

A central limit theorem for Latin hypercube sampling with dependence and application to exotic basket option pricing

Christoph Aistleitner* Markus Hofer† Robert Tichy ‡

Abstract

We consider the problem of estimating $\mathbb{E}[f(U^1, \dots, U^d)]$, where (U^1, \dots, U^d) denotes a random vector with uniformly distributed marginals. In general, Latin hypercube sampling (LHS) is a powerful tool for solving this kind of high-dimensional numerical integration problem. In the case of dependent components of the random vector (U^1, \dots, U^d) one can achieve more accurate results by using Latin hypercube sampling with dependence (LHSD). We state a central limit theorem for the d -dimensional LHSD estimator, by this means generalising a result of Packham and Schmidt. Furthermore we give conditions on the function f and the distribution of (U^1, \dots, U^d) under which a reduction of variance can be achieved. Finally we compare the effectiveness of Monte Carlo and LHSD estimators numerically in exotic basket option pricing problems.

1 Introduction

In this article we consider the problem of reducing the variance of a Monte Carlo (MC) estimator for special functionals of a random vector with dependent components. Several different techniques can be used for this kind of problem, with different advantages and shortcomings (for a detailed comparison, see [Glasserman, 2004, Section 4]). A well-known technique is *Latin hypercube sampling* (LHS), which is a multi-dimensional version of the *stratified sampling* method and has been introduced by [McKay et al., 1979]. Although this method is well applicable to many different types of problems, it cannot deal with dependence structures among the components of random vectors. Therefore, we consider *Latin hypercube sampling with dependence* (LHSD), which was introduced by [Stein, 1987] and provides variance reduction for many problems, especially in financial mathematics.

Consider the problem of estimating $\mathbb{E}[f(U^1, \dots, U^d)]$ for a Borel-measurable and C -integrable function $f : [0, 1]^d \rightarrow \mathbb{R}$, where (U^1, \dots, U^d) is a random vector with uniformly distributed marginals and copula C . Let (U_i^1, \dots, U_i^d) , $1 \leq i \leq n$, denote an i.i.d. sample from this distribution. The standard Monte Carlo estimator, which is given by $1/n \sum_{i=1}^n f(U_i^1, \dots, U_i^d)$, is strongly consistent, and by the central limit theorem for sums of independent random variables the distribution of the scaled estimator converges

*Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: aistleitner@math.tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23.

†Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: markus.hofer@tugraz.at. Research supported by the Austrian Research Foundation (FWF) and by the DOC [Doctoral Fellowship Programme of the Austrian Academy of Sciences], Project S9603-N23.

‡Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: tichy@tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23.

Mathematics Subject Classification: 62H20 (62G05 62G20 62K15 62P05 91B82)

Keywords: Monte Carlo, Variance reduction techniques, Latin hypercube sampling, option pricing, variance gamma, probabilistic methods

to a normal distribution, ie:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n [f(U_i^1, \dots, U_i^d) - \mathbb{E}[f(U^1, \dots, U^d)]] \xrightarrow{\mathcal{D}} N(0, \sigma_{MC}^2),$$

where $\sigma_{MC}^2 = \text{Var}(f(U^1, \dots, U^d))$. In particular this means that the standard deviation of the estimator converges to zero with rate $\frac{1}{\sqrt{n}}$.

The aim of this paper is to establish a similar result for the LHSD estimator, under some additional conditions on the copula C and the function f . This has already been done in the bivariate case by [Packham and Schmidt, 2010] by using a result of [Fermanian et al., 2004]. Packham and Schmidt [2010, Proposition 5.9] also showed that under more restrictive conditions on the copula function C , the variance of the bivariate LHSD estimator does not exceed the variance of the standard Monte Carlo estimator.

An important application of Monte Carlo integration techniques lies in the field of financial mathematics. Many problems in finance result in the numerical computation of high-dimensional integrals, for which MC methods provide an efficient solution. Two examples are the pricing of Asian and discrete lookback options on several possibly correlated assets. We will investigate these special derivatives in numerical examples in the last section.

This paper is organised as follows: in the second section we introduce the main ideas of LHSD and recall some important results. Our main results are presented in the third section, where we state a central limit theorem and show under which conditions a reduction of variance, compared to the standard Monte Carlo method, is possible. The last section is dedicated to a comparison of the effectiveness of LHSD and MC in numerical examples.

2 Preliminaries

In this section, we recall the concept of stratified sampling and its extensions to Latin hypercube sampling and Latin hypercube sampling with dependence. We also state a consistency result, which was proved by [Packham and Schmidt, 2010].

2.1 Stratified sampling and LHS

Suppose that we want to estimate $\mathbb{E}(f(U))$, where U is a uniformly distributed random variable on the interval $[0, 1]$ (from now on denoted by $U([0, 1])$), and where $f : [0, 1] \rightarrow \mathbb{R}$ is a Borel-measurable and integrable function. By the simple fact that

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

where the intervals A_1, \dots, A_n (the so-called *strata*) form a partition of $[0, 1]$, we get an estimator for $\mathbb{E}(f(U))$ by sampling U conditionally on the events $\{U \in A_i\}, i = 1, \dots, n$. Choosing strata of the form $A_i = [\frac{i-1}{n}, \frac{i}{n})$ we can simply transform independent samples U^1, \dots, U^n from $U([0, 1])$ by setting

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

which implies $V_i \in A_i, i = 1, \dots, n$. The resulting estimator for $\mathbb{E}(f(U))$ given by $\frac{1}{n} \sum_{i=1}^n f(V_i)$ is consistent, and by the central limit theorem for sums of independent random variables the limit variance is smaller than the limit variance of a standard Monte Carlo estimator. For a more detailed analysis of stratified sampling techniques, see [Glasserman, 2004, Section 4.3.1].

This approach can be extended to the multivariate case in different ways. If we require that there has to be exactly one sample in every stratum, we need to draw n^d samples, which is not feasible for a high number of dimensions d . One way to avoid this problem is Latin hypercube sampling. Assume we want to estimate $\mathbb{E}(f(U^1, \dots, U^d))$, where $f : [0, 1]^d \rightarrow \mathbb{R}$ is a Borel-measurable and integrable function. For fixed n we generate n independent samples denoted by (U_i^1, \dots, U_i^d) , $i = 1, \dots, n$, where the U_i^j , $j = 1, \dots, d$ are uniformly distributed on $[0, 1]$. Additionally, we generate d independent permutations of $\{1, \dots, n\}$, denoted by π_1, \dots, π_d , drawn from a discrete uniform distribution on the set of all possible permutations. Denote by π_i^j the value to which i is mapped by the j -th permutation. Then the j -th component of a Latin hypercube sample is given by

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

By fixing a dimension j , the components (V_1^j, \dots, V_n^j) form a stratified sample with strata of equal length. It can be shown that the resulting estimator for $\mathbb{E}(f(U))$ is consistent, and by assuming that $f(U^1, \dots, U^d)$ has a finite second moment it follows that the variance of the LHS estimator

$$\frac{1}{n} \sum_{i=1}^n f(V_i^1, \dots, V_i^d)$$

is smaller than the variance of the standard MC estimator, provided the number of sample points is sufficiently large, see [Stein, 1987]. If f is bounded a central limit theorem for the LHS estimator can be shown, see [Owen, 1992]. Berry-Esseen-type bounds are also known, see [Loh, 1996]. A detailed discussion of LHS is given in [Glasserman, 2004, Section 4.4].

This technique is not suitable for dealing with random vectors with dependent components since the random variables V_i^j , $j = 1, \dots, d$, are independent. One way to extend the LHS method to random vectors with dependent components is to apply LHS to independent components and then introduce dependencies through a transformation of the LHS points. Such a procedure is tedious in general, and we will not pursue this approach any further.

2.2 Latin hypercube sampling with dependence

In this subsection, we introduce Latin hypercube sampling with dependence. The main difference to the LHS method is that instead of random permutations π_i we use rank statistics, which are defined as follows:

Definition 2.1 (Rank statistics) *Let X_1, \dots, X_n be i.i.d. random variables with a continuous distribution function. Denote the ordered random variables by $X_{(1)} < \dots < X_{(n)}$, \mathbb{P} -a.s. We call the index of X_i within $X_{(1)} < \dots < X_{(n)}$ the i -th rank statistic, given by*

$$r_{i,n} = r_{i,n}(X_1, \dots, X_n) := \sum_{k=1}^n \mathbf{1}_{\{X_k \leq X_i\}}. \quad (1)$$

Consider a random vector $U = (U^1, \dots, U^d)$, where every component U^j is uniformly distributed on $[0, 1]$ and the dependence structure of U is modeled by a copula C . Let (U_i^1, \dots, U_i^d) , $i = 1, \dots, n$ denote a sequence of independent samples of (U^1, \dots, U^d) , and let $r_{i,n}^j$ be the i -th rank statistic of (U_1^j, \dots, U_n^j) for $i = 1, \dots, n$ and $j = 1, \dots, d$. Then a LHSD is given by

$$V_{i,n}^j := \frac{r_{i,n}^j - 1}{n} + \frac{\eta_{i,n}^j}{n}, \quad i = 1, \dots, n, \forall j = 1, \dots, d, \quad (2)$$

where $\eta_{i,n}^j$ are random variables in $[0, 1]$. It is clear that $(V_{1,n}^j, \dots, V_{n,n}^j)$ forms a stratified sampling in every dimension j , where every stratum has equal length.

Packham and Schmidt [2010] consider different choices for $\eta_{i,n}^j$ to obtain special properties. For example, by choosing all $\eta_{i,n}^j$ uniformly distributed on $[0, 1]$ and independent of U_i^j , the distribution of the $V_{i,n}^j$ within their strata is uniform. This choice has the disadvantage of necessitating the generation of $2n$ random variables instead of only n . An effective choice in terms of computation time is $\eta_{i,n}^j = 1/2$, which means that every $V_{i,n}^j$ is located exactly in the centre of its stratum. In the remainder of this section, we briefly recall a result of [Packham and Schmidt, 2010] concerning the consistency of the LHSD estimator for $\mathbb{E}(f(U))$, which is defined by

$$\frac{1}{n} \sum_{i=1}^n f(V_{i,n}^1, \dots, V_{i,n}^d). \quad (3)$$

The usual law of large numbers for sums of independent random variables does not apply in this case for two reasons: firstly in each dimension the samples fail to be independent because of the application of the rank statistic, and secondly, increasing the samples size n by one changes every term of the sum instead of just adding one. Nevertheless, it can be shown that the following consistency result holds, see [Packham and Schmidt, 2010, Proposition 4.1]:

Proposition 2.1 *Let $f : [0, 1]^d \rightarrow \mathbb{R}$ be bounded and continuous C-a.e. . Then the LHSD estimator (3) is strongly consistent, ie :*

$$\frac{1}{n} \sum_{i=1}^n f(V_{i,n}^1, \dots, V_{i,n}^d) \xrightarrow{\mathbb{P} \text{ a.s.}} \mathbb{E}(f(U^1, \dots, U^d)), \quad \text{as } n \rightarrow \infty.$$

3 Central limit theorem and variance reduction

In this section we investigate the speed of convergence of the LHSD estimator and discuss situations in which the use of LHSD results in a reduction of variance. This has already been done for the bivariate case by [Packham and Schmidt, 2010]. They have also guessed the higher-dimensional version of the main theorem, but no rigorous proof was given. Because of the fact that most problems in finance for which Monte Carlo techniques are suitable are high-dimensional integration problems, it is reasonable to investigate the speed of convergence and the (asymptotic) value of the variance also in the multivariate case.

In the sequel, let \bar{C}_n denote the empirical distribution of the LHSD sample given by

$$\bar{C}_n(u^1, \dots, u^d) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{V_{i,n}^1 \leq u^1, \dots, V_{i,n}^d \leq u^d\}},$$

which is a distribution function. Furthermore, we define C_n as

$$C_n(u^1, \dots, u^d) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{F_n^1(U_i^1) \leq u^1, \dots, F_n^d(U_i^d) \leq u^d\}}, \quad (4)$$

where

$$F_n^j(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{U_i^j \leq u\}}, \quad u \in [0, 1],$$

are the one-dimensional empirical distribution functions based on U_1^j, \dots, U_n^j for $j = 1, \dots, d$. To formulate a central limit theorem we will need some regularity conditions on the integrand f and the copula C .

Definition 3.1 (Hardy-Krause bounded variation) A function $f : [0, 1]^d \rightarrow \mathbb{R}$ is of bounded variation (in the sense of Hardy-Krause) if $V(f) < \infty$ with

$$V(f) = \sum_{k=1}^d \sum_{1 \leq i_1 < \dots < i_k \leq d} V^{(k)}(f; i_1, \dots, i_k).$$

Here, the functional $V^{(k)}(f)$ denotes the variation in the sense of Vitali of f restricted to the k -dimensional face $F^{(k)}(i_1, \dots, i_k) = \{(u_1, \dots, u_d) \in [0, 1]^d : u_j = 1 \text{ for } j \neq i_1, \dots, i_k\}$. The variation of a function f in the sense of Vitali is defined by

$$V^{(k)}(f; i_1, \dots, i_k) = \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}(i_1, \dots, i_k)} |\Delta(f; J)|,$$

where the supremum is extended over all partitions $\mathcal{P}(i_1, \dots, i_k)$ of $F^{(k)}(i_1, \dots, i_k)$ into subintervals J and $\Delta(f; J)$ denotes the alternating sum of the values of f at the vertices of J . For more information on this topic, see [Owen, 2005].

Definition 3.2 A function $f : [0, 1]^d \rightarrow \mathbb{R}$ is right continuous if for any sequence $(u_n^1, u_n^2, \dots, u_n^d)_{n \in \mathbb{N}}$ with $u_n^j \downarrow u^j, j = 1, \dots, d$,

$$\lim_{n \rightarrow \infty} f(u_n^1, u_n^2, \dots, u_n^d) = f(u^1, u^2, \dots, u^d).$$

The next statement concerning the convergence of random sequences will be used to prove Proposition 3.1 and Theorem 3.2. For more details see eg [Jacod and Protter, 2003, Theorem 18.8].

Lemma 3.1 Let $(X_n)_{n \geq 1}$ and $(Y_n)_{n \geq 1}$ be sequences of \mathbb{R} -valued random variables, with $X_n \xrightarrow{\mathcal{D}} X$ and $|X_n - Y_n| \xrightarrow{\mathbb{P}} 0$. Then $Y_n \xrightarrow{\mathcal{D}} X$.

The following proposition of [Tsukahara, 2005] is a generalization of earlier results of [Stute, 1984] and [Fermanian et al., 2004]. It is the essential ingredient in proofs of our main theorems.

Proposition 3.1 Assume that C is differentiable with continuous partial derivatives $\partial_j C(u^1, \dots, u^d) = \frac{\partial C(u^1, \dots, u^d)}{\partial u^j}$ for $j = 1, \dots, d$. Then

$$\sqrt{n} \left(\tilde{C}_n(u^1, \dots, u^d) - C(u^1, \dots, u^d) \right) \xrightarrow{\mathcal{D}} G_C(u^1, \dots, u^d),$$

where

$$\tilde{C}_n(u^1, \dots, u^d) = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{U_k^1 \leq F_n^{1-}(u^1), \dots, U_k^d \leq F_n^{d-}(u^d)\}},$$

denotes the empirical copula function and F_n^{j-} denote the generalised quantile functions of F_n^j for $j = 1, \dots, d$, defined by

$$F_n^{j-}(u) = \inf\{x \in \mathbb{R} | F_n^j(x) \geq u\}.$$

Furthermore, G_C is a centred Gaussian random field given by

$$G_C(u^1, \dots, u^d) = B_C(u^1, \dots, u^d) - \sum_{j=1}^d \partial_j C(u^1, \dots, u^d) B_C(1, \dots, 1, u^j, 1, \dots, 1), \quad (5)$$

B_C is a d -dimensional pinned Brownian sheet on $[0, 1]^d$ with covariance function

$$\mathbb{E}[B_C(u^1, \dots, u^d) \cdot B_C(\bar{u}^1, \dots, \bar{u}^d)] = C((u^1, \dots, u^d) \wedge (\bar{u}^1, \dots, \bar{u}^d)) - C(u^1, \dots, u^d)C(\bar{u}^1, \dots, \bar{u}^d), \quad (6)$$

where $(u^1, \dots, u^d) \wedge (\bar{u}^1, \dots, \bar{u}^d)$ denotes the componentwise minimum.

We can formulate a similar result for the sequence C_n .

Proposition 3.2 *Under the conditions of Proposition 3.1,*

$$\sqrt{n} \left(C_n(u^1, \dots, u^d) - C(u^1, \dots, u^d) \right) \xrightarrow{\mathcal{D}} G_C(u^1, \dots, u^d) \quad (7)$$

holds, where all definitions are as in Proposition 3.1 and $C_n(u^1, \dots, u^d)$ is given in (4).

Proof:

We only have to show that the supremum of the difference of C_n and \tilde{C}_n vanishes for $n \rightarrow \infty$ to apply Lemma 3.1, which completes the proof. Note that C_n and \tilde{C}_n coincide on the grid $\{(i_1/n, \dots, i_d/n), 1 \leq i_1, \dots, i_d \leq n\}$. It follows that

$$\begin{aligned} & \sup_{u^1, \dots, u^d} |\tilde{C}_n(u^1, \dots, u^d) - C_n(u^1, \dots, u^d)| \\ & \leq \max_{1 \leq i^1, \dots, i^d \leq n} \left| \tilde{C}_n\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right) - \tilde{C}_n\left(\frac{i_1-1}{n}, \dots, \frac{i_d-1}{n}\right) \right| \leq \frac{d}{n}. \end{aligned}$$

Thus, $\sup_{u^1, \dots, u^d} |\tilde{C}_n(u^1, \dots, u^d) - C_n(u^1, \dots, u^d)| \rightarrow 0$ for $n \rightarrow \infty$ and (7) follows. \square

In the sequel, all $U^i, i = 1, \dots, d$ are uniformly distributed random variables on $[0, 1]$ and all integrals have to be understood in the sense of Lebesgue-Stieltjes. Note that the next theorem is an extension of [Fermanian et al., 2004, Theorem 6] from the case of bivariate to the case of multi-variate random vectors $U = (U^1, \dots, U^d)$.

Theorem 3.1 *Let the copula C of (U^1, \dots, U^d) have continuous partial derivatives and let $f : [0, 1]^d \rightarrow \mathbb{R}$ be a right-continuous function of bounded variation in the sense of Hardy-Krause. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) - \mathbb{E}[f(U^1, \dots, U^d)] \right) \xrightarrow{\mathcal{D}} \int_{[0,1]^d} G_C(u^1, \dots, u^d) d\hat{f}(u^1, \dots, u^d),$$

where the function $\hat{f} : [0, 1]^d \rightarrow \mathbb{R}$ is defined by:

$$\hat{f}(u^1, \dots, u^d) = \begin{cases} 0 & \text{if at least one } u^j = 1, \text{ for } j = 1, \dots, d, \\ f(u^1, \dots, u^d) & \text{otherwise.} \end{cases} \quad (8)$$

Furthermore, the limit distribution is Gaussian.

Proof:

By definition \hat{f} is right-continuous and of bounded variation in the sense of Hardy-Krause. Furthermore, it follows that almost surely

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) - \mathbb{E}[f(U^1, \dots, U^d)] \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\widehat{f}(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) - \mathbb{E}[\widehat{f}(U^1, \dots, U^d)] \right),$$

by the fact that C is continuous on $[0, 1]^d$.

We use a multidimensional integration-by-parts technique proposed by [Zaremba, 1968, Proposition 2]. Using the notation of [Zaremba, 1968] we get

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\widehat{f}(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) - \mathbb{E}[\widehat{f}(U^1, \dots, U^d)] \right) \\ &= \sqrt{n} \int_{[0,1]^d} \widehat{f}(u^1, \dots, u^d) d(C_n - C)(u^1, \dots, u^d) \\ &= \sqrt{n} \sum_{k=0}^d (-1)^k \sum_{1, \dots, d; k} \Delta_{j_{k+1}, \dots, j_d}^* \int_{[0,1]^k} (C_n - C)(u^1, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d). \quad (9) \end{aligned}$$

Here $\sum_{1, \dots, d; k}$ denotes the sum over all possible partitions of the set $\{j_1, \dots, j_d\}$ into two subsets $\{j_1, \dots, j_k\}$ and $\{j_{k+1}, \dots, j_d\}$ of k respectively $d - k$ elements, where each partition is taken exactly once. In the cases $k = 0$ and $k = d$, the sum is interpreted as being reduced to one term.

Furthermore, the operator d_{j_1, \dots, j_k} indicates that the integral only applies to the variables j_1, \dots, j_k . Note that after the application of the integral with respect to $d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d)$, the integrated function is a function in $d - k$ variables. Furthermore for a function g of $d - k$ variables, the operator $\Delta_{j_{k+1}, \dots, j_d}^*$ is given by

$$\Delta_{j_{k+1}, \dots, j_d}^* g(j_{k+1}, \dots, j_d) = \sum_{\{i_1, \dots, i_{d-k}\} \in \{0,1\}^{d-k}} (-1)^m g(i_1, \dots, i_{d-k}),$$

where m denotes the number of zeros in $\{i_1, \dots, i_{d-k}\}$. This means that, for $j \notin \{j_1, \dots, j_k\}$

$$\begin{aligned} & \Delta_j^* \int_{[0,1]^{d-k}} (C_n - C)(u^1, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d) \\ &= \int_{[0,1]^{d-k}} (C_n - C)(u^1, \dots, u^{j-1}, 1, u^{j+1}, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^{j-1}, 1, u^{j+1}, \dots, u^d) \\ &\quad - \int_{[0,1]^{d-k}} (C_n - C)(u^1, \dots, u^{j-1}, 0, u^{j+1}, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^{j-1}, 0, u^{j+1}, \dots, u^d) \end{aligned}$$

and

$$\Delta_{j_{k+1}, \dots, j_d}^* = \Delta_{j_{k+1}}^* \dots \Delta_{j_d}^*.$$

Thus

$$\begin{aligned} & \sqrt{n} \sum_{k=0}^d (-1)^k \sum_{1, \dots, d; k} \Delta_{j_{k+1}, \dots, j_d}^* \int_{[0,1]^k} (C_n - C)(u^1, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d) \\ &= \sqrt{n} \sum_{k=0}^{d-1} (-1)^k \sum_{1, \dots, d; k} \Delta_{j_{k+1}, \dots, j_d}^* \int_{[0,1]^k} (C_n - C)(u^1, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d) \\ &\quad + \sqrt{n} (-1)^d \int_{[0,1]^d} (C_n - C)(u^1, \dots, u^d) d\widehat{f}(u^1, \dots, u^d) \\ &= \sqrt{n} (-1)^d \int_{[0,1]^d} (C_n - C)(u^1, \dots, u^d) d\widehat{f}(u^1, \dots, u^d). \end{aligned}$$

The term

$$\sqrt{n} \sum_{k=0}^{d-1} (-1)^k \sum_{1, \dots, d; k} \Delta_{j_{k+1}, \dots, j_d}^* \int_{[0,1]^k} (C_n - C)(u^1, \dots, u^d) d_{j_1, \dots, j_k} \widehat{f}(u^1, \dots, u^d)$$

vanishes because each of its terms is equal to zero due to at least one of the following two reasons: firstly, at least one $u^j, j = 1, \dots, d$ is equal to one and therefore $\widehat{f}(u^1, \dots, u^d) = 0$ by definition, or, secondly, at least one $u^j, j = 1, \dots, d$ is equal to zero, hence $C_n(u^1, \dots, u^d) = C(u^1, \dots, u^d) = 0$. Thus, by the continuous mapping theorem and (7), it follows that

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) - \mathbb{E}[f(U^1, \dots, U^d)] \right) \\ &= (-1)^d \sqrt{n} \int_{[0,1]^d} (C_n - C)(u^1, \dots, u^d) d\widehat{f}(u^1, \dots, u^d) \\ &\xrightarrow{\mathcal{D}} \int_{[0,1]^d} G_C(u^1, \dots, u^d) d\widehat{f}(u^1, \dots, u^d). \end{aligned}$$

Since $\int_{[0,1]^d} G_C(u^1, \dots, u^d) d\widehat{f}(u^1, \dots, u^d)$ is a continuous, linear transformation of a tight Gaussian process, it follows that the limiting distribution is Gaussian. \square

Remark 3.1 *The reason for using the function \widehat{f} instead of f is that the integrals of dimension $k = 2, \dots, d - 1$ in (9) are in general not vanishing. The one-dimensional integrals are zero for every right-continuous function of bounded variation f because of special properties of the function C_n , for more details see [Fermanian et al., 2004]. In particular, this means that in the two-dimensional case it is sufficient to assume*

$$\widehat{f}(x) = f(x), \quad x \in \mathbb{R}^2.$$

With this assumption instead of (8) and $d = 2$, Theorem 3.1 is equivalent to [Fermanian et al., 2004, Theorem 6]. We use the function \widehat{f} to get a more convenient representation for the limit variance of the LHSD technique, which we state in the next theorem.

Theorem 3.2 *Under the assumptions and notations of Theorem 3.1, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(f(V_{i,n}^1, \dots, V_{i,n}^d) - \mathbb{E}[f(U^1, \dots, U^d)] \right) \xrightarrow{\mathcal{D}} N(0, \sigma_{LHSD}^2), \quad (10)$$

where

$$\sigma_{LHSD}^2 = \int_{[0,1]^{2d}} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] d\widehat{f}(u^1, \dots, u^d) d\widehat{f}(\bar{u}^1, \dots, \bar{u}^d). \quad (11)$$

Proof:

We want to apply Theorem 3.1 together with Lemma 3.1, so we have to show that

$$\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left[f(V_{i,n}^1, \dots, V_{i,n}^d) - f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) \right] \right| \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

By [Leonov, 1998, Corollary 1]

$$\left| \sum_{i=1}^n \left[f(V_{i,n}^1, \dots, V_{i,n}^d) - f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) \right] \right| \leq V(f) < \infty,$$

where $V(f)$ is the Hardy-Krause variation of f . Hence

$$\frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \left[f(V_{i,n}^1, \dots, V_{i,n}^d) - f(F_n^1(U_i^1), \dots, F_n^d(U_i^d)) \right] \right| \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

which, together with Lemma 3.1 and Theorem 3.1, proves equation (10).

To derive equation (11) we apply Fubini's theorem to $\mathbb{E}[(\int_{[0,1]^d} G_C(u^1, \dots, u^d) d\hat{f}(u^1, \dots, u^d))^2]$. By [Leonov, 1998, Theorem 3] a function of bounded variation \hat{f} can always be written as the difference of two completely monotone functions g, h and therefore an integral with respect to \hat{f} can be written as a difference of two integrals with respect to positive measures g, h . Thus

$$\begin{aligned} & \mathbb{E} \left[\left(\int_{[0,1]^d} G_C(u^1, \dots, u^d) d\hat{f}(u^1, \dots, u^d) \right)^2 \right] = \\ &= \mathbb{E} \left[\left(\int_{[0,1]^d} G_C(u^1, \dots, u^d) d\hat{f}(u^1, \dots, u^d) \right) \cdot \left(\int_{[0,1]^d} G_C(\bar{u}^1, \dots, \bar{u}^d) d\hat{f}(\bar{u}^1, \dots, \bar{u}^d) \right) \right] \\ &= \mathbb{E} \left[\left(\int_{[0,1]^d} G_C(u^1, \dots, u^d) dg(u^1, \dots, u^d) - \int_{[0,1]^d} G_C(u^1, \dots, u^d) dh(u^1, \dots, u^d) \right) \right. \\ & \quad \cdot \left. \left(\int_{[0,1]^d} G_C(\bar{u}^1, \dots, \bar{u}^d) dg(\bar{u}^1, \dots, \bar{u}^d) - \int_{[0,1]^d} G_C(\bar{u}^1, \dots, \bar{u}^d) dh(\bar{u}^1, \dots, \bar{u}^d) \right) \right] \\ &= \mathbb{E} \left[\left(\int_{[0,1]^{2d}} G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) dg(u^1, \dots, u^d) dg(\bar{u}^1, \dots, \bar{u}^d) \right. \right. \\ & \quad - \int_{[0,1]^d} G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) dh(u^1, \dots, u^d) dg(\bar{u}^1, \dots, \bar{u}^d) \\ & \quad - \int_{[0,1]^d} G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) dg(u^1, \dots, u^d) dh(\bar{u}^1, \dots, \bar{u}^d) \\ & \quad \left. \left. + \int_{[0,1]^d} G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) dh(u^1, \dots, u^d) dh(\bar{u}^1, \dots, \bar{u}^d) \right) \right] \\ &= \int_{[0,1]^{2d}} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] dg(u^1, \dots, u^d) dg(\bar{u}^1, \dots, \bar{u}^d) \\ & \quad - \int_{[0,1]^d} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] dh(u^1, \dots, u^d) dg(\bar{u}^1, \dots, \bar{u}^d) \\ & \quad - \int_{[0,1]^d} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] dg(u^1, \dots, u^d) dh(\bar{u}^1, \dots, \bar{u}^d) \\ & \quad + \int_{[0,1]^d} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] dh(u^1, \dots, u^d) dh(\bar{u}^1, \dots, \bar{u}^d) \\ &= \int_{[0,1]^{2d}} \mathbb{E} \left[G_C(u^1, \dots, u^d) G_C(\bar{u}^1, \dots, \bar{u}^d) \right] d\hat{f}(u^1, \dots, u^d) d\hat{f}(\bar{u}^1, \dots, \bar{u}^d), \end{aligned}$$

where the use of Fubini's theorem is justified since \hat{f} is bounded and $\mathbb{E}[XY] < \infty$ for two jointly normal random variables X and Y . \square

Remark 3.2 Note that by (5) and (6) the expression for σ_{LHSD}^2 in equation (11) can be represented in terms of C . Additionally, further simplifications can be given for the following terms:

$$\begin{aligned} & \mathbb{E}[B_C(u^1, \dots, u^d) \cdot B_C(1, \dots, 1, \bar{u}^j, 1, \dots, 1)] \\ &= C((u^1, \dots, u^{j-1}, u^j \wedge \bar{u}^j, u^{j+1}, \dots, u^d)) - C(u^1, \dots, u^d) \bar{u}^j, \\ & \mathbb{E}[B_C(1, \dots, 1, u^i, 1, \dots, 1) \cdot B_C(1, \dots, 1, \bar{u}^j, 1, \dots, 1)] \\ &= C((1, \dots, 1, u^i, 1, \dots, 1, \bar{u}^j, 1, \dots, 1)) - u^i \bar{u}^j, \\ & \mathbb{E}[B_C(1, \dots, 1, u^j, 1, \dots, 1) \cdot B_C(1, \dots, 1, \bar{u}^j, 1, \dots, 1)] = u^j \wedge \bar{u}^j - u^j \bar{u}^j, \end{aligned}$$

since $C(1, \dots, 1, u^j, 1, \dots, 1) = u^j$ for all $j = 1, \dots, d$.

It is important to know if the LHSD estimator has a smaller variance than the Monte Carlo estimator. The variance of a standard Monte Carlo estimator is given by

$$\sigma_{MC}^2 = \int_{[0,1]^d} f(u^1, \dots, u^d)^2 dC(u^1, \dots, u^d) - \left(\int_{[0,1]^d} f(u^1, \dots, u^d) dC(u^1, \dots, u^d) \right)^2.$$

We use this fact to establish a relation between σ_{MC}^2 and σ_{LHSD}^2 .

Proposition 3.3 Let the copula C of (U^1, \dots, U^d) have continuous partial derivatives, let $f : [0, 1]^d \rightarrow \mathbb{R}$ be a right-continuous function of bounded variation in the sense of Hardy-Krause and let \hat{f} be as defined in Theorem 3.1. Set $\partial_j C(u^1, \dots, u^d) = \frac{\partial C(u^1, \dots, u^d)}{\partial u^j}$ and

$$C_{i,j}(u^i, \bar{u}^j) = \begin{cases} C(1, \dots, 1, u^i, 1, \dots, 1, \bar{u}^j, 1, \dots, 1), & i \neq j \\ u^i \wedge \bar{u}^j, & i = j. \end{cases}$$

Then

$$\begin{aligned} \sigma_{LHSD}^2 &= \sigma_{MC}^2 \\ &+ \int_{[0,1]^{2d}} 2 \sum_{j=1}^d \partial_j C(u^1, \dots, u^d) \left(C(\bar{u}^1, \dots, \bar{u}^d) u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) \right) \\ &+ \sum_{j=1}^d \sum_{i=1}^d \partial_j C(\bar{u}^1, \dots, \bar{u}^d) \partial_i C(u^1, \dots, u^d) \left(C_{i,j}(u^i, \bar{u}^j) - u^i \bar{u}^j \right) d\hat{f}(u^1, \dots, u^d) d\hat{f}(\bar{u}^1, \dots, \bar{u}^d). \end{aligned} \tag{12}$$

Proof:

Note that

$$\int_{[0,1]^d} f(u^1, \dots, u^d)^2 dC(u^1, \dots, u^d) = \int_{[0,1]^{2d}} f(u^1, \dots, u^d) f(\bar{u}^1, \dots, \bar{u}^d) dC(u^1 \wedge \bar{u}^1, \dots, u^d \wedge \bar{u}^d),$$

and that the function $C(u^1 \wedge \bar{u}^1, \dots, u^d \wedge \bar{u}^d)$ is also a copula, which follows by observing that

$$\begin{aligned} C(u^1 \wedge \bar{u}^1, \dots, u^d \wedge \bar{u}^d) &= \mathbb{P}(U^1 \leq u^1 \wedge \bar{u}^1, \dots, U^d \leq u^d \wedge \bar{u}^d) \\ &= \mathbb{P}(U^1 \leq u^1, U^1 \leq \bar{u}^1, \dots, U^d \leq u^d, U^d \leq \bar{u}^d) \end{aligned}$$

is a joint probability distribution with uniform marginals.

By integration-by-parts like in Theorem 3.1 it follows for the variance of the Monte Carlo estimator that

$$\begin{aligned}
\sigma_{MC}^2 &= \int_{[0,1]^d} f(u^1, \dots, u^d)^2 dC(u^1, \dots, u^d) - \left(\int_{[0,1]^d} f(u^1, \dots, u^d) dC(u^1, \dots, u^d) \right)^2 \\
&= \int_{[0,1]^{2d}} f(u^1, \dots, u^d) f(\bar{u}^1, \dots, \bar{u}^d) dC\left((u^1, \dots, u^d) \wedge (\bar{u}^1, \dots, \bar{u}^d)\right) \\
&\quad - \int_{[0,1]^{2d}} f(u^1, \dots, u^d) f(\bar{u}^1, \dots, \bar{u}^d) dC(u^1, \dots, u^d) dC(\bar{u}^1, \dots, \bar{u}^d) \\
&= \int_{[0,1]^{2d}} C\left((u^1, \dots, u^d) \wedge (\bar{u}^1, \dots, \bar{u}^d)\right) d\hat{f}(u^1, \dots, u^d) d\hat{f}(\bar{u}^1, \dots, \bar{u}^d) \\
&\quad - \int_{[0,1]^{2d}} C(u^1, \dots, u^d) C(\bar{u}^1, \dots, \bar{u}^d) d\hat{f}(u^1, \dots, u^d) d\hat{f}(\bar{u}^1, \dots, \bar{u}^d).
\end{aligned}$$

The proof is completed by using equations (5), (6), (11) and Remark 3.2. \square

Theorem 3.3 *Let C and f satisfy the assumptions in Theorem 3.1 and let \hat{f} be defined as in Theorem 3.1. Furthermore let the function f be monotone non-decreasing in each argument and $\max_{x \in [0,1]^d} (f(x)) \leq 0$. Moreover assume that C satisfies the following conditions:*

$$\frac{C(u^1, \dots, u^d)}{u^j} \geq \partial_j C(u^1, \dots, u^d), \quad j \in \{1, \dots, d\}, \quad (13)$$

$$\sum_{i=1, i \neq j}^d \frac{C_{i,j}(u^j, \bar{u}^i)}{\bar{u}^j} \leq (d-2)u^j + \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)}, \quad (14)$$

where $u^j \in [0, 1]$, $(\bar{u}^1, \dots, \bar{u}^d), (u^1, \dots, u^d) \in [0, 1]^d$.

Then $\sigma_{LHSD}^2 \leq \sigma_{MC}^2$.

Proof:

By the assumptions on f it follows that \hat{f} is right-continuous, of bounded variation in the sense of Hardy-Kraus and monotone non-decreasing in each argument. Thus by (12) it is sufficient to show that

$$\begin{aligned}
&2 \sum_{j=1}^d \partial_j C(u^1, \dots, u^d) \left(C(\bar{u}^1, \dots, \bar{u}^d) u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) \right) \\
&+ \sum_{j=1}^d \sum_{i=1}^d \partial_i C(\bar{u}^1, \dots, \bar{u}^d) \partial_j C(u^1, \dots, u^d) \left(C_{i,j}(u^j, \bar{u}^i) - u^j \bar{u}^i \right) \leq 0
\end{aligned}$$

for all $(u^1, \dots, u^d), (\bar{u}^1, \dots, \bar{u}^d) \in [0, 1]^d$.

This is true if

$$2 \left(C(\bar{u}^1, \dots, \bar{u}^d) u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) \right) \leq \sum_{i=1}^d \partial_i C(\bar{u}^1, \dots, \bar{u}^d) \left(u^j \bar{u}^i - C_{i,j}(u^j, \bar{u}^i) \right)$$

holds for every $j \in \{1, \dots, d\}$ and all $u^j \in [0, 1]$, $(\bar{u}^1, \dots, \bar{u}^d) \in [0, 1]^d$.

First we show that

$$C(\bar{u}^1, \dots, \bar{u}^d) u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) \leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) \left(u^j \bar{u}^j - u^j \wedge \bar{u}^j \right).$$

Note that this is always true if $u^j \wedge \bar{u}^j \in \{0, 1\}$. Now assume that $0 < \bar{u}^j \leq u^j < 1$, then

$$\begin{aligned} C(\bar{u}^1, \dots, \bar{u}^d)u^j - C(\bar{u}^1, \dots, \bar{u}^d) &\leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) (u^j \bar{u}^j - \bar{u}^j) \\ C(\bar{u}^1, \dots, \bar{u}^d)(u^j - 1) &\leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) \bar{u}^j (u^j - 1) \\ \frac{C(\bar{u}^1, \dots, \bar{u}^d)}{\bar{u}^j} &\geq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) \end{aligned}$$

which is true by assumption (13). Next assume that $0 < u^j < \bar{u}^j < 1$, then

$$\begin{aligned} C(\bar{u}^1, \dots, \bar{u}^d)u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) &\leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) (u^j \bar{u}^j - u^j) \\ C(\bar{u}^1, \dots, \bar{u}^d)u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) &\leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) u^j (\bar{u}^j - 1) \\ C(\bar{u}^1, \dots, \bar{u}^d) - \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{u^j} &\leq \partial_j C(\bar{u}^1, \dots, \bar{u}^d) (\bar{u}^j - 1) \\ C(\bar{u}^1, \dots, \bar{u}^d) - \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{u^j} &\leq \frac{C(\bar{u}^1, \dots, \bar{u}^d)}{\bar{u}^j} (\bar{u}^j - 1) \\ \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{u^j} &\geq \frac{C(\bar{u}^1, \dots, \bar{u}^d)}{\bar{u}^j}, \end{aligned}$$

which holds since assumption (13) implies that $\frac{C(u^1, \dots, u^d)}{u^j}$ is non-increasing in u^j for all $u^j \in [0, 1]$, $(u^1, \dots, u^d) \in [0, 1]^d$.

Let $C(\bar{u}^1, \dots, \bar{u}^d) > 0$ then

$$\begin{aligned} C(\bar{u}^1, \dots, \bar{u}^d)u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) &\leq \sum_{\substack{i=1 \\ i \neq j}}^d \partial_i C(\bar{u}^1, \dots, \bar{u}^d) (u^j \bar{u}^i - C_{i,j}(u^j, \bar{u}^i)) \\ C(\bar{u}^1, \dots, \bar{u}^d)u^j - C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d) &\leq \sum_{\substack{i=1 \\ i \neq j}}^d \frac{C(\bar{u}^1, \dots, \bar{u}^d)}{\bar{u}^i} (u^j \bar{u}^i - C_{i,j}(u^j, \bar{u}^i)) \\ (d-2)u^j + \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)} &\geq \sum_{\substack{i=1 \\ i \neq j}}^d \frac{C_{i,j}(u^j, \bar{u}^i)}{\bar{u}^i} \end{aligned}$$

which is true by assumption (14). The case $C(\bar{u}^1, \dots, \bar{u}^d) = 0$ follows by the fact that $\frac{C(\bar{u}^1, \dots, \bar{u}^d)}{\bar{u}^i} \leq 1$ for all $(\bar{u}^1, \dots, \bar{u}^d) \in [0, 1]^d$. \square

Remark 3.3 Note that in the two-dimensional case, assumption (13) is equivalent to the left tail increasing property which implies a positive quadrant dependence of the copula C . Loosely speaking this means that the components of C are more likely to be simultaneously small or simulatneously large than in the independent case. More information on different dependence properties can be found in [Joe, 1997] and [Nelsen, 1999].

In the following two remarks we give examples of copula distributions which satisfy the assumptions of Theorem 3.3.

Remark 3.4 Consider a multi-dimensional, one-parametric extension of the Farlie-Gumbel-Morgenstern (FGM) copula given by

$$C(u^1, \dots, u^d) = \left(\prod_{i=1}^d u^i \right) \left(\alpha \prod_{i=1}^d (1 - u^i) + 1 \right)$$

where $\alpha \in [-1, 1]$. Simple calculations show that the assumption (13) is true if $\alpha \in [0, 1]$. Now consider the right hand-side of (14)

$$\begin{aligned} \sum_{i=1, i \neq j}^d \frac{C_{i,j}(u^j, \bar{u}^i)}{\bar{u}^i} &= \sum_{i=1, i \neq j}^d \frac{u^j \bar{u}^i}{\bar{u}^i} \\ &= (d-1)u^j. \end{aligned}$$

Finally assumption (14) holds since

$$\begin{aligned} &\frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)} \\ &= \min \left(1, \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)} \right) \\ &= \min \left(1, \frac{\left(\prod_{i=1, i \neq j}^d \bar{u}^i \right) u^j \left(\alpha \prod_{i=1, i \neq j}^d (1 - \bar{u}^i) (1 - u^j) + 1 \right)}{\left(\prod_{i=1}^d \bar{u}^i \right) \left(\alpha \prod_{i=1}^d (1 - \bar{u}^i) + 1 \right)} \right) \\ &= \min \left(1, u^j \frac{\left(\alpha \prod_{i=1, i \neq j}^d (1 - \bar{u}^i) (1 - u^j) + 1 \right)}{\bar{u}^j \left(\alpha \prod_{i=1}^d (1 - \bar{u}^i) + 1 \right)} \right) \\ &\geq u^j \end{aligned}$$

for $\alpha \in [0, 1]$.

Note that the independence copula $C(u^1, \dots, u^d) = \prod_{i=1}^d u^i$ is the special case of the FGM copula with $\alpha = 0$, therefore Theorem 3.3 holds also for the independence copula.

Remark 3.5 A multi-dimension version of the Ali-Mikhail-Haq (AMH) copula is given by

$$C(u^1, \dots, u^d) = \frac{\prod_{i=1}^d u^i}{1 - \alpha \prod_{i=1}^d (1 - u^i)}$$

where $\alpha \in [-1, 1]$. As in the previous example it is easy to see that (13) is fulfilled if $\alpha \in [0, 1]$. To prove (14) consider again the term on the right hand-side

$$\begin{aligned} \sum_{i=1, i \neq j}^d \frac{C_{i,j}(u^j, \bar{u}^i)}{\bar{u}^i} &= \sum_{i=1, i \neq j}^d \frac{u^j \bar{u}^i}{\bar{u}^i} \\ &= (d-1)u^j. \end{aligned}$$

Furthermore Theorem 3.3 applies since

$$\frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, \bar{u}^j \wedge u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)}$$

$$\begin{aligned}
&= \min \left(1, \frac{C(\bar{u}^1, \dots, \bar{u}^{j-1}, u^j, \bar{u}^{j+1}, \dots, \bar{u}^d)}{C(\bar{u}^1, \dots, \bar{u}^d)} \right) \\
&= \min \left(1, u^j \frac{\left(\prod_{i=1, i \neq j}^d \bar{u}^i \right) \left(1 - \alpha \prod_{i=1}^d (1 - \bar{u}^i) \right)}{\left(\prod_{i=1}^d \bar{u}^i \right) \left(1 - \alpha \prod_{i=1}^d (1 - \bar{u}^i) (1 - u_j) \right)} \right) \\
&\geq u^j
\end{aligned}$$

4 Application to option pricing

In this section we illustrate the effectiveness of Latin hypercube sampling with dependence in basket option pricing problems. The derivatives which we consider are Asian and lookback basket options. Let $(S_t)_{t \geq 0}$ be a d -dimensional vector of asset price processes and let $(S_t^j)_{t \geq 0}$ denote its j -th component. Then the price of an Asian basket call option is given by

$$\text{ABC} = \mathbb{E} \left[e^{-rT} \left(\frac{1}{m} \sum_{j=1}^m \frac{1}{d} \sum_{i=1}^d S_{t_j}^i - K \right)^+ \right],$$

where $K > 0$ denotes the fixed strike price, d is the number of underlying assets, $0 = t_0 < t_1 < t_2 < \dots < t_m = T$ denote the observation points, T is the maturity of the option and r denotes the risk free interest rate. Similarly, the price of a discrete lookback basket call option is given by

$$\text{DLC} = \mathbb{E} \left[e^{-rT} \left(\max_{j=1, \dots, m} \frac{1}{d} \sum_{i=1}^d S_{t_j}^i - K \right)^+ \right].$$

As a model for the asset price process $(S_t^j)_{t \geq 0}$ of each asset $j = 1, \dots, d$, we use

$$S_t^j = S_0^j e^{(w^j - r)t + X_t^j}, \quad j = 1, \dots, d, t \geq 0,$$

where $w^j \in \mathbb{R}$ are constants, $S_0^j > 0$ denote the constant initial asset values and X_t^j are variance gamma (VG) processes for $j = 1, \dots, d$. The VG process $(X_t^j)_{t \geq 0}$ with parameters $(\theta^j, \sigma^j, c^j)$, which was first introduced by [Madan and Seneta, 1990], is defined as a subordinated Brownian motion by

$$X_t^j = X_t^j(\theta^j, \sigma^j, c^j) = B_{G_t^j(c^j, 1)}^j(\theta^j, \sigma^j), \quad j = 1, \dots, d, t \geq 0, \quad (15)$$

where $B_t^j(\theta^j, \sigma^j)$ are independent Brownian motions with drift parameters θ^j and volatility parameters σ^j , $j = 1, \dots, d$, and $G_t^j(c^j, 1)$ are independent gamma processes independent of B^j , $j = 1, \dots, d$ with drift equal to one and volatility $c^j > 0$. To ensure that the discounted value of a portfolio invested in the asset is a martingale, we choose

$$w^j = \log(1 - \mu^j c^j - (\sigma^j)^2 c^j / 2) / c^j, \quad j = 1, \dots, d.$$

By [Madan et al., 1998] a VG process can also be represented as the difference of two independent gamma processes, ie $X_t^j = G_t^{+,j} - G_t^{-,j}$, $j = 1, \dots, d$. Let (μ_+^j, ν_+^j) and (μ_-^j, ν_-^j) denote the parameters of the gamma processes $G^{+,j}$, $G^{-,j}$, respectively. These pairs of parameters can be easily calculated from the parameters in equation (15) through

$$\mu_{\pm}^j = (\sqrt{(\theta^j)^2 + 2(\sigma^j)^2 / c^j} \pm \theta^j) / 2, \quad \nu_{\pm}^j = (\mu_{\pm}^j)^2 c^j, \quad j = 1, \dots, d.$$

Due to the fact that a gamma process has non-decreasing paths, $G_t^{+,j}$ corresponds to the positive movements of X_t^j and $G_t^{-,j}$ corresponds to the negative movements of X_t^j . Our assumption is that all positive movements of components of $X_t = (X_t^1, \dots, X_t^d)$ are dependent and all negative movements of components of X_t are dependent, but positive (negative) movements of the j -th component are independent of negative (positive) movements of all other components, for all $j = 1, \dots, d$. The dependence structure between positive and negative movements will be modelled by copulae C^\pm , respectively. Summarising, the increment of the d -dimensional gamma processes in the interval $[t_{i-1}, t_i]$ given by $(G_{t_i}^{\pm,1} - G_{t_{i-1}}^{\pm,1}, \dots, G_{t_i}^{\pm,d} - G_{t_{i-1}}^{\pm,d})$ has cumulative distribution function $C^\pm(F_{1,\pm}^{-1}, \dots, F_{d,\pm}^{-1})$, where $F_{j,\pm}^{-1}$ is the inverse cumulative distribution function of a gamma distribution with the specific parameters of the j -th asset.

4.1 Numerical results

In this subsection, we compare the performance of LHSD with a standard Monte Carlo method in option pricing problems.

Parameters of the numerical examples

VG parameters:	
$\mu_j, j = 1, \dots, d$	-0.2859
$\sigma_j, j = 1, \dots, d$	0.1927
$c_j, j = 1, \dots, d$	0.2505
Option parameters:	
number of assets d	10
maturity T	1
initial asset price $S_0^j, j = 1, \dots, d$	100
risk free interest rate r	0.05
number of monitoring points k	4
time between monitoring points $t_i - t_{i-1}, i = 1, \dots, k$	0.25
Simulation parameters:	
number of simulated option prices per estimator n	8000
number of simulations of the estimators m	100
choice of parameters $\eta_{i,n}^j, j = 1, \dots, d, i = 1, \dots, n$	0.5

Table 1: Parameters sets for the VG processes, the options and the simulations.

The parameters of the underlying VG processes are stated in Table 1 and are the same for all components of $(S_t)_{t \geq 0}$. The parameter values are taken from a calibration of the VG process against options on the S&P 500 index by [Hirsa and D.B.Madan, 2004]. We observed in price valuations, which we do not state here in detail, that the computation of one LHSD estimator took about 1.4 times of the computation time of a corresponding Monte Carlo estimator. Nevertheless in our concrete implementation the most time consuming part was the transformation of uniformly distributed random variables into gamma distributed random variables. This has to be done only once for all LHSD estimations since by (2) where $\eta_{i,n}^j = 1/2, j = 1, \dots, d, i = 1, \dots, n$ one only needs fixed quantiles of the gamma distribution. Therefore computation of 4000 LHSD estimators was about five times faster than the computation of 4000 Monte Carlo estimators. On the other hand for the Monte Carlo estimator, one has to perform the transformation dn times for each estimator.

Using the parameters of Table 1, the evaluation of each of the option values included the computation of an 80-dimensional integral. Standard deviation and variance were computed based on the $m = 100$

runs of the LHSD and MC estimators. The ratios in columns 6 and 7 of each table were computed as the quotient of MC value and LHSD value.

It is obvious that the effectiveness of LHSD compared to MC decreases with increasing strike price K . The same phenomenon was also observed by [Packham and Schmidt, 2010] in a multi-dimensional Black-Scholes model for the LHSD estimator and by [Glasserman, 2004] for the standard LHS estimator.

Prices of Asian basket call options with varying strike price K

α	K	Price LHSD	Price MC	Std. Dev. LHSD	Std. Dev. MC	Std. Dev. ratio	Var. ratio
0.5	80	22.0542	22.0448	0.00071	0.00748	10.419	108.575
0.5	90	12.5511	12.5419	0.00080	0.00748	9.270	85.944
0.5	100	3.79294	3.78732	0.00241	0.00621	2.577	6.642
0.5	110	0.17227	0.17210	0.00119	0.00140	1.174	1.379
0.5	120	0.00024	0.00024	0.000040	0.000041	1.009	1.018

Table 2: Prices of Asian basket call options, where the dependence structure of positive and negative movements are modelled by a FGM copula with parameter α .

Prices of Lookback basket call options with varying strike price K

α	K	Price LHSD	Price MC	Std. Dev. LHSD	Std. Dev. MC	Std. Dev. ratio	Var. ratio
0.5	80	25.662	25.658	0.00294	0.00839	2.850	8.125
0.5	90	16.151	16.147	0.00294	0.00839	2.850	8.125
0.5	100	6.893	6.890	0.00322	0.00760	2.356	5.553
0.5	110	1.192	1.192	0.00305	0.00406	1.332	1.775
0.5	120	0.060	0.060	0.00086	0.00089	1.029	1.060

Table 3: Prices of Lookback basket call options, where the dependence structure of positive and negative movements are modelled by a FGM copula with parameter α .

Bibliography

- J.-D. Fermanian, D. Radulovic, and M. Wegkamp. Weak convergence of empirical copula processes. *Journal of the Bernoulli Society*, 10(5):847–860, 2004.
- P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer Verlag, New York, 2004.
- A. Hirska and D.B.Madan. Pricing American options under variance gamma. *Journal of Computational Finance*, 7(2):63–80, 2004.
- J. Jacod and P. Protter. *Probability Essentials, 2nd edn*. Springer, Berlin, 2003.
- H. Joe. *Multivariate Models and Dependence Concepts*. Chapman and Hall, London, 1997.
- A. Leonov. On the total variation for functions of several variables and a multidimensional analog of Helly’s selection principle. *Mathematical Notes*, 63(1):61–71, 1998.
- W.-L. Loh. On Latin hypercube sampling. *The annals of statistics*, 24(5):2058–2080, 1996.

- D.B. Madan and E. Seneta. The variance gamma (v.g.) model for share market returns. *The Journal of Business*, 63: 511–524, 1990.
- D.B. Madan, P.P. Carr, and E.C. Chang. The variance gamma process and option pricing. *European Finance Review*, 1:39–55, 1998.
- M.D. McKay, R.J. Beckman, and W.J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245, 1979.
- R.B. Nelsen. *An Introduction to Copulas*. Springer, New York, 1999.
- A.B. Owen. A central limit theorem for Latin hypercube sampling. *Journal of the Royal Statistical Society Series B* 54, 13:541–551, 1992.
- A.B. Owen. Multidimensional variation for quasi-Monte Carlo. In J. Fan and G. Li (Eds.), *International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday*, 2005.
- N. Packham and W.M. Schmidt. Latin hypercube sampling with dependence and application in finance. *Journal of Computational Finance*, 13(3):81–111, 2010.
- M. Stein. Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.
- W. Stute. The oscillation behaviour of empirical processes: the multivariate case. *The Annals of Probability*, 12: 361–379, 1984.
- H. Tsukahara. Semiparametric estimation in copula models. *The Canadian Journal of Statistics*, 33:357–375, 2005.
- S.K. Zaremba. Some applications of multidimensional integration by parts. *Ann. Polon. Math.*, 21:85–96, 1968.