Applications and Applied Mathematics: An International Journal (AAM)

Volume 10 | Issue 1 | Article 4

6-2015

Generating Random Vectors Using Transformation with Multiple Roots and its Applications

Qidi Peng
Claremont Graduate University

Henry Schellhorn
Claremont Graduate University

Lu Zhu
University of Wisconsin–Eau Claire

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466
Vol. 10, Issue 1 (June 2015), pp. 50 – 70

Applications and Applied Mathematics:
An International Journal
(AAM)

Generating Random Vectors Using Transformation with Multiple Roots and its Applications

[1]Qidi Peng, [2]Henry Schellhorn, and [3]Lu Zhu

[1]Corresponding author, Institute of Mathematical Sciences
Claremont Graduate University
710 College Avenue, Claremont, CA 91711, U.S.
Qidi.Peng@cgu.edu
&
[2]Institute of Mathematical Sciences, Claremont Graduate University
Henry.Schellhorn@cgu.edu
&
[3]Department of Accounting and Finance, University of Wisconsin–Eau Claire, U.S.
zhul@uwec.edu

Received: November 2, 2014; Accepted: April 21, 2015

Abstract

An approach is proposed to generate random vectors using transformation with multiple roots. This approach generalizes the one-dimensional inverse transformation with multiple roots method to higher dimensions, i.e., to random vectors with or without densities. In this approach, multiple roots of the transformation and probabilities of selecting each of the roots are derived. The strategies for constructing such a transformation are discussed and several examples are presented to motivate this simulation approach.

Keywords: Multiple roots; random vector; simulation

MSC 2010 No.: 65C10; 62H10; 60G15

1. Introduction

A general dynamical system can be mathematically represented by

$$V = g(X), \tag{1}$$

where g denotes some transformation (or an operator), X is the input signal and V is the output signal. When X and V are one-dimensional random variables and only V is observed, the problems regarding the probability distribution of X have been heavily studied (see, for instance, (Bonarini and Bontempi, 1994; Macháček, 1983; Rahman, 2009)). Nowadays, the interests of multivariate dynamical systems have grown fast thanks to their complexity and flexibility in practice. In this paper, we study the problem of generating X when both X and V are random vectors of dimension $n \geq 2$ and $g: \mathbb{R}^n \to \mathbb{R}^n$ is some transformation having a discrete set of roots, namely, either of the following two conditions holds:

(a) The set of roots of Equation (1) is finite.
(b) The set of roots of Equation (1) contains infinite number of isolated points. Recall that a point $p \in \mathbb{R}^n$ is called an isolated point in a set $B \subset \mathbb{R}^n$ if there exists an open neighborhood of p, which does not contain any other point in B (the set of all rational numbers in $\mathbb{R}$ is a counterexample.).

We call Model (1) a multivariate random dynamical system (e.g., (Bhattacharya and Majumdar, 2004; Bhattacharya and Majumdar, 2007)). Suppose that the output signal V can be observed or it is easier to generate V than X, then an intuitively appealing idea to generate X would be to first generate the random vector V, and then apply the inverse transformation of g to determine X. This method is trivial when g is invertible (see (Devroye, 1986)), but less so when g is a transformation with multiple roots (also called many-to-one transformation). To apply the inverse transformation method, the key problem is to determine which root will be mostly accepted. (Michael et al., 1976) introduced a general approach to generate one random variable provided that the density is known, by using a transformation with multiple roots, i.e., the case when the dimension $n = 1$.

We extend (Michael et al., 1976)'s approach to generate random vectors X (with or without densities) in Sections 2 and 3. When applying the multiple roots transformation method, the choice of g is crucial. Then Section 4 is devoted to discussing some strategies for constructing such a proper transformation.

In Section 5, two applications are discussed in simulating general uniform random variables and financial stochastic modeling respectively. In Example 4, we study the problem of generating random variables uniformly distributed over irregular domains. While generating uniform random variables over a convex polytope has received much attention in the literature, research on uniform distributions over more irregular domains such as quadratic curves has been sparse. We show in this example that transformation with multiple roots can reduce the complexity of using conditional density methods, as the dimension of irregular domain increases. Example 5 is another excellent application of the generating method, where it is used to simulate first exit times of correlated Brownian motions. The simulation algorithm has particular interests in describing the behavior of default time by n dependent firms in credit analysis, option pricing and risk management (for more on this topic, see (Zhou, 2001; Packham et al., 2013; Fok, 2013; Hashorva and Ji, 2014)).

2. Methodology and general result

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in which the probability measure $\mathbb{P}$ is as general as possible (i.e., its cumulative probability distribution function is not necessarily differentiable). For example, $\mathbb{P}$ can be discrete, continuous with or without density, or some mixture. Assume that a random vector $X = (X_1, \ldots, X_n)$'s joint distribution is given and further define a transformation $g: \mathbb{R}^n \to \mathbb{R}^n$, so that with probability 1,

$$g(X) = \big(g_1(X), \ldots, g_n(X)\big) = V, \tag{2}$$

where for $i = 1, \ldots, n$, $g_i : X(\Omega) \to \mathbb{R}$ is almost everywhere differentiable and $V := (V_1, \ldots, V_n)$. Assume that, for any specific observation $V = v$, the set of roots of Equation (2) is discrete. Then we denote its cardinality by $K \in \mathbb{N}^* \cup \{+\infty\}$ and the roots by: $\{r_k\}_{k=1,\ldots,K}$. Hence, the key of our approach for generating the random vector X lies in determining the probability of selecting each root r_k. For a specific value $x \in \mathbb{R}$, we denote by

$$\mathbb{P}(X \in \mathrm{d}x) := \mathrm{d}\mathbb{P}(X \le x),$$

the differential of the cumulative distribution function of X on x. Thus the probability of selecting the kth root r_k when $V = v$ is observed is given by $\mathbb{P}\big(X \in \mathrm{d}r_k | V = v\big)$. Note that

$$\mathbb{P}\big(X \in \mathrm{d}r_k | V = v\big) = \frac{\mathbb{P}(X \in \mathrm{d}r_k,\ V \in \mathrm{d}v)}{\mathbb{P}(V \in \mathrm{d}v)}. \tag{3}$$

On the one hand, the fact that $X \in \mathrm{d}r_k$ implies $V \in \mathrm{d}v$ shows

$$\mathbb{P}(X \in \mathrm{d}r_k,\ V \in \mathrm{d}v) = \mathbb{P}(X \in \mathrm{d}r_k). \tag{4}$$

On the other hand, since $\{r_k\}_{k=1,\ldots,K}$ is a discrete set, then $\{X \in \mathrm{d}r_k\}_{k=1,\ldots,K}$ are disjoint events. It follows that

$$\mathbb{P}\big(V \in \mathrm{d}v\big) = \mathbb{P}\Big(\bigcup_{k=1}^{K} \{X \in \mathrm{d}r_k\}\Big) = \sum_{k=1}^{K} \mathbb{P}(X \in \mathrm{d}r_k). \tag{5}$$

It results from (3), (4) and (5) that the probabilities of selecting each root behave as a multinomial probability distribution: for $k = 1, \ldots, K$,

$$\mathbb{P}\big(X \in \mathrm{d}r_k | V = v\big) = \frac{\mathbb{P}(X \in \mathrm{d}r_k)}{\sum_{i=1}^{K} \mathbb{P}(X \in \mathrm{d}r_i)}. \tag{6}$$

The generating method can be hence summarized as the following theorem.

Theorem 1 *Assume that the transformation g has $K \in \mathbb{N}^* \cup \{+\infty\}$ distinct isolated roots, the* *following equality holds in distribution:*

$$X \sim \sum_{k=1}^{K} R_k \quad_{U\in I_k},$$

where R_k is the kth distinct root of equation $V = g(X)$; U denotes a uniform random variable $Unif([0,1])$ independent of V and the subintervals $\{I_k\}_{k=1,\dots,K}$ are chosen as any partition of $[0,1]$ such that

$$\mathbb{P}\big(U \in I_k | V = v\big) = \frac{\mathbb{P}(X \in \mathrm{d}r_k)}{\sum\limits_{i=1}^{K} \mathbb{P}(X \in \mathrm{d}r_i)},$$

where $\{r_k\}_{k=1,\dots,K}$ are all the distinct roots of equation $v = g(x)$ for any specific observation $V = v$.

Remark that, in Theorem 1, the partition $\{I_k\}_{k=1,\dots,K}$ is generally a set of random elements. They depend on V and can be chosen as K subintervals of $[0,1]$ equal in size to the probabilities of selecting each of the K distinct roots. The examples of explicitly constructing these intervals will be provided in Sections 4 and 5. We also remark that, when applying Theorem 1, the K roots don't necessarily need to be distinct. Each root can be surplus.

When the random vectors X and V follow discrete probability distributions, the selection probabilities in Theorem 1 can be straightforwardly expressed as multinomial probabilities:

$$\mathbb{P}(U \in I_k | V = v) = \frac{\mathbb{P}(X = r_k)}{\sum\limits_{i=1}^{K} \mathbb{P}(X = r_i)}.$$

When the random vectors are absolutely continuous with a density, the problem of determining selection probabilities becomes more complicated than in a discrete probability space. We consider the probability on $X = r_k$ as that of X lying in an arbitrarily small sized rectangle $(r_k - \mathrm{d}r, r_k + \mathrm{d}r)$ in $\mathbb{R}^n$. The ratio of discrete probabilities becomes the ratio of densities. In this way, we derive the corresponding simulating approach in the next section.

3. Generate random vector with density

Assume that X is a continuous random vector and has joint density f_X. Let $g = (g_1, \dots, g_n) : \mathbb{R}^n \to \mathbb{R}^n$ be defined as: for $i \in \{1, \dots, n\}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ is an almost everywhere differentiable function over $X(\Omega)$ and their derivatives verify: for the K distinct roots $\{r_k\}_{k=1,\dots,K}$ of Equation (2) and any $k, l \in \{1, \dots, K\}$,

$$\prod_{i=1}^{n} \left| \frac{\partial_i g_i(r_k)}{\partial_i g_i(r_l)} \right| < +\infty, \tag{7}$$

where ∂_i denotes the partial derivative with respect to the ith coordinate. Let $h = (h_1, \dots, h_n)$, with $h_i > 0$ for $i \in \{1, \dots, n\}$. Since the K distinct roots are isolated, when h_i is very small, the reciprocal image of the neighborhood of v, $(v - h, v + h)$, consists of K disjoint neighborhoods of the K distinct roots. The neighborhood of v is hence given as:

$$\big(v-h, v+h\big) := \big(v_1-h_1, v_1+h_1\big) \times \big(v_2-h_2, v_2+h_2\big) \times \ldots \times \big(v_n-h_n, v_n+h_n\big).$$

Denote the part of reciprocal image including the kth root $r_k = (r_1^{(k)}, \ldots, r_n^{(k)})$ by $\prod_{i=1}^n \big(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)}\big)$. Hence, for $k = 1, \ldots, K$ and $i = 1, \ldots, n$,

$$g_i\Big(\prod_{l=1}^{n}(y_{l,\text{low}}^{(k)}, y_{l,\text{up}}^{(k)})\Big) = \big(v_i - h_i, v_i + h_i\big). \tag{8}$$

Given that v is in the neighborhood $(v-h, v+h)$, Theorem 1 entails that the chance that the kth root is selected, $p_k^h(v)$, is

$$\begin{aligned} p_k^h(v) &:= \frac{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)})\Big)}{\sum_{l=1}^{K}\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(l)}, y_{i,\text{up}}^{(l)})\Big)} \\ &= \left(1 + \sum_{l\in\{1,\ldots,K\}\setminus\{k\}} \frac{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(l)}, y_{i,\text{up}}^{(l)})\Big)}{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)})\Big)}\right)^{-1}. \end{aligned} \tag{9}$$

Notice that, in the sense of set-theoretic limit,

$$\lim_{h\to 0^+}\prod_{i=1}^{n}\big(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)}\big) = \{r_k\}, \text{ and } \lim_{h\to 0^+}\big(v-h, v+h\big) = \{v\}.$$

Therefore the kth root should be selected with probability $p_k(v) := \lim_{h\to 0^+} p_k^h(v)$. More precisely, from (9) we have

$$\begin{aligned} p_k(v) &= \lim_{h\to 0^+} p_k^h(v) \\ &= \left(1 + \sum_{l\in\{1,\ldots,K\}\setminus\{k\}} \lim_{h\to 0^+} \frac{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(l)}, y_{i,\text{up}}^{(l)})\Big)}{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)})\Big)}\right)^{-1} \\ &= \left(1 + \sum_{l\in\{1,\ldots,K\}\setminus\{k\}} \lim_{h\to 0^+} \frac{\frac{\mathbb{P}\big(X\in\prod_{i=1}^{n}(y_{i,\text{low}}^{(l)}, y_{i,\text{up}}^{(l)})\big)}{\prod_{i=1}^{n}(y_{i,\text{up}}^{(l)} - y_{i,\text{low}}^{(l)})}\prod_{i=1}^{n}\frac{(y_{i,\text{up}}^{(l)} - y_{i,\text{low}}^{(l)})}{2h_i}}{\frac{\mathbb{P}\big(X\in\prod_{i=1}^{n}(y_{i,\text{low}}^{(k)}, y_{i,\text{up}}^{(k)})\big)}{\prod_{i=1}^{n}(y_{i,\text{up}}^{(k)} - y_{i,\text{low}}^{(k)})}\prod_{i=1}^{n}\frac{(y_{i,\text{up}}^{(k)} - y_{i,\text{low}}^{(k)})}{2h_i}}\right)^{-1}. \end{aligned} \tag{10}$$

Since g_i's are almost everywhere differentiable, the limit in (10) exists and can be expressed in terms of densities. To this end we observe that, on the one hand,

$$\frac{\mathbb{P}\Big(X \in \prod_{i=1}^{n}(y_{i,\mathrm{low}}^{(l)}, y_{i,\mathrm{up}}^{(l)})\Big)}{\prod_{i=1}^{n}(y_{i,\mathrm{up}}^{(l)} - y_{i,\mathrm{low}}^{(l)})} \xrightarrow[h\to 0^+]{} f_X(r_l); \tag{11}$$

on the other hand, for $l \in \{1,\dots,n\}$, using (8),

$$\frac{2h_i}{(y_{i,\mathrm{up}}^{(l)} - y_{i,\mathrm{low}}^{(l)})} = \frac{\lambda\big((v_i - h_i, v_i + h_i)\big)}{\lambda\big((y_{i,\mathrm{low}}^{(l)}, y_{i,\mathrm{up}}^{(l)})\big)} = \frac{\lambda\Big(g_i\Big(\prod_{j=1}^{n}(y_{j,\mathrm{low}}^{(l)}, y_{j,\mathrm{up}}^{(l)})\Big)\Big)}{\lambda\big((y_{i,\mathrm{low}}^{(l)}, y_{i,\mathrm{up}}^{(l)})\big)}$$
$$\xrightarrow[h\to 0^+]{} \big|\partial_i g_i(r_l)\big|, \tag{12}$$

where λ denotes the Lebesgue measure on $\mathbb{R}$.

Finally, it follows from (10), (11) and (12) that, for $k \in \{1,\dots,K\}$,

$$p_k(v) = \left(1 + \sum_{l\in\{1,\dots,K\}\setminus\{k\}} \frac{f_X(r_l)}{f_X(r_k)}\Big(\prod_{i=1}^{n}\Big|\frac{\partial_i g_i(r_k)}{\partial_i g_i(r_l)}\Big|\Big)\right)^{-1}. \tag{13}$$

The sequence of probabilities in (13) together with Equation (1) yields the following result:

Proposition 1 *Under the above assumptions on X, V and g, the following equality holds in distribution:*

$$X \sim \sum_{k=1}^{K} R_k \ {}_{U\in I_k},$$

where $\{R_k\}_{k=1,\dots,K}$ are the multiple roots of equation $V = g(X)$; $U \sim Unif([0,1])$ is independent of V and conditional on $V = v$, $\{I_k\}_{k\in\{1,\dots,K\}}$ is a partition of interval $[0,1]$ satisfying

$$\mathbb{P}\big(U \in I_k | V = v\big) = \left(1 + \sum_{l\in\{1,\dots,K\}\setminus\{k\}} \frac{f_X(r_l)}{f_X(r_k)}\Big(\prod_{i=1}^{n}\Big|\frac{\partial_i g_i(r_k)}{\partial_i g_i(r_l)}\Big|\Big)\right)^{-1}.$$

4. Construction of transformation with multiple roots

When applying the multiple roots transformation approach, the main question is how to choose the proper transformation g. In the following we mainly discuss random vectors with densities. The strategies of choosing g for generating discrete random vectors are quite similar. We first present a useful result. Given f_X, the density of the random vector X, and a multiple roots transformation g, the fundamental theorem (see (Papoulis, 1991)) below provides a representation of the density of $V = g(X)$.

Theorem 2 (An extension of the change of variable theorem) *Let g be the transformation defined in Equation (2), let f_X be the joint density of $X = (X_1,\dots,X_n)$. Denote by $r_1,\dots,r_K \in \mathbb{R}^n$*

the K isolated roots of equation $v = g(x)$, then the joint density f_V of $V = g(X)$ is given as: for $v \in \mathbb{R}^n$,

$$f_V(v) = \sum_{k=1}^{K} \frac{f_X(r_k)}{|det(J_g(r_k))|},$$

where $det(J_g(r_k))$ denotes the determinant of the Jacobian matrix of g on r_k.

We note that this determinant $det(J_g(r_k))$ does not vanish for any r_k, since g is assumed to be locally invertible in neighborhood of r_k, for $k = 1, \ldots, K$.

Now we introduce some situations, where the multiple roots transformation method could be applied.

Case 1V is uniformly distributed over a regular domain.

In this case we discuss two examples: the density f_X is a linear transformation of symmetric function and f_X is periodic.

Case 1,f_X is a linear transformation of some symmetric function.

Definition 1 *A function $S : \mathbb{R}^n \to \mathbb{R}$ is said to be symmetric with respect to some point $t \in \mathbb{R}^n$, if $S(x) = -S(2t - x)$ for all $x \in \mathbb{R}^n$. For example, the mapping $(x_1, x_2) \longmapsto (\cos(x_1), \cos(x_2))$ is symmetric with respect to $(\frac{\pi}{2}, \frac{\pi}{2})$.*

Let $D \subset \mathbb{R}^n$, $t = (t_1, \ldots, t_n) \in D$ and let $X = (X_1, \ldots, X_n)$ be a random vector taking values over D with joint density f_X satisfying:

$$f_X(x) = aS(x) + b, \tag{14}$$

where $a \neq 0$, $b > 0$ and S is symmetric with respect to t over D.

In this case, we can apply the absolute value transformation for simulating X. The method involves setting

$$V = g(X) = \big(|X_1 - t_1|, |X_2 - t_2|, \ldots, |X_n - t_n|\big).$$

g verifies the conditions in Equation (2). Observe that the equation $g(x) = v = (v_1, \ldots, v_n)$ has 2^n isolated roots $\{r_k\}_{k \in \{1,\ldots,2^n\}}$. Theorem 2 together with (14) entails that the joint density of V can be written as: for $v \in g(D)$,

$$f_V(v) = \sum_{k=1}^{2^n} f_X(r_k) = a \sum_{k=1}^{2^n} S(r_k) + 2^n b.$$

By construction of g, if r_k is a root of the equation $g(x) = v$, then $2t - r_k$ is also a root. Therefore by the property of symmetry of S, we get $\sum_{k=1}^{2^n} S(r_k) = 0$. It follows that

$$f_V(v) = 2^n b \tag{15}$$

for $v \in g(D)$, which turns out to be a uniform probability density with support $g(D)$. This absolute value transformation is particularly useful when $g(D)$ is a regular domain, such as rectangle, triangle, disk, ...

Example 1 (The absolute value transformation for symmetric density)

Suppose the target density is given as: for $x = (x_1, \dots, x_n) \in [0, \pi]^n$,

$$f_X(x) = \frac{1}{n\pi^n}\Big(n + \sum_{i=1}^{n} \cos(x_i)\Big),$$

$f_X(x) = 0$ for $x \notin [0, \pi]^n$. Observe that the function $S(x) = \sum_{i=1}^{n} \cos(x_i)$ is symmetric over $[0, \pi]^n$ with respect to $t = (\frac{\pi}{2}, \dots, \frac{\pi}{2})$. Thus according to (15) with $b = \pi^{-n}$, after transforming X to absolute value

$$V = g(X) = \Big(\big|X_1 - \frac{\pi}{2}\big|, \dots, \big|X_n - \frac{\pi}{2}\big|\Big),$$

the density of V is given as: for $v \in [0, \frac{\pi}{2}]^n$, $f_V(v) = \big(\frac{2}{\pi}\big)^n$; $f_V(v) = 0$ if $v \notin [0, \frac{\pi}{2}]^n$. Therefore V is a uniformly distributed random vector over $g([0, \pi]^n) = [0, \frac{\pi}{2}]^n$. Applying Proposition 1, X can be generated as:

$$X \sim \sum_{k_1, \dots, k_n \in \{1,2\}} \Big((-1)^{k_1} V_1 + \frac{\pi}{2}, \dots, (-1)^{k_n} V_n + \frac{\pi}{2}\Big)_{U \in I_{k_1 \dots k_n}},$$

where $V_1, \dots, V_n \sim Unif([0, \frac{\pi}{2}])$ and $U \sim Unif([0, 1])$ are $n + 1$ independent uniform random variables and $\{I_{k_1 \dots k_n}\}_{k_1, \dots, k_n \in \{1,2\}}$ satisfy: for $v \in [0, \frac{\pi}{2}]^n$,

$$\mathbb{P}\big(U \in I_{k_1 \dots k_n} | V = v\big) = \frac{1}{n2^n}\Big(n + \sum_{i=1}^{n} (-1)^{k_i+1} \sin(v_i)\Big).$$

Here is an example to explicitly construct the intervals $I_{k_1 \dots k_n}$. One defines a bijection:

$$\mathcal{H}: \ \{1, 2\}^n \longrightarrow \{1, 2, \dots, 2^n\}.$$

Also define, $p_0 = 0$, and for $k \in \{1, \dots, 2^n\}$, the random variable

$$p_k = \frac{1}{n2^n}\Big(n + \sum_{i \in \{1, \dots, n\}, \ (k_1, \dots, k_n) = \mathcal{H}^{-1}(k)} (-1)^{k_i+1} \sin(V_i)\Big).$$

Finally one can set, for $(k_1, \dots, k_n) \in \{1, 2\}^n$,

$$I_{k_1 \dots k_n} = \Big[\sum_{k=0}^{\mathcal{H}(k_1, \dots, k_n)-1} p_k, \sum_{k=0}^{\mathcal{H}(k_1, \dots, k_n)} p_k\Big).$$

Case 1.2 f_X is periodic.

Definition 2 *A density f_X is called periodic over D if there exists $p \in \mathbb{R}^n$ such that $f_X(x) = f_X(x + p)$ for all x satisfying $x, x + p \in D$.*

The problem of simulating random variables from a periodic density is motivated in the study of directional statistics (we refer to (Mardia, 1972; Bahlmann, 2006)). It also has a wide range of applications in engineering such as physical modeling, harmonic oscillator diffusion systems and signal processing. For example, (Brenner et al., 1988) showed there exists a unique time-periodic probability density distribution arising in a sedimentation-diffusion problem, provided that the initial spatial distribution recurs after one complete period of the flipping motion; (Bishop and Legleye, 1994) used a mixture of the Von Mises distributions to model the distribution of a velocity vector in two dimensions; in modern dynamic signal and FFT-analyzers the frequency response function of a system can be measured using periodic excitations (see (Pintelon et al., 2003)). We remark that, by the multivariate Fourier analysis, any periodic probability density f_X with support D has the following n-dimensional trigonometric form Fourier series expansion (see (Hsu, 1967; Tolstov, 1962)): there exist sequences of strictly positive values $(\alpha_{k,n})_{k\geq 1}$ and real numbers $(a_{i_k})_{k\geq 1, 1\leq i_k\leq n}$, $(b_{i_k})_{k\geq 1, 1\leq i_k\leq n}$, so that for $x = (x_1, \ldots, x_n) \in D \subset \mathbb{R}^n$,

$$f_X(x) = \sum_{k=1}^{+\infty} \alpha_{k,n} \prod_{i_k=1}^{n} \sin(a_{i_k} x_{i_k} + b_{i_k}).$$

Take $\beta_n = \sum\limits_{k=1}^{+\infty} \alpha_{k,n}^2$. β_n is strictly positive, thanks to Parseval's formula. Then set $\gamma_{k,n} = \frac{\alpha_{k,n}^2}{\beta_n}$ and $c_{k,n} = \frac{\beta_n}{\alpha_{k,n}}$ for $k \in \mathbb{N}^*$, one gets

$$f_X(x) = \sum_{k=1}^{+\infty} \gamma_{k,n} \Big(c_{k,n} \prod_{i_k=1}^{n} \sin(a_{i_k} x_{i_k} + b_{i_k}) \Big).$$

Then observe that the above density can be further written under the form

$$f_X(x) = \sum_{k=1}^{+\infty} \gamma_{k,n} \tilde{f}_k(x),$$

with the coefficients $\gamma_{k,n}$ summing up to 1: $\sum\limits_{k=1}^{+\infty} \gamma_{k,n} = 1$. Then X can be simulated by applying the composition approach, provided that data from densities $\tilde{f}_k(x) = c_{k,n} \prod\limits_{i_k=1}^{n} \sin(a_{i_k} x_{i_k} + b_{i_k})$ can be generated for all $k \geq 1$. In the following example, without loss of generality, we only discuss generating random vectors from the densities of the type

$$f_X(x) = c_n \prod_{i=1}^{n} \big| \sin\big(f_i(x)\big)\big| \; {}_{x\in D},$$

where $c_n > 0$ and the Jacobian matrix of $(f_1, \ldots, f_n)$ is non-singular.

If one takes $g(\cdot) = \big(\cos(f_1(\cdot)), \ldots, \cos(f_n(\cdot))\big)$, then using Theorem 2 and the chain rule:

$$|J_g(x)| = |J_{(f_1 \ldots, f_n)}(x)| \prod_{i=1}^{n} |\sin(f_i(x))|, \text{ for } x \in D,$$

one gets

$$f_V(v) = \sum_{k=1}^{K} \frac{c_n \prod_{i=1}^{n} \left|\sin\left(f_i(r_k)\right)\right|}{|J_{(f_1\ldots,f_n)}(r_k)| \prod_{i=1}^{n} \left|\sin(f_i(r_k))\right|} \mathbb{1}_{r_k \in D}$$

$$= c_n \sum_{k=1}^{K} |J_{(f_1\ldots,f_n)}(r_k)|^{-1} \mathbb{1}_{r_k \in D}. \tag{16}$$

We remark that when f_i's are polynomials of degree 1, the Jacobian $|J_{(f_1\ldots,f_n)}(r_k)|$ reduces to some constant, thus V follows a uniform distribution over $g(D)$: $V \sim Unif(g(D))$.

Example 2 (Periodic incomplete density function)

In an n-dimensional harmonic oscillator diffusion system, one assumes the probability density has been observed as: for $x \in [0, 2\pi)^n$,

$$f_X(x) = c_n \prod_{i=1}^{n} \sin(a_i x_i + b_i) \mathbb{1}_{\prod_{i=1}^{n} \sin(a_i x_i + b_i) \geq 0},$$

where $a_i \in \mathbb{N}^*$, $b_i \in \mathbb{R}$ are observed constants and $c_n > 0$ is an unknown constant which only depends on the dimension n.

Let $D := \{x \in [0, 2\pi)^n : \prod_{i=1}^{n} \sin(a_i x_i + b_i) \geq 0\}$. Now let's take, for $i = 1, \ldots, n$, $f_i(x) = a_i x_i + b_i$ and $g(x) = \big(\cos(f_1(x)), \ldots, \cos(f_n(x))\big)$. Observe that $|J_{(f_1,\ldots,f_n)}(x)| = \prod_{i=1}^{n} a_i$ and the transformation g has all the isolated roots as:

$$r_{k_1\ldots k_n} := \Big(\frac{\arccos(v_1) - b_1 - 2k_1\pi}{a_1}, \ldots, \frac{\arccos(v_n) - b_n - 2k_n\pi}{a_n}\Big),$$

where, for $i = 1, \ldots, n$, the index k_i takes all the possible integer values such that $\frac{\arccos(v_i) - b_i - 2k_i\pi}{a_i} \in [0, 2\pi)$; the period of the function $x_i \mapsto \cos(a_i x_i + b_i)$ is $\frac{2\pi}{a_i}$, thus it has a_i distinct roots over $[0, 2\pi)$. It results that the total number of r_k's is $K = \prod_{i=1}^{n} a_i$. It follows from (16) that the density of $V = g(X)$ is:

$$f_V(v) = c_n \prod_{i=1}^{n} \frac{a_i}{a_i} \mathbb{1}_{v \in g(D)} = c_n \mathbb{1}_{v \in g(D)}.$$

As a consequence, $V \sim Unif(g(D))$. Moreover, since $a_i \geq 1$, then $x_i \mapsto \cos(a_i x_i + b_i)$ has period less than 2π. Therefore, for any $v_i \in [-1, 1]$, there always exist two x_i (one entails $\sin(a_i x_i + b_i) \geq 0$, the other one entails $\sin(a_i x_i + b_i) < 0$) such that $\cos(a_i x_i + b_i) = v_i$. It follows that $g(D) = [-1, 1]^n$. Now we are in position to simulate X. First, simulate $V = (V_1, \ldots, V_n)$ from distribution $Unif([-1, 1]^n)$. Then by using Proposition 1,

$$X \sim \sum_{\text{for } i=1,\ldots,n,\ k_i \in \left(\frac{arccos(V_i) - b_i}{2\pi} - a_i, \frac{arccos(V_i) - b_i}{2\pi}\right] \cap \mathbb{Z}} R_{k_1\ldots,k_n} \mathbb{1}_{U \in I_{k_1\ldots k_n}},$$

where

$$R_{k_1\dots k_n} := \Big(\frac{\arccos(V_1) - b_1 - 2k_1\pi}{a_1}, \dots, \frac{\arccos(V_n) - b_n - 2k_n\pi}{a_n}\Big)$$

and

$$\{I_k\}_{k=1,\dots,K} := \Big\{I_{k_1\dots k_n}\Big\}_{i=1,\dots,n,\ k_i \in \big(\frac{arccos(V_i)-b_i}{2\pi} - a_i, \frac{arccos(V_i)-b_i}{2\pi}\big] \cap \mathbb{Z}}.$$

This partition of $[0,1]$ verifies for $k = 1, \dots, K$,

$$\mathbb{P}\big(U \in I_k | V = v\big) = \Bigg(1 + \sum_{l \in \{1,\dots,K\}\setminus\{k\}} \frac{f_X(r_l)}{f_X(r_k)} \Big(\prod_{i=1}^{n} \Big|\frac{\partial_i g_i(r_k)}{\partial_i g_i(r_l)}\Big|\Big)\Bigg)^{-1} = \frac{1}{K}.$$

Therefore the chance of selecting each root r_k is equally likely. It suffices to define an arbitrary bijection

$$\mathcal{H}:\ \mathbb{Z}^n \bigcap \prod_{i=1}^{n} \Big(\frac{arccos(V_i) - b_i}{2\pi} - a_i, \frac{arccos(V_i) - b_i}{2\pi}\Big] \longrightarrow \{1, \dots, K\}$$

and take, for all possible $(k_1, \dots, k_n)$,

$$I_{k_1\dots k_n} = \Big[\frac{\mathcal{H}(k_1, \dots, k_n) - 1}{K}, \frac{\mathcal{H}(k_1, \dots, k_n)}{K}\Big).$$

Remark 1 *In Example 2, a straightforward computation based on the fact that $\int_{[0,2\pi)^n} f_X(x)\,\mathrm{d}x = 1$ shows the normalizing constant $c_n = 2^{-n}$. However, one does not necessarily need this information for simulation.*

Case 2:f_X is product of independent blocks.

Now assume that the joint density f_X verifies: for any $x \in \mathbb{R}^n$,

$$f_X(x) = \prod_{i=1}^{n} f_i(g_i(x)),$$

where for $k = 1, \dots, n$, g_k denotes a function from $\mathbb{R}^n$ to $\mathbb{R}$ and f_k from $\mathbb{R}$ to $\mathbb{R}$. Denote by $g = (g_1, \dots, g_n)$. Note that g is not necessarily invertible. We also assume that

$$0 < |det(J_g(x))| = \prod_{i=1}^{n} H_i(g_i(x)).$$

By such a choice as $V = g(X)$, one can see that all the Cartesian components of V are independent. To show this fact we rely on Theorem 2. The density of V, denoted by f_V, is given as: for any $v = (v_1, \dots, v_n) \in \mathbb{R}^n$,

$$f_V(v) = \sum_{k=1}^{K} \frac{f_X(r_k)}{|det(J_g(r_k))|} = \sum_{k=1}^{K} \prod_{i=1}^{n} \frac{f_i(g_i(r_k))}{H_i(g_i(r_k))}.$$

Observe that $g_i(r_k) = v_i$ for any $i = 1, \dots, n$ and any $k = 1, \dots, K$. Then one can write

$$f_V(v) = K \prod_{i=1}^{n} \frac{f_i(v_i)}{H_i(v_i)}.$$

Therefore the marginal distributions of $V = (V_1, \dots, V_n)$ are independent. One can simulate V by generating V_i, respectively for $i = 1, \dots, n$. The problem of simulating random vectors reduces to univariate simulations.

Example 3 (Transformation to independence)

Consider a random vector $X = (X_1, X_2)$ having the probability density below:

$$f_X(x_1, x_2) = 3(x_1 + x_2)x_2^2 \;{}_{x_1 \in [-x_2, 1-x_2], x_2 \in [-1,1]}.$$

In order to simulate X, one sets

$$g(x_1, x_2) = \big(g_1(x_1, x_2), g_2(x_1, x_2)\big) = \big(3(x_1 + x_2) \;{}_{x_1+x_2 \in [0,1]}, x_2^2 \;{}_{x_2 \in [-1,1]}\big).$$

Observe that $f_X(x_1, x_2) = g_1(x_1, x_2)g_2(x_1, x_2)$ and for $x_2 \in [-1, 1]$, $x_1 + x_2 \in [0, 1]$,

$$|det(J_g(x_1, x_2))| = 6|x_2|.$$

Let $V = g(X)$, using Theorem 2, one gets: for $v \in \mathbb{R}^2$, the density of V is

$$f_V(v) = \frac{1}{3} v_1 \sqrt{v_2} \;{}_{v_1 \in [0,3], v_2 \in [0,1]}.$$

f_V is product of two marginal densities. Hence one can easily simulate V by independently generating marginal variables V_1, V_2 respectively with densities $f_{V_1}(v_1) = \frac{2}{9} v_1 \;{}_{v_1 \in [0,3]}$ and $f_{V_2}(v_2) = \frac{3}{2}\sqrt{v_2} \;{}_{v_2 \in [0,1]}$. Using the inverse transform sampling method,

$$V \sim \big(3\sqrt{U_1}, U_2^{2/3}\big),$$

where U_1, U_2 are two independent uniform random variables following the distribution $Unif([0, 1])$. Then by Proposition 1,

$$X \sim \Big(\sqrt{U_1} - U_2^{1/3}, U_2^{1/3}\Big) \;{}_{U \in I_1} + \Big(\sqrt{U_1} + U_2^{1/3}, -U_2^{1/3}\Big) \;{}_{U \in I_2},$$

where the uniform random variable $U \sim Unif([0, 1])$ is independent of U_1, U_2 and the partition $\{I_1, I_2\}$ of $[0, 1]$ verifies, for any observation $V = v \in [0, 3] \times [0, 1]$,

$$\mathbb{P}\big(U \in I_1 | V = v\big) = \mathbb{P}\big(U \in I_2 | V = v\big) = \frac{1}{2}.$$

One option is to take: $I_1 = [0, \frac{1}{2})$ and $I_2 = [\frac{1}{2}, 1]$.

5. Applications

The generating method using multiple roots transformation has the potential to expose a large number of applications in multivariate statistics and finance. In this section, we introduce two of them.

Example 4 (Uniform random variable over irregular domain)

Many special techniques have been used to tackle the problem of generating uniform random variables over an irregular domain. We refer to (Devroye, 1986) for the special approaches to generate uniform variables over some convex polytope. In the latter approaches, the key idea is to introduce some linear transformation. This algorithm fails whenever the domain is not a convex polytope (saying a convex closed curve). We note in this paper that non-linear transformations together with conditional density method can be applied in this case. As an example, consider an irregular closed domain $D = \{(x_1, x_2) \in \mathbb{R}^2 : \ x_2^2 \leq -x_1(x_1 - 1)\}$, our goal is to generate $X = (X_1, X_2) \sim Unif(D)$. The density of X has the form

$$f_X(x_1, x_2) = c \quad {}_{(x_1,x_2)\in D},$$

where $c > 0$ is the normalizing constant.

Let $V = g(X_1, X_2) = (-X_1(X_1 - 1), X_2^2)$, therefore, by Theorem 2, the density of V can be obtained:

$$f_V(v_1, v_2) \quad = \quad \frac{c}{\sqrt{(1 - 4v_1)v_2}} \quad {}_{0\leq v_2\leq v_1\leq 1/4}.$$

Observe that the component V_1 has density:

$$f_{V_1}(v_1) = \frac{2c\sqrt{v_1}}{\sqrt{1 - 4v_1}} \quad {}_{v_1\in[0,1/4]}.$$

Equivalently, $4V_1$ follows a Beta distribution with parameters $(3/2, 1/2)$: $4V_1 \sim Beta(3/2, 1/2)$. Moreover, given V_1, the conditional distribution $V_2/V_1 \sim Beta(1/2, 1)$. Thus conditional density method can be easily applied to generate $V = (V_1, V_2)$.

The equation $v = g(x)$ has 4 roots

$$r_1 = \left(\frac{1 + \sqrt{1 - 4v_1}}{2}, \sqrt{v_2}\right), \ r_2 = \left(\frac{1 - \sqrt{1 - 4v_1}}{2}, \sqrt{v_2}\right),$$
$$r_3 = \left(\frac{1 + \sqrt{1 - 4v_1}}{2}, -\sqrt{v_2}\right), \ r_4 = \left(\frac{1 - \sqrt{1 - 4v_1}}{2}, -\sqrt{v_2}\right).$$

By Proposition 1, the probability of selecting the kth root is

$$p_k(v) = \left(1 + \sum_{l\in\{1,\ldots,4\}\setminus\{k\}} \frac{f_X(r_l)}{f_X(r_k)} \Big(\prod_{i=1}^{2} \Big|\frac{\partial_i g_i(r_k)}{\partial_i g_i(r_l)}\Big|\Big)\right)^{-1} = \frac{1}{4}.$$

Finally, X can be generated as

$$X \sim \sum_{k=1}^{4} r_k \quad {}_{U\in I_k},$$

with

$$\mathbb{P}(U \in I_k | V = v) = \frac{1}{4}.$$

It is worth noting that, in Example 4, an alternative is to apply conditional density method to generate X straightforwardly. However, multiple roots transformation method shows its advantage when simulating high dimensional random vectors, since it allows to reduce the computational complexity. Let X be uniformly distributed over a closed domain with finite volume in $\mathbb{R}^n$:

$$D = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \ h_1(x_1) + \ldots + h_n(x_n) \le 0\}.$$

If all h_k are almost everywhere differentiable and some $h_k(x_k) = v_k$ have multiple roots, then it is complicated to apply conditional density method on the simulation of X directly. However by Theorem 2, a prior transformation

$$V = g(X_1, \ldots, X_n) = (h_1(X_1), \ldots, h_n(X_n))$$

implies the density of V is necessary of the form

$$f_V(v_1, \ldots, v_n) \ = \ \sum_{g(r)=v} \frac{c \ _{r \in D}}{\prod_{i=1}^n h_i'(r^{(i)})},$$

where $c > 0$ is the normalizing constant, $r^{(i)}$ is the ith component of r. Moreover if one assumes that for $i = 1, \ldots, n$, there exists $\theta_i : \ \mathbb{R} \to \mathbb{R}$ such that $|h_i'(x)| = |\theta_i(h_i(x))|$ (this is satisfied by most of the elementary functions) for all $x \in \mathbb{R}$, then

$$f_V(v_1, \ldots, v_n) = \frac{c \ _{v_1 + \ldots + v_n \le 0, \ (v_1, \ldots, v_n) \in g(\mathbb{R}^n)}}{\prod_{i=1}^n |\theta_i(v_i)|}.$$

Since $g(\mathbb{R}^n)$ is union of disjoint rectangles in $\mathbb{R}^n$, then applying the conditional density method to generate V turns out to be much easier than to generate X. In addition, when the set of roots is finite, all the selection probabilities are equal: for $k = 1, \ldots, K$ with $K \in \mathbb{N}^*$ being the number of distinct roots of $g(x) = v$, then $p_k(v) = \frac{1}{K}$.

Example 5 (Joint first exit times of correlated Brownian motions)

We provide a new approach to simulate the joint first exit times of Brownian motions, using multiple roots transformation and NORTA approach. The simulation of the joint first exit times of Brownian motions has attracted the interests of both researchers in probability, statistics and financial mathematics. In this example we only consider a bivariate first exit time vector, since its explicit joint probability density function is currently known. Recall that (Iyengar, 1985; Metzler, 2010) have given an approach to find the density of the bivariate first exit times of correlated drifted Brownian motions (where they supposed $x > 0$, $b = 0$ in Definition 3 as below). However, neither author has given the explicit form of the joint density. We complete their work and generalize it to $x, b \in \mathbb{R}$ here.

Definition 3 *Let $\{X(t)\}_{t \ge 0}$ be a Brownian motion with drift $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. The* *first exit time τ starting from $x \in \mathbb{R}$ to the barrier $b \in \mathbb{R}$, $b \neq x$ is defined as*

$$\tau = \inf\big\{t \geq 0 : \ X(t) = b | X(0) = x\big\}.$$

Through the remaining part of this example we suppose that $\frac{\mu}{b-x} \geq 0$, so that the density of τ is not defective (if $\frac{\mu}{b-x} < 0$, this density of first exit time is defective, i.e., $\mathbb{P}(\tau < +\infty) < 1$ (see for example (Karlin and Taylor, 1981)).

Now let us consider (τ_1, τ_2) as the bivariate first exit times of the correlated two-dimensional Brownian motion $\{(X_1(t), X_2(t))\}_{t \geq 0}$. The following proposition provides an explicit joint density of (τ_1, τ_2), which will be useful to compute the selection probabilities.

Proposition 2 *Let $(X_1(t), X_2(t))$ be a two-dimensional Brownian motion respectively with drifts $\mu_1, \mu_2 \in \mathbb{R}$ and variances $\sigma_1^2, \sigma_2^2 > 0$. Denote by $\rho \in (-1, 1)$ the correlation coefficient of $X_1(t), X_2(t)$ for $t > 0$. Then starting from $(x_1, x_2) \in \mathbb{R}^2$, the joint density of the first exit times (τ_1, τ_2) to the barriers $(b_1, b_2) \in \mathbb{R}^2$ with $b_1 \neq x_1$, $b_2 \neq x_2$ is given as*

(1) *Let $(\gamma_1, \gamma_2) = (\frac{\sigma_2\mu_1 - \sigma_1\mu_2\rho}{\sigma_1\sigma_2\sqrt{1-\rho^2}}, \frac{\mu_2}{\sigma_2})$, and $(\tilde{\mu_1}, \tilde{\mu_2}) = \big((sgn(x_1 - b_1))\gamma_1, (sgn(x_2 - b_2))\gamma_2\big)$. For $0 < s < t$,*

$$f(s,t) = \sqrt{\frac{\pi}{2}} \frac{\sin\alpha}{\alpha^2 s\sqrt{(t-s)^3}} e^{-r_0(\frac{r_0}{2s} + \tilde{\mu_1}\cos\theta_0 + \tilde{\mu_2}\sin\theta_0) - \frac{\tilde{\mu_1}^2 s + \tilde{\mu_2}^2 t}{2}}$$
$$\times \sum_{n=1}^{+\infty} n \sin\big(\frac{n\pi(\alpha - \theta_0)}{\alpha}\big) \int_0^{+\infty} e^{r\tilde{\mu_1}\cos\alpha - r^2(\frac{t - s\cos^2\alpha}{2s(t-s)})} I_{n\pi/\alpha}(\frac{rr_0}{s})\, \mathrm{d}r.$$

(2) *For $0 < t < s$,*

$$f(s,t) = \sqrt{\frac{\pi}{2}} \frac{\sin\alpha}{\alpha^2 t\sqrt{(s-t)^3}} \exp\Big(-r_0(\frac{r_0}{2t} + \tilde{\mu_1}\cos\theta_0 + \tilde{\mu_2}\sin\theta_0)$$
$$-\frac{(\tilde{\mu_1}^2 + \tilde{\mu_2}^2)t}{2} - \frac{(\tilde{\mu_1}\sin\alpha - \tilde{\mu_2}\cos\alpha)^2(s-t)}{2}\Big) \sum_{n=1}^{+\infty} n \sin\big(\frac{n\pi\theta_0}{\alpha}\big)$$
$$\times \int_0^{+\infty} e^{-r\big(\tilde{\mu_1}\cos^2\alpha + \tilde{\mu_2}\sin\alpha\cos\alpha\big) - r^2\big(\frac{s - t\cos^2\alpha}{2t(s-t)}\big)} I_{n\pi/\alpha}(\frac{rr_0}{t})\, \mathrm{d}r,$$

where

$$
\begin{aligned}
\tilde{\rho} &= \Big(sgn(\frac{b_1 - x_1}{b_2 - x_2})\Big)\rho \text{ with } sgn(\cdot) \text{ being the sign function},\\
\alpha &= \begin{cases} \pi + \tan^{-1}\Big(-\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\Big) & \text{if } \tilde{\rho} > 0,\\ \frac{\pi}{2} & \text{if } \tilde{\rho} = 0,\\ \tan^{-1}\Big(-\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\Big) & \text{if } \tilde{\rho} < 0, \end{cases}\\
r_0 &= \frac{1}{\sigma_1\sigma_2}\sqrt{\frac{(b_1-x_1)^2\sigma_2^2 + (b_2-x_2)^2\sigma_1^2 - 2|(b_1-x_1)(b_2-x_2)|\tilde{\rho}\sigma_1\sigma_2}{1-\tilde{\rho}^2}},\\
\theta_0 &= \begin{cases} \pi + \tan^{-1}\Big(\frac{\sigma_1|b_2-x_2|\sqrt{1-\tilde{\rho}^2}}{|b_1-x_1|\sigma_2 - \tilde{\rho}|b_2-x_2|\sigma_1}\Big) & \text{if } |b_1-x_1|\sigma_2 < \tilde{\rho}|b_2-x_2|\sigma_1,\\ \frac{\pi}{2} & \text{if } |b_1-x_1|\sigma_2 = \tilde{\rho}|b_2-x_2|\sigma_1,\\ \tan^{-1}\Big(\frac{\sigma_1|b_2-x_2|\sqrt{1-\tilde{\rho}^2}}{|b_1-x_1|\sigma_2 - \tilde{\rho}|b_2-x_2|\sigma_1}\Big) & \text{if } |b_1-x_1|\sigma_2 > \tilde{\rho}|b_2-x_2|\sigma_1, \end{cases}
\end{aligned}
$$

and I_β *denotes the modified Bessel function of the first kind of order* $\beta > 0$.

Notice that in Proposition 2, we derive the idea from the seminal work by (Iyengar, 1985). Unfortunately, this work contains errors. (Metzler, 2010) provided the correct formula, however the formula for joint density with drift is not explicitly given. Proposition 2 extends the joint density in (Metzler, 2010) of first exit times to the case where the barriers could be any real values, due to a shift operator.

The following lemma is the key to the construction of g in the simulation approach. It provides an exact relationship between the first exit times (τ_1, τ_2) and some bivariate Chi-squared random vector. Note that this result can be easily extended to $n > 2$ dimensions.

Lemma 1 *((Shuster, 1968)) Let* τ_1, τ_2 *be two first exit times with the same parameters as in Proposition 2. Then there exists a bivariate Chi-squared random vector* (χ_1^2, χ_2^2) *such that*

$$g(\tau_1, \tau_2) = \Big(\frac{\big(\mu_1\tau_1 - (b_1 - x_1)\big)^2}{\sigma_1^2\tau_1}, \frac{\big(\mu_2\tau_2 - (b_2 - x_2)\big)^2}{\sigma_2^2\tau_2}\Big) \sim \big(\chi_1^2, \chi_2^2\big).$$

It seems impossible to generate accurately the random vector (χ_1^2, χ_2^2) since its exact distribution is difficult to compute and moreover its copula is not Gaussian (see (Metzler, 2008)). Instead, we consider an approximating simulation by conserving its marginal distribution and covariance matrix, i.e., the so-called NORTA method (see for example (Ghosh and Henderson, 2002; Ghosh and Henderson, 2003; Yuan et al., 2006)). Let us introduce some definitions.

Definition 4 *Two second order (i.e. with finite variance) random vectors* (X_1, X_2) *and* (Y_1, Y_2) *are said to be approximately identically distributed if for* $i \in \{1, 2\}$, $X_i \sim Y_i$ *and their covariance matrices are equal:*

$$Cov\big((X_1, X_2)\big) = Cov\big((Y_1, Y_2)\big).$$

We denote this relationship by

$$(X_1, X_2) \overset{approx}{\sim} (Y_1, Y_2).$$

First, by using the exact relationship between Spearman correlation and Bravais-Pearson correlation (see (Hotelling and Pabst, 1936) the following lemma holds:

Lemma 2 *Let U_1, U_2 be two uniform random variables following the distribution $Unif(0,1)$ with correlation ρ_U, then there exists a normal random vector (Z_1, Z_2) with standard normal marginal distributions and correlation ρ_Z satisfying*

$$\rho_Z = 2\sin\big(\frac{\pi}{6}\rho_U\big),$$

and

$$(U_1, U_2) \overset{approx}{\sim} \big(\varphi(Z_1), \varphi(Z_2)\big),$$

where φ is the cumulative distribution function of the standard normal random variable.

By using Lemma 2, we can generate approximately the joint distribution of (χ_1^2, χ_2^2) starting from (Z_1, Z_2).

Proposition 3 *Let (τ_1, τ_2) be the first exit times corresponding to the bivariate Brownian motions $(X_1(t), X_2(t))$ with correlation $\rho \in (-1, 1)$. Set (Z_1, Z_2) as a Gaussian vector of two standard normal random variables with:*

$$\begin{aligned}&Corr(Z_1, Z_2)\\ &= 2\sin\Big(\frac{\pi}{6}Corr\Big(2\varphi\Big(\frac{|\mu_1\tau_1 - (b_1 - x_1)|}{\sigma_1\sqrt{\tau_1}}\Big) - 1, 2\varphi\Big(\frac{|\mu_2\tau_2 - (b_2 - x_2)|}{\sigma_2\sqrt{\tau_2}}\Big) - 1\Big)\Big).\end{aligned}$$

Let

$$(\widetilde{\chi_1}^2, \widetilde{\chi_2}^2) = \Big(\big(\varphi^{-1}(\frac{\varphi(Z_1)+1}{2})\big)^2, \big(\varphi^{-1}(\frac{\varphi(Z_2)+1}{2})\big)^2\Big), \tag{17}$$

then, $(\widetilde{\chi_1}^2, \widetilde{\chi_2}^2) \overset{approx}{\sim} (\chi_1^2, \chi_2^2)$.

Finally, the simulation of the first exit times for two correlated Brownian motions with non-zero drifts are given in the following theorem.

Proposition 4 *Suppose for $i = 1, 2$, $\frac{\mu_i}{b_i - x_i} \geq 0$, hence the densities of first exit times are not defective. Let the bivariate Chi-squared vector $(\widetilde{\chi_1}^2, \widetilde{\chi_2}^2)$ verify (17).*
For $i = 1, 2$,

- *if $\mu_i \neq 0$, set*

$$(R_{i1}, R_{i2}) = \Big(\frac{b_i - x_i}{\mu_i} + \frac{\sigma_i^2\widetilde{\chi_i}^2}{2\mu_i^2} - \frac{\sigma_i|\widetilde{\chi_i}|}{2\mu_i^2}\sqrt{4\mu_i(b_i - x_i) + \sigma_i^2\widetilde{\chi_i}^2}, \frac{(b_i - x_i)^2}{\mu_i^2 R_{i1}}\Big) \tag{18}$$

- *if* $\mu_i = 0$, *set*

$$R_{i1} = R_{i2} = \frac{(b_i - x_i)^2}{\sigma_i^2 \widetilde{\chi}_i^2}. \tag{19}$$

For $u, v \in \{1, 2\}$, *define*

$$p_{uv} = \left(1 + \sum_{(i,j)\in\{1,2\}^2\setminus\{(u,v)\}} \left|\frac{R_{1i}R_{2j}}{R_{1u}R_{2v}}\right|^2 \frac{f(R_{1i}, R_{2j})}{f(R_{1u}, R_{2v})} \prod_{l=1}^{2} \left|\frac{(\mu_l X_{lu})^2 - (b_l - x_l)^2}{(\mu_l X_{li})^2 - (b_l - x_l)^2}\right|\right)^{-1},$$

where f *is the joint density of* (τ_1, τ_2). *Therefore a random partition of interval* $[0, 1]$ *can be constructed as:*

$$\begin{aligned} &I_{11} = [0, p_{11}), \;\; I_{12} = [p_{11}, p_{11} + p_{12}), \\ &I_{21} = [p_{11} + p_{12}, p_{11} + p_{12} + p_{21}), \\ &I_{22} = [p_{11} + p_{12} + p_{21}, 1]. \end{aligned} \tag{20}$$

Under this choice of $\{I_{ij}\}_{i,j\in\{1,2\}}$, (τ_1, τ_2) *can be generated by*

$$\sum_{i,j\in\{1,2\}} \big(R_{1i}, R_{2j}\big)_{\;U\in I_{ij}}, \tag{21}$$

in the sense that

$$g(\tau_1, \tau_2) \overset{approx}{\sim} g\Big(\sum_{i,j\in\{1,2\}} \big(R_{1i}, R_{2j}\big)_{\;U\in I_{ij}}\Big), \tag{22}$$

where $U \sim Unif(0, 1)$ *is a uniform random variable independent of* χ_1^2 *and* χ_2^2.

Proof. The multiple roots in (18) and (19) of (τ_1, τ_2) are deduced from (17). Then by using Theorem 1 and the construction of the partition as in (20), allows us to generate (τ_1, τ_2) starting from joint Chi-squared distribution defined by (17). Finally relation (22) follows from (21). □

It is worth noting that, although the simulation of the bivariate first exit times is approximate, (Overbeck and Schmidt, 2005) and (McLeish, 2004) have proved that the accuracy of approximation is good enough for practical purpose, since the true copula of (χ_1^2, χ_2^2) is quite similar to that of $(\widetilde{\chi_1}^2, \widetilde{\chi_2}^2)$. Compared to the Monte Carlo simulation, the Kolmogorov-Smirnov test also shows the two samples are not different in distribution at significant level $\alpha = 5\%$. The following illustration is a comparison of (τ_1, τ_2)'s joint density simulated by methods respectively using transformations with multiple roots and its true density (since the true density has singularities on $s = t$, in order to clearly compare with the estimator we set $f(s, s) = 0$ instead.). One can see from Figure 1 that they fit very well, which supports the result by this method.

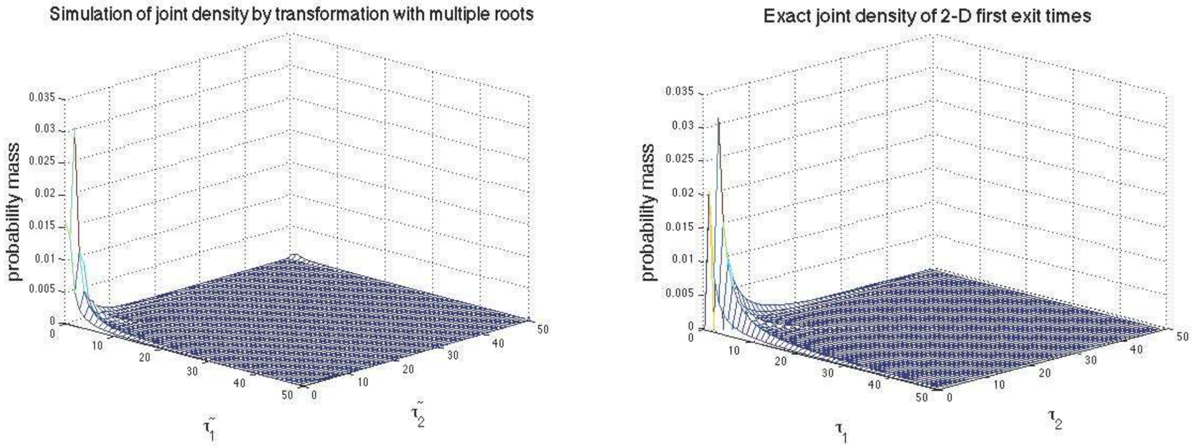

Fig. 1: Numerical estimate of the density of (τ_1, τ_2) and the exact density of (τ_1, τ_2) with parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2, x_1, x_2, b_1, b_2, \rho) = (0, 0, 1, 1, 0, 0, 0.5, 0.5, 0.7)$.

6. Conclusions

Our work extends the inverse multiple roots transformation approach from one dimensional to high dimensional and multiple roots transformations. We also discuss some strategies of choosing such a transformation. The inverse transformation approach can be widely used. It may not be as fast as some particular approaches for special classical distributions such as multivariate normal, multivariate Gamma, etc. However it is the fastest and the most accurate one among the existing general simulation approaches. The reason is, to apply inverse transformation approach, most of the work such as constructing transformation, inverting it, determining probabilities of selection, can be done off line. The algorithm's speed mainly depends on the number of multiple roots and the complexity of the selecting probabilities. It is almost independent of the dimension of the target random vector (one only needs to simulate $n + 1$ independent uniform random variables). We compare this approach to some existing classical methods, such as conditional distribution approach and NORTA approach (see for example (Devroye, 1986)). We see that, the conditional distribution approach is a completely general approach, but it has at least two inconveniences: at each step, it could be hard to find the inverse marginal cumulative distribution function; and it requires much more input information than inverse transformation approach. Indeed, the CPU running time of the former approach increases as the dimension of random vector increases. For NORTA approach, it is quite general and attractive for simulating multivariate data, however, it is at most time an approximation.

Acknowledgments

The authors are thankful to the anonymous referees and the Editor-in-Chief Professor Aliakbar Montazer Haghighi for valuable comments and suggestions towards the improvement of this paper.

REFERENCES

Bahlmann, C. (2006). Directional features in online handwriting recognition. *Pattern Recognition*, 39(1):115–125.

Bhattacharya, R. and Majumdar, M. (2004). Random dynamical systems: A review. *Economic Theory*, 23:13–38.

Bhattacharya, R. and Majumdar, M. (2007). *Random Dynamical Systems: Theory and Applications*. Cambridge University Press.

Bishop, C. and Legleye, C. (1994). Estimating conditional probability densities for periodic variables. *Advances in Neural Information Processing Systems*, 7:641–648.

Bonarini, A. and Bontempi, G. (1994). A qualitative simulation approach for fuzzy dynamical models. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 4(4):285–313.

Brenner, H., Nitsche, L., and Nitsche, J. (1988). Existence, uniqueness and regularity of a time-periodic probability density distribution arising in a sedimentation-diffusion problem. *SIAM J. Math. Anal.*, 19(1):153–166.

Devroye, L. (1986). *Non-uniform Random Variate Generation*. Springer-Verlag, New York.

Fok, P. (2013). Drift reconstruction from first passage time data using the levenberg-marquardt method. *Inverse Prob. Sci. Eng.*, 21(8).

Ghosh, S. and Henderson, S. (2002). Properties of the norta method in higher dimensions. *In: Proceedings of the 2002 Winter Simulation Conference*.

Ghosh, S. and Henderson, S. (2003). Behavior of the norta method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(43):276–294.

Hashorva, E. and Ji, L. (2014). Extremes and first passage times of correlated fractional brownian motions. *Stoch. Models*, 30(3):272–299.

Hotelling, H. and Pabst, M. (1936). Rank correlation and tests of significance involving no assumption of normality. *The Annals of Mathematical Statistics*, 7(1):29–43.

Hsu, H. (1967). *Outline of Fourier Analysis*. Associate Educational Services, New York.

Iyengar, S. (1985). Hitting lines with two-dimensional brownian motion. *SIAM Journal on Applied Mathematics*, 45:983–989.

Karlin, S. and Taylor, H. (1981). *A Second Course in Stochastic Processes*. Academic Press.

Macháček, M. (1983). Stationary probability distribution for nonlinear dynamical systems with noise. *Czechoslovak Journal of Physics B*, 33(4):365–376.

Mardia, K. (1972). *Statistics of Directional Data, in: Volume 13 of Probability and Mathematical Statistics Quantum Electronics–Principles and Applications*. University of Minnesota, Academic Press.

McLeish, D. (2004). Estimating the correlation of processes using extreme values. *Fields Institute Communications*, 44:447–467.

Metzler, A. (2008). *Multivariate First Passage Models in Credit Risk*. Ph.D. Dissertation, University of Waterloo.

Metzler, A. (2010). On the first passage problem for correlated brownian motion. *Statistics and Probability Letters*, 80:277–284.

Michael, J., Schucany, W., and Haas, R. (1976). Generating random variates using transformations with multiple roots. *The American Statistician*, 30:88–90.

Overbeck, L. and Schmidt, W. (2005). Modeling default dependence with threshold models. *Journal of Derivatives, Summer*, pages 10–19.

Packham, N., Schloegl, L., and Schmidt, W. (2013). Credit gap risk in a first passage time model with jumps. *Quant. Finance*, 13(12):1871–1889.

Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes, 3rd edition, Communications and Signal Processing*. Mcgraw-Hill Series in Electrical Engineering.

Pintelon, R., Rolain, Y., and Moer, W. V. (2003). Probability density function for frequency response function measurements using periodic signals. *IEEE Transactions on Instrumentation and Measurement*, 52(1):61–68.

Rahman, S. (2009). Probability distributions of natural frequencies of uncertain dynamic systems. *AIAA Journal*, 47(6):1579–1589.

Shuster, J. (1968). On the inverse gaussian distribution function. *Journal of the American Statistical Association*, 63:1514–1516.

Tolstov, G. (1962). *Fourier Series*. Dover Publications, New York.

Yuan, H., Hayashi, K., and Yanagihara, H. (2006). A class of population covariance matrices for monte carlo simulation. *Multivariate Behavioral Research*, 34:181–197.

Zhou, C. (2001). An analysis of default correlations and multiple defaults. *The Review of Financial Studies*, 14:555–576.