IMPROVING POINT SELECTION IN CUBATURE BY A NEW DISCREPANCY*

JIANBING CHEN† AND SHENGHAN ZHANG†

Abstract. Reasonable point set selection is of paramount importance to the accuracy of high-dimensional integrals that will be encountered in various disciplines. In the present paper, to improve the point selection and to overcome the computational complexity of evaluating classical discrepancies, the concept of extended F-discrepancy (EF-discrepancy) and generalized F-discrepancy (GF-discrepancy) of a point set is introduced and justified by comparative studies with other existing discrepancies. Meanwhile, the extensions of Koksma-Hlawka inequality for EF-discrepancy are proved and a conjecture for GF-discrepancy is put forward and discussed. This GF-discrepancy is then employed as the objective function when selecting the optimal rotation angles in the rotation transform of the quasi-symmetric point method (Q-SPM). Meanwhile, it is also proved that the rotation transform will keep the degree of algebraic accuracy. A genetic algorithm is adopted to solve the optimization problem. Several numerical examples are elaborated, demonstrating that the GF-discrepancy is a reasonable index in judging the goodness of a point set, and that the optimal rotation of Q-SPM will greatly improve the accuracy of stochastic analysis of nonlinear structures. The proposed GF-discrepancy and the resulting rotational Q-SPM point sets could be applied directly to other problems of uncertainty quantification. Problems to be further studied are discussed.

Key words. cubature, generalized F-discrepancy, quasi-symmetric point method, rotation transform; genetic algorithm, probability density evolution method

AMS subject classifications.

1. Introduction.

High-dimensional integrals are encountered in various science and engineering disciples, particularly when stochastic phenomena are addressed, e.g., finance [13], uncertainty quantification of various physical and engineering systems [26] and stochastic dynamics [3, 14]. A variety of methods, usually called cubature formulae [8], have been developed in the past decades. The cubature formulae generally adopt a weighted summation of a series of function values at a specified point set as an approximation of the target high-dimensional integral. Basically, they could be classified into two types: (1) the type in which the point set is determined by prescribing the degree of algebraic accuracy, e.g., the Gaussian quadrature, the sparse grid method [22] and the Quasi-Symmetric Point Method (Q-SPM) [23], and (2) the type in which the point set is generated by some sampling techniques to achieve uniformity to a degree. The latter includes methods adopting random points, e.g., Monte Carlo method [20] and their improvements, and methods adopting deterministic points, e.g., the Number Theoretical Method (NTM, also known as Quasi-Monte Carlo method in literature) [12, 13, 17].

In many cases, the NTM point sets take advantages over Monte Carlo simulations because of their deterministic and lower discrepancy, which in turn guarantees their deterministic and lower error bound in the sense of the Koksma-Hlawka inequality [7, 17]. In this context, the discrepancy plays an important role in point selection. However, the efforts of computing discrepancies, except some analytical elegant results for special cases, are usually prohibitively large due to the so-called curse of dimen-

*This work was supported by the National Natural Science Foundation of China (Grant No.11172210), the Fundamental Fund for Central Universities, the Shuguang Plan of Shanghai City and the State Key Laboratory of Disaster Reduction in Civil Engineering (Grant Nos. SLDRCE08-A-01 and SLDRCE10-B-02).

†State Key Laboratory of Disaster Reduction in Civil Engineering & School of Civil Engineering, Tongji University 1239 Siping Road, Shanghai 200092, P.R.China (chenjb@tongji.edu.cn).

sionality. For instance, it was shown that the computation of star discrepancy is an NP-hard problem [7]. In addition, mostly equal weights are used in the construction of formulae based on NTM point sets. The possibility of improving the accuracy by adopting non-equal weights is seldom explored.

In the present paper, the concept of generalized F-discrepancy (GF-discrepancy), where the assigned probabilities proposed by Chen et al [2] are involved to replace the equal weights to capture the effect of non-uniformity in a way, is introduced and justified by some comparative studies with existing discrepancies. The GF-discrepancy is then adopted as the objective function in optimally determining the rotation angles in the rotation transform of Q-SPM. Meanwhile, it is proved that the rotation transform will not reduce the algebraic accuracy of the point set. A genetic algorithm is employed to solve the optimization problem. It is shown that by performing an appropriate rotation, the marginal probability density could be better reproduced and numerical examples show that the accuracy of stochastic response analysis can be improved considerably. Rotational quasi-symmetric point method (RQ-SPM) maintains high accuracy even in the numerical cases where Q-SPM does not behave very well. The results presented in this paper show that there is indeed room for improving the accuracy of numerical integral or uncertainty quantification by minimizing GF-discrepancy through rearranging existing cubature point sets or other kind of good point sets.

2. Generalized F-discrepancy. In this section, the concepts of F-discrepancy and extended F-discrepancy are reviewed first. Then, the generalized F-discrepancy is introduced, and the error bounds associated with the extended F-discrepancy and generalized F-discrepancy are studied. Different discrepancies of points sets selected by various methods are exemplified to justify the concept of generalized F-discrepancy.

2.1. Generalized F-discrepancy. The discrepancy of a point set is defined by

$$D(\mathcal{P}_n) = \sup_{x \in C^s} \left[\frac{N(\mathcal{P}_n, [\mathbf{0}, \mathbf{x}))}{n} - V([\mathbf{0}, \mathbf{x})) \right], \tag{2.1}$$

where $C^s = [0,1]^s$ is the s-dimensional unit hyper-cube, $\mathcal{P}_n = \{\mathbf{x}_q = (x_{q,1}, x_{q,2}, \ldots, x_{q,n}), q = 1, 2, \ldots, n\}$ is the point set with cardinal number n, $N(\mathcal{P}_n, [\mathbf{0}, \mathbf{x}))$ is the number of the points scattered in the hyper-rectangle $[\mathbf{0}, \mathbf{x})$, and $V([\mathbf{0}, \mathbf{x}]) = \prod_{j=1}^{s} x_j$ is the volume of the hyper-rectangle $[\mathbf{0}, \mathbf{x})$. Such defined discrepancy is called Weyl's discrepancy (also called star discrepancy, e.g., in [17]) and is usually adopted to measure the uniformity of a point set [12]. What is very important is that this discrepancy gives an error bound for any function f with bounded variation by the celebrated Koksma-Hlawka inequality [12]

$$\left| \int_{C^s} f(\mathbf{x}) d\mathbf{x} - \frac{1}{n} \sum_{q=1}^{n} f(\mathbf{x}_q) \right| \leq D(\mathcal{P}_n) \cdot \mathrm{TV}(f), \tag{2.2}$$

where

$$\mathrm{TV}(f) = \sum_{\alpha_1 + \cdots + \alpha_s = 1}^{s} \int_0^1 \cdots \int_0^1 \left| \frac{\partial^{\alpha_1 + \cdots + \alpha_s} f}{\partial x_1^{\alpha_1} \cdots \partial x_s^{\alpha_s}} \right| dx_1 \cdots dx_s \tag{2.3}$$

is the total variation of the function f in the sense of Hardy and Krause, which essentially characterizes the irregularity of the hyper-surface $f(\mathbf{x})$ [17]. Generally,

if $f(\mathbf{x})$ is smoother, $\mathrm{TV}(f)$ is smaller. In the discretized case, for a 2-dimensional function $\mathrm{TV}(f)$ reads [12]

$$\mathrm{TV}(f) = \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} |\triangle_{11} f(x_i, y_i)| + \sum_{i=0}^{l-1} |\triangle_{10} f(x_i, 1)| + \sum_{j=0}^{l-1} |\triangle_{01} f(1, y_j)|\,, \tag{2.4}$$

where $\triangle_{10} f(x_i, y) = f(x_{i+1}, y) - f(x_i, y)$, $\triangle_{01} f(x, y_j) = f(x, y_{j+1}) - f(x, y_j)$, $\triangle_{11} f(x_i, y_j) = f(x_i, y_j) - f(x_{i+1}, y_j) - f(x_i, y_{j+1}) + f(x_{i+1}, y_{j+1})$, and $(l+1)$ is the number of discrete nodes in each direction of the discrete grid used for evaluating the variation.

Clearly, 2.1 could only be applied to the cases when the random variables follow uniform distributions. For non-uniform distributions, it could be extended to the F-discrepancy defined by [9]

$$D_F(\mathcal{P}_n) = \sup_{\mathbf{x}\in\mathbb{R}^s} |F_n(\mathbf{x}) - F(\mathbf{x})|\,, \tag{2.5}$$

where $F(\mathbf{x})$ is the cumulative distribution function (CDF) of the random vector $\mathbf{X}$ and $F_n(\mathbf{x})$ is the empirical CDF given by

$$F_n(\mathbf{x}) = \frac{1}{n}\sum_{q=1}^{n} I\{\mathbf{x_q} \le \mathbf{x}\}, \tag{2.6}$$

where $I\{\cdot\}$ is the indicator function whose value is one if the event is true and otherwise zero. Clearly, for uniform distributions 2.5 and 2.6 reduce to 2.1.

The definition in 2.6 indicates that all the points have the same weight $\frac{1}{n}$. Because the points are usually not adequately uniformly scattered, to characterize the non-uniformity to a degree by introducing the concept of assigned probability is more reasonable. To this end, the space $\Omega_{\mathbf{X}}$ is partitioned by adopting the point sets $\mathbf{x}_q, q = 1, 2, \cdots, n$ as the nucleus of the Voronoi cells [6]. Denote the Voronoi cell of $\mathbf{x}_q$ by Ω_q. We have $\cup_{q=1}^{n}\Omega_q = \Omega_{\mathbf{X}}$ and $\Omega_q \cap \Omega_k = \emptyset$, for $q \neq k$. Then for each point $\mathbf{x}_q$ the probability

$$P_q = \int_{\Omega_q} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad q = 1, 2, \cdots, n \tag{2.7}$$

is assigned.

The F-discrepancy in 2.5 could be further extended by replacing the empirical CDF in 2.6 by [14]

$$F_n(\mathbf{x}) = \sum_{q=1}^{n} P_q \cdot I\{\mathbf{x_q} \le \mathbf{x}\}, \tag{2.8}$$

where P_q's are the assigned probabilities defined in 2.7.

It is expected that the smaller the extended F-discrepancy (which could be denoted by $D_{EF}(\mathcal{P}_n)$ and shortly called EF-discrepancy), the better performance the point set will exhibit (in the sense of charactering the joint probability distribution). However, the efforts of computing the extended F-discrepancy defined by 2.5 and 2.8 are the same as star discrepancy, which is an NP-hard problem [7]. To avoid this trouble, a generalized F-discrepancy (GF-discrepancy) could be proposed as the maxima of all the marginal F-discrepancies

$$D_{GF}(\mathcal{P}_n) = \max_{1\le i\le s} \{D_{F,i}(\mathcal{P}_n)\}, \tag{2.9}$$

where

$$D_{F,i}(\mathcal{P}_n) = \sup_{x\in\mathbb{R}} |F_{n,i}(x) - F_i(x)|, \quad i = 1, 2, \cdots, s \tag{2.10}$$

is the marginal F-discrepancies in the direction X_i. Here, $F_i(x)$ is the marginal CDF of X_i and $F_{n,i}(x)$ is the empirical marginal CDF of X_i defined by

$$F_{n,i}(x) = \sum_{q=1}^{n} P_q \cdot I\{x_{q,i} \le x\}, \tag{2.11}$$

where $x_{q,i}$ is the i-th component of $\mathbf{x}_q$ and P_q is the corresponding assigned probability. We stress here that even in the marginal discrepancies the cross-variate information is involved in some way because the assigned probabilities are related to cross-variate spatial position.

The computation of $D_{GF}(\mathcal{P}_n) = \max_{1\le i\le s}\{D_{F,i}(\mathcal{P}_n)\}$ involves s one-dimensional empirical CDF evaluations of cost M for a total cost of order sM, whereas the computation of $D_{EF}(\mathcal{P}_n)$ defined by 2.5 and 2.8 involves an s-dimensional empirical CDF evaluation whose computational effort is of order M^s. Thus, the employment of $D_{EF}(\mathcal{P}_n)$ is usually unfeasible for multi-random variables but the computation of $D_{GF}(\mathcal{P}_n)$ is very efficient. The following theoretical discussions and numerical exemplifications will justify the concept of EF- and GF-discrepancies.

2.2. Extension of Koksma-Hlawka inequality. Because the concept of discrepancies is extended to take into account the effect of assigned probabilities, it is expected that the Koksma-Hlawka inequality 2.2 could also be extended. This is true due to the following two theorems.

THEOREM 2.1. (*extended Koksma inequality for one-dimensional problems*) *If f has bounded variation* $\mathrm{TV}(f)$ *on* $[0,1]$*, then for any $x_1, x_2, \cdots, x_n \in [0,1]$ whose assigned probabilities are $P_1, P_2, \cdots, P_n$, respectively, where $P_q > 0$ and $\sum_{q=1}^{n} P_q = 1$, we have*

$$\left| \int_0^1 f(x)dx - \sum_{q=1}^{n} P_q f(x_q) \right| \le D_{EF}(\mathcal{P}_n)\mathrm{TV}(f), \tag{2.12}$$

where $D_{EF}(\mathcal{P}_n)$ is defined by 2.5 and 2.8 with $F(x) = x$ because uniform distribution is considered here. Note that in this case $D_{GF}(\mathcal{P}_n) = D_{EF}(\mathcal{P}_n)$.

Proof. The theorem is an extension of the Koksma inequality and the proof is analogous to the proof of the Koksma inequality provided by [17].

Without loss of generality, we let $x_1 < x_2 < \cdots < x_n$. Put $x_0 = 0$, $P_0 = 0$ and $x_{n+1} = 1$. Using summation by parts and denoting $f_q = f(x_q)$, we obtain

$$\begin{aligned} &\sum_{q=1}^{n} P_q f(x_q) \\ &= -\left[P_0(f_1 - f_0) + (P_0 + P_1)(f_2 - f_1) + \cdots + \left(\sum_{q=0}^{n} P_q \right) (f_{n+1} + f_n) \right] \\ &= f(1) - \sum_{j=0}^{n} \left[\left(\sum_{q=0}^{j} P_q \right) (f_{j+1} - f_j) \right]. \end{aligned} \tag{2.13}$$

Then using integration by parts, we have $\int_0^1 f(x)dx = f(1) - \int_0^1 xdf(x)$, which is followed by

$$\begin{aligned}\int_0^1 f(x)dx - \sum_{q=1}^{n} P_q f(x_q) &= \sum_{j=0}^{n}\left\{\left(\sum_{q=1}^{j} P_q\right)(f_{j+1}-f_j)\right\} - \int_0^1 xdf(x)\\ &= \sum_{j=0}^{n}\int_{x_j}^{x_{j+1}}\left(\sum_{q=1}^{j} P_q - x\right)df(x)\\ &= \sum_{j=0}^{n}\int_{x_j}^{x_{j+1}}\left(\sum_{q=1}^{n} P_q I\{x_q \le x\} - x\right)df(x).\end{aligned} \tag{2.14}$$

Considering 2.5 and 2.8 and noting $F(x) = x$, we have

$$\left|\sum_{q=1}^{n} P_q I\{x_q \le x\} - x\right| \le D_{EF}\left(\mathcal{P}_n\right). \tag{2.15}$$

Substituting 2.15 in 2.14 yields

$$\begin{aligned}\left|\int_0^1 f(x)dx - \sum_{q=1}^{n} P_q f(x_q)\right| &= \left|\sum_{j=0}^{n}\int_{x_j}^{x_{j+1}}\left(\sum_{q=1}^{n} P_q I\{x_q \le x\} - x\right)df(x)\right|\\ &\le D_{EF}(\mathcal{P}_n)\sum_{j=0}^{n}\int_{x_j}^{x_{j+1}}|df(x)|\\ &\le D_{EF}(\mathcal{P}_n)\mathrm{TV}(f).\end{aligned} \tag{2.16}$$

This completes the proof of Theorem 2.1. □

It should be mentioned that this theorem has once been implied in [19] proved in a different way.

THEOREM 2.2. (*extended Koksma-Hlawka inequality for multi-dimensional problems*) *If a multi-variable function f has bounded variation $\mathrm{TV}(f)$ on $[0,1]^s$, then for any $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n \in [0,1]^s$ whose assigned probabilities are $P_1, P_2, \cdots, P_n$, respectively, where $P_q > 0$ and $\sum_{q=1}^{n} P_q = 1$, we have*

$$\left|\int_0^1 \cdots \int_0^1 f(\mathbf{x})d\mathbf{x} - \sum_{q=1}^{n} P_q f(\mathbf{x}_q)\right| \le D_{EF}(\mathcal{P}_n)\mathrm{TV}(f), \tag{2.17}$$

where $D_{EF}(\mathcal{P}_n)$ is defined by 2.5 and 2.8 and $F(\mathbf{x}) = \prod_{i=1}^{s} x_i$ for the uniform distribution.

Proof. Theorem 2.2 is a generalization of the classical Koksma-Hlawka inequality (2.2) from equal weights to non-equal weights. By analogy to the proof given in [12], Theorem 2.2 can be proved. We only consider the proof of the 2-dimensional case here. Detailed proof for higher-dimensional ($s \ge 3$) situation is provided in the Appendix to avoid lengthiness in this section.

According to Theorem 5.1 in [12], any bounded function could be represented as a difference of two generalized monotonic functions. Thus, we have

$$f(x,y) = f_1(x,y) - f_2(x,y), \tag{2.18}$$

where $f_1(x,y)$ and $f_2(x,y)$ are two generalized monotonic functions. A generalized monotonic function $f_1(x,y)$ is such a function that the following three conditions are met: (1) $f_1(x',y)-f_1(x,y)$ has the same sign or equals zero; (2) $f_1(x,y')-f_1(x,y)$ has the same sign or equals zero; and (3) $f_1(x,y)-f_1(x',y)-f_1(x,y')+f_1(x',y')$ has the same sign or equals zero for any given x,y,x',y' satisfying $0\le x\le x'\le 1$, $0\le y\le y'\le 1$.

In the area $\frac{i-1}{q}\le x\le\frac{i}{q}$, $\frac{j-1}{q}\le y\le\frac{j}{q}$, $i\le 1$, $j\le q$, $f_1(x,y)$ is no larger than $f_1\left(\frac{i}{q},\frac{j}{q}\right)$ considering the fact that f_1 is a generalized monotonic (non-decreasing) function. Thus we have

$$
\begin{aligned}
S_1 &= \sum_{k=1}^{n} P_k f_1\left(x_1(k),x_2(k)\right)\\
&\le \sum_{i=1}^{q}\sum_{j=1}^{q}\left(\sum_{\frac{i-1}{q}\le x_1(k)<\frac{i}{q},\frac{j-1}{q}\le x_2(k)<\frac{j}{q}} P_k\right) f_1\left(\tfrac{i}{q},\tfrac{j}{q}\right)\\
&= \sum_{i=1}^{q-1}\sum_{j=1}^{q-1}\left(\sum_{\mathbf{x}_k<(\frac{i}{q},\frac{j}{q})} P_k\right)\\
&\quad\times\Big[f_1\left(\tfrac{i}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i+1}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i}{q},\tfrac{j+1}{q}\right)+f_1\left(\tfrac{i+1}{q},\tfrac{j+1}{q}\right)\Big] \qquad (2.19)\\
&\quad+\sum_{i=1}^{q-1}\left(\sum_{\mathbf{x}_k<(\frac{i}{q},1)} P_k\right)\Big[f_1\left(\tfrac{i}{q},1\right)-f_1\left(\tfrac{i+1}{q},1\right)\Big]\\
&\quad+\sum_{j=1}^{q-1}\left(\sum_{\mathbf{x}_k<(1,\frac{j}{q})} P_k\right)\Big[f_1\left(1,\tfrac{j}{q}\right)-f_1\left(1,\tfrac{j+1}{q}\right)\Big]\\
&\quad+f_1(1,1).
\end{aligned}
$$

For convenience, denote $F_n(x,y)=\sum_{\mathbf{x}_k<(x,y)}P_k$ which is the empirical CDF defined in 2.8. 2.19 could be re-written as

$$
\begin{aligned}
S_1 &\le \sum_{i=1}^{q-1}\sum_{j=1}^{q-1} F_n\left(\tfrac{i}{q},\tfrac{j}{q}\right)\Big[f_1\left(\tfrac{i}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i+1}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i}{q},\tfrac{j+1}{q}\right)+f_1\left(\tfrac{i+1}{q},\tfrac{j+1}{q}\right)\Big]\\
&\quad+\sum_{i=1}^{q-1} F_n\left(\tfrac{i}{q},1\right)\Big[f_1\left(\tfrac{i}{q},1\right)-f_1\left(\tfrac{i+1}{q},1\right)\Big]\\
&\quad+\sum_{j=1}^{q-1} F_n\left(1,\tfrac{j}{q}\right)\Big[f_1\left(1,\tfrac{j}{q}\right)-f_1\left(1,\tfrac{j+1}{q}\right)\Big] \qquad (2.20)\\
&\quad+f_1(1,1).
\end{aligned}
$$

Recalling the definition of $D_{EF}(\mathcal{P}_n)$ in 2.5 and 2.8, the following holds

$$
F_n(x,y)=\sum_{q=1}^{n} P_q I\{\mathbf{x}_q<\mathbf{x}\}=xy+\vartheta(x,y)D_{EF}(\mathcal{P}_n), \qquad (2.21)
$$

where $|\vartheta(x,y)| \leq 1$. Substituting 2.21 in 2.19 yields

$$
\begin{aligned}
S_1 \leq & \sum_{i=1}^{q-1}\sum_{j=1}^{q-1}\left(\tfrac{ij}{q^2}+\vartheta\left(\tfrac{i}{q},\tfrac{j}{q}\right)D_{EF}(\mathcal{P}_n)\right) \\
& \times\left[f_1\left(\tfrac{i}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i+1}{q},\tfrac{j}{q}\right)-f_1\left(\tfrac{i}{q},\tfrac{j+1}{q}\right)+f_1\left(\tfrac{i+1}{q},\tfrac{j+1}{q}\right)\right] \\
& +\sum_{i=1}^{q-1}\left(\tfrac{i}{q}+\vartheta\left(\tfrac{i}{q},1\right)D_{EF}(\mathcal{P}_n)\right)\left[f_1\left(\tfrac{i}{q},1\right)-f_1\left(\tfrac{i+1}{q},1\right)\right] \\
& +\sum_{j=1}^{q-1}\left(\tfrac{j}{q}+\vartheta\left(1,\tfrac{j}{q}\right)D_{EF}(\mathcal{P}_n)\right)\left[f_1\left(1,\tfrac{j}{q}\right)-f_1\left(1,\tfrac{j+1}{q}\right)\right] \\
& +f_1(1,1) \\
\leq & \frac{1}{q^2}\sum_{i=1}^{q}\sum_{j=1}^{q}f_1\left(\tfrac{i}{q},\tfrac{j}{q}\right)+\mathrm{TV}(f_1)D_{EF}(\mathcal{P}_n),
\end{aligned}
\tag{2.22}
$$

where $\mathrm{TV}(f_1)$ is the total variation in the sense of Hardy and Krause defined in 2.4.

Let $q \to \infty$. It follows from 2.22

$$
S_1 \leq \int_0^1\int_0^1 f_1(x_1,x_2)dx_1dx_2 + D_{EF}(\mathcal{P}_n)\mathrm{TV}(f_1). \tag{2.23}
$$

Likewise, we also have

$$
S_1 \geq \int_0^1\int_0^1 f_1(x_1,x_2)dx_1dx_2 - D_{EF}(\mathcal{P}_n)\mathrm{TV}(f_1). \tag{2.24}
$$

Therefore,

$$
S_1 = \int_0^1\int_0^1 f_1(x_1,x_2)dx_1dx_2 + \vartheta_1 D_{EF}(\mathcal{P}_n)\mathrm{TV}(f_1), \tag{2.25}
$$

where $|\vartheta_1| \leq 1$.

Similarly, the following equation holds

$$
\begin{aligned}
S_2 &= \sum_{k=1}^{n} P_k f_2\left(x_1(k),x_2(k)\right) \\
&= \int_0^1\int_0^1 f_2(x_1,x_2)dx_1dx_2 + \vartheta_2 D_{EF}(\mathcal{P}_n)\mathrm{TV}(f_2),
\end{aligned}
\tag{2.26}
$$

where $|\vartheta_2| \leq 1$.

It can be proved that $\mathrm{TV}(f) = \mathrm{TV}(f_1) + \mathrm{TV}(f_2)$ [12]. Then 2.25 and 2.26 immediately lead to

$$
\begin{aligned}
S_1 + S_2 &= \sum_{k=1}^{n} P_k f\left(x_1(k),x_2(k)\right) \\
&= \int_0^1\int_0^1 f(x_1,x_2)dx_1dx_2 + D_{EF}(\mathcal{P}_n)\left[\vartheta_1\mathrm{TV}(f_1)+\vartheta_2\mathrm{TV}(f_2)\right],
\end{aligned}
\tag{2.27}
$$

which finally results in

$$
\begin{aligned}
&\left|\sum_{k=1}^{n} P_k f\left(x_1(k), x_2(k)\right) - \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2\right| \\
&= D_{EF}(\mathcal{P}_n)\left[\vartheta_1 \mathrm{TV}(f_1) + \vartheta_2 \mathrm{TV}(f_2)\right] \\
&\leq D_{EF}(\mathcal{P}_n)\left[\mathrm{TV}(f_1) + \mathrm{TV}(f_2)\right] \\
&= D_{EF}(\mathcal{P}_n)\mathrm{TV}(f).
\end{aligned} \tag{2.28}
$$

This completes the proof of Theorem 2.2 for $s = 2$. The situation for $s \geq 3$ is similar whereas more lengthy and will be provided in the Appendix. □

2.3. Quantitative relation between EF- and GF-discrepancy and error bounds. Both Theorem 2.1 and Theorem 2.2 are the error bounds in terms of EF-discrepancy. Analogous to these theorems, we have the following conjecture regarding GF-discrepancy

$$
\left|\int_0^1 \cdots \int_0^1 f(\mathbf{x}) d\mathbf{x} - \sum_{q=1}^{n} P_q f(\mathbf{x}_q)\right| \leq \alpha(s) D_{GF}(\mathcal{P}_n)\mathrm{TV}(f), \tag{2.29}
$$

where $D_{GF}(\mathcal{P}_n)$ is the GF-discrepancy defined by 2.9 and 2.10, $\alpha(s) = \mathcal{O}(s)$ and s is the dimension. We should note that for the one-dimensional problem $D_{GF}(\mathcal{P}_n) = D_{EF}(\mathcal{P}_n)$ and thus the inequality 2.29 reduces to inequality 2.12, which has been proved in Theorem 2.1. For multi-dimensional cases the rigorous proof of 2.29 is still open. Nonetheless, we can say something more about the reasonability of 2.29 by considering the quantitative reliability between EF- and GF-discrepancy.

According to the definition of EF-discrepancy we have

$$
F_n(x_1, x_2, \cdots, x_s) - F(x_1, x_2, \cdots, x_s) = \vartheta(x_1, x_2, \cdots, x_s) D_{EF}(\mathcal{P}_n), \tag{2.30}
$$

where $|\vartheta(x_1, x_2, \cdots, x_s)| \leq 1$. It is noted that 2.21 and 6.8 in the Appendix is actually the special case of this equation as $F(x_1, x_2, \cdots, x_s) = \prod_{j=1}^{s} x_j$. Following 2.30 there is

(2.31)

$$
F_n(1, \cdots, 1, x_j, 1, \cdots, 1) - F(1, \cdots, 1, x_j, 1, \cdots, 1) = \vartheta(1, \cdots, 1, x_j, 1, \cdots, 1) D_{EF}(\mathcal{P}_n)
$$

for any $1 \leq j \leq s$.

Substituting 2.31 in the marginal F-discrepancy as defined in 2.10 yields

$$
\begin{aligned}
D_{F,j}(\mathcal{P}_n) &= \sup_{x_j \in \mathbb{R}} \left|F_n(1, \cdots, 1, x_j, 1, \cdots, 1) - F(1, \cdots, 1, x_j, 1, \cdots, 1)\right| \\
&= \left|F_n(1, \cdots, 1, x_j^*, 1, \cdots, 1) - F(1, \cdots, 1, x_j^*, 1, \cdots, 1)\right| \\
&= \left|\vartheta(1, \cdots, 1, x_j^*, 1, \cdots, 1)\right| D_{EF}(\mathcal{P}_n),
\end{aligned} \tag{2.32}
$$

where x_j^* is the point at which $|F_n(1, \cdots, 1, x_j, 1, \cdots, 1) - F(1, \cdots, 1, x_j, 1, \cdots, 1)|$ reaches its maxima.

According to the definition of GF-discrepancy $D_{GF}(\mathcal{P}_n) = \max_{1 \leq j \leq s} D_{F,j}(\mathcal{P}_n)$, there is

$$
D_{GF}(\mathcal{P}_n) = \left\{\max_{1 \leq j \leq s} \left|\vartheta(1, \cdots, 1, x_j^*, 1, \cdots, 1)\right|\right\} D_{EF}(\mathcal{P}_n). \tag{2.33}
$$

Because $\vartheta(x_1, x_2, \cdots, x_n) \leq 1$ for all $x_1, x_2, \cdots, x_n$, we have immediately

$$D_{GF}(\mathcal{P}_n) \leq D_{EF}(\mathcal{P}_n). \tag{2.34}$$

On the other hand, we note that the EF-discrepancy characterizes the global discrepancy involving all s directions, whereas the GF-discrepancy seems to taking only marginal information separately, but it should be stressed that the computation of assigned probabilities involves global distribution information. Thus, there should be more connections between EF- and GF-discrepancy. Intuitively, we guess that EF-discrepancy should not exceed s times of GF-discrepancy, i.e.

$$D_{EF}(\mathcal{P}_n) \leq sD_{GF}(\mathcal{P}_n). \tag{2.35}$$

Substituting inequality 2.35 in the extended Koksma-Hlawka inequality 2.17 immediately yields inequality 2.29 provided $\alpha(s) = s$.

It follows from inequalities 2.34 an 2.35 that

$$\frac{1}{s}D_{EF}(\mathcal{P}_n) \leq D_{GF}(\mathcal{P}_n) \leq D_{EF}(\mathcal{P}_n) \text{ or } D_{GF}(\mathcal{P}_n) \leq D_{EF}(\mathcal{P}_n) \leq sD_{GF}(\mathcal{P}_n). \tag{2.36}$$

Again we stress that part of the above inequality is rigorous proved by inequality 2.34 but the rest part is a conjecture. For uniform distributions and $s = 2, 3, 4$ and 5, Fig. 2.1 shows the pairs of EF- and GF-discrepancy of different point sets in logarithm coordinate system. It is seen that inequality 2.36 holds very well without exception. Of course, to be more confident, we may replace s in inequality 2.36 by $\alpha(s) = \mathcal{O}(s)$.

2.4. Justification of the concept of EF- and GF-discrepancy by numerical examples. It should be emphasized that the employment of the assigned probabilities rather than $1/n$ is essential for the definition of $D_{GF}(\mathcal{P}_n)$ (2.9). Otherwise, there might be severe deficiency. For comparison, we could define an index by

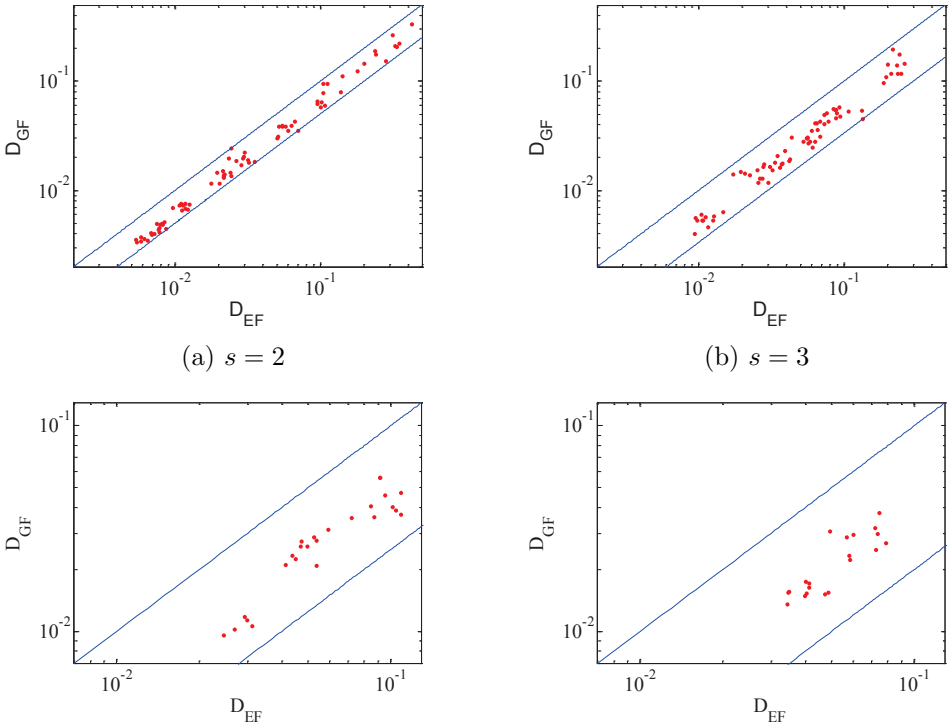
$$D_{CF} = \max_{1 \leq i \leq s} \left\{ \sup_{x \in \mathbb{R}} \left| \tilde{F}_{n,i}(x) - F_i(x) \right| \right\}. \tag{2.37}$$

The only difference between 2.9 and 2.37 is that the marginal empirical CDF in 2.37 is defined as usual by

$$\tilde{F}_{n,i}(x) = \sum_{q=1}^{n} \frac{1}{n} \cdot I\{x_{i,q} < x\}. \tag{2.38}$$

instead of 2.11, i.e., the assigned probabilities P_q's are replaced by $1/n$. We will show that the index D_{CF} in 2.37, if adopted to replace D_{GF}, will give misleading results even for some very simple cases.

Suppose X, Y are two independent random variables uniformly distributed over $[0, 1] \times [0, 1]$. To illustrate the problem, we consider 3 different point sets with the same cardinal number: one is uniformly scattered on the diagonal line, one is generated by the number theoretical method (NTM) by $x_{1,q} = \frac{2q-1}{2n}$, $x_{2,q} = \frac{2F_{m-1}q-1}{2n} - \text{int}\left(\frac{2F_{m-1}q-1}{2n}\right)$ for $q = 1, 2, \cdots, n$, $n = F_m$, where $F_m = F_{m-1} + F_{m-2}$, $F_0 = 1, F_1 = 1$ is the Fibonacci sequence, and int$(\cdot)$ indicates the greatest integer smaller than the bracketed value [12], the third point set is generated by Monte Carlo sampling (MCS). For $n = 13$ the three point sets are shown in Figs. 2.2(a), 2.2(b) and



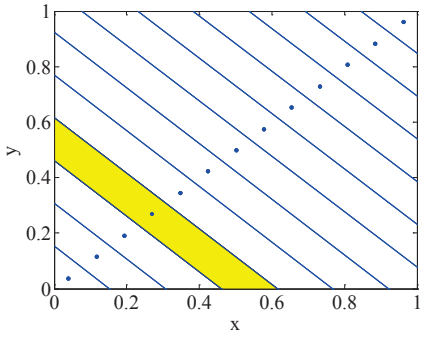
(a) $s = 2$ (b) $s = 3$

(c) $s = 4$ (d) $s = 5$

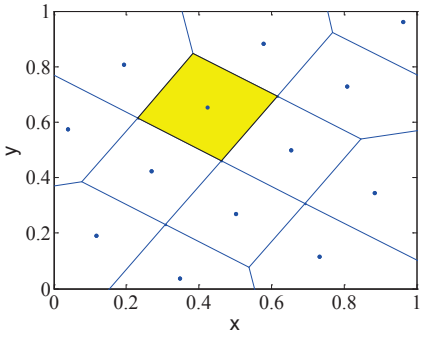
FIG. 2.1. *The pair of EF- and GF-discrepancy of different point sets in logarithm coordinate system (Annotation: The lower line is* $D_{GF}(\mathcal{P}_n) = \frac{1}{s} D_{EF}(\mathcal{P}_n)$*, the upper line is* $D_{GF}(\mathcal{P}_n) = D_{EF}(\mathcal{P}_n)$*).*

2.2(c), respectively. Simultaneously, the corresponding Voronoi cells, whose areas are the assigned probabilities for the uniform distribution, are also pictured in Fig. 2.2.

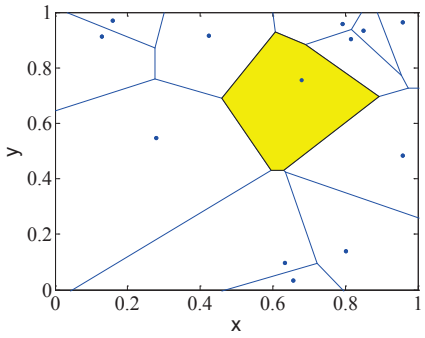
The comparison among CDFs of X is shown in Fig. 2.3. Because of the symmetry the CDFs of Y are identical to those of X, respectively. Intuitively it is seen that the NTM point set is far more uniform over $[0,1] \times [0,1]$ than the Diagonal and MCS point set (Figs. 2.2(a), 2.2(b) and 2.2(c)). This is verified by the discrepancy (F-discrepancy): for the point sets in Fig. 2.2 with $n = 13$, $D_F(\mathcal{P}_{\text{Diagonal}}) = 0.2884$ and $D_F(\mathcal{P}_{\text{MCS}}) = 0.4993$ (by 2.5 and 2.6), which are greater than $D_F(\mathcal{P}_{\text{NTM}}) = 0.1405$ (Table 2.1). In the case the assigned probabilities are employed to replace $1/n$, the EF-discrepancies for the point sets are, respectively, $D_{EF}(\mathcal{P}_{\text{Diagonal}}) = 0.3831$, $D_{EF}(\mathcal{P}_{\text{MCS}}) = 0.3451$, $D_{EF}(\mathcal{P}_{\text{NTM}}) = 0.1813$ and (computed by 2.5 and 2.8), again the former two are greater than the third one. This means that both D_F and D_{EF} could reasonably character the uniformity property of the point sets. Besides, it is noted that in the above point sets for $n = 13$ the MCS point set in Fig. 2.2(c) of course seems more uniform than the Diagonal point set in Fig. 2.2(a), but the F-discrepancies are $D_F(\mathcal{P}_{\text{Diagonal}}) = 0.2884 < D_{EF}(\mathcal{P}_{\text{MCS}}) = 0.4993$ which gives misleading result. Considering the EF-discrepancy where the effect of assigned probability is involved, there is $D_F(\mathcal{P}_{\text{Diagonal}}) = 0.3831 > D_{EF}(\mathcal{P}_{\text{MCS}}) = 0.3451$, which yields a reasonable judgement for the degree of uniformity of the point sets. This means that in essence D_{EF} is an index more reasonable than D_F, i.e., introducing the effect of assigned



(a) Voronoi cells for Diagonal point set ($n = 13$)



(b) Voronoi cells for NTM point set ($n = 13$)



(c) Voronoi cells for MCS point set ($n = 13$)

FIG. 2.2. *The point set and the corresponding Voronoi cells in a unit square.*

TABLE 2.1

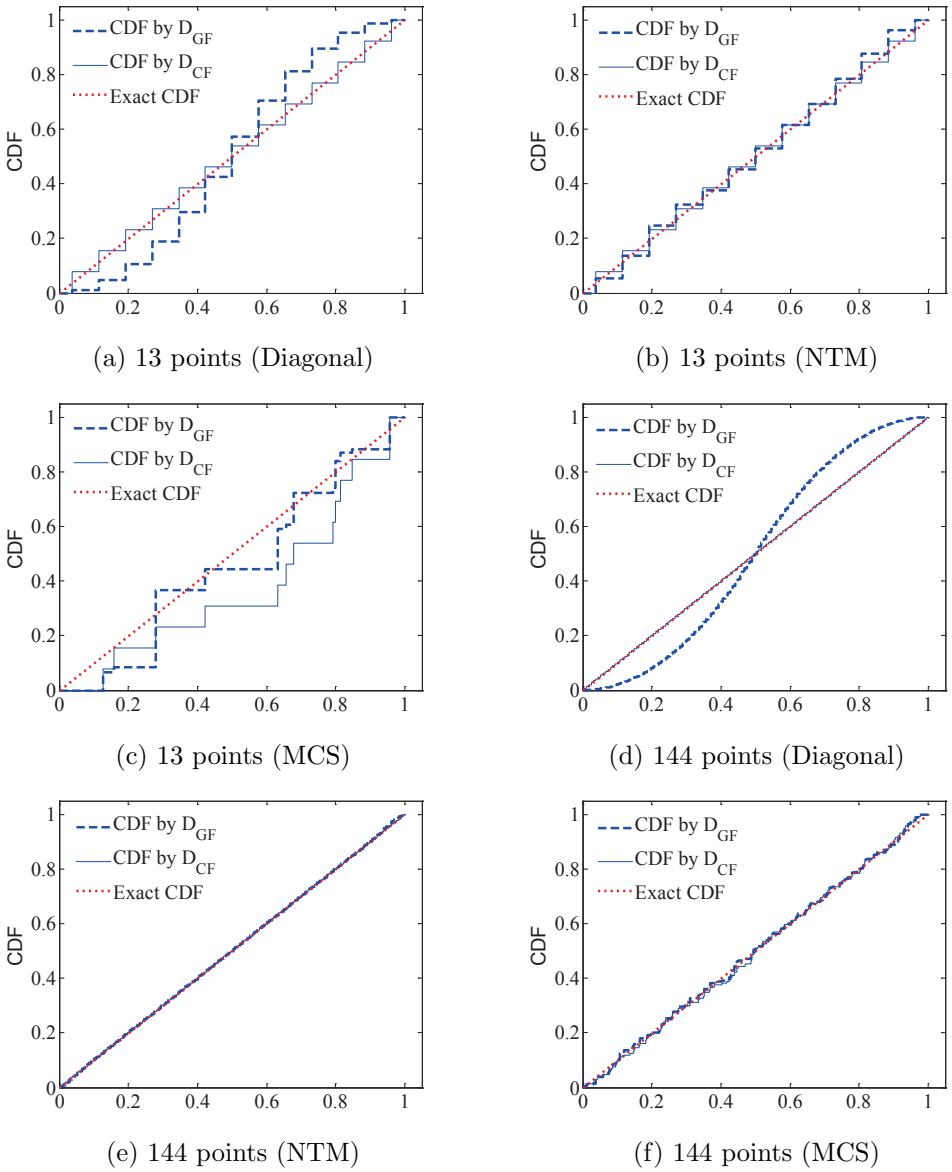
Different F-discrepancies of 3 types of point sets with different number as shown in Figs. 2.2 and 2.3.

Methods	Number of points	D_F	D_{CF}	D_{EF}	D_{GF}
Diagonal	13	0.2884	0.0385	0.3831	0.1627
	144	0.2534	0.0035	0.3379	0.1285
	1597	0.2503	0.0003	0.3337	0.1253
NTM	13	0.1405	0.0385	0.1813	0.0793
	144	0.0180	0.0035	0.0257	0.0122
	1597	0.0021	0.0003	0.0052	0.0026
MCS	13	0.4993	0.3247	0.3451	0.1952
	144	0.0911	0.0354	0.0494	0.0275
	1597	0.0263	0.0217	0.0093	0.0061

probability will improve the adequacy of the discrepancy for judging uniformity, especially in the cases the discrepancy is relatively large.

Now we consider the maximum marginal discrepancies. In this case, employing 2.37 and 2.38 where $1/n$ is adopted in the computation of empirical CDF (thus the effect of assigned probability is not taken into account), we have $D_{CF}(\mathcal{P}_{\text{Diagonal}}) = 0.0385$ and $D_{CF}(\mathcal{P}_{\text{NTM}}) = 0.0385$. They are exactly equal! Actually this is expected because the projections of the point sets $\mathcal{P}_{\text{Diagonal}}$ and $\mathcal{P}_{\text{NTM}}$ in X_l are exactly the same uniformly spaced and the weights are all identical to $1/n$. The exact values of

D_{CF} of the two point sets are all $1/(2n)$. This means that D_{CF} could not distinguish the degree of uniformity of these two point sets. But if we make use of 2.9 and 2.11 where the effect of assigned probability is involved, the GF-discrepancies of the two point sets could be obtained as $D_{GF}(\mathcal{P}_{\text{Diagonal}}) = 0.1627$ and $D_{GF}(\mathcal{P}_{\text{NTM}}) = 0.0793$. The former is greater than the latter, implying that the NTM point set is more uniform than the Diagonal point set. This judgement is of course reasonable. The distinction of D_{CF} and D_{GF} could also be seen clearly from Figs. 2.3(a)-2.3(f), where the comparisons between different point sets are illustrated. From these figures, it is seen that D_{GF} is much more sensitive than D_{CF} to the degree of uniformity of



(a) 13 points (Diagonal) (b) 13 points (NTM)

(c) 13 points (MCS) (d) 144 points (Diagonal)

(e) 144 points (NTM) (f) 144 points (MCS)

FIG. 2.3. *The CDFs defined by different F-discrepancies.*

the point sets. The reason could be seen from Fig. 2.2(a). It is observed clearly that the assigned probabilities are far from $1/n$. Another interesting property of $D_{GF}(\mathcal{P}_{\text{Diagonal}})$ is that the GF-discrepancy of Diagonal point set could not decrease to zero as the number of points n increases to infinity. Actually, considering the Voronoi cells and the assigned probabilities, it is easy to prove that the saturated limit value is $\lim_{n\to\infty} D_{GF}(\mathcal{P}_{\text{Diagonal}}) = \max_{0\le x\le \frac{1}{2}}(x - 2x^2) = \frac{1}{8} = 0.125$. The tendency of approaching this value could be observed very clearly from Table 2.1. This means that because all the points are located on the diagonal line, even infinite points could not improve the degree of uniformity of the point set over the square. On the other hand, for the NTM point set, the GF-discrepancy will tend to zero as n increases to infinity, which means that the degree of uniformity could of course be improved for NTM point set when the number of points increases. Clearly, such judgements based on GF-discrepancy are reasonable.

Viewed from the above tables and figures, the employment of the GF-discrepancy D_{GF} defined in 2.9 and 2.11 could be justified. Compared to the EF-discrepancy involving the effect of assigned probability whose computation is an NP-hard problem, the computation of GF-discrepancy is much more efficient and increases only linearly against the dimension. This makes GF-discrepancy applicable to higher-dimensional problems.

3. Applications of GF-discrepancy in rotational Q-SPM. In this section, the GF-discrepancy will be applied to the optimal selection of rotation angles for the improvement of a class of asymmetric cubature formulae, which could be called the quasi-symmetric point method (Q-SPM) [27], developed for symmetric measures by Victoir [23]. The necessity of the rotational transform on Q-SPM is firstly discussed. The algebraic accuracy of the new point set is proved to be identical to Q-SPM, but the GF-discrepancy of the new point set is greatly reduced and the accuracy is thus obviously improved.

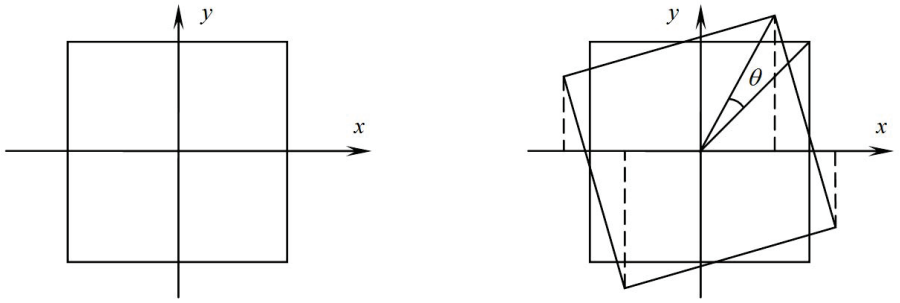
3.1. Q-SPM and rotational Q-SPM. Using the invariant theory and orthogonal arrays, Victoir constructed some cubature formulae for Gaussian weighted multi-dimensional integrals with 5th degree of algebraic accuracy [23]. These points are called the quasi-symmetric points [27]. If $f(\mathbf{x})$ is a polynomial whose degree is no more than 5, the following equation holds exactly

$$\frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{x}^T\mathbf{x}/2} f(\mathbf{x}) dx_1 \cdots dx_s = \sum_{k=1}^{N} a_k f(\mathbf{x}_k) \tag{3.1}$$

with $\mathbf{x}_k$ being the quasi-symmetric points and a_k being the corresponding weights. Two classes of cubature formulae with all positive weights are given in Q-SPM. The number of points needed by the second class of Q-SPM for the dimensions from 3 through 16 is less than 300, from 17 to 20 is limited to a little greater than 500 whereas from 21 to 24 is only nearly 1000 [23, 27]. This is extremely exciting for the algebraic accuracy of up to 5th degree and the dimension up to 24. The second class of these integral points is given by

$$\mathbf{x}_0 = (hr, 0, \cdots, 0), \quad \mathbf{x}_1 = (h\eta, \cdots, h\eta) \tag{3.2}$$

$$r^2 = (s+2)/2, \quad \eta^2 = (s+2)/(s-2) \tag{3.3}$$



(a) (b)

FIG. 3.1. *Quasi-symmetric points and rotation of the quasi-symmetric points.*

where s is the dimension and h is the permutation of ± 1 . The sums of the weights for these points read

$$w_0 = 8s/(s+2)^2, \quad w_1 = (s-2)^2/(s+2)^2 \tag{3.4}$$

Clearly $w_0 + w_1 = 1$. By this, the weights of the points could be determined. For the points on the coordinates, the weights are w_0/n_1 , for the other points the weights are w_1/n_2, where n_1, n_2 are specified for different s in Q-SPM.

Although Q-SPM behaves well in the application to stochastic response analysis of nonlinear structures [27], there are also some deficiencies in this method. Particularly, because of the quasi-symmetry and sparseness of Q-SPM, the information of the marginal probability density function of the basic random variables, which could be intuitively characterized by the projection ratio [5], could not be captured sufficiently. Thus the application of Q-SPM is limited in some cases. To be clearer, we use an illustrative example to clarify the problem and to introduce the basic idea of resolving it. It can be shown from 3.2 that there are two sets of points for the quasi-symmetric points. One set is located on the coordinate axes and the other set includes some of the vertexes of a hyper-cube which can be illustrated by the vertexes of the square in Fig. 3.1. The overlapping of the projection of the four points in the x-axis and y-axis can be easily observed. For instance, in Fig. 3.1(a), consider all the points on the coordinate axes and the vertex points, there are totally 8 points, whereas the projection of these 8 points on both x and y axis contains only 3 different points, thus, the projection ratios on both axes are only 3/8. Such overlapping is more severe in the higher-dimensional space and would result in very few effective points on each axis. In other words, the projection ratio of Q-SPM is usually quite small. Actually, as $s = 24$, when the point set in 3.2 is employed, the total number of points is 1072 and the projection ratios on all the 24 axes are now $5/1072 \approx 1/200$, which is very small.

To resolve the matter of overlapping and to improve the projection ratios, a natural idea is to rotate these quasi-symmetric points (Fig. 3.1(b)). The accuracy of the integral is expected to not decrease because of the rotational symmetry of 3.1 (actually, we would show later that the accuracy is much higher) and the problem of projection overlapping is expected to be resolved by rotation. By doing so, the projection ratio could be increased to nearly or exactly 1, as could be seen intuitively clearly from Fig. 3.1. But a more rigorous criterion should resort to GF-discrepancy.

The corresponding equation to Fig. 3.1(b) is

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.5}$$

This idea could be extended to high-dimensional spaces by the Givens transform [25]. For instance, rotating a vector $\mathbf{p} = (p_1, p_2, \cdots, p_s)$ in the (i, j) plane by an angle of θ (in rad) counterclockwise will yield

$$\mathbf{p}' = \mathbf{G}_{ij}(\theta)\mathbf{p}, \tag{3.6}$$

where $\mathbf{p}'$ is the point after rotation and

$$\mathbf{G}_{ij}(\theta) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos\theta & \cdots & -\sin\theta & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \sin\theta & \cdots & \cos\theta & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ i\text{-th row} \\ \\ j\text{-th row} \\ \\ \\ \end{matrix}, \tag{3.7}$$

$$\qquad\qquad\qquad\qquad\quad i\text{-th column} \qquad j\text{-th column}$$

therefore, any rotation of the point in the space can be expressed as

$$\mathbf{p}' = \prod_{i=1}^{s} \prod_{j=i+1}^{s} \mathbf{G}_{ij}(\theta_{ij})\mathbf{p} = \mathbf{R}\mathbf{p}, \tag{3.8}$$

which includes $s(s-1)/2$ times of plane rotations. In 3.8 $\mathbf{R} = \prod_{i=1}^{s} \prod_{j=i+1}^{s} \mathbf{G}_{ij}(\theta_{ij})$. It is easy to verify that $\mathbf{R}$ is an orthogonal matrix [18], i.e.,

$$\mathbf{R}^T\mathbf{R} = \mathbf{I}, |\mathbf{R}| = 1, \tag{3.9}$$

in which $|\mathbf{R}|$ is the determinant of $\mathbf{R}$.

Clearly, for a Q-SPM point set $\mathcal{P}_{\mathrm{Q-SPM}}$ the rotation angles θ_{ij}'s determine the rotational Q-SPM (RQ-SPM). Thus, two questions arise: (1) Will the algebraic accuracy of $\mathcal{P}_{\mathrm{Q-SPM}}$ be kept in the rotational transform? and (2) How to specify θ_{ij}'s optimally to improve the properties of $\mathcal{P}_{\mathrm{Q-SPM}}$? The answer to the first question is positive and will be elaborated soon. To answer the second question, some type of objective function $J(\mathcal{P}_{\mathrm{RQ-SPM}}) = J(\theta_{ij})$, e.g., the GF-discrepancy defined in the preceding section, could be adopted and thus an optimization problem is encountered. Genetic algorithm could be employed in this step.

3.2. Degree of algebraic accuracy of rotational Q-SPM. Intuitively, a rotation is a linear transform and does not change the degree of the polynomials of the integrand function. Besides, the weight function in 3.1 is rotationally symmetric. It is thus expected that the rotational transform will not reduce the algebraic accuracy in comparison to the Q-SPM point set. This assertion will be proven rigorously as follows.

Let $\mathbf{R}$ be the rotation matrix and $f(\mathbf{x})$ be a polynomial whose degree is less than or equal to 5. Let $\mathbf{y} = \mathbf{R}\mathbf{x}$, then it follows that $\mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{x}$ (cf. 3.9) and

$d\mathbf{x} = |\mathbf{R}^{-1}|d\mathbf{y} = d\mathbf{y}$, therefore we have

$$\begin{aligned} &\frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{x}^T\mathbf{x}/2} f(\mathbf{R}\mathbf{x}) dx_1 \cdots dx_s \\ &= \frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{x}^T\mathbf{x}/2} f(\mathbf{x}) dx_1 \cdots dx_s. \end{aligned} \tag{3.10}$$

Employing the Q-SPM point set $\mathcal{P}_{\mathrm{Q-SPM}} = \{\mathbf{x}_k, k = 1, 2, \cdots, N\}$ with the weights a_k's, 3.1 holds exactly for $f(\mathbf{x})$ with the degree of polynomials less than or equal to 5. As mentioned, the rotational transform is a linear transform and thus does not change the degree of polynomials, i.e., $f(\mathbf{R}\mathbf{x})$ has the same polynomial degree as $f(\mathbf{x})$. Thus the following cubature holds exactly

$$\frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{x}^T\mathbf{x}/2} f(\mathbf{R}\mathbf{x}) dx_1 \cdots dx_s = \sum_{k=1}^{N} a_k f(\mathbf{R}\mathbf{x}_k). \tag{3.11}$$

Incorporating 3.1, 3.10 and 3.11 yields

$$\frac{1}{(2\pi)^{s/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathbf{x}^T\mathbf{x}/2} f(\mathbf{x}) dx_1 \cdots dx_s = \sum_{k=1}^{N} a_k f(\mathbf{R}\mathbf{x}_k) = \sum_{k=1}^{N} a_k f(\mathbf{y}_k), \tag{3.12}$$

where $\mathbf{y}_k = \mathbf{R}\mathbf{x}_x$ are the points after rotational transform, i.e., $\mathcal{P}_{\mathrm{RQ-SPM}} = \{\mathbf{y}_k = \mathbf{R}\mathbf{x}_k, k = 1, 2, \cdots, N\}$. This demonstrates that the rotational Q-SPM point set $\mathcal{P}_{\mathrm{RQ-SPM}} = \{\mathbf{y}_k = \mathbf{R}\mathbf{x}_k, k = 1, 2, \cdots, N\}$ has the same degree of algebraic accuracy as the original Q-SPM point set $\mathcal{P}_{\mathrm{Q-SPM}} = \{\mathbf{x}_k, k = 1, 2, \cdots, N\}$ when the corresponding weights are invariant.

It is easy to verify that the proof is also applicable to other cubature formulae whose domain of integration and weight function both exhibit rotational symmetry.

3.3. Genetic algorithm for optimal selection of rotational angles. Employing the GF-discrepancy as the objective function, we encounter the following optimization problem:

$$\begin{aligned} &\min\{J(\theta_{i,j})\} = \min\{D_{GF}\left(\mathcal{P}_{\mathrm{RQ-SPM}}(\theta_{1,2}, \theta_{1,3}, \cdots, \theta_{i,j}, \cdots, \theta_{s-1,s})\right)\} \\ &s.t. \qquad 0 \le \theta_{i,j} < 2\pi, 0 < i < j \le s. \end{aligned} \tag{3.13}$$

where $\mathcal{P}_{\mathrm{RQ-SPM}}(\theta_{1,2}, \theta_{1,3}, \cdots, \theta_{i,j}, \cdots, \theta_{s-1,s})$ is the resulting RQ-SPM when the transform 3.8 is performed in sequences on the original Q-SPM point set $\mathcal{P}_{\mathrm{Q-SPM}}$. In this sense, $\mathcal{P}_{\mathrm{Q-SPM}}$ could be regarded as a special case of $\mathcal{P}_{\mathrm{RQ-SPM}}$ when all the rotation angles $\theta_{i,j}$'s are zeroes.

This is a problem of multi-variable optimization where the inputs are the rotation angles and the objective function is the GF-discrepancy. Exhaustive search is unfeasible because the computational effort grows exponentially against the dimension. Nor is analytical optimization feasible because of the complexity of the problem. A genetic algorithm (GA) will be employed here. We should also note that once the optimized RQ-SPM is determined, it could be applied to different problems of uncertainty quantification or stochastic dynamics. In other words, such an optimization problem could be solved once and for ever, and the tabulated results could be employed by all the users in various fields. Thus, the rotational transform will not increase the computational efforts when solving practical problems.

GA was developed by Holland [11] firstly and consummated owing to the work of De Jong and Goldberg. Based on the principles of genetics and natural selection, GA allows a population composed of many individuals to evolve under specified selection rules to a state that maximizes the fitness (i.e., minimize D_{GF} in the present case). The employment of GA in the above problem will be outlined briefly. For more information, refer to [10]. A string of binary values is used here to represent an angle ranging from 0 to 2π. Although the angle is discretized in this way, the difference is negligible if the string is long enough. For the above problem, $s(s-1)/2$ strings are needed to represent the rotation angles. These strings are then connected from head to tail to obtain a longer one, referred to as a chromosome

$$\text{chromosome} = \left[\underbrace{01\ldots0}_{\theta_{1,2}}\underbrace{11\ldots0}_{\theta_{1,3}}\ldots\underbrace{00\ldots0}_{\theta_{s-1,s}}\right]. \tag{3.14}$$

The process of changing the original set of angles into a long string is called encoding chromosomes.

GA begins by generating a number of chromosomes randomly. Each chromosome, composed by 0 and 1, corresponds to a set of angles which can be decoded along the length of chromosome as they were encoded. Each specified set of angles decoded from a chromosome are inserted into 3.13 to evaluate the objective function respectively. A fitness value is then assigned to each chromosome based on these objective function values (i.e. the fitness could be calculated based on the rank of the objective function values).

Chromosomes that have a lower fitness are more likely to be discarded which imitates the procedure of natural selection. The surviving chromosomes undergo a mating or recombination process to produce offspring by combining the information from two or more parent chromosomes. A mutation operator, which means a bit-flip in our case, is also introduced in this procedure to keep GA from converging too early.

After the mutations take place, the fitness associated with the offspring and mutated chromosomes is computed once again. The process described is then iterated until the quality of the strings evolves to an acceptable level or the number of iteration reaches the prescribed value.

The program is based on the Matlab toolbox developed at Sheffield University [21]. The size of the population is chosen to be 40. The generation gap is 0.95. The length of the string of binary values for a single angle is 20. The probability for mutation is taken to be 0.01 and the crossover rate is 0.7. Fig. 3.2 shows the efficacy of GA. Experiences show that the iteration number will be usually more than 3000 to make the GF-discrepancy small enough.

4. Numerical examples. In this section, two numerical examples are presented to show the effectiveness of GF-discrepancy and the rotational transform on Q-SPM. In the first example, it is shown that the probability information in the initial random variables can be characterized much better by RQ-SPM than by Q-SPM. The second numerical examples include a series of structural dynamics analyses involving random parameters which demonstrate that due to smaller GF-discrepancy, RQ-SPM has considerably higher accuracy than Q-SPM.

4.1. Case 1: Probability information characterized by RQ-SPM. In this case the 20-dimensional point set is employed to illustrate the effectiveness of RQ-SPM. Similar results can be observed in problems in different dimensions. The GF-discrepancy is adopted to be the objective function and genetic algorithm is performed

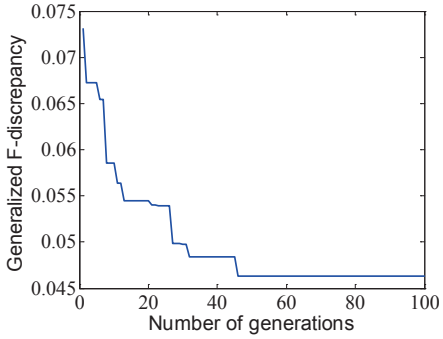


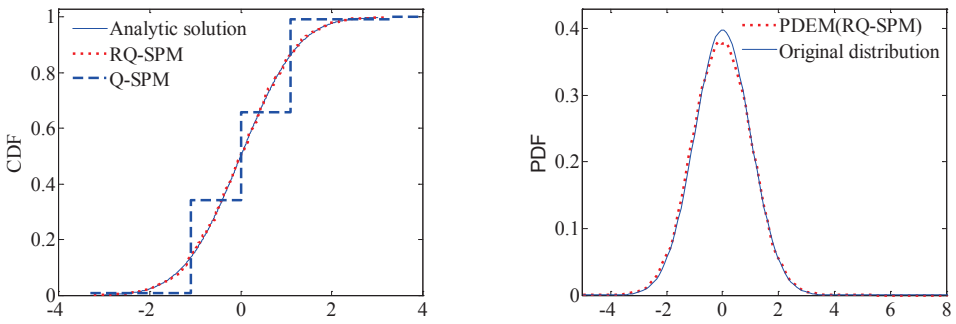
FIG. 3.2. *The GF-discrepancy (objective function value) against number of generations.*

to choose the rotation angles. Each dimension of RQ-SPM (or Q-SPM) represents an independent random variable with normal distribution $N(\mu, \sigma)$. The empirical marginal CDF of a typical dimension calculated by Q-SPM and RQ-SPM are shown in Fig. 4.1(a). It can be seen that the result obtained by RQ-SPM shows perfect agreement with the exact CDF whereas the overlapping of projections in Q-SPM can again be observed resulting in very few effective points. By rotation the GF-discrepancy is reduced from 0.2085 (Q-SPM) to 0.0318 (RQ-SPM).

To show how the point selection will affect the output probability density function in a clear way, the computation of the extreme value distribution (EVD) is illustrated here. It is known that the EVD can be obtained via the probability density evolution method (PDEM) by constructing a virtual stochastic process [4]. For clarity, we let

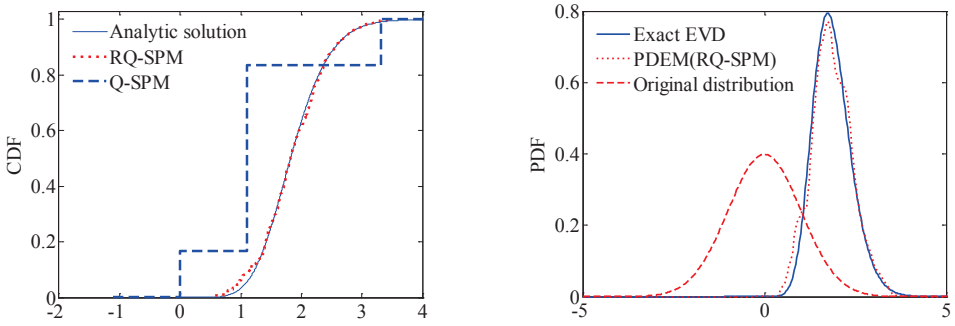
$$Y = \max(X_1, X_2, \cdots, X_{20}), \tag{4.1}$$

where $X_1, X_2, \cdots, X_{20}$ are 20 i.i.d. random variables with normal distribution $N(\mu, \sigma)$. The representative points of Y are the maximum of the representative points of $X_1, X_2, \cdots, X_{20}$. Because of the overlapping of projections of Q-SPM point set (as shown in 3.2 and 3.3, and schematically shown in Fig. 3.1), on each axis there are only 5 different points, which can also be observed in Fig. 4.1(a). According to the rule of Q-SPM, by 4.1 Y could only take 3 possible values. This is shown clearly by the step



(a) CDF (b) PDF

FIG. 4.1. *The CDF and PDF of the original random variables.*



(a) CDF (b) PDF

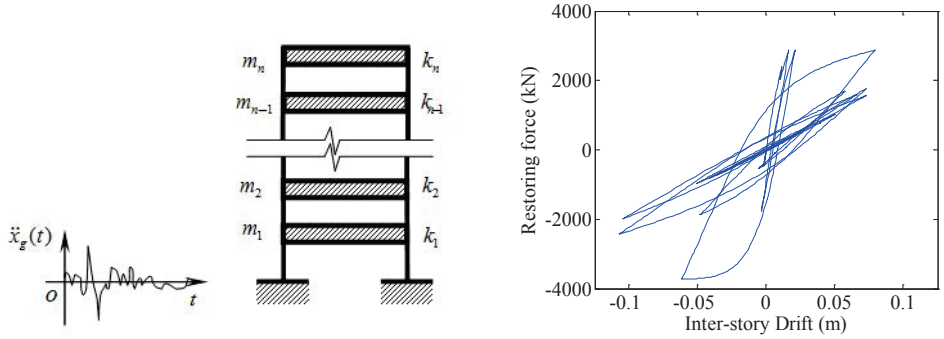
FIG. 4.2. *The empirical CDF, CDF and PDF of the extreme value by Q-SPM and RQ-SPM.*

curve labeled Q-SPM in Fig. 4.2(a), where the empirical CDF is the step curve, which is far from the exact CDF. Thereby, employing Q-SPM will definitely fail to obtain EVD with fair accuracy. On the other hand, this problem of overlapping of projection is almost completely avoided by employing RQ-SPM (Figs. 4.1(a) and 4.1(b)). Clearly, such resulted EVD accords fairly well with the exact distribution (Fig. 4.2(b)), of which the PDF is given by [1]

$$f_Y(y) = \frac{s}{\sqrt{2\pi}\sigma}\left[\Phi\left(\frac{y-\mu}{\sigma}\right)\right]^{s-1} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{4.2}$$

with $\mu = 0$, $\sigma = 1$ and $s = 20$ in the present case. Here $\Phi(\cdot)$ is the CDF of standard normal distribution. The CDF can be obtained by integrating over the above function. The empirical CDF of Y calculated by RQ-SPM is also shown in Fig. 4.2(a), showing good agreement with the analytical CDF. This is very impressive considering the fact that we did not require the empirical CDF of the extreme value approach the analytical solutions in the objective function. The PDF of the extreme value obtained by RQ-SPM also accords well with the exact extreme value distribution (EVD) (Fig. 4.2(b)).

4.2. Case 2: Structural analysis employing RQ-SPM. In this subsection, a variety of multi-degree-of-freedom (MDOF) nonlinear systems are analyzed. We start with a 9-story shear frame structure (Fig. 4.3(a)) with 20 random variables. All the lumped masses and the initial lateral inter-story stiffness are regarded as random variables. The mean values of mass from bottom to top are 3.5, 3.3, 3.0, 2.7, 2.7, 2.7, 2.7, 2.7 and 2.7 ($\times 10^5$kg), respectively. The mean values of the initial lateral inter-story stiffness are 2.7, 3.0, 3.0, 3.0, 3.0, 3.0, 3.0, 3.0 and 2.7 ($\times 10^5$kN/m) in turn from bottom to top. The extended Bouc-Wen model is adopted for the restoring forces [16, 24]. In total 13 parameters are included in this model [16], taking the values $A = 1$, $n = 1$, $q = 0.25$, $p = 1000$, $\lambda = 0.5$, $\phi = 0.05$, $d_\phi = 5$, $d_\upsilon = 2000$, $d_\eta = 2000$ and $\zeta = 0.99$. β and γ in the Bouc-Wen model are also regarded as independent random variables with mean values of 30 and 10, respectively. All the random variables are normally distributed and the coefficients of variation (COVs) are all 0.2. The El Centro accelerogram in E-W direction is adopted as the ground motion input. The first inter-story drift is the response of interest for all the cases. A typical sample of the restoring force v.s. inter-drift is shown in Fig. 4.3(b), which



(a) n-story shear frame (b) Restoring force

FIG. 4.3. *The shear frame and a typical sample of the restoring force.*

exhibits strong nonlinearity.

The evaluation of the mean and standard deviation of the responses involves a high-dimensional integral with respect to the basic random variables, which could be implemented by adopting the resulting point set by the proposed method in preceding sections. In this example 20 random variables are involved and 552 points were adopted in both Q-SPM and RQ-SPM. The comparison between RQ-SPM and Q-SPM is shown in Fig. 4.4 where RQ-SPM has accuracy higher than Q-SPM. To check the error quantitatively, we introduce the 2-norm relative errors of the mean and of the standard deviation as

$$e_{\|\mu\|} = \frac{\|\mu_{\mathrm{PDEM}}(t) - \mu_{\mathrm{MCS}}(t)\|_2}{\|\mu_{\mathrm{MCS}}(t)\|_2}, \quad e_{\|\sigma\|} = \frac{\|\sigma_{\mathrm{PDEM}}(t) - \sigma_{\mathrm{MCS}}(t)\|_2}{\|\sigma_{\mathrm{MCS}}(t)\|_2}, \tag{4.3}$$

respectively, where MCS denotes the results given by 9999 times of Monte Carlo simulation, $\|\cdot\|_2$ means 2-norm, i.e. $\|z(t)\| = \sqrt{\sum_{i=1}^{k} z^2(t_i)}$.

Computations show that Q-SPM does not behave well for some nonlinear cases. For instance, the accuracy of Q-SPM deteriorates quickly with the growth of the coefficient of variation (COV) of stiffness in some nonlinear cases. Listed in Fig. 4.5 are the relative errors of Q-SPM and RQ-SPM in terms of the mean and standard deviation. The structural parameters except the COV of stiffness have already been given above. It is observed that the relative errors grow rapidly with the increase of the COV of stiffness for Q-SPM whereas the accuracy of RQ-SPM is almost invariant. Similar results can be observed in Fig. 4.6 where the COV of all random variables varies from 0 to 0.2. Although the RQ-SPM and Q-SPM share the same degree of algebraic accuracy, we see that the accuracy of both mean and standard deviation by RQ-SPM is obviously higher than that by Q-SPM. The reason of such room of improvement for Q-SPM is that the response surface of a nonlinear structure is usually strongly nonlinear in terms of the basic random variables, and the nonlinearity is much stronger than the polynomials of 5th degree.

This problem of accuracy deterioration for Q-SPM can be observed clearly. Table 4.1 shows the results of several cases with different number of basic variables ranging from 15 to 20. The parameters in the Bouc-Wen model and the Rayleigh damping remain the same for all examples and all the coefficients of variation are specified to be 0.2. The parameters for the 20-dimensional example are already given above. In the case with 19 random variables, the parameters are the same as the

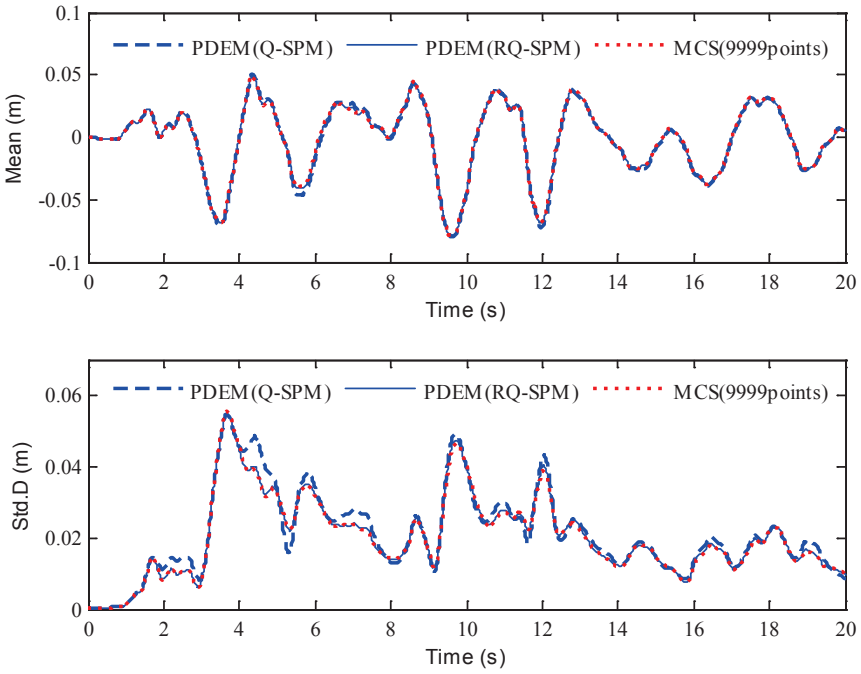
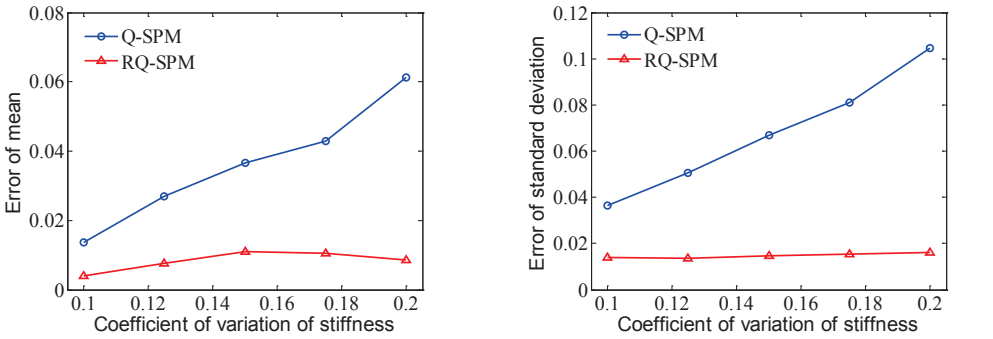


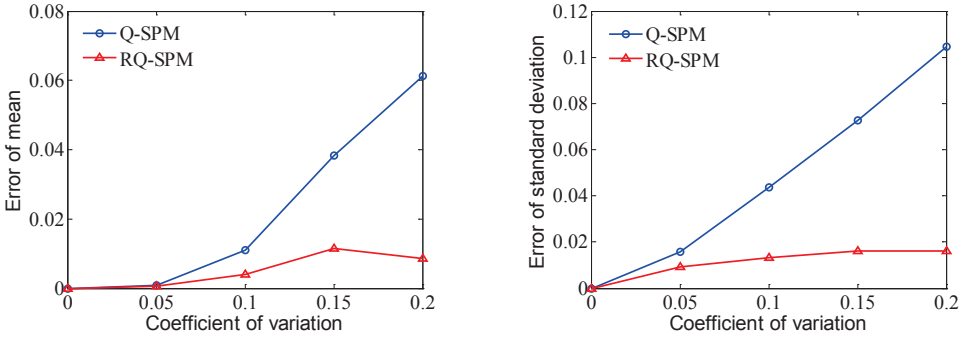
FIG. 4.4. *Comparison between different methods (Case 2) (RQ-SPM:* $e_{\|\mu\|} = 0.0085$, $e_{\|\sigma\|} = 0.0160$*; Q-SPM:* $e_{\|\mu\|} = 0.0612$, $e_{\|\mu\|} = 0.1045$*).*



(a) Relative error of mean (b) Relative error of standard deviation

FIG. 4.5. *The relative error of Q-SPM and RQ-SPM v.s. coefficient of variation of stiffness.*

20-dimensional one except that the top lumped mass is taken to be deterministic. Similarity, an 8-story shear frame structure is investigated for the numerical cases with 17 and 18 random variables. The lumped masses and the lateral stiffness from bottom to top are the same as the previous case from bottom to the 8th floor, respectively. For the purpose of examining if different inputs affect the accuracy, the ground motion is changed to be El Centro accelerogram in N-S direction. To verify the effectiveness of RQ-SPM in the numerical cases with 15 and 16 random variables,



(a) Relative error of mean (b) Relative error of standard deviation

FIG. 4.6. *The relative error of Q-SPM and RQ-SPM v.s. coefficient of all random variables.*

TABLE 4.1
Errors of Q-SPM, RQ-SPM and MCS with the same number of points.

Dimension			15	16	17	18	19	20
Number of points			286	288	546	548	550	552
Q-SPM	Error of	Mean	0.0453	0.063	0.1084	0.1283	0.0549	0.0612
		Std.D	0.0928	0.097	0.0808	0.0781	0.0958	0.1045
	D_{GF}		0.1938	0.1864	0.1925	0.1982	0.2035	0.2085
RQ-SPM	Error of	Mean	0.0246	0.0279	0.0183	0.0278	0.0088	0.0085
		Std.D	0.0377	0.0387	0.0191	0.0251	0.0202	0.016
	D_{GF}		0.0358	0.0406	0.0442	0.0381	0.0361	0.0318
MCS	Mean	Mean	0.0435	0.0525	0.0322	0.038	0.0264	0.0357
	Error of	Std.D	0.0793	0.0851	0.0297	0.034	0.045	0.0562
	COV of	Mean	0.45	0.41	0.61	0.62	0.57	0.45
	Error of	Std.D	0.37	0.55	0.40	0.48	0.32	0.74
	Mean of D_{GF}		0.0792	0.081	0.0638	0.0595	0.0625	0.0611
	0.95 quantile	Mean	0.0757	0.0879	0.0645	0.0768	0.0512	0.0621
	Error of	Std.D	0.1276	0.1621	0.0492	0.0608	0.0687	0.1246

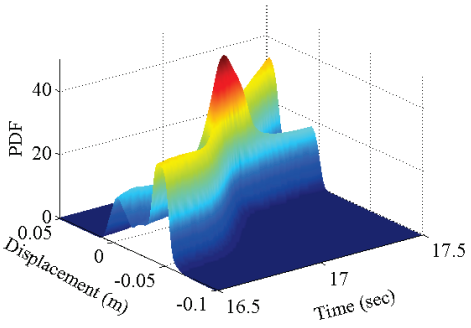
a 7-story shear frame structure is adopted and the parameters are chosen similarly. In these two cases the ground motion is changed to Kobe accelerogram.

In Table 4.1 for each case the results obtained by Monte Carlo simulation (MCS), Q-SPM, and RQ-SPM with the same cardinal number are listed. Note that Q-SPM and RQ-SPM are deterministic point sets, but MCS with the same cardinal number is a random point set. The errors of the mean and standard deviation shown in Table 4.1 for Q-SPM and RQ-SPM are deterministic. However, the errors of the mean and standard deviation for MCS with the same cardinal number is essentially random, thus, shown in Table 4.1 is the mean and coefficient of variation of the relative errors of n results. The COV of the relative error for MCS is computed by

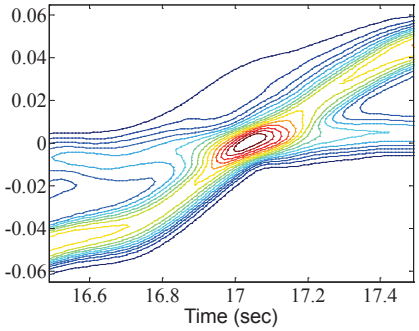
$$\delta = \frac{\sqrt{\frac{1}{n-1}\left\{\sum_{j-1}^{n} e_j^2 - \frac{1}{n}\left(\sum_{j-1}^{n} e_j\right)^2\right\}}}{\frac{1}{n}\sum_{j=1}^{n} e_j}, \tag{4.4}$$

where e_j's are the errors $e_{\|\mu\|}$ or $e_{\|\sigma\|}$ in the j-th round of MCS. In the present paper we take $n = 10$. Likewise, GF-discrepancies of MCS shown in Table 4.1 are the mean values. It is seen from Table 4.1 that the COVs of errors of the mean and standard deviation of MCS are usually as high as between 0.4 and 0.75, showing large

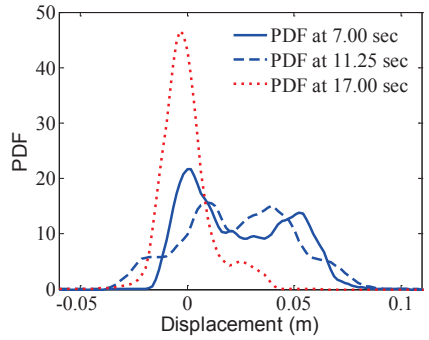
variability. To be clearer, we also listed the 95 % quantile errors of MCS, which are the mean plus 1.645 times standard deviation. It is observed clearly that, for each case, the GF-discrepancy of RQ-SPM is considerably smaller than those of Q-SPM and MCS, and simultaneously the RQ-SPM have considerably lower errors compared to the Q-SPM and MCS with the same number of points. In addition, it is noticeable from this table that the accuracy of Q-SPM is rather low although it has algebraic accuracy of 5th degree. This make it clear that the degree of nonlinearity involved in these stochastic response analyses is far beyond polynomials of 5th degree, which is also why there is room of improving accuracy by rotation of Q-SPM. The above



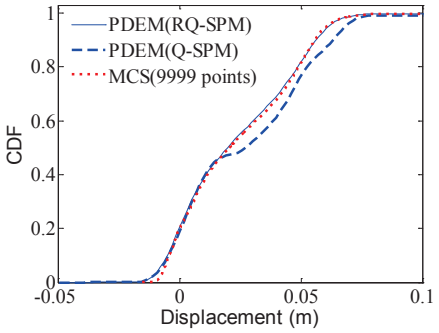
(a) The PDF evolving against time



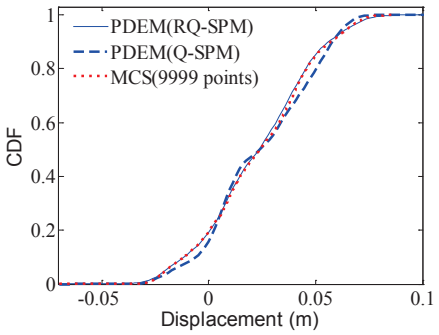
(b) The contour to the PDF surface



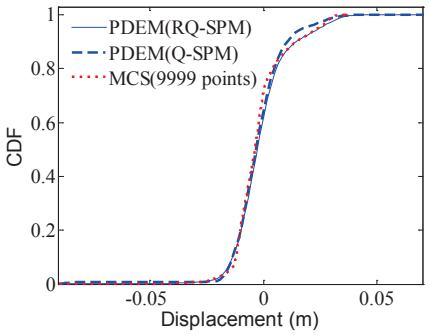
(c) PDFs at typical instants of time



(d) CDFs at 7.00 s

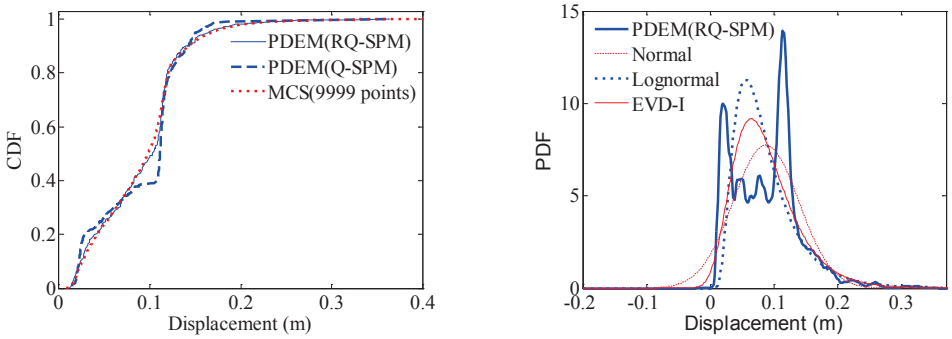


(e) CDFs at 11.25 s



(f) CDFs at 17.00 s

FIG. 4.7. *The probabilistic information of the first inter-story drift.*



(a) CDF (b) PDF

FIG. 4.8. *The CDF and PDF of extreme values of first inter-story drift ($T = 20s$).*

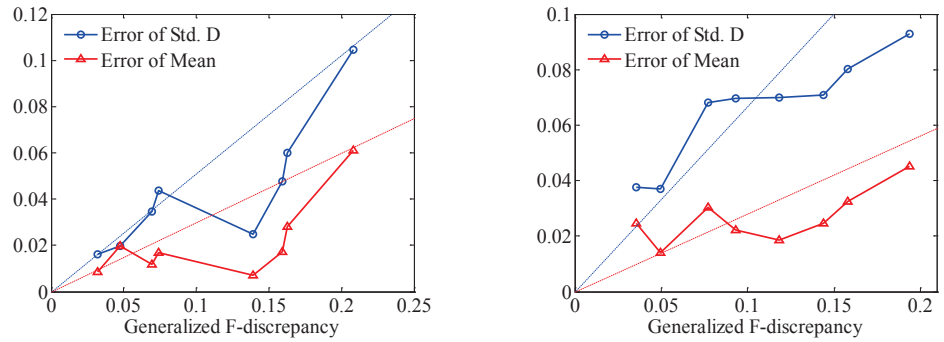
examples enhance our confidence that GF-discrepancy works well even for systems depending on random parameters in a very complex nonlinearity.

We now come back to the initial 20-dimensional example. The instantaneous PDFs and typical PDFs at some specified instants of time could be abtained by PDEM (RQ-SPM) [14, 15]. The instantaneous PDFs over the interval [16.5, 17.5]s are shown in Fig. 4.7(a). Viewing Fig. 4.7(a) together with Fig. 4.7(b), we can see more clearly that the two peaks (two modes) of the PDF at 16.5s are merged at about 17s and then separate again from each other at 17.5s. PDFs at three typical instants of time are shown in Fig. 4.7(c) where highly irregular distribution of PDF is found at 7s and 11.25s. The CDFs by different point sets at these time instants are shown in Figs. 4.7(d), (e) and (f), from which it is observed that the results obtained by RQ-SPM and MCS are in good agreement.

The CDFs and PDFs of the extreme value of the displacement of the first floor involves reliability information of the system and could be obtained also by PDEM [4]. The CDFs over the time interval [0, 20]s by different methods are pictured in Fig. 4.8(a). Here, good accordance is observed between RQ-SPM and MCS whereas Q-SPM deviates from the other two methods again. The PDF of the maximum displacement of the first floor is shown in Fig. 4.8(b), where the feature of the two peaks in the EVD cannot be captured by the other regular distributions as shown in the figure.

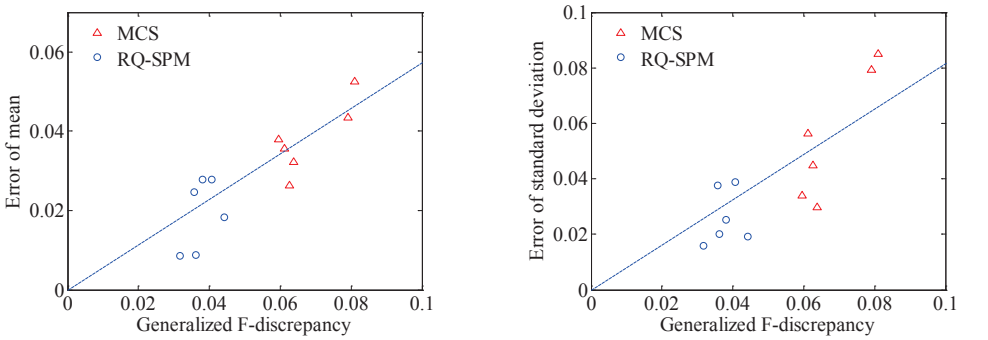
For the same 9-story structure involving 20 basic random variables, the GF-discrepancy of different point sets v.s. the relative error is shown in Fig. 4.9(a). A trend of error reduction against the reduction of GF-discrepancy is clear. However, the relationship is not rigorously monotonic, because the error is not only related to the GF-discrepancy, but also related to the total variation of the response surface (cf. inequalities 2.12 and 2.29, i.e., the sensitivity in terms of the basic random variables. But if the order of magnitude of GF-discrepancy is reduced, it is almost sure that the order of magnitude of the errors will be reduced. The above observations could also be verified in Fig. 4.10(b) for the 7-story structure involving 15 basic random variables as studied in Table 4.1.

Pictured in Fig. 4.10 are all the data in Table 4.1, for the problem of different dimensions and for different methods (MCS and RQ-SPM) where the two lines are fitted by the method of least squares. It is seen from these two figures that the trend of the reduction of error against the reduction of GF-discrepancy is still very clear. In a



(a) 20 dimension (b) 15 dimension

FIG. 4.9. *The relative error v.s. GF-discrepancy.*



(a) The error of mean (b) The error of the standard deviation

FIG. 4.10. *The relative error of response v.s. GF-discrepancy.*

sense, this gives a strong demonstration, though not rigorous proof, of the correctness of inequality 2.29, which does not depend on the dimension of the problem, depending only on the GF-discrepancy and the total variation of the integrand (response) in the sense of Hardy and Krause. The GF-discrepancy, as an important index of optimally select representative points, is thus strongly justified.

5. Concluding remarks. Smart selection of point sets in high-dimensional cubature formulae is of paramount importance and is being extensively studied. In the present paper, the concept of extended F-discrepancy (EF-discrepancy) and generalized F-discrepancy (GF-discrepancy) of a point set is introduced and justified by the comparison with other existing discrepancies. The extensions of Koksma-Hlawka inequality for EF-discrepancy are proved. Quantitative relationship between EF- and GF-discrepancy is studied. As an application, the quasi-symmetric point method is employed and the rotation transforms are performed to minimize the GF-discrepancy. It is rigorously proved that the rotation transform will keep the degree of algebraic accuracy. A genetic algorithm is adopted to solve the optimization problem of finding optimal rotation angles with GF-discrepancy as the objective function. Numerical examples are studied, demonstrating that: (1) The GF-discrepancy can be adopted as an effective index for judging the goodness of a point set; (2) The error of both mean and standard deviation decreases following a clear trend with the reduction

of GF-discrepancy; and (3) The accuracy of the rotational Q-SPM is improved in comparison to the original Q-SPM. The numerical examples also show that despite possessing algebraic accuracy of 5th degree, Q-SPM performs poor for stochastic response analysis of nonlinear systems but RQ-SPM with lower GF-discrepancy works well, which indicates that GF-discrepancy could be applied to problems depending on random parameters in a strong nonlinear way far beyond polynomials of 5th degree.

It is expected that the GF-discrepancy would also be valuable in other uncertainty quantification applications. Problems to be further studied include:

1. Rigorous proof of the error bound by GF-discrepancy (inequality 2.29) and the quantitative relationship between EF- and GF-discrepancy (inequality 2.35);

2. To generalize the idea of improving accuracy by reducing GF-discrepancy to other cubature formulae. Actually, based on the elaboration in the paper, GF-discrepancy itself is essential in measure the goodness of a point set. Thus, there is a possibility of generating good point set independent of existing cubature formulae by minimizing GF-discrepancy;

3. The adaptivity of RQ-SPM to non-normal random-variate spaces. In this aspect, the proof of the extended Koksma-Hlawka inequality for general non-uniform distribution is possible and needs to be explored. To this end, GF-discrepancy will play an important role in point selection beyond Q-SPM for general non-uniform, non-Gaussian distributions; and

4. The degree of nonlinearity of the problems that GF-discrepancy and cubature rules could perform well. As demonstrated by examples, GF-discrepancy works well for the problems depending on random parameters in a nonlinear way far beyond polynomials of 5th degree, where Q-SPM performs poor, because what is really important for the applicability of GF-discrepancy is related to the total variation of the integrand. To obtain a rational criterion that could guarantee the application of GF-discrepancy is a very important open problem.

Acknowledgments. The authors are grateful to the supports of the National Natural Science Foundation of China (Grant No.11172210), the Fundamental Fund for Central Universities, the Shuguang Plan of Shanghai City and the State Key Laboratory of Disaster Reduction in Civil Engineering (Grant Nos. SLDRCE08-A-01 and SLDRCE10-B-02). Prof. Jie Li at Tongji University is appreciated for his constructive suggestions and discussions. Mr. Jun Xu is also appreciated for his help in the studies.

6. Appendix: Proof of extended Koksma-Hlawka inequality in higher dimension ($s \geq 3$). Proof of Theorem 2.2 (extended Koksma-Hlawka inequality) in higher dimensions for $s \geq 3$ is similar to the case $s = 2$.

Firstly, according to Theorem 5.1 in [12], any bounded function could be represented as a difference of two generalized monotonic functions. Thus, we have, analogous to 2.18,

$$f(x_1, x_2, \cdots, x_s) = f_1(x_1, x_2, \cdots, x_s) - f_2(x_1, x_2, \cdots, x_s), \tag{6.1}$$

where $f_1(x_1, x_2, \cdots, x_s)$, $f_2(x_1, x_2, \cdots, x_s)$ are two generalized monotonic functions. Because $f_1(x_1, x_2, \cdots, x_s)$ is a generalized monotonic (non-decreasing) function, in the hyper-rectangle area $\prod_{j=1}^{s} \left(\frac{i_j - 1}{q} \leq x_j < \frac{i_j}{q} \right)$, $f_1(x_1, x_2, \cdots, x_s)$ is no greater than

$f_1\left(\frac{i_1}{q}, \frac{i_2}{q}, \cdots, \frac{i_s}{q}\right)$. Thereby we have

$$
\begin{aligned}
S_1 &= \sum_{k=1}^{n} P_k f_1\left(x_1(k), x_2(k), \cdots, x_s(k)\right) \\
&\le \sum_{i_1=1}^{q} \cdots \sum_{i_s=1}^{q} \left(\sum_{\frac{i_1-1}{q} \le x_1(k) < \frac{i_1}{q}, \cdots, \frac{i_s-1}{q} \le x_s(k) < \frac{i_s}{q}} P_k \right) f_1\left(\frac{i_1}{q}, \frac{i_2}{q}, \cdots, \frac{i_s}{q}\right), \qquad (6.2) \\
&= \sum_{j=1}^{s-1} R_j + f_1(1, 1, \cdots, 1)
\end{aligned}
$$

where $R_1, \cdots, R_s$ are summations given by

$$
\begin{aligned}
R_1 = &\sum_{j=1}^{s} \sum_{i_j=1}^{q-1} \left(\sum_{\mathbf{x}_k < (1, \cdots, 1, \frac{i_s}{q}, 1, \cdots, 1)} P_k \right) \\
&\times \left[f_1\left(1, \cdots, 1, \frac{i_s}{q}, 1, \cdots, 1\right) - f_1\left(1, \cdots, 1, \frac{i_s+1}{q}, 1, \cdots, 1\right) \right], \qquad (6.3)
\end{aligned}
$$

$$
\begin{aligned}
R_2 = &\sum_{1 \le j < k \le s} \left\{ \sum_{i_j=1}^{q-1} \sum_{i_k=1}^{q-1} \left(\sum_{\mathbf{x}_k < (1, \frac{i_j}{q}, 1, \cdots, 1, \frac{i_k}{q}, \cdots, 1)} P_k \right) \right. \\
&\left. \times \begin{bmatrix} f_1\left(1, \frac{i_j}{q}, 1, \cdots, 1, \frac{i_k}{q}, \cdots, 1\right) \\ -f_1\left(1, \frac{i_j+1}{q}, 1, \cdots, 1, \frac{i_k}{q}, \cdots, 1\right) \\ -f_1\left(1, \frac{i_j}{q}, 1, \cdots, 1, \frac{i_k+1}{q}, \cdots, 1\right) \\ +f_1\left(1, \frac{i_j+1}{q}, 1, \cdots, 1, \frac{i_k+1}{q}, \cdots, 1\right) \end{bmatrix} \right\}, \qquad (6.4)
\end{aligned}
$$

$$
\begin{aligned}
R_3 = &\sum_{1 \le j < k < l \le s} \sum_{i_j=1}^{q-1} \sum_{i_k=1}^{q-1} \sum_{i_l=1}^{q-1} \left(\sum_{\mathbf{x}_k < (1, \cdots, \frac{i_j}{q}, \cdots, \frac{i_k}{q}, \cdots, \frac{i_l}{q}, \cdots, 1)} P_k \right) \\
&\times \begin{bmatrix} f_1\left(1, \cdots, \frac{i_j}{q}, \cdots, \frac{i_k}{q}, \cdots, \frac{i_l}{q}, \cdots, 1\right) \\ -f_1\left(1, \cdots, \frac{i_j+1}{q}, \cdots, \frac{i_k}{q}, \cdots, \frac{i_l}{q}, \cdots, 1\right) \\ +f_1\left(1, \cdots, \frac{i_j+1}{q}, \cdots, \frac{i_k+1}{q}, \cdots, \frac{i_l}{q}, \cdots, 1\right) \\ -f_1\left(1, \cdots, \frac{i_j}{q}, \cdots, \frac{i_k+1}{q}, \cdots, \frac{i_l}{q}, \cdots, 1\right) \\ +f_1\left(1, \cdots, \frac{i_j+1}{q}, \cdots, \frac{i_k}{q}, \cdots, \frac{i_l+1}{q}, \cdots, 1\right) \\ -f_1\left(1, \cdots, \frac{i_j}{q}, \cdots, \frac{i_k}{q}, \cdots, \frac{i_l+1}{q}, \cdots, 1\right) \\ +f_1\left(1, \cdots, \frac{i_j}{q}, \cdots, \frac{i_k+1}{q}, \cdots, \frac{i_l+1}{q}, \cdots, 1\right) \\ -f_1\left(1, \cdots, \frac{i_j+1}{q}, \cdots, \frac{i_k+1}{q}, \cdots, \frac{i_l+1}{q}, \cdots, 1\right) \end{bmatrix}. \qquad (6.5)
\end{aligned}
$$

Actually, for any $1 \leq \mu \leq s$, we have

$$R_\mu = \sum_{1\leq j_1<j_2<\cdots<j_\mu\leq s} \sum_{i_{j_1}=1}^{q-1}\sum_{i_{j_2}=1}^{q-1}\cdots\sum_{i_{j_\mu}=1}^{q-1}\left[\left(\sum_{\mathbf{x}_k<(1,\cdots,\frac{i_{j_1}}{q},\cdots,\frac{i_{j_2}}{q},\cdots,\frac{i_{j_\mu}}{q},\cdots,1)} P_k\right)\right. \tag{6.6}$$
$$\left.\times \sum_{I_{j_1},I_{j_2},\cdots,I_{j_\mu}=0}^{1} G(I_{j_1},I_{j_2},\cdots,I_{j_\mu})\right].$$

in which

$$G(I_{j_1},I_{j_2},\cdots,I_{j_\mu}) = (-1)^{I_{j_1}+I_{j_2}+\cdots+I_{j_\mu}} \tag{6.7}$$
$$\times f_1\left(1,\cdots,\tfrac{i_{j_1}+I_{j_1}}{q},\cdots,\tfrac{i_{j_2}+I_{j_2}}{q},\cdots,\tfrac{i_{j_\mu}+I_{j_\mu}}{q},\cdots,1\right),$$

where I_{j_r}, $r = 1, 2, \cdots, \mu$ could take value only either 0 or 1. In 6.6, it is seen that in the most inner summation symbol 2^μ terms are involved, while in the most outside summation symbol $C_s^\mu = \frac{s!}{\mu!(s-\mu)!}$ terms are involved. Clearly, 6.3-6.5 are the special cases of 6.6 as μ=1, 2 and 3, respectively. Note that the entity in $[\cdot]$ in 6.4 is not a vector but a pure quantity, so is for 6.5. To proceed, we note that the empirical CDF $F_n(x_1, x_2, \cdots, x_n)$, and according to the definition of EF-discrepancy in 2.5 and 2.8, there is, analogous to 2.21,

$$F_n(x_1,x_2,\cdots,x_s) = \sum_{q=1}^{N} P_q I\{\mathbf{x}_q \leq \mathbf{x}\} = x_1x_2\cdots x_s + \vartheta D_{EF}(n), \tag{6.8}$$

where $|\vartheta(x_1, x_2, \cdots, x_n)| \leq 1$. Replacing $\sum_{\mathbf{x}_k<(x_1,x_2,\cdots,x_s)} P_k$ by $F_n(x_1, x_2, \cdots, x_s)$ and Substituting 6.8 in 6.2 yields, analogous to 2.22,

$$S_1 \leq \frac{1}{q^s}\sum_{i_1=1}^{q}\cdots\sum_{i_s=1}^{q} f_1\left(\tfrac{i_1}{q},\cdots,\tfrac{i_s}{q}\right) + \mathrm{TV}(f_1)D_{EF}(n), \tag{6.9}$$

where $\mathrm{TV}(f_1)$ is the discretized version of the total variation in the sense of Hardy and Krause defined in 2.3 for higher-dimension.

According to 6.9, Letting $q \to \infty$ yields

$$S_1 \leq \int_0^1\int_0^1 f_1(x_1,x_2,\cdots,x_s)dx_1dx_2\cdots dx_s + D_{EF}(n)\mathrm{TV}(f_1). \tag{6.10}$$

The next steps follow 2.24 through 2.28 in a trivial way and will not be repeated here. The above derivation completes the proof of Theorem 2.2 for any dimensional case $s \geq 1$.

REFERENCES

[1] A. H-S. Ang and W. H. Tang, *Probability Concepts in Engineering Planning and Design (Vol. 2)*, John Wiley & Sons, 1984.

[2] J. B. Chen, R. Ghanem, and J. Li, *Partition of the probability-assigned space in probability density evolution analysis of nonlinear stochastic structures*, Probabilist. Eng. Mech., 24 (2009), pp. 27–42.

[3] J. B. Chen, W. L. Sun, J. Li, and J. Xu, *Stochastic harmonic function representation of stochastic process*, J. Appl. Mech., 80 (2013), pp. 011001-1-11.
[4] J. B. Chen and J. Li, *The extreme value distribution and dynamic reliability analysis of nonlinear structures with uncertain parameters*, Struct. Saf., 29 (2007), pp. 77–93.
[5] J. B. Chen and J. Li, *Strategy for selecting representative points via tangent spheres in the probability density evolution method*, Int. J. Numer. Meth. Eng., 74 (2008), pp 1988–2014.
[6] J. H. Conway and N. J. A. Sloan, *Sphere Packings, Lattices and Groups*, Third ed., Springer, 1999.
[7] J. Dick and F. Pillichshammer, *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, 2010.
[8] H. Engels, *Numerical Quadrature and Cubature*, Academic Press, 1980.
[9] K. T. Fang, Y. Wang, *Number-theoretic Methods in Statistics*, Chapman & Hall, London, 1994.
[10] R. L. Haupt, S. E. Haupt, *Practical Genetic Algorithms*, Second ed., John Wiley & Sons, 2004.
[11] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, 1975.
[12] L. K. Hua and Y. Wang, *Applications of Number Theory to Numerical Analysis*, Springer-Verlag, Berlin, 1981.
[13] P. L'Ecuyer, *Quasi-Monte Carlo methods with applications in finance*, Financ. Stoch., 13 (2009), pp. 307–349.
[14] J. Li and J. B. Chen , *Stochastic Dynamics of Structures*, John Wiley & Sons, 2009.
[15] J. Li, J. B. Chen, W. L. Sun, and Y. B. Peng, *Advances of probability density evolution method for nonlinear stochastic systems*, Probabilist. Eng. Mech., 28 (2012), pp. 132–142.
[16] F. Ma, H. Zhang, A. Bockstedte, G. C. Foliente, and P. Paevere, *Parameter analysis of the differential model of hysteresis*, J. Appl. Mech., 71 (2004), pp. 342–349.
[17] H. Niederreiter, *Random number generation and Quasi-Monte Carlo methods*, SIAM, 1992.
[18] D. L. Peter, *Linear Algebra and Its Application*, Second ed., John Wiley & Sons, 2007.
[19] P. D. Proinov, *Generalization of two results of the theory of uniform distribution*, Proc. Amer. Math. Soc., 95 (1985), pp. 527–534.
[20] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*, John Wiley & Sons, New York, 1981.
[21] F. Shi, H. Wang, L. Yu, and F. Hu, *30 Case Studies of Intelligent Algorithm Based on MATLAB*, Beihang University Press, Beijing, 2011. (in Chinese).
[22] S. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Soviet Math. Dokl., 4 (1963), pp. 240–243.
[23] V. Victoir, *Asymmetric cubature formulae with few points in high dimension for symmetric measures*, SIAM J. Numer. Anal., 42 (2004), pp. 209–227.
[24] Y. K. Wen, *Method for random vibration of hysteretic systems*, J. Eng. Mech., 102 (1976), pp. 249–263.
[25] H. P. William, P. F. Brain, and A. T. San, *Numerical Recipes: the Art of Scientific Computing*, Cambridge University Press, London, 1986.
[26] D. B. Xiu, *Fast numerical methods for stochastic computations*, Commun. Comput. Phys., 5 (2009), pp. 242–272.
[27] J. Xu, J. B. Chen, and J. Li, *Probability density evolution analysis of engineering structures via cubature points*, Comput. Mech, 50 (2012), pp. 135–156.