	1.75	1.78	1.82
CPU (Sobol + LHSD)/Sobol	1.47	1.52	1.54	1.55	1.56

Note: CPU time ratios involve the generation of random samples only. Adding the CPU time required for computing the payoff decreases the ratios accordingly.

7.2. *Example: Valuing an Asian basket option*

We now consider pricing an Asian basket option[c], whose payoff depends on the sum of several underlying assets monitored at various points in time. As this is a path-dependent option in a high-dimensional setting, simulation is a standard valuation approach. Following [Imai and Tan (2007)], the payoff may be formulated as a function of a matrix product whose dimensions depend on the number of assets and time points monitored.

Assume a basket of m assets, with S_t^i the price of the i-th asset at time t, $i = 1, \ldots, m$. Fixing a maturity T, a strike K, a set of n monitoring time points $0 < t_1 < t_2 < \ldots < t_n = T$ and weights w^{ij}, $i = 1, \ldots m$, $j = 1, \ldots n$, $\sum_{i,j} w^{ij} = 1$, the payoff of the Asian basket call option on the m-asset basket is

$$\max\left(\sum_{i=1}^{m}\sum_{j=1}^{n} w^{ij} S_{t_j}^i - K, 0\right). \tag{16}$$

We assume that asset prices follow a Geometric Brownian motion, i.e., $S^1, \ldots, S^m$ is the solution of the stochastic differential equation (SDE)

$$\mathrm{d}S_t^i = r S_t^i \,\mathrm{d}t + \sigma^i S_t^i \,\mathrm{d}W_t^i, \quad i = 1, \ldots, m,$$

where r is the risk-free interest rate, σ^i is the volatility of the i-th asset and $(W^1, \ldots, W^m)$ is an m-dimensional Brownian motion, whose components W^i, W^k are correlated with ρ^{ik}, $1 \le i, k \le m$. The solution of the SDE is given by

$$S_t^i = S_0^i \mathbf{e}^{(r - (\sigma^i)^2/2)t + \sigma^i W_t^i}, \quad i = 1, \ldots, m. \tag{17}$$

Pricing the option requires simulating the paths of each asset at the monitoring time points. Assume that the time points $t_1, \ldots, t_n$ are equidistant and let $\Delta t = T/n$ so that $t_j = j\Delta t$. Let Σ be an $m \times m$ covariance matrix given by $\Sigma = (\rho^{ik}\sigma^i\sigma^k \,\Delta t)_{i,k=1,\ldots,m}$. Let $\tilde{\Sigma}$ be the $nm \times nm$-matrix

[c] This is also known as an arithmetic average Asian option.

Table 3. Parameters of Asian basket option

Parameter	Value
Maturity	$T = 1$ (years)
Number of assets	$m = 10$
Number of time steps	$n = 250$
Weights	$w^{ij} = 1/(nm)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$
Initial asset value	$S_0^j = 100$, $j = 1, \ldots, m$
Asset volatility	$\sigma^j = 0.1 + (j-1)/(m-1) \cdot 0.4$, $j = 1, \ldots, m$
Correlation	$\rho^{ij} = 0.4$, $1 \le i < j \le m$
Interest rate	$r = 0.04$
Strike	$K = 90, 100, 110$

generated from Σ via

$$\tilde{\Sigma} = \begin{pmatrix} \Sigma & \Sigma & \cdots & \Sigma \\ \Sigma & 2\Sigma & \cdots & 2\Sigma \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma & 2\Sigma & \cdots & n\Sigma \end{pmatrix}.$$

The asset prices may be simulated according to Equation (17) with $\tilde{W} = (\sigma^1 W_{t_1}^1, \ldots, \sigma^m W_{t_1}^m, \sigma^1 W_{t_2}^1, \ldots, \sigma^m W_{t_n}^m)'$ derived via

$$\tilde{W} = \tilde{C} Z,$$

where $\tilde{C}$ is such that $\tilde{C}\tilde{C}' = \tilde{\Sigma}$ and Z is a vector of nm independent standard Normal random variables. The payoff at time T of the Asian basket option can then be written as

$$\max(g(\tilde{W}) - K, 0)$$

with

$$\begin{aligned} g(\tilde{W}) &= \sum_{k=1}^{mn} \mathbf{e}^{\mu^k + \tilde{W}_k} \\ \mu^k &= \ln(w_{k_1 k_2} S_{k_1}(0)) + \left(r - (\sigma^{k_1})^2/2\right) t_{k_2}, \quad \text{where} \\ k_1 &= (k-1) \mod m + 1 \\ k_2 &= \lfloor (k-1)/m \rfloor + 1, \quad k = 1, \ldots, mn. \end{aligned}$$

In this approach, simulation of option payoffs involves the computation of products of high-dimensional matrices. For $\tilde{C}$ we choose the Cholesky decomposition of $\tilde{\Sigma}$ (i.e., $\tilde{C}$ is lower triangular with $\tilde{C}\tilde{C}' = \tilde{\Sigma}$). Typical choices of $\tilde{C}$ other than the Cholesky decomposition yield a reduction of the dimension of the matrix multiplication, while keeping the error introduced small. For example, the simulation technique of [Imai and Tan (2007)] reduces the effective dimension of the problem by determining $\tilde{C}$ as the solution of an optimization problem.

Based on the data set of [Imai and Tan (2007)], we simulate $\tilde{W}$ using a Cholesky decomposition and introducing an LHSD step in each dimension over all simulations. The parameters of the option are given in Table 3. As in [Imai and Tan (2007)], we computed 10 runs of 4096 simulations and 10 runs of 8192 simulations. In addition to performing standard Monte Carlo simulation and LHSD, we applied LHS to independent uniform random numbers, which were then correlated. The prices and standard deviations for Monte Carlo, LHS and LHSD are given in Table 4. The implementation was done in C++ using `QuantLib`, see [QuantLib (2008)], and using matrix multiplication routines from the Fortran code of `GNU Octave`, see [GNU Octave (2008)]. The results show that LHSD outperforms the standard Monte Carlo simulator by factors of 2.5 to 15 based on standard deviations (resp. 9 and 200 in terms of variance); the computing time consumed by LHSD increases by a factor of approximately 1.4. In comparison, LHS outperforms Monte Carlo simulation by factors of 2 to 5

Table 4. Simulated prices of an Asian basket option (parameters in Table 3) for strikes $K \in \{90, 100, 110\}$. The results are based on 10 runs with 4096 and 8192 simulations each. The numbers in parentheses denote the sample standard deviation based on the 10 runs. The CPU time ratios of LHS versus MC were 1.31 CPU seconds (4096 and 8192 simulations). The CPU time ratios of LHSD versus MC were 1.40 CPU seconds (4096 simulations) and 1.44 CPU seconds (8192 simulations).

	sim. size	$K = 90$		$K = 100$		$K = 110$	
MC	4096	12.3045	(0.1930)	5.6726	(0.1402)	2.0574	(0.0916)
MC+LHS	4096	12.3322	(0.0392)	5.6549	(0.0580)	2.0372	(0.0471)
MC+LHSD	4096	12.3283	(0.0130)	5.6567	(0.0187)	2.0288	(0.0316)
MC	8192	12.3481	(0.1602)	5.6697	(0.1041)	2.0413	(0.0633)
MC+LHS	8192	12.3221	(0.0340)	5.6417	(0.0458)	2.0169	(0.0506)
MC+LHSD	8192	12.3253	(0.0150)	5.6535	(0.0166)	2.0302	(0.0261)
MC/(MC+LHS)	4096		4.9235		2.4172		1.9448
MC/(MC+LHSD)	4096		14.8462		7.4973		2.8987
MC/(MC+LHS)	8192		4.7118		2.2729		1.2510
MC/(MC+LHSD)	8192		10.6800		6.2711		2.4253

(resp. 4 to 25 in term of variance). The pickup in accuracy depends strongly on the strike of the option and decreases with increasing strike. The same observation is made in an example from [Glasserman (2004), p. 242-243], where an Asian call option is priced using standard LHS. There, this behaviour is attributed to the fact that LHS is more effective the more the function to be estimated is "additive"; this is observed at lower strikes, where the option payoffs are more linear.

To benchmark their method, [Imai and Tan (2007)] simulated the Asian basket option using a Quasi-Monte Carlo method together with a technique called Latin supercube sampling. The latter method avoids sampling low discrepancy sequences in high dimensions, see [Owen (1998)]. The standard error of LHSD is comparable to that of the QMC technique of [Imai and Tan (2007)], the latter being between 0.00905 and 0.0144. By comparison, recall that LHSD is a very simple and practicable technique. Finally, it should be noted that our results do not keep up with standard errors obtained from the dimension reduction technique of [Imai and Tan (2007)], but we conjecture that combinations of LHSD together with dimension-reduction techniques may be effective.

Appendix A. Integration by parts formula

We derive the integration by parts formula as it is used in the paper. An integration by parts formula for two dimensions is given in [Gill *et al.* (1995)]; a version for $\mathbb{R}^k$ is found in [Gill and Johansen (1990), p. 1530].

Let us recall some well-known concepts and facts, see e.g. [von Neumann (1950), Chapter X.5]. Let $H : [0,1]^d \to \mathbb{R}$ be a right-continuous function. For rectangles $B = (a_1, b_1] \times \cdots \times (a_d, b_d] \subset [0,1]^d$, define

$$V_H(B) := \sum_{c \text{ vertex of } B} \operatorname{sgn}(c) H(c),$$

where

$$\operatorname{sgn}(c) = \begin{cases} 1, & \text{if } c_k = a_k \text{ for an even number of } k\text{'s}, \\ -1, & \text{if } c_k = a_k \text{ for an odd number of } k\text{'s}. \end{cases}$$

If in addition $V_H(B) \geq 0$ for all rectangles B, then H is called *quasi-monotone*. If H is quasi-monotone and right-continuous, it determines a σ-additive, nonnegative measure, which we also denote by H, via

$$\int_B \mathrm{d}H = V_H(B), \tag{A.1}$$

for all rectangles B. If H is of bounded variation and right-continuous, then it is the difference of two quasi-monotone, right-continuous functions, and hence determines a σ-additive, signed measure via the relationship (A.1).

Proposition 18. *Let* $H, G : [0,1]^4 \to \mathbb{R}$ *be of bounded variation and right-continuous, with at least one of* H, G *continuous and such that* $\iiiint H \,\mathrm{d}G$ *exists. Then,*

$$\begin{aligned}\iiiint H(u_1,u_2,u_3,u_4)\,\mathrm{d}G(u_1,u_2,u_3,u_4) &= \iiiint V_G((u_1,1]\times\cdots\times(u_4,1])\,\mathrm{d}H(u_1,u_2,u_3,u_4)\\ &+ \iiiint \big\{H(0,u_2,u_3,u_4)+H(u_1,0,u_3,u_4)+H(u_1,u_2,0,u_4)+H(u_1,u_2,u_3,0)\\ &\quad - H(0,0,u_3,u_4)-H(0,u_2,0,u_4)-H(0,u_2,u_3,0)\\ &\quad - H(u_1,0,0,u_4)-H(u_1,0,u_3,0)-H(u_1,u_2,0,0)\\ &\quad + H(0,0,0,u_4)+H(0,0,u_3,0)+H(0,u_2,0,0)+H(u_1,0,0,0)\\ &\quad - H(0,0,0,0)\big\}\,\mathrm{d}G(u_1,u_2,u_3,u_4).\end{aligned} \tag{A.2}$$

Proof. By Equation (A.1),

$$\begin{aligned}H(u_1,\dots,u_4) = &\int_{(0,u_1]\times\cdots\times(0,u_4]} \mathrm{d}H(x_1,\dots,x_4)\\ &+ H(0,u_2,u_3,u_4)+H(u_1,0,u_3,u_4)+H(u_1,u_2,0,u_4)+H(u_1,u_2,u_3,0)\\ &- H(0,0,u_3,u_4)-H(0,u_2,0,u_4)-H(0,u_2,u_3,0)\\ &- H(u_1,0,0,u_4)-H(u_1,0,u_3,0)-H(u_1,u_2,0,0)\\ &+ H(0,0,0,u_4)+H(0,0,u_3,0)+H(0,u_2,0,0)+H(u_1,0,0,0)\\ &- H(0,0,0,0).\end{aligned}$$

Insert this expression into Equation (A.2) and apply Fubini's theorem to the first term, for which we then obtain

$$\iiiint V_G([u_1,1]\times\cdots\times[u_4,1])\,\mathrm{d}H(u_1,u_2,u_3,u_4).$$

The statement follows by observing that from the continuity of one of H and G,

$$\int_{(0,1]^1\times\cdots\times(0,1]^4}\int_{(0,1]^1\times\cdots\times(0,1]^4}\sum_{i=1}^4 \mathbf{1}_{\{u_i=x_i\}}\prod_{j=1,j\neq i}^4 \mathbf{1}_{\{u_j\geq x_j\}}\mathrm{d}G(u_1,\dots,u_4)\,\mathrm{d}H(x_1,\dots,x_4)=0. \quad \square$$

References

Adams, C. and Clarkson, J. (1934). Properties of functions $f(x,y)$ of bounded variation, *Trans. Amer. Math. Soc* **36**, 4, pp. 711–730.

Billingsley, P. (1968). *Convergence of probability Measures* (Wiley, New York).

Clarkson, J. and Adams, C. (1933). On definitions of bounded variation for functions of two variables, *Trans. Amer. Math. Soc.* **35**, 4, pp. 824–854.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés, *Bulletin de l'Académie Royale de Belgique, Classe des Sciences* **5**, 65, pp. 274–292.

Deheuvels, P. (1981). Multivariate tests of independence, in *Analytical Methods in Probability Theory, Proceedings of the conference held at Oberwolfach, June 1980*, Lecture Notes in Mathematics (Springer), pp. 42–50.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd edn. (John Wiley & Sons, New York).

Fermanian, J.-D., Radulovic, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes, *Journal of the Bernoulli Society* **10**, 5, pp. 847–860.

Gaenssler, P. and Stute, W. (1987). *Seminar on Empirical Processes*, DMV seminar, volume 9 (Birkhäuser).

Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika* **82**, 3, pp. 543–552.

Gill, R. and Johansen, S. (1990). A survey of product-integration with a view toward application in survival analysis, *Ann. Statist* **18**, 4, pp. 1501–1555.

Gill, R. D., van der Laan, M. J. and Wellner, J. A. (1995). Inefficient estimators of the bivariate survival function for three models, *Ann. Inst. Henri Poincaré* **31**, 3, pp. 545–597.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering* (Springer).

GNU Octave (2008). A high-level language, primarily intended for numerical computations, version 2.9.9, URL `http://www.octave.org`, http://www.octave.org.

Hobson, E. (1921). *The Theory of Functions of a Real Variable Vol. I*, 2nd edn. (Cambridge University Press).

Hull, J. and White, A. (2004). Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation, *Journal of Derivatives* **12**, 2 (Winter 2004), pp. 8–23.

Imai, J. and Tan, K. S. (2007). A general dimension reduction technique for derivative pricing, *Journal of Computational Finance* **10**, 2, pp. 129–155.

Jacod, J. and Protter, P. (2003). *Probability Essentials*, 2nd edn. (Springer).

Joe, H. (1997). *Multivariate Models and Dependence Concepts* (Chapman & Hall/CRC).

Kallenberg, O. (2001). *Foundations of Modern Probability*, 2nd edn. (Springer).

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics* **21**, pp. 239–245.

Nelsen, R. B. (1999). *An Introduction to Copulas* (Springer).

Owen, A. B. (1992). A central limit theorem for Latin hypercube sampling, *Journal of the Royal Statistical Society* **Series B 54, 13**, pp. 541–551.

Owen, A. B. (1998). Latin supercube sampling for very high-dimensional simulations, *ACM Transactions on Modeling and Computer Simulation* **8**, 1, pp. 71–102.

Press, W. H., Teukolsky, S. A., Vetterling, V. and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. (Cambridge University Press).

QuantLib (2008). A free/open-source library for quantitative finance, version 0.9.0, URL `http://quantlib.org`, http://quantlib.org.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics, *Ann. Statist.* **4**, 5, pp. 912–923.

Ruymgaart, F. H., Shorack, G. R. and van Zwet, W. R. (1972). Asymptotic normality of nonparametric tests for independence, *Annals of Mathematical Statistics* **43**, 4, pp. 1122–1135.

Schmidt, W. and Ward, I. (2002). Pricing default baskets, *RISK* **15**, 1.

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling, *Technometrics* **29**, 2, pp. 143–151, correction in 32:367.

Vaart, A. W. v. d. and Wellner, J. A. (1996). *Weak convergence and empirical processes* (Springer).

von Neumann, J. (1950). *Functional operators. Volume I: Measures and integrals* (Princeton University Press).