Metrika (2012) 75:939–951
DOI 10.1007/s00184-011-0360-x

A conditional distribution approach to uniform sampling on spheres and balls in L_p spaces

Vladimír Lacko · Radoslav Harman

Received: 20 December 2010 / Published online: 16 June 2011

Abstract Liang and Ng (Metrika 68:83–98, 2008) proposed a componentwise conditional distribution method for L_p-uniform sampling on L_p-norm n-spheres. On the basis of properties of a special family of L_p-norm spherical distributions we suggest a wide class of algorithms for sampling uniformly distributed points on n-spheres and n-balls in L_p spaces, generalizing the approach of Harman and Lacko (J Multivar Anal 101:2297–2304, 2010), and including the method of Liang and Ng as a special case. We also present results of a numerical study proving that the choice of the best algorithm from the class significantly depends on the value of p.

Keywords L_p-norm · n-sphere · n-ball · Uniform distribution · p-generalized normal distribution · Monte Carlo simulation · Probabilistic robustness analysis

1 Introduction

Sampling uniformly distributed points on n-dimensional Euclidean balls and spheres has been a widely studied problem since 1950's; we refer the reader to the paper by Harman and Lacko (2010) for a brief survey and a classification of existing methods. In this paper we deal with the problem of sampling from the L_p-uniform distribution on n-dimensional spheres and balls in L_p spaces, that is, instead of the Euclidean norm we consider the L_p-norm defined by $\|\mathbf{x}\| = (\sum_{i=1}^{n} |x_i|^p)^{1/p}, \mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n$,

V. Lacko (✉) · R. Harman
Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovak Republic
e-mail: lacko@pc2.iam.fmph.uniba.sk; lackovladimir@gmail.com

R. Harman
e-mail: harman@fmph.uniba.sk

where $p > 0$ is fixed. Note that the L_p-uniform distribution is uniform with respect to the L_p-metrics induced by the L_p-norm.

Since the L_p-metrics is translation invariant, for any p the L_p-volume is a constant multiple of the Lebesgue measure. Therefore, the L_p-uniform distributions on L_p-norm n-balls are the same as the standard L_2-uniform distributions on the L_p-norm n-balls. Nevertheless, on manifolds of dimensions lower than n such as n-spheres (that is, the surfaces of n-balls), the L_p-uniform distributions do not need to coincide for different values of p. For instance, the L_p-uniform distribution on the unit L_p-norm n-sphere is defined by the density of its $n-1$ dimensional subvector $(X_1, \ldots, X_{n-1})'$ with respect to the Lebesgue measure as

$$f(x_1, \ldots, x_{n-1}) = \frac{p^{n-1}\Gamma(n/p)}{2^{n-1}\Gamma^n(1/p)} \left(1 - \sum_{i=1}^{n-1} |x_i|^p\right)^{1/p-1},$$

where $-1 < x_i < 1, i = 1, \ldots, n-1, \sum_{i=1}^{n-1} |x_i|^p < 1$ (see Gupta and Song (1997)), and $\Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\mathrm{d}t$ is the Gamma function.

In addition to general applications such as stochastic optimization and Monte Carlo integration on the L_p-norm balls, and generating from the L_p-norm spherical distributions (Gupta and Song 1997), L_p-uniform sampling on L_p-norm spheres and balls has found its applications in probabilistic robustness analysis (Calafiore et al. 2000).

There are three distinct approaches to generating L_p-uniformly distributed points on L_p-norm spheres and balls.

The first and the most basic approach is the simple rejection method: we generate a sample $\mathbf{x}$ uniformly on the n-dimensional cube $[-1, 1]^n$, and if its L_p-norm $\|\mathbf{x}\|$ is less than 1, we accept $\mathbf{x}$ as a sample from the L_p-uniform distribution on the unit L_p-norm n-ball and $\mathbf{x}/\|\mathbf{x}\|$ as a sample from the L_p-uniform distribution on the unit L_p-norm n-sphere. We note that the acceptance ratio

$$\alpha(n, p) = \frac{\Gamma^n(1/p + 1)}{\Gamma(n/p + 1)}$$

decreases to zero for a fixed p and $n \to \infty$, and increases to 1 for a fixed n and $p \to \infty$. Therefore, depending on the values of n and p, the rejection method might and might not be a practical choice.

The second approach, see Calafiore et al. (1998), is based on the fact that if an n-dimensional random vector $\mathbf{Y}$ has any L_p-norm spherical distribution, then $\mathbf{Y}/\|\mathbf{Y}\|$ has the L_p-uniform distribution on n-dimensional L_p-norm sphere; cf. Theorem 1.1 in Gupta and Song (1997). More precisely, the method uses a generator from the multivariate p-generalized normal distribution with uncorrelated components (see Goodman and Kotz (1973)); therefore, it can be viewed as an L_p analogy of the method of Muller (1959) based on the (Box and Muller 1958) generator of random normal variates. Moreover, if U has the uniform distribution on the unit interval $(0, 1)$, then $U^{1/n}\mathbf{Y}/\|\mathbf{Y}\|$ has the uniform distribution on the n-dimensional L_p-norm ball, which is simple to show.

The third is an approach of Liang and Ng (2008) who, making use of results of Gupta and Song (1997), based their method on the marginal and conditional distributions of the L_p-uniform distribution on L_p-norm n-spheres. Note that this approach is an L_p generalization of the method of Fang et al. (2001).

The main aim of this paper is to propose a wide class of algorithms for sampling from the L_p-uniform distributions on L_p-norm n-spheres and n-balls, which is based on a special family of multivariate distributions described and analyzed in Sects. 2 and 3. The class of algorithms and its relations to the componentwise method of Liang and Ng (2008) are discussed in Sect. 4. We make use of the flexibility of the proposed class of algorithms and the knowledge of p to increase the speed of simulation, which is demonstrated in Sect. 5.

Notation: Let $r > 0$ be the radius, $n \in \mathbb{N}$ be the dimension, and $p > 0$ be the fixed L_p-norm parameter. The symbol $\mathscr{S}_n(r)$ denotes the L_p-norm n-dimensional sphere (n-sphere) with radius r and the center $\mathbf{0}_n$, i.e., $\mathscr{S}_n(r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = r\}$, where $\|\cdot\|$ is the L_p-norm. By $\mathscr{S}_n$ we denote the L_p-norm unit n-sphere, that is, $\mathscr{S}_n = \mathscr{S}_n(1)$. Analogously, by $\mathscr{B}_n(r)$ we denote the L_p-norm n-dimensional ball (n-ball) with radius r and the center $\mathbf{0}_n$, that is, $\mathscr{B}_n(r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \le r\}$, and $\mathscr{B}_n = \mathscr{B}_n(1)$ is the L_p-norm unit n-ball. For a set $\mathscr{X}$, by $\mathrm{U}_{\mathscr{X}}$ we denote the L_p-uniform distribution on $\mathscr{X}$. We emphasize that all spheres, balls, norms and distributions are understood to be "L_p", though p is not explicit in the notation. We also note that $\|\cdot\|$ has the properties of a norm only for $p \ge 1$, but the results stated in this paper hold for any $p > 0$.

2 The p-generalized $\mathscr{B}^{\mathrm{B}}$ distribution family

In this section we present a direct generalization of the $\mathscr{B}^{\mathrm{B}}$ distribution family defined by Harman and Lacko (2010). The family of $\mathscr{B}^{\mathrm{B}}$ distributions that we define below is based on the properties of the beta distribution $\mathrm{B}_{a,b}$ and the marginals of the uniform distribution on $\mathscr{S}_n$.

For $a > 0$ and $b > 0$ the random variable $V \sim \mathrm{B}_{a,b}$ is continuous with density

$$f(v; a, b) = \frac{1}{B(a, b)} v^{a-1}(1 - v)^{b-1}, \quad 0 < v < 1,$$

where $B(a, b) = \int_0^1 v^{a-1}(1-v)^{b-1}\mathrm{d}v = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the Beta function. We assume that the random variable $V \sim \mathrm{B}_{a,0}$ is equal to 1 almost surely, that is, $\mathrm{Prob}[V = 1] = 1$. By the symbol $\mathrm{B}_{a,b}^{1/p}$ we denote the distribution of the random variable $W = V^{1/p}$, where $V \sim \mathrm{B}_{a,b}$.

The first part of Theorem 3.1 in Gupta and Song (1997) implies that if $\mathbf{X} \sim \mathrm{U}_{\mathscr{S}_n}$ and $\{i_1, \ldots, i_m\} \subseteq \{1, \ldots, n\}$, then the m-dimensional subvector $(X_{i_1}, \ldots, X_{i_m})'$ of $\mathbf{X} = (X_1, \ldots, X_n)'$ has the distribution of the random vector $R\mathbf{S}_m$, where $R \sim \mathrm{B}_{m/p,(n-m)/p}^{1/p}$ and $\mathbf{S}_m \sim \mathrm{U}_{\mathscr{S}_m}$ are independent. This fact motivates us to define the following family of L_p-norm spherical distributions:

Definition 1 Let $\mathbf{X}$ be a random vector satisfying $\mathbf{X} = rR\mathbf{S}_n$, where $R \sim \mathrm{B}^{1/p}_{n/p,d/p}$ and $\mathbf{S}_n \sim \mathrm{U}_{\mathscr{S}_n}$ are independent, $r > 0$ and $d \geq 0$. We denote this by $\mathbf{X} \sim \mathscr{B}^{\mathrm{B}}_n(r, d)$, where r is the radius and d is the shape parameter.

There are two important special cases of the $\mathscr{B}^{\mathrm{B}}_n(r, d)$ distributions. Firstly, for $d = 0$ the distribution $\mathscr{B}^{\mathrm{B}}_n(1, 0)$ corresponds to the distribution $\mathrm{U}_{\mathscr{S}_n}$. In particular, $\mathscr{B}^{\mathrm{B}}_1(1, 0)$ corresponds to the random sign. Secondly, note that $\mathrm{B}_{n/p,1}$ is the distribution of the random variable $U^{p/n}$, where $U \sim \mathrm{U}_{(0,1)}$; consequently, the distribution $\mathscr{B}^{\mathrm{B}}_n(1, p)$ corresponds to the distribution $\mathrm{U}_{\mathscr{B}_n}$.

If $d > 0$ then by applying Lemma 2.1 in Gupta and Song (1997) we get that the probability density function of the distribution $\mathscr{B}^{\mathrm{B}}_n(r, d)$ with respect to the Lebesgue measure on $\mathbb{R}^n$ is

$$f(\mathbf{x}; r, d) = \frac{\Gamma\left(\frac{n+d}{p}\right)}{2^n r^n \Gamma^n\left(\frac{1}{p}+1\right)\Gamma\left(\frac{d}{p}\right)}\left(1 - \frac{\|\mathbf{x}\|^p}{r^p}\right)^{\frac{d}{p}-1}, \quad 0 \leq \|\mathbf{x}\| < 1.$$

Figure 1 depicts density functions of the distributions $\mathscr{B}^{\mathrm{B}}_2(1, d)$ for different values of p and the shape parameter d.

To state further theoretical results we need to formulate the family of $\mathscr{B}^{\mathrm{B}}$ distributions in terms of "beta-independent" stochastic representation. The reader is referred to Hashorva (2008) for more details about beta-independent random vectors. As a direct consequence of Theorem 1 in Gupta and Song (1997), and Definition 1 we obtain

Lemma 1 *Let $\mathbf{X} \sim \mathscr{B}^{\mathrm{B}}_n(r, d)$, and let $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are n_1-dimensional and n_2-dimensional subvectors of $\mathbf{X}$ (i.e., $n = n_1 + n_2$). Then the "beta-independent" stochastic representation of $\mathbf{X}$ is*

$$\mathbf{X} = \left(rRW\mathbf{S}_1', rR(1 - W^p)^{1/p}\mathbf{S}_2'\right)',$$

where $R \sim \mathrm{B}^{1/p}_{n/p,d/p}$, $W \sim \mathrm{B}^{1/p}_{n_1/p,n_2/p}$, $\mathbf{S}_1 \sim \mathrm{U}_{\mathscr{S}_{n_1}}$, and $\mathbf{S}_2 \sim \mathrm{U}_{\mathscr{S}_{n_2}}$ are independent.

Theorem 1 in the paper by Jambunathan (1954), and Definition 1 yield that the family of $\mathscr{B}^{\mathrm{B}}$ distributions is closed with respect to multiplication by suitable beta distributed random variables:

Lemma 2 *Let $\mathbf{X} \sim \mathscr{B}^{\mathrm{B}}_m(r, d_1)$ and $R \sim \mathrm{B}^{1/p}_{(m+d_1)/p,d_2/p}$ are independent. Then $R\mathbf{X} \sim \mathscr{B}^{\mathrm{B}}_m(r, d_1 + d_2)$.*

For $d_1 = 0$ and $d_2 = p$ we obtain the fact that if $\mathbf{S}_n \sim \mathrm{U}_{\mathscr{S}_n}$ and $U \sim \mathrm{U}_{(0,1)}$, then $U^{1/n}\mathbf{S}_n \sim \mathrm{U}_{\mathscr{B}_n}$, which we discussed previously.

Lemma 2 determines a property of the family of $\mathscr{B}^{\mathrm{B}}$ distributions, which, together with Lemma 1, enables us to make use of Theorem 2.5 by Hashorva (2008) to obtain

Theorem 1 *Let $\mathbf{X} \sim \mathscr{B}^{\mathrm{B}}_n(r, d)$, and let $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are n_1-dimensional and n_2-dimensional subvectors of $\mathbf{X}$ (i.e., $n = n_1 + n_2$). Then* **(i)**

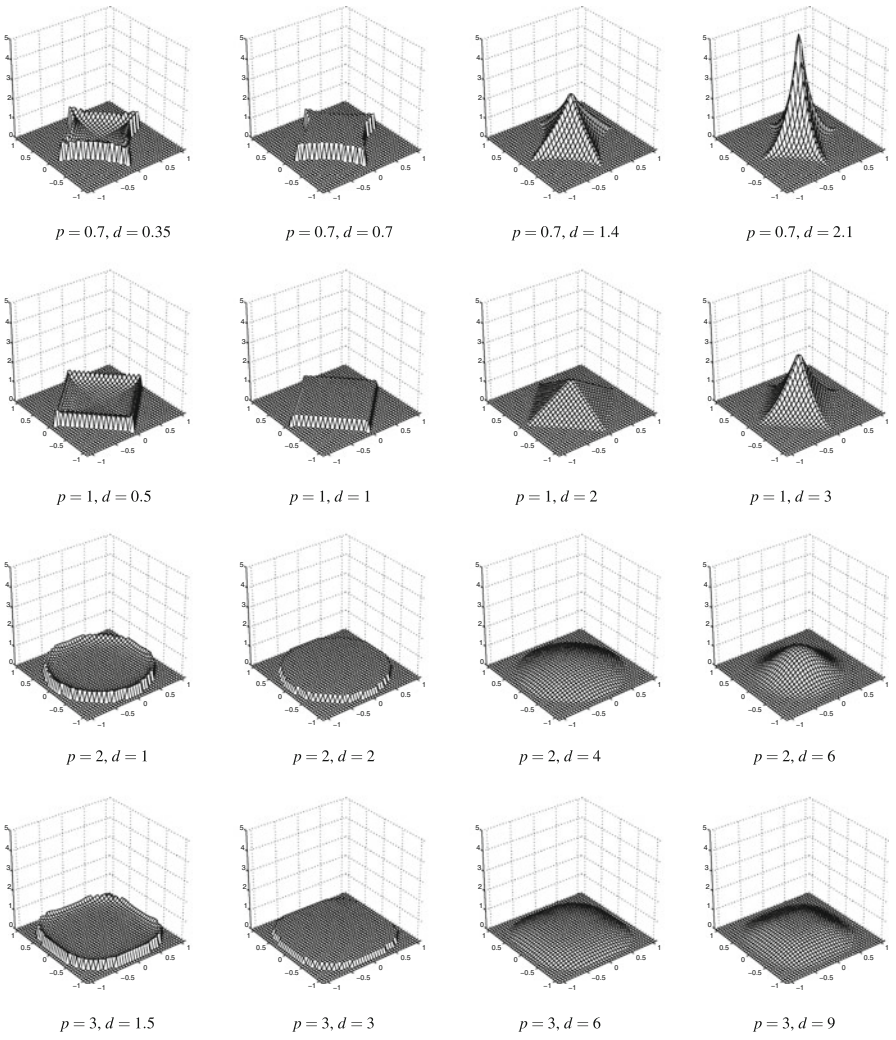

Fig. 1 The probability density function of $\mathscr{B}_2^{\mathrm{B}}(1, d)$ distributions for different values of p and the shape parameter d. Since for $d < p$ (the first column of plots) the actual density functions are unbounded, we display only trimmed versions of the densities

$\mathbf{X}_1 \sim \mathscr{B}_{n_1}^{\mathrm{B}}(r, d + n_2)$, *and* **(ii)** *if* $\mathbf{x}_1 \in \mathbb{R}^{n_1}$ *satisfies* $\|\mathbf{x}_1\| < r$, *then* $\mathbf{X}_2 \mid \mathbf{X}_1 = \mathbf{x}_1 \sim \mathscr{B}_{n_2}^{\mathrm{B}}((r^p - \|\mathbf{x}_1\|^p)^{1/p}, d)$.

Theorem 1 states that the family of $\mathscr{B}^{\mathrm{B}}$ distributions is closed with respect to forming marginal and conditional distributions, which are crucial properties for the family to be a sufficient tool for constructing the whole class of algorithms for sampling uniformly distributed points on spheres and balls in L_p spaces. Before presenting the class of algorithms and the numerical results, we mention some properties of the $\mathscr{B}^{\mathrm{B}}$ family and its relations to the p-generalized normal distribution.

From Theorem 5.4 in Gupta and Song (1997) we obtain the following proposition:

Proposition 1 *Let* $\mathbf{X} \sim \mathscr{B}_n^{\mathrm{B}}(r, d)$. *Then*

$$(\mathbf{i})\ \mathrm{E}[\mathbf{X}] = \mathbf{0}_n, \quad \textit{and} \quad (\mathbf{ii})\ \mathrm{Var}[\mathbf{X}] = r^2 \frac{B\left(\frac{3}{p}, \frac{n+d}{p}\right)}{B\left(\frac{1}{p}, \frac{n+d+2}{p}\right)} \mathbf{I}_n. \tag{1}$$

Note that the components of the distribution $\mathscr{B}_n^{\mathrm{B}}(r, d)$ are uncorrelated, but dependent.

3 The relation between the $\mathscr{B}^{\mathrm{B}}$ and the normal distributions

In this section we discuss relations between the family of $\mathscr{B}^{\mathrm{B}}$ distributions and the p-generalized normal distribution. Goodman and Kotz (1973) defined the p-generalized normal distribution as a continuous distribution with the density of the form

$$f(x) = K \exp\{-C|x - \mu|^p\}.$$

If we set

$$C = \left[\frac{\Gamma(3/p)}{\Gamma(1/p)}\right]^{p/2} \quad \text{and} \quad K = \frac{p\Gamma^{1/2}(3/p)}{2\Gamma^{3/2}(1/p)}, \tag{2}$$

then we get the unit variance. Moreover, for the zero mean we obtain that the n independent random variables with p-generalized standard normal distribution have the joint density

$$f(\mathbf{x}) = K^n \exp\{-C\|\mathbf{x}\|^p\}.$$

We will denote the p-generalized multivariate standard normal distribution by the symbol $\mathrm{N}_n(\mathbf{0}_n, \mathbf{I}_n; p)$. For $p = 2$ the distribution $\mathrm{N}_n(\mathbf{0}_n, \mathbf{I}_n; p)$ is rotationally invariant, for $p \neq 2$ it is only a (L_p-norm) spherical distribution.

Theorem 1 yields the following relation between the $\mathscr{B}^{\mathrm{B}}$ distributions and the standard normal distribution.

Corollary 1 *Let* $k \in \mathbb{N} \cup \{0\}$. *Let* $\mathbf{Y} \sim \mathrm{N}_{n+k}(\mathbf{0}_{n+k}, \mathbf{I}_{n+k}; p)$ *and* $\mathbf{X} = \mathbf{Y}_*/\|\mathbf{Y}\|$, *where* $\mathbf{Y}_*$ *consists of* n *components of* $\mathbf{Y}$. *Then* $\mathbf{X} \sim \mathscr{B}_n^{\mathrm{B}}(1, k)$.

In particular, if p is a positive integer, then we can use the previous corollary with $k = p$ to obtain a simple method for generating from the uniform distribution on $\mathscr{B}_n$, which is an alternative to the method of Calafiore et al. (1998).

In the following proposition we show that the family of $\mathscr{B}^{\mathrm{B}}$ distributions converges to the standard multivariate p-generalized normal distribution.

Proposition 2 *Let $n \in \mathbb{N}$ be fixed, and let*

$$r_d = \frac{B^{1/2}\left(\frac{1}{p}, \frac{n+d+2}{p}\right)}{B^{1/2}\left(\frac{3}{p}, \frac{n+d}{p}\right)}.$$

Then the distribution $\mathscr{B}_n^{\mathrm{B}}(r_d, d)$ converges to $\mathrm{N}_n(\mathbf{0}_n, \mathbf{I}_n; p)$ for $d \to \infty$.

Proof Clearly,

$$\lim_{d\to\infty}\left(1 - \frac{\|\mathbf{x}\|^p}{r_d^p}\right)^{d/p-1} = \lim_{d\to\infty}\left(\left(1 - \frac{C\|\mathbf{x}\|^p}{\kappa(d)}\right)^{\kappa(d)}\right)^{(d/p-1)/\kappa(d)},$$

where

$$\kappa(d) = \frac{\Gamma^{p/2}[(n+d+2)/p]}{\Gamma^{p/2}[(n+d)/p]},$$

and C is defined in (2). From formula 6.1.46 in Abramowitz and Stegun (1972) we obtain that $\lim_{d\to\infty} p\kappa(d)/d = 1$ and, evidently, $\lim_{d\to\infty} 1/\kappa(d) = 0$. Therefore, $(d/p - 1)/\kappa(d)$ converges to 1 for $d \to \infty$. Consequently, the limiting density is proportional to $\exp\{-C\|\mathbf{x}\|^p\}$. □

In the previous proposition we chose the value of r_d such that the components of $\mathbf{X} \sim \mathscr{B}_n^{\mathrm{B}}(r_d, d)$ are standardized to unit variance (cf. (1)). A similar result is stated in the next proposition where we show that the family of $\mathscr{B}^{\mathrm{B}}$ distributions converges to the standard normal distribution also with increasing n.

Proposition 3 *Let $k \in \mathbb{N}$ and $d \geq 0$ be fixed. Let $(\mathbf{X}^{(n)})_{n=k}^{\infty}$ be a sequence of random vectors such that $\mathbf{X}^{(n)} \sim \mathscr{B}_n^{\mathrm{B}}(r_n, d)$, where $r_n = (\frac{n}{Cp})^{1/p}$ and C is defined in (2). Take the sequence of k-dimensional random vectors $(\mathbf{W}^{(n)})_{n=k}^{\infty}$ such that $\mathbf{W}^{(n)} = (X_1^{(n)}, \ldots, X_k^{(n)})'$ for all $n \geq k$. Then the distribution of random vectors $(\mathbf{W}^{(n)})_{n=k}^{\infty}$ converges to the distribution $\mathrm{N}_k(\mathbf{0}_k, \mathbf{I}_k; p)$.*

Proof Clearly, the definition of $\mathscr{B}^{\mathrm{B}}$ distributions and Theorem 1.1 in Gupta and Song (1997) imply that if $R_n \sim \mathrm{B}_{n/p,d/p}^{1/p}$ and $\mathbf{Y}_n \sim \mathrm{N}_n(\mathbf{0}_n, \mathbf{I}_n; p)$, then the random vector $r_n R_n \mathbf{Y}_n / \|\mathbf{Y}_n\|$ has the distribution $\mathscr{B}_n^{\mathrm{B}}(r_n, d)$. It is obvious that $R_n \to 1$ in distribution as $n \to \infty$. It can be shown that $\|\mathbf{Y}_n\|^p$ has the gamma distribution with the shape parameter $k = n/p$ and the scale parameter $\theta = 1/C$. From the expectation and variance of gamma distributions we obtain that the expectation of $\|\mathbf{Y}_n\|/r_n$ converges to 1 and the variance of $\|\mathbf{Y}_n\|/r_n$ converges to zero for $n \to \infty$, i.e., $\|\mathbf{Y}_n\|/r_n$ converges to 1 in distribution for $n \to \infty$. That is, using Corollary 1 of Theorem 5.1 in Billingsley (1968), the distribution of $\mathbf{W}^{(n)}$ converges to $\mathrm{N}_k(\mathbf{0}_k, \mathbf{I}_k; p)$ as $n \to \infty$. □

Proposition 3 gives us a method for sampling random vectors with approximately standard normal distribution. For instance, consider that we need a large amount, say

m, of samples from the k-dimensional standard p-generalized normal distribution. If m is large enough, then we can generate $(m \times k)$-dimensional sample from the uniform distribution on $\mathscr{S}_n(r_n)$ or $\mathscr{B}_n(r_n)$, where $n = m \times k$ and r_n is defined in Proposition 3. If we split this sample into m subvectors, then these subvectors approximate samples from the k-dimensional standardized p-generalized normal distribution.

4 A conditional distribution method for sampling from the $\mathscr{B}^{\mathrm{B}}$ distributions

In this part of the paper we introduce a class of algorithms for sampling from the $\mathscr{B}^{\mathrm{B}}$ distributions. Since the uniform distributions on $\mathscr{S}_n$ $(d = 0)$ and $\mathscr{B}_n$ $(d = p)$ are special cases of the $\mathscr{B}_n^{\mathrm{B}}(r, d)$ family of distributions, the algorithms can be applied to obtain uniformly distributed points on $\mathscr{S}_n$ and $\mathscr{B}_n$.

First we note that it is possible to generate from the given $\mathscr{B}^{\mathrm{B}}$ distribution by a direct method. Recall that if $\mathbf{X} \sim \mathscr{B}_n^{\mathrm{B}}(r, d)$ then $\mathbf{X} = rV^{1/p}\mathbf{S}_n$, where $V \sim \mathrm{B}_{n/p,d/p}$ and $\mathbf{S}_n \sim \mathrm{U}_{\mathscr{S}_n}$. The random variable V can be generated using wide variety of algorithms for sampling from the beta distribution (see Fishman (1996) or McLeish (2005)), and a realization of the random vector $\mathbf{S}_n$ can be obtained by a simple rejection. The efficiency of this method rapidly decreases for increasing dimension n.

The algorithm that we will propose decomposes a high-dimensional problem into analogous problems of lower dimensions that can already be effectively solved using direct methods. More precisely, if we choose $n_1 > 0$ and $n_2 > 0$ such that $n_1 + n_2 = n$, then Theorem 1 splits the problem of generating from the distribution $\mathscr{B}_n^{\mathrm{B}}(r, d)$ into problem of generating from two $\mathscr{B}^{\mathrm{B}}$ distributions with lower dimensions:

Scheme 1
if parameters n and p are "suitable" for direct sampling (see the next section)
    return a random sample from $\mathscr{B}_n^{\mathrm{B}}(r,d)$ generated by the direct method
else
    Choose $n_1 \in \{1, \ldots, n-1\}$ and set $n_2 = n - n_1$
    Generate $\mathbf{x}_1 :\sim \mathscr{B}_{n_1}^{\mathrm{B}}(r, d + n_2)$
    Generate $\mathbf{x}_2 :\sim \mathscr{B}_{n_2}^{\mathrm{B}}((r^p - \|\mathbf{x}_1\|^p)^{1/p}, d)$
return $\mathbf{x} = (\mathbf{x}_1', \mathbf{x}_2')'$

We can apply this idea again to sampling subvectors $\mathbf{x}_1$ and $\mathbf{x}_2$. Thus, we can rewrite the method to a nonrecursive form as follows: Let $\mathbf{X} = (\mathbf{X}_1', \ldots, \mathbf{X}_k')' \sim \mathscr{B}_n^{\mathrm{B}}(r, d)$, where $\mathbf{X}_i$ is an n_i-dimensional subvector of $\mathbf{X}$, $n_i > 0$, $i = 1, \ldots, k$, and $n_1 + \cdots + n_k = n$. Then

$$\mathbf{X}_i \sim \mathscr{B}_{n_i}^{\mathrm{B}}\left(\left(r^p - \sum_{j=1}^{i-1} \|\mathbf{x}_j\|^p\right)^{1/p}, d + n - \sum_{j=1}^{i} n_j\right),$$

where $\mathbf{x}_j$ is a realization of $\mathbf{X}_j$, $j = 1, \ldots, i-1$. In this scheme all subvectors are generated directly, which results in the following algorithm (by the symbol := we denote the standard operation of assignment, and by :$\sim$ we denote the assignment of a random sample from a given distribution):

Algorithm 1

Target: *Generate a sample from the distribution* $\mathscr{B}_n^{\mathrm{B}}(r, d)$.

Input: *dimensions of subvectors* $n_1, \ldots, n_k$*;*
radius r *and shape parameter* d*;*
parameter p.

Set $\rho := r^p$
Set $\delta := d + n$
for $i = 1, \ldots, k$
Set $\delta := \delta - n_i$
Generate $\mathbf{S}_i :\sim \mathrm{U}_{\mathscr{S}_{n_i}}$
Generate $R_i :\sim \mathrm{B}_{n_i/p,\delta/p}$
Set $\mathbf{x}_i := (\rho R_i)^{1/p}\mathbf{S}_i$
Set $\rho := \rho - \rho R_i$
end
return $\mathbf{x} = (\mathbf{x}_1', \ldots, \mathbf{x}_k')'$

We note that for $d = 0$ and $n_i = 1$ for all i we obtain a componentwise method of Liang and Ng (2008) for generating from the uniform distribution on n-spheres in L_p spaces. However, the general algorithm is more flexible since it permits using higher-dimensional direct methods (e.g., $n_i = 2, 3, \ldots$), which can, in some cases, enhance computational speed. We demonstrate this fact in the next section.

Although Algorithm 1 gives a universal recipe for any programming language, in Appendix we provide a simple source code for the statistical software R (see, R Development Core Team (2008)).

Notice that if $\mathbf{X}_1 \sim \mathscr{B}_{n_1}^{\mathrm{B}}(r, d)$ and $H \sim \mathrm{B}^{1/p}_{(n_1+d)/p,n_2/p}$, then $H\mathbf{X}_1 \sim \mathscr{B}_{n_1}^{\mathrm{B}}(r, d+n_2)$ (cf. Lemma 2). Thus we obtain another scheme:

Scheme 2
if parameters n and p are "suitable" for direct sampling
return a random sample from $\mathscr{B}_n^{\mathrm{B}}(r,d)$ generated by the direct method
else
Choose $n_1 \in \{1,\ldots,n-1\}$ and set $n_2 = n - n_1$
Generate $\mathbf{x}_1 :\sim \mathscr{B}_{n_1}^{\mathrm{B}}(r,d)$ and $H :\sim \mathrm{B}^{1/p}_{(n_1+d)/p,n_2/p}$
Set $\mathbf{x}_1 := H\mathbf{x}_1$
Generate $\mathbf{x}_2 :\sim \mathscr{B}_{n_2}^{\mathrm{B}}((r^p - \|\mathbf{x}_1\|^p)^{1/p}, d)$
return $\mathbf{x} = (\mathbf{x}_1', \mathbf{x}_2')'$

A disadvantage of this scheme, if used for a componentwise sampling, is its quadratic complexity caused by the assignment $\mathbf{x}_1 := H\mathbf{x}_1$ in each iteration (the Scheme 1 is linear, not counting the complexity of the auxiliary generators). Although it is possible to rewrite Scheme 2 in a linear form, usually the large number of operations makes this approach less efficient compared to Scheme 1.

Nonetheless, under some circumstances Scheme 2 can be useful. For instance, consider the situation that we need a sample $\mathbf{x}$ from the distribution $\mathrm{U}_{\mathscr{S}_{n_1+n_2}}$, but we already have a sample $\mathbf{x}_1$ from $\mathrm{U}_{\mathscr{S}_{n_1}}$. Then it is sufficient to generate a sample $\mathbf{x}_2$ from $\mathrm{U}_{\mathscr{S}_{n_2}}$ by, e.g., Scheme 1, and merge it with a transformed $\mathbf{x}_1$ in accord with Scheme 2.

In general, this approach is computationally simpler than using Scheme 1 to generate a new sample from $\mathrm{U}_{\mathscr{S}_{n_1+n_2}}$.

5 Optimal choice of dimensions $\mathbf{n_i}$ and numerical results

Algorithm 1 gives us a general scheme, which requires us to choose the dimensions n_i of subvectors $\mathbf{x}_i$, $i = 1, \ldots, k$. Note that an important factor influencing the optimal choice of n_i's is the value of p. For example, for a large value of p an n-dimensional L_p-norm ball is close to an n-cube, which indicates that in this case the use of rejection method to generate from $\mathrm{U}_{\mathscr{S}_n}$ can be efficient. This idea implies that for a large p it is more efficient to choose larger dimensions n_i, which eventually decreases the number of loops (and, therefore, the number of samples we have to generate from the beta distribution and the uniform distribution on n_i-spheres). In this section we give results of a numerical study of the speed of the algorithm for varying dimensions n_i and values of p. Note, however, that the optimal choice of the dimensions n_i depends also on the choice of the method for generating from $\mathrm{U}_{\mathscr{S}_{n_i}}$, on the generator from the beta distribution $\mathrm{B}_{n_i/p,\delta/p}$, as well as specific programming language, hardware and other factors.

In our numerical study, we used the simple rejection method on $[-1, 1]^{n_i}$ to generate from $\mathrm{U}_{\mathscr{S}_{n_i}}$. We note that a rejection method on a suitable multiple of the L_1-norm sphere could also be used for some situations, especially if p is close to 1. Generating from the uniform distribution on the L_1-norm sphere can be done using specialized algorithms utilizing generators on the unit simplex (see, e.g., Fishman (1996)).

To generate samples from $\mathrm{B}_{n_i/p,\delta/p}$ we implemented a standard generalization of the classical method of von M D Jöhnk (1964), which can be found in, e.g., Theorem 17 in McLeish (2005): Let $U, V \sim \mathrm{U}_{(0,1)}$ be independent. Then $U^{1/a}/(U^{1/a} + V^{1/b})$ given that $U^{1/a} + V^{1/b} \leq 1$ has distribution $\mathrm{B}_{a,b}$. Naturally, there exist more efficient (but substantially more complicated) methods for generating from the beta distribution, see, for instance, Section 3.15 in Fishman (1996).

For simplicity of comparison of generators, we only generated uniform samples from the 60-dimensional unit L_p-norm sphere, so that the dimension can be partitioned into subdimensions $n_i \in \{1, \ldots, 5\}$ without remainder.

The simulation results are exhibited in Table 1. The computations were performed on the Intel Pentium Dual-Core processor with 1.86 GHz clock frequency and 2 GB of RAM. The algorithms were written in the C language and implemented in Bloodshed Dev C++ v. 4.9.9.2 environment. In the case of other choices of the generators from $\mathrm{U}_{\mathscr{S}_{n_i}}$ and $\mathrm{B}_{n_i/p,\delta/p}$ we recommend running preliminary simulations to choose the optimal values of the dimensions n_i, since they can be different from the optimal values in Table 1. Note also that for $p = 1$ (i.e., on the L_1-norm sphere), and for $p = 2$ we can use specialized and more efficient implementations (see Harman and Lacko (2010)).

6 Conclusion

In the presented paper we proposed a flexible class of a conditional-distribution algorithms for sampling from L_p-norm spheres and balls based on properties of a special

Table 1 Mean computation times (in seconds) of 100 simulation runs of Algorithm 1

p	Number of samples	**n***				
		5	**4**	**3**	**2**	**1**
1	10^3	–	–	83.0917	8.1777	**0.8316**
2	10^4	24.4561	10.7550	4.8136	2.2663	**1.2829**
3	10^4	2.4191	1.6213	1.1476	0.8829	**0.7935**
4	10^5	9.1772	7.4455	6.4232	**6.0439**	6.3345
5	10^5	5.6704	5.0835	**4.8632**	5.0557	5.7501
6	10^5	4.3030	**4.0877**	4.1324	4.5326	5.3662
7	10^5	3.6333	**3.5751**	3.7363	4.2335	5.1371
8	10^5	**3.2240**	**3.2386**	3.4617	3.9959	4.9059
9	1.5×10^5	**4.4904**	4.5892	4.9792	5.8422	7.2927

For a pair (p, n^*) the corresponding number in the table is the mean time of generating the given number of samples from the uniform distribution on the L_p-norm 60-sphere by an implementation of Algorithm 1 with $n_1 = \cdots = n_{60/n^*} = n^*$. The case $n^* = 1$ is an implementation of the componentwise method of Liang and Ng (2008). Note that the optimal choice of n^* and the corresponding computation times, indicated by boldface, depend on the value of p

family of random vectors, which extends the recent results of Harman and Lacko (2010) and Liang and Ng (2008). An advantage of the class of algorithms is that it permits combining efficient generators on low-dimensional L_p-norm spheres and balls. In our numerical experiments we employed simple rejection for generating on low-dimensional L_p-norm spheres and balls. Despite the fact that this is the simplest possible method, its implementation in our algorithm can lead to a significant decrease of the computation speed compared to existing methods, especially for larger values of p.

Acknowledgments The research was supported by the Slovak VEGA Grant No. 1/0077/09 and VEGA Grant No. 1/0381/09. The authors thank the anonymous referees for their valuable comments on the first version of the paper.

Appendix

Below we provide a transcription of Algorithm 1 for the statistical software R. The code is written with emphasis on clarity rather than efficiency. Thus, the reader can create an efficient implementation of the procedure in a low-level programming language such as C++, which we recommend for large-scale simulations.

```
rSphere <- function(n, p){
     repeat{
          S <- runif(n, min=-1, max=1); npS <- sum(abs(S)^p)
          if(npS<1) break
     }
     S <- S/(npS)^(1/p); S
}
```

```
rBeta <- function(a, b){
     if(b == 0) V <- 1
     else V <- rbeta(1, shape1=a, shape2=b)
     V
}
rBB <- function(N, Ni, r, d, p){
     k <- length(Ni); n <- sum(Ni)
     Samples <- matrix(nrow=N, ncol=n)
     for(j in 1:N){
          rho <- r^p; delta <- d + n; cntr <- 0;
          for(i in 1:k){
               delta <- delta - Ni[i]
               Si <- rSphere(Ni[i],p)
               Ri <- rBeta(Ni[i]/p, delta/p)
               Samples[j,(cntr+1):(cntr+Ni[i])] <- (rho*Ri)^(1/p)*Si
               rho <- rho - rho*Ri
               cntr <- cntr + Ni[i]
          }
     }
     Samples
}
```

The function `rBB` returns `N` random samples from the distribution $\mathscr{B}_n^{\mathrm{B}}(r, d)$. The parameters `r`, `d` and `p` correspond to the radius, the shape parameter and the L_p-norm parameter, respectively. The parameter `Ni` is a vector of dimensions of the subvectors $\mathbf{x}_i$ in Algorithm 1, so the sum of the components of `Ni` is equal to the desired dimension n.

The auxiliary function `rSphere` returns a random sample from the L_p-uniform distribution on the unit L_p-norm sphere of the dimension `n`. It implements the simplest naive rejection method. Note that for some values `n` and `p` the function `rSphere` can be substituted by a more efficient user-defined procedure. The function `rBeta` returns a random sample from the beta distribution with parameters `a` > 0 and `b` ≥ 0.

References

Abramowitz M, Stegun IA (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables, 10th edn. No. 55 in Applied Mathematics Series, National Bureau of Standards

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

Box GEP, Muller ME (1958) A note on the generation of random normal deviates. Ann Math Stat 29:610–611

Calafiore G, Dabbene F, Tempo R (1998) Uniform sample generation in ℓ_p balls for probabilistic robustness analysis. In: Proceedings of 37th IEEE

Calafiore GC, Dabbene F, Tempo R (2000) Randomized algorithms for probabilistic robustness with real and complex structured uncertainty. IEEE T Automat Contr 45:2218–2235

Fang KT, Yang Z, Kotz S (2001) Generation of multivariate distributions by vertical density representation. Statistics 35:281–293

Fishman GS (1996) Monte Carlo: concepts, algorithms, and applications. Springer, New York

Goodman IR, Kotz S (1973) Multivariate θ-generalized normal distributions. J Multivar Anal 3:204–219

Gupta AK, Song D (1997) l_p-norm spherical distribution. J Stat Plan Infer 60:241–260
Harman R, Lacko V (2010) On decompositional algorithms for uniform sampling from n-spheres and n-balls. J Multivar Anal 101:2297–2304
Hashorva E (2008) Conditional limiting distribution of beta-independent random vectors. J Multivar Anal 99:1438–1459
Jambunathan MV (1954) Some properties of beta and gamma distributions. Ann Math Stat 25:401–405
Liang J, Ng KW (2008) A method for generating uniformly scattered points on the l_p-norm unit sphere and its applications. Metrika 68:83–98
McLeish DL (2005) Monte Carlo simulation and finance. Wiley, Hoboken
Muller ME (1959) A note on a method for generating points uniformly on n-dimensional spheres. Commun ACM 2:19–20
R Development Core Team (2008) R: A language and environment for statistical computing. R Found Stat Comput, Vienna
von Jöhnk KMD (1964) Erzeugung von betaverteilten und gammaverteilten Zufallszahlen. Metrika 8:5–15