

Analysis and Applications
 © World Scientific Publishing Company

On the Probability Estimates of Quadrature Rules from Uniformly Sampled Points on Spheres

Ruigang Zheng

*School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, China
 ruigzheng2-c@my.cityu.edu.hk*

Xiaosheng Zhuang*

*Department of Mathematics, City University of Hong Kong, Hong Kong, China
 xzhuang7@cityu.edu.hk*

In this paper, we provide detailed discussions on the probability estimates of quadrature rules from uniformly sampled points on spheres. Besides gathering relevant lemmas in the literature to derive probability estimates on the existence of exact quadrature rules for spherical harmonics, we provide additional estimates with finer characterizations based on probabilistic quantities related to the measure and the diameter of Voronoi cells. Specifically, our estimates provide additional affirmative answers to certain relations between the number of sampled points and the degree of spherical harmonics. We further investigate the problem of setting the number of points to be of order t^d . Simple analysis based on our estimates suggests that the constant in the order cannot be fixed for all t and should increase as t increases. This is empirically verified in our experiments.

Keywords: d -dimensional sphere; uniform distribution; Voronoi cell; exact quadrature rule; probability estimate; spherical designs; nested structure.

Mathematics Subject Classification 2020: 60D05, 65D30, 65D32

1. Introduction and motivation

Various methods for function approximation on d -dimensional unit sphere $\mathbb{S}^d$ such as hyperinterpolation [34], multi-scale analysis [13], localized systems [22, 31], sketching for noisy data fitting [25], and spherical framelets [23, 35], are based on exact quadrature rules for spherical harmonics. Apart from the numerical applications, recent works in machine learning on spheres (see e.g., [9]), which generalize analysis on Euclidean spaces (see e.g., [27]), are also deeply related to exact quadrature rules on spheres. Intuitively, points in quadrature rules serve as sampling locations. However, unlike 1- and 2-dimensional signals defined on regular grids, commonly used quadrature rules on the d -dimensional unit sphere $\mathbb{S}^d := \{\mathbf{x} \in \mathbb{R}^{d+1} : |\mathbf{x}| = 1\}$, e.g., the Gauss–Legendre tensor product rule [16], generally do not support the direct up- and down-sampling operations [7, 15, 26]. This is due to the non-nested

*Corresponding author.

property of the quadrature rules, i.e., the points in the quadrature rule for low-degree spherical harmonics are not contained in the quadrature rule for high-degree spherical harmonics. More precisely, for n numbers $w_1, \dots, w_n \in \mathbb{R}$ and n points $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{S}^d$, if

$$\int_{\mathbb{S}^d} P(\mathbf{x}) d\mu_d(\mathbf{x}) = \sum_{i=1}^n w_i P(\mathbf{x}_i),$$

holds for any $(d+1)$ -variate polynomial P (equivalent to any spherical harmonic) with total degree up to $t \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, where μ_d is the uniform probability measure on $\mathbb{S}^d$, then we call the finite sequence $\mathcal{X}_t := ((w_i, \mathbf{x}_i))_{i=1}^n$ a *t-exact quadrature rule* (for $\mathbb{S}^d$), in which the numbers $w_1, \dots, w_n$ are called *quadrature weights* and the points $\mathbf{x}_1, \dots, \mathbf{x}_n$ are called *quadrature nodes*. Moreover, if two finite sequences $\mathcal{X}_{t_1} = ((w_i, \mathbf{x}_i))_{i=1}^n$ and $\mathcal{X}_{t_2} = ((\tilde{w}_i, \mathbf{x}_i))_{i=1}^{n+m}$ (i.e., $(\mathbf{x}_i)_{i=1}^n$ is a subsequence of $(\mathbf{x}_i)_{i=1}^{n+m}$), are t_1 -exact and t_2 -exact ($t_1 < t_2$), respectively, then the two quadrature rules are said to be *nested*. One can have a sequence $\mathcal{X}_{t_1}, \mathcal{X}_{t_2}, \dots, \mathcal{X}_{t_k}, \dots$ of *nested quadrature rules*, similarly, if $\mathcal{X}_{t_i}$ and $\mathcal{X}_{t_{i+1}}$ are nested for $i = 1, 2, \dots$.

In practice, when approximating functions using spherical harmonics from low degree to high degree, e.g., in numerical integration, multiple quadrature rules have to be adopted. The computation and storage burden are directly related to the total number of nodes in such quadrature rules. If the quadrature rules are *non-nested*, then the computation and storage burden could be quite large, not to mention the inconvenience in the down- and up-sampling operations for spherical signal processing. Therefore, it is desirable to have nested quadrature rules for which the number of points is as small as possible.

One special case of the *t*-exact quadrature rules is the so-called spherical *t*-designs [1], where each of the n points $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{S}^d$ has the same (quadrature) weight $w_i \equiv \frac{1}{n}$. Since the weights are the same, one only needs to consider the storage of the points. The well-known conjecture regarding the optimal order of spherical *t*-designs is resolved in [2], which proves that for any $t \in \mathbb{N}_0$, there exists a spherical *t*-design such that its number of nodes is $C_d t^d$ and the constant C_d depends only on d . Hence, spherical *t*-designs are nice quadrature rules having a simple quadrature weight with the optimal number of nodes (in a certain sense). It is thus desirable to consider spherical designs with the extra desirable property of nesting. Our recent work [37] addresses such nested spherical designs and proves several fundamental results on their existence and estimates. To elaborate, given a spherical t_1 -design with points $\mathbf{x}_1, \dots, \mathbf{x}_n$, it is confirmed that one can append extra points $\mathbf{y}_1, \dots, \mathbf{y}_m$ such that all the points form a spherical *t*-design ($t > t_1$). Though spherical designs enjoy many nice properties, it is, however, rather time-consuming to obtain spherical *t*-designs with large degree *t*. It is shown in [36] that for $\mathbb{S}^2$ and $t > 1000$ it takes months to complete the corresponding numerical optimizations. Moreover, in [37], we proved that the spherical *t*-design that extends a spherical t_1 -design could be of order t^{2d+1} . Thus, in terms of the number of points, nested spherical designs in general could be too large. For practical considerations, we need

to consider other types of quadrature rules that allow nested structure with easy derivation, efficient computation, and effective applications.

Other than the spherical t -designs that are with fixed constant weights, one simple and practical way is to consider the derivation of t -exact quadrature rules from uniformly sampled points on $\mathbb{S}^d$ with *nearly* constant weights. In fact, by the law of large numbers, given n points $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n \in \mathbb{S}^d$ sampled from the uniform distribution μ_d , we have (with n sufficiently large)

$$\int_{\mathbb{S}^d} P(\mathbf{x}) d\mu_d(\mathbf{x}) \approx \frac{1}{n} \sum_{i=1}^n P(\tilde{\mathbf{x}}_i).$$

It is thus very natural to expect that one can obtain a quadrature rule by slightly adjusting the weight $\frac{1}{n}$ for each point. Most importantly, by continuing the uniform sampling process, it is straightforward to obtain sequences of uniformly sampled points that are nested. Nonetheless, compared to the spherical t -designs, the weights here still require extra storage.

Since there is no obstacle in obtaining nested sequences of uniformly sampled points, the problem of finding nested quadrature rules is reduced to showing the existence of quadrature weights to form a quadrature rule from a given uniformly sampled point set. On the other hand, due to the presence of noise in spherical signal processing, it is also important to consider quadrature rules with bounded weights. Thus, we are interested in answering the following problems:

- Q1) Given n uniformly sampled points on the sphere, what is the probability (lower bound) to obtain a t -exact quadrature rule with bounded weights?
- Q2) Suppose that for some constant $C_d > 0$, it is enough with $C_d t_1^d$ number of uniformly sampled points on the sphere to obtain a t_1 -exact quadrature rule with bounded weights. Then, for any $t > t_1$, are $C_d t^d$ uniformly sampled points enough to guarantee the existence of a t -exact quadrature rule with high probability?
- Q3) If the relation $n = C_d t^d$ is not sufficient to guarantee the existence of a t -exact quadrature rule with high probability, is there any other relation between the number of sampled points n and the degree t that does?

There are three closely related works [10, 14, 22], in which results related to the M-Z inequalities or exact quadrature rules on uniformly sampled points are given. In [22], a method for finding quadrature rules on $\mathbb{S}^d$ is proposed. Given a set $\{\mathbf{x}_i\}_{i=1}^n$ of points on the sphere, finding the quadrature weights $\{a_i\}_{i=1}^n$ can be reduced to solve a system of linear equations that involves a positive definite Gram matrix, which is induced by an alternative inner-product for the spherical harmonics. An estimate on the numbers of required points is also given in [22], which is of order $O(t^6 \log t)$ for $\mathbb{S}^2$ (though the numerical experiments in [22] shown that at most $4(t+1)^2$ points are needed with $t \leq 140$). Moreover, on such a point set, the computed quadrature weights have high probability to be of order $O(\frac{1}{n})$. However, since the points are random, there should also be an estimate on the probability of

such a point sequence satisfying the M–Z inequalities, which is missing in [22]. On the other hand, in [14], by utilizing numerical optimizations, the empirical study shows successful results for t -exact quadrature rules on $\mathbb{S}^2$ with the number of uniformly sampled points of order $(t+1)^2$. However, as opposed to [22], there is no theoretical guarantee on the number of required points. Finally, in [10], it is shown that for sufficiently large n , with probability at least $1 - \frac{1}{n}$, the sampled points satisfy the M–Z inequalities. Therefore, the existence of exact quadrature rules with the same probability estimate easily follows.

To answer question Q1, we first summarize relevant results in the current literature and provide estimates that are straightforward to obtain by following our summarized framework. In detail, our summarized framework consists of the implication from the deterministic statements of the Marcinkiewicz–Zygmund (M–Z) inequalities to the existence of exact quadrature rules [12, 28, 29, 31], and the implication from probabilistic quantities to the M–Z inequalities. We also give additional probability estimates based on the probabilistic quantities related to the measure and diameter of Voronoi cells induced by the sampled points, for which we follow [8] to obtain generalized results on $\mathbb{S}^d$. Based on these quantities, we obtain a more specific probability estimate of satisfying the M–Z inequalities. Consequently, we obtain additional estimates on the probability lower bound of the existence of t -exact quadrature rules with finer characterizations. In numerical applications, our results together with those derived from the literature provide upper bounds on the number of points that allow finding t -exact quadrature rules with high probability.

On the other hand, the question Q2 is motivated by the optimal order of spherical t -designs. In practice, to obtain nested rules of order t^d via uniformly sampled points, we may first obtain a t_1 -exact quadrature rule with $C_d t_1^d$ points, and the constant C_d is as small as possible. Then, after sampling extra points, we can numerically check whether it is possible to find a t -exact rule ($t > t_1$) with $C_d t^d$ points. By simply taking the limit with respect to t in our estimates, the answer to Q2 is possibly negative since the lower bound diverges to $-\infty$ and we thus lose the probability guarantee. Moreover, a smaller constant will speed up this process. Therefore, it is suggested that the aforementioned procedure will not be successful. This is empirically tested in numerical experiments in which setting a smaller constant leads to drastic changes in the time consumed and outcomes of optimizations for relatively large t . Therefore, when dealing with very large t , instead of setting a fixed constant C_d , one might have to increase the number C in $n = C t^d$ as t grows to ensure an efficient optimization with acceptable results.

Finally, question Q3 aims to find other possible relations between the number of sampled points n and the degree t of spherical harmonics. Based on our estimates, we show that both relations

$$\frac{n}{\ln n} = C_d t^d, \quad n = C_d t^d \ln t$$

guarantee the existence of exact quadrature rules with high probability, which is consistent with the results derived from the literature. Moreover, as we show later,

our estimates give asymptotically higher probability under such relations.

The paper is organized as follows. In Section 2, we state our main results together with the estimates derived from the literature. In Section 3, numerical results of finding quadrature rules on $\mathbb{S}^2$ are demonstrated. The lemmas, theorems, and proofs of our main result are postponed to Section 4. Finally, we give conclusions and further remarks in Section 5.

2. Main Results

In this section, we introduce our main results on the measure and the diameter of Voronoi cell with respect to a sequence of uniform sampled points on the sphere as well as the existence of quadrature weights for such a sequence. Proofs of our main results will be postponed to Section 4.

The usual surface measure m_d on $\mathbb{S}^d$ satisfies

$$\omega_d := \int_{\mathbb{S}^d} dm_d = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)},$$

where Γ is the Gamma function, e.g., $\omega_2 = 4\pi$. For the uniform distribution μ_d on $\mathbb{S}^d$, we have $\mu_d(A) = \omega_d^{-1}m_d(A)$ for a measurable set A on $\mathbb{S}^d$. In the following, unless explicitly stated, we assume that all functions are real-valued. The L^p -norm of a measurable function on $\mathbb{S}^d$ is defined with respect to the measure m_d : $\|\cdot\|_p := \|\cdot\|_{p,m_d} := (\int_{\mathbb{S}^d} |\cdot|^p dm_d)^{1/p}$ for $1 \leq p < \infty$ and $\|\cdot\|_\infty$ is the essential supremum. The space $L^p(\mathbb{S}^d, m_d)$ is then the usual L_p -space with respect to such a L^p -norm.

The Laplace-Beltrami operator on $\mathbb{S}^d$ is defined as:

$$\Delta f := -\operatorname{div}(\nabla f)$$

where $f \in C^\infty(\mathbb{S}^d)$ and $\operatorname{div}, \nabla$ are the divergence and function gradient on $\mathbb{S}^d$. The eigenfunctions of Δ , or so called spherical harmonics, are the restrictions of real-valued homogeneous harmonic polynomials on $\mathbb{S}^d$. These eigenfunctions form an orthonormal basis for $L^2(\mathbb{S}^d, m_d)$. Its distinct eigenvalues are $\lambda_\ell^* := \ell(\ell + d - 1)$, $\ell \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, where ℓ corresponds to the degree of the polynomial. For each λ_ℓ^* , the dimension of the associated eigenspace $\mathbf{H}_\ell^d$ is

$$\dim_\ell := \begin{cases} \frac{2\ell + d - 1}{\ell + d - 1} \binom{\ell + d - 1}{\ell} & \text{if } \ell \geq 1, \\ 1 & \text{if } \ell = 0. \end{cases}$$

We denote the sorted eigenvalues counted with multiplicities as λ_ℓ and an orthonormal basis in $\mathbf{H}_\ell^d$ as $\{Y_{\ell,k} : \ell \in \mathbb{N}_0, k = 1, \dots, \dim_\ell\}$. The space Π_t^d of polynomials on the sphere $\mathbb{S}^d$ up to degree t is then given by the direct sum of $\mathbf{H}_\ell^d$, $\ell \leq t$, i.e., $\Pi_t^d := \bigoplus_{\ell=0}^t \mathbf{H}_\ell^d$. We denote the dimension of Π_t^d as $N_t = \sum_{\ell=0}^t \dim_\ell$. The N_t spherical harmonics that form an orthonormal basis for Π_t^d can be sorted as $\{y_i : i = 1, \dots, N_t\}$. For more details, we refer to, e.g., [6, 30].

The geodesic distance of two points $\mathbf{x}, \mathbf{y}$ on $\mathbb{S}^d$ is $\rho(\mathbf{x}, \mathbf{y}) := \arccos(\mathbf{x} \cdot \mathbf{y})$. Denote $B_{\mathbf{x},r} := \{\mathbf{y} \in \mathbb{S}^d : \rho(\mathbf{x}, \mathbf{y}) \leq r\}$ as the spherical cap with radius $r > 0$ and center $\mathbf{x} \in \mathbb{S}^d$. Given n points $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{S}^d$, define the *Voronoi tessellation* $\mathbf{m}_{n,d}$ as $\mathbf{m}_{n,d} := \{\mathcal{C}_{1,n}, \dots, \mathcal{C}_{n,n}\}$, where

$$\mathcal{C}_{i,n} := \left\{ \mathbf{x} \in \mathbb{S}^d \mid \rho(\mathbf{x}, \mathbf{x}_i) \leq \min_{j=\{1,\dots,n\} \setminus \{i\}} \rho(\mathbf{x}, \mathbf{x}_j) \right\}, \quad i \in \{1, \dots, n\}.$$

As mentioned in [18], the Voronoi cells $\mathcal{C}_{i,n}$, $i = 1, \dots, n$, are spherical polytopes covering $\mathbb{S}^d$. If $\mathbf{x}_1, \dots, \mathbf{x}_n$ are distinct, then each $\mathbf{x}_i$ is associated with $\mathcal{C}_{i,n}$ and lies in the non-empty and disjoint interior of $\mathcal{C}_{i,n}$. The partition norm of $\mathbf{m}_{n,d}$ is defined as

$$\|\mathcal{R}(\mathbf{m}_{n,d})\| := \max_i \text{diam}(\mathcal{C}_{i,n}), \quad \text{diam}(\mathcal{C}_{i,n}) := \sup_{\mathbf{x}, \mathbf{y} \in \mathcal{C}_{i,n}} \rho(\mathbf{x}, \mathbf{y}).$$

Let $X_1, X_2, \dots, X_n$ be n independent, identically, and uniformly distributed random vectors taking values in $\mathbb{S}^d$. Given a point $\mathbf{x} \in \mathbb{S}^d$ (or X_1), we denote the μ_d -measure of the (random) Voronoi cell of $\mathbf{x}$ (or X_1) with respect to the Voronoi tessellation of $\mathbf{x}$ (or X_1) and $X_2, \dots, X_n$ as $M_n(\mathbf{x})$. In what follows, the expectation $\mathbb{E}[M_n(\mathbf{x})]$ is taken with respect to the joint distribution of $X_2, \dots, X_n$ while $\mathbb{E}[M_n(X_1)]$ is taken with respect to the joint distribution of $X_1, X_2, \dots, X_n$. The expectation $\mathbb{E}[M_n(X_i)]$, $i \in \{1, \dots, n\}$ can be similarly defined, which are also taken with respect to the joint distribution of $X_1, X_2, \dots, X_n$. We have omitted the dependencies for brevity. The variance of $M_n(X_i)$ is denoted as $\text{Var}[M_n(X_i)]$. The following result gives the estimates related to the moment information of $M_n(X_i)$.

Theorem 2.1. *Let $X_1, X_2, \dots, X_n$ be n independent, identically, and uniformly distributed random vectors taking values in $\mathbb{S}^d$, and $\mathcal{C}_{1,n}, \dots, \mathcal{C}_{n,n}$ be the Voronoi cells determined by $X_1, \dots, X_n$. Let $M_n(X_i)$ denote the μ_d -measure of the Voronoi cell associated with X_i for $i = 1, \dots, n$. Then, for all $i \in \{1, \dots, n\}$, we have $\mathbb{E}[M_n(X_i)] = \frac{1}{n}$, $\mathbb{E}[M_n(X_i)^2] < \frac{2}{n^2}$, and $\text{Var}[M_n(X_i)] < \frac{1}{n^2}$. Moreover, for any $c > 1$, with probability at least $1 - \frac{1}{c^2}$, we have*

$$\max_i |M_n(X_i) - \mathbb{E}[M_n(X_i)]| = \max_i \left| M_n(X_i) - \frac{1}{n} \right| < \frac{c}{\sqrt{n}}.$$

We use $D_n(\mathbf{x})$ to denote the diameter of the Voronoi cell associated with $\mathbf{x}$ with respect to the Voronoi tessellation of $\mathbf{x}, X_2, \dots, X_n$. Thus, the probabilities of events of $D_n(\mathbf{x})$, e.g., $\{D_n(\mathbf{x}) > \delta\}$, involve the joint distribution of $X_2, \dots, X_n$. For $D_n(X_1)$, the probabilities of the events of $D_n(X_1)$, e.g., $\{D_n(X_1) > \delta\}$, involve the joint distribution of $X_1, X_2, \dots, X_n$. Similarly, $D_n(X_2), \dots, D_n(X_n)$ can be defined likewise. We use $\mathbb{P}(P)$ to denote the probability of an event P . We have the following result show that the diameters of the Voronoi cells possess an exponential tail, which further implies the probability of the tail event $\{\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta\}$.

Theorem 2.2. *Let $X_1, X_2, \dots, X_n$ be n independent, identically, and uniformly distributed random vectors taking values in $\mathbb{S}^d$, and $\mathcal{C}_{1,n}, \dots, \mathcal{C}_{n,n}$ be the Voronoi*

cells determined by $X_1, \dots, X_n$. Let $D_n(X_i)$ denote the diameter of the Voronoi cell associated with X_i for $i = 1, \dots, n$ and $\|\mathcal{R}(\mathbf{m}_{n,d})\|$ be the partition norm of the Voronoi tessellation $\mathbf{m}_{n,d} = \{\mathcal{C}_{1,n}, \dots, \mathcal{C}_{n,n}\}$. Then, for each $\delta > 0$, there exist two constants $c_{1,d} > 0$ and $c_{2,d,\delta} > 0$ such that

$$\mathbb{P}(\{D_n(X_i) > \delta\}) \leq c_{1,d}(1 - c_{2,d,\delta})^{n-1} \quad \forall i \in \{1, \dots, n\}$$

and

$$\mathbb{P}(\{\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta\}) \leq nc_{1,d}(1 - c_{2,d,\delta})^{n-1}, \quad (2.1)$$

where the constants $c_{1,d}$ depends on d and $c_{2,d,\delta}$ depends on d, δ .

With the above results and the results on the existence of quadrature rules derived from the Marcinkiewicz–Zygmund (M–Z) inequalities (details are given in Section 4.3), we have the following theorem which provides the probability estimate on the existence of t -exact quadrature rules deriving from a finite sequence of uniformly sampled points on the sphere $\mathbb{S}^d$.

Theorem 2.3. *Suppose integers $d \in \mathbb{N}$ and $t \geq 1$. Let $\tau_d := 6(3^{\frac{d}{2}}\pi + 2d + 3)$. Let $X_1, X_2, \dots, X_n$ be n independent, identically, and uniformly distributed random vectors taking values in $\mathbb{S}^d$. Then the following statements hold.*

- (i) *For any $c > 1$ and $\eta \in (0, 1)$ with probability at least $1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1} - \frac{1}{c^2}$, there exist quadrature weights $a_1, \dots, a_n \in \mathbb{R}$ such that $((a_i, X_i))_{i=1}^n$ forms a t -exact quadrature rule satisfying*

$$|a_i| \leq \frac{c\sqrt{n} + 1}{n}(1 - \eta)^{-1}, \quad i \in \{1, \dots, n\}. \quad (2.2)$$

- (ii) *For any $\eta \in (0, \frac{1}{2})$, with probability at least $1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1}$, there exist quadrature weights $b_1, \dots, b_n \in \mathbb{R}$ such that $((b_i, X_i))_{i=1}^n$ is a t -exact quadrature rule satisfying*

$$0 \leq b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}, \quad i \in \{1, \dots, n\}, \quad (2.3)$$

where $\delta := \frac{\eta}{\tau_d(t+d^2)}$, the constant $c_{1,d}$ depends on d , and the constant $c_{2,d,\delta}$ depends on d, δ .

We shall provide the proofs of the main results in Section 4. In comparison, we provide alternative estimates that directly follows from [10, Theorem 5.1] and Lemma 4.3 Section 4.

Theorem 2.4. *Suppose integers $d \in \mathbb{N}$ and $t \geq 1$. Let $\tau_d := 6(3^{\frac{d}{2}}\pi + 2d + 3)$ and $\alpha_d := 8 \left(\frac{\omega_d d}{\omega_{d-1}} \right)^{\frac{1}{d}}$. Let $X_1, X_2, \dots, X_n$ be n independent, identically, and uniformly distributed random vectors taking values in $\mathbb{S}^d$. Let $\eta \in (0, 1)$, if*

$$6\tau_d\alpha_d \left(\frac{n}{4 \ln n} \right)^{-\frac{1}{d}} (t + d^2) < \eta, \quad (2.4)$$

then the following statements hold.

- (i) With probability at least $1 - \frac{1}{n}$, there exist quadrature weights $a_1, \dots, a_n \in \mathbb{R}$ such that $((a_i, X_i))_{i=1}^n$ forms a t -exact quadrature rule satisfying

$$|a_i| \leq \frac{4 \ln n}{n} (1 - \eta)^{-1}, \quad i \in \{1, \dots, n\}, \quad (2.5)$$

- (ii) If $\eta \in (0, \frac{1}{2})$, then with probability at least $1 - \frac{1}{n}$, there exist quadrature weights $b_1, \dots, b_n \in \mathbb{R}$ such that $((b_i, X_i))_{i=1}^n$ is a t -exact quadrature rule satisfying

$$0 \leq b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}, \quad i \in \{1, \dots, n\}, \quad (2.6)$$

We give below some remarks with quantitative and qualitative analysis based on the above main results.

Remark 2.1. The constants $c_{1,d}$ and $c_{2,d,\delta}$ in Theorem 2.2 can be explicitly computed or estimated, e.g., for $\mathbb{S}^2$, $c_{1,d} = 8$ and $c_{2,d,\delta} = \frac{1}{32\pi} \cdot m_2(B_{\mathbf{x}, \sqrt{2}\delta}) = \frac{1}{16} \cdot (1 - \cos(\sqrt{2}\delta))$, where $m_2(B_{\mathbf{x}, \sqrt{2}\delta})$ is the m_2 -measure of a spherical cap with radius $\sqrt{2}\delta$. The definitions of the constants will be presented in the proof of Theorem 2.2.

Remark 2.2. By Theorem 2.3, given a fixed number of points, η and bounds on quadrature weights, we can explicitly compute the probability estimate of the existence of quadrature rules under prescribed conditions, since all the constants are explicitly defined. However, this estimate is still pessimistic due to that the constant τ_d in Lemma 4.4 (see Section 4.3) is not well-estimated and the upper bound of the tail event in Theorem 2.2 is over-estimated. Take $\mathbb{S}^2$ and $\eta = \frac{1}{2}$ for example, when $t = 10$ and $n = 4(t+1)^2$, $\delta \approx 1/689.85$ and we have

$$nc_{1,d} (1 - c_{2,d,\delta})^{n-1} \approx 3872 \left(1 - \frac{1 - \cos\left(\frac{\sqrt{2}}{689.85}\right)}{16} \right)^{484}$$

and

$$\ln \left[nc_{1,d} (1 - c_{2,d,\delta})^{n-1} \right] \approx \ln 3872 + 483 \ln (1 - c_{2,d,\delta}) \approx \ln 8.262,$$

where $c_{2,d,\delta}$ involves the area of a spherical cap with radius $\sqrt{2}\delta$, which is $2\pi(1 - \cos(\sqrt{2}\delta))$ [24]. The absolute value of the second addendum is approximately less than $1\text{E-}7$, and thus the whole term is dominated by $\ln 3872$. In this case, the probability lower bound $1 - nc_{1,d} (1 - c_{2,d,\delta})^{n-1}$ becomes negative and therefore we lose the probability guarantee. On the other hand, given the same values of t , n and η , the right hand side of (2.4) is far greater than η . Thus, we also can not obtain probability guarantee from Theorem 2.4. However, as in the experimental results mentioned in [22] and also ours shown later in Section 3, the number $4N_t = 4(t+1)^2$ of points on the 2-sphere is enough for finding quadrature rules for t up to 512.

Therefore, there are empirically very large gaps between the estimates and the numerical findings.

Remark 2.3. Though the estimates are not empirically ideal, we can still provide qualitative assertions based on Theorem 2.3. To investigate Q2 mentioned in Section 1, we first fix the dimension d , the constant η and the constant c in Theorem 2.3. Then, let $n = C_d t^d$ for each $t \in \mathbb{N}$. Since $\mu_d(B_{\mathbf{x}, t^{-1}}) \sim (1/t)^d$ for any $\mathbf{x} \in \mathbb{S}^d$ and $t \in \mathbb{N}$ [16], we have

$$\lim_{t \rightarrow \infty} (1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1} - c^{-2}) = -\infty \Leftrightarrow \lim_{t \rightarrow \infty} t^d(1 - Ct^{-d})^{t^d-1} = \infty, \quad (2.7)$$

where C is some constant not depending on t . The right hand side is obvious since $(1 - Ct^{-d})^{t^d-1}$ converges to e^{-C} and t^d diverges. Therefore, it is suggested that we might not be able to set $n = C_d t^d$ with small C_d if we attempt to find exact quadrature rules successfully for large t . This will be empirically validated in the following experiments in Section 3.

Remark 2.4. For fixed η , by (2.4) in Theorem 2.4, when n and t satisfy the relation

$$\frac{n}{\ln n} = C_d t^d \quad (2.8)$$

where C_d is some positive number independent of t , there exist exact quadrature rules for Π_t^d with probability $1 - \frac{1}{n}$. We would like to compare the probability estimates from Theorem 2.3 under the same relation. We first show that

$$\lim_{t \rightarrow \infty} \frac{nc_{1,d}(1 - c_{2,d,\delta})^{n-1}}{1/n} = \lim_{n \rightarrow \infty} \frac{nc_{1,d}(1 - c_{2,d,\delta})^{n-1}}{1/n} = 0 \text{ given } \frac{n}{\ln n} = C_d t^d. \quad (2.9)$$

For simplicity, we denote $c_1 := c_{1,d}$ and $c_3 := C_d$. As in the previous remark, by the definition of δ and that $\mu_d(B_{\mathbf{x}, t^{-1}}) \sim (1/t)^d$, it is equivalent to show that

$$\lim_{n \rightarrow \infty} \frac{nc_1(1 - c_2 t^{-d})^{n-1}}{1/n} = 0 \text{ given } \frac{n}{\ln n} = c_3 t^d,$$

where c_2 is some positive constant independent of t . Take the logarithm of $n^2 c_1(1 - c_2 t^{-d})^{n-1}$ and combine with $t^{-d} = c_3 \ln n / n$, then it is equivalent to show that

$$2 \ln n + n \ln \left(1 - c_2 c_3 \frac{\ln n}{n} \right) \rightarrow -\infty, \quad n \rightarrow \infty.$$

When n is sufficiently large, by the inequality $\ln(1+x) \leq x, x > -1$ we have

$$2 \ln n + n \ln \left(1 - c_2 c_3 \frac{\ln n}{n} \right) \leq 2 \ln n - c_2 c_3 \ln n.$$

Note that by definition $c_3 = C_d$ and thus c_3 is undetermined and adjustable. Therefore, (2.9) is true when $c_2 c_3 > 2$. Therefore, in terms of the probability estimates of the existence of quadrature rules with non-negative weights, under the relation (2.8), Theorem 2.3 provides a lower bound with asymptotically faster convergence to 1, and overall high probability guarantee for sufficiently large t .

Remark 2.5. Similarly, we would like to investigate the relation

$$n = C_d t^d \ln t, \quad (2.10)$$

in which the logarithm factor is related to t instead. To show that the estimates in Theorem 2.3 give high probability guarantees for sufficiently large t , we prove that

$$\lim_{t \rightarrow \infty} n c_1 (1 - c_2 t^{-d})^{n-1} = 0, \quad n = c_3 t^d \ln t \quad (2.11)$$

where c_1, c_2, c_3 are defined as in the previous remark. Substitute n with $c_3 t^d \ln t$, then it is equivalent to show that

$$d \ln t + \ln(\ln t) + c_3 t^d (\ln t) \ln(1 - c_2 t^{-d}) \rightarrow -\infty, \quad t \rightarrow \infty.$$

Note that $\lim_{t \rightarrow \infty} t^d \ln(1 - c_2 t^{-d}) = -c_2$. Therefore, when t is sufficiently large, we have

$$d \ln t + \ln(\ln t) + c_3 (\ln t) (t^d \ln(1 - c_2 t^{-d})) \leq (d+1) \ln t - \frac{1}{2} c_2 c_3 \ln t.$$

Thus, when $\frac{1}{2} c_2 c_3 > d+1$, we have (2.11).

Remark 2.6. Despite the estimate $1 - n c_{1,d} (1 - c_{2,d,\delta})^{n-1}$ from Theorem 2.3 possessing faster convergence, there are still some drawbacks compared with (2.5) of Theorem (2.4). By (2.2), the amount needed for adjusting $\frac{1}{n}$ to obtain a t -exact quadrature rule is of order $O(\frac{1}{\sqrt{n}})$. However, (2.5) provides a lower order, which is $O(\frac{\ln n}{n})$. Moreover, our estimates have an extra term $-\frac{1}{c^2}$, which limits the overall probability estimate. These drawbacks are induced by the randomness of the Voronoi cells. In contrast, [10, Theorem 5.1] involves deterministic and equal-area partitions, which are free from such randomness. It can be seen from the proofs in Section 4.1 that the order $O(\frac{1}{\sqrt{n}})$ is largely due to trivially handling the n events of the Voronoi cells. It would be ideal to show an order smaller than $O(\frac{\ln n}{n})$ or even to be $O(\frac{1}{n})$.

Remark 2.7. Since the mesh norm (also called covering radius) of a point set on the sphere is comparable to the partition norm induced by the associated Voronoi tessellation [19], an alternative way to obtain probability estimates is to estimate the mesh norm with respect to the sampled points. One can follow [33, (5.1), (5.2)] which are stated for a certain class of smooth manifolds, and obtain estimates that essentially differ from the right hand side of (2.7) by multiplying $-(1 - C t^d)$. In contrast, our approach is more constructive and provides explicit constants.

3. Numerical Experiments

We also would like to mention some details of finding exact quadrature rules on $\mathbb{S}^2$ from [14]. Such rules were obtained numerically by applying fast spherical Fourier and adjoint transforms [20, 21] and solving the following optimization problem:

$$\min_{w \geq \mathbf{0}} \|Y^* w - \sqrt{4\pi} e_0\|_2 \quad (3.1)$$

where $Y \in \mathbb{C}^{n \times N_t}$, $Y_{ij} := \tilde{y}_j(X_i)$ ($\tilde{y}_i$ are complex-valued spherical harmonics), $w, e_0 \in \mathbb{R}^n$ with $e_0 := [1, 0, \dots, 0]^\top$. A 500-exact rule using 10^6 sampled points with residual 4.843589E-13 was found. We follow [14] and consider the equivalent problem of finding quadrature rules for Π_t^2 :

$$\min_{w \in \mathbb{R}^n} \mathcal{L} := \|Aw - \sqrt{4\pi}e_0\|_2 \quad (3.2)$$

where $A := [A_1, A_2]^\top \in \mathbb{R}^{(t+1)^2 \times n}$ with

$$\begin{aligned} A_1 &:= \Re(Y_k^l(X_i))_{i=0, \dots, n; l=0, \dots, t, 0 \leq k \leq l} \in \mathbb{R}^{n \times \frac{(t+1)(t+2)}{2}}, \\ A_2 &:= \Im(Y_k^l(X_i))_{i=1, \dots, n; l=1, \dots, t, -l \leq k \leq -1} \in \mathbb{R}^{n \times \frac{t(t+1)}{2}}. \end{aligned}$$

Here we use Y_k^l to denote the complex valued spherical harmonics in $\mathbb{S}^2$ [30]. The operations $\Re(f)$, $\Im(f)$ are the real and imaginary part of a complex number (function) f , respectively.

Incorporating the lower bounds and upper bounds in (2.2) and (2.3), we therefore consider the following constrained convex optimizations adapted from (3.2):

$$\min_{L_1 \leq w \leq R_1} \|Aw - \sqrt{4\pi}e_0\|_2, \quad (3.3)$$

$$\min_{0 \leq w \leq R_2} \|Aw - \sqrt{4\pi}e_0\|_2, \quad (3.4)$$

where

$$L_1 := -\omega_2 \frac{9\sqrt{n} + 1}{n} \cdot 2 \quad (c = 9, \eta = \frac{1}{2}), R_1 := -L_1, R_2 := \frac{\omega_2}{N_{\lfloor \frac{t}{2} \rfloor}}.$$

The problems (3.3) and (3.4) are then numerically solved using the MATLAB Optimization Toolbox with algorithm `trust-region-reflective` [3, 4] (see also [36]) and custom functions for computing the gradient $\nabla \mathcal{L}$ and Hessian of $\mathcal{L}$.^a Such functions adopted the fast spherical Fourier and adjoint transforms in [20, 21]. All experiments were done on a desktop PC with an Intel Core i7-12700KF CPU and 16GB of RAM. The maximum number of iterations is set to be 201.

Since $(t+1)^2 \sim t^2$, it is equivalent to set $n = c(t+1)^2$ where c is the constant to be tested. We set $t \in \{16, 32, 64, 128, 256, 512\}$. The initial value of each component of w to be $\frac{4\pi}{n}$. The results associated with (3.3) and (3.4) are shown in Table 1 and 2.

We give some detailed comments from our numerical results .

Remark 3.1. Exp 1 and Exp 2 in Table 1 show that for both (3.3) and (3.4), a large enough fixed constant c enables the successful findings of t -exact quadrature rules with a fixed range of t .

Remark 3.2. However, by lowering the constant c , the performance behaves differently. Exp 1 and Exp 3 in Table 1 show the difference for (3.3). When c is

^aSee <https://github.com/zrgcityu/usq-code> for implementation details

Table 1. Numerical Results of 3 Settings [objective, n]: Exp 1 [(3.3), $n = 4(t+1)^2$], Exp 2 [(3.4), $n = 6(t+1)^2$], Exp 3 [(3.3), $n = 1.2(t+1)^2$].

Exp 1	For (3.3), $n = 4(t+1)^2$			
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	3.134581E-15	3.063353E-15	1.788075s	15
32	3.291015E-14	3.388295E-14	0.807776s	16
64	1.841470E-13	1.872858E-13	6.399274s	17
128	2.637861E-13	3.969358E-13	18.665806s	20
256	3.995873E-13	6.811824E-13	52.704175s	18
512	4.320490E-13	1.262332E-12	238.550729s	18
Exp 2	For (3.4), $n = 6(t+1)^2$			
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	1.875412E-15	3.118444E-15	2.047900s	19
32	3.570558E-14	4.459392E-14	1.295156s	22
64	1.049723E-13	1.410922E-13	18.936084s	29
128	2.005977E-13	3.013737E-13	138.984551s	31
256	2.633198E-13	4.254067E-13	761.470559s	30
512	3.420550E-13	5.841217E-13	12563.455464s	46
Exp 3	For (3.3), $n = \lceil 1.2(t+1)^2 \rceil$			
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	3.993703E-14	1.353421E-14	9.872654s	32
32	2.262766E-12	9.221247E-13	14.793752s	23
64	1.574769E-11	9.349000E-12	509.541738s	23
128	1.626569E-11	7.383318E-12	1095.151538s	23
256	2.613153E-11	1.885719E-11	3440.798370s	20
512	5.795588E-11	2.279233E-11	33192.043693s	22

reduced from 4 to 1.2, we can still find rules with losses $\mathcal{L}$ very close to 0. Nonetheless, the orders in the magnitudes rise to 1E-11 for most t and the time consumed is increased, especially for large t .

Remark 3.3. The effect of lowering c is more obvious for (3.4), where quadrature weights are required to be non-negative. When c is reduced from 6 to 4, in all 4 attempts in Table 2, we see that not only the time consumed is greatly increased, but also failures (3 out of 4 total trials) in the optimizations for $t = 512$, in which the final losses $\mathcal{L} \approx 1\text{E-}2$ are far from 0.

Remark 3.4. Compared with Exp 1, the result in Exp 2 corresponds to a large constant c . Compared with Exp 3, Table 2 displays more drastic changes. Note that the constants c in Exp 1 and Table 2 are identical. Therefore, in contrast to (3.3), we can say that (3.4) is more difficult to be solved numerically. This is consistent with Lemma 4.3 in Section 4.3, in which the existence of exact quadrature rules

with non-negative weights requires a stronger condition of $\eta \in (0, 1/2)$, suggesting that the number of points should be larger.

Remark 3.5. Our experiments have shown some discouraging results by which the idea of setting a small fixed constant C_d in $n = C_d t^d$ for all t is empirically denied. In detail, we cannot choose C_d by trial and error on small t and then adopting the smallest possible C_d . Though Table 1 show successful findings for the given t , we expect that when t is large enough, failures in the optimizations will arise as those in Table 2. For fixing a constant C_d , more reasonable choices would be $\frac{n}{\ln n} = C_d t^d$ and $n = C_d t^d \ln t$ as discussed in Remarks 2.4 and 2.5, respectively.

4. Proofs of the Main Theorems

In this section, we provide detailed proofs of our main results in Theorem 2.1, Theorem 2.2, and Theorem 2.3. We first prove the results concerning the measure and the diameter of Voronoi cells on the sphere in Section 4.1, Section 4.2, respectively. Similar results for Voronoi cells on Euclidean spaces can be found in [8]. We then adopt the results on M–Z inequalities to our setting in Section 4.3 and derive the main results by combining those results together in Section 4.4.

4.1. The measure of Voronoi cells

Recall that $X_1, X_2, \dots, X_n$ are n independent, identically and uniformly distributed random vectors taking values in $\mathbb{S}^d$. The μ_d -measure of the (random) Voronoi cell of $\mathbf{x}$ (or X_1) with respect to the Voronoi tessellation of $\mathbf{x}$ (or X_1) and $X_2, \dots, X_n$ is denoted as $M_n(\mathbf{x})$. The expectation $\mathbb{E}[M_n(\mathbf{x})]$ is taken with respect to the joint distribution of $X_2, \dots, X_n$ and the expectation $\mathbb{E}[M_n(X_i)]$ is taken with respect to the joint distribution of $X_1, X_2, \dots, X_n$.

The proof of Theorem 2.1 can be done by proving the following two lemmas.

Lemma 4.1. *For all $i \in \{1, \dots, n\}$, we have*

$$n\mathbb{E}[M_n(X_i)] = 1, \quad n^2\mathbb{E}[M_n(X_i)^2] < 2, \quad \text{and} \quad n^2\text{Var}[M_n(X_i)] < 1.$$

Proof. For the first moment, without loss of generality, we first show that $n\mathbb{E}[M_n(X_1)] = 1$. This directly follows from the fact that $\sum_{i=1}^n M_n(X_i) = 1$ and that $M_n(X_i), i \in \{1, \dots, n\}$ are identically distributed.

As for the second moment, we fix $\mathbf{x} \in \mathbb{S}^d$ and observe that

$$\begin{aligned} \mathbb{E}[M_n(\mathbf{x})^2] &= \mathbb{P}(\{X \in \mathcal{C}_{1,n}, X' \in \mathcal{C}_{1,n} | X_1 = \mathbf{x}\}) \\ &= \mathbb{P}(\{\cap_{i=2}^n \{X_i \notin B_{X,\rho(X,\mathbf{x})} \cup B_{X',\rho(X',\mathbf{x})}\}\}) \\ &= \mathbb{E}[(1 - Z_2(\mathbf{x}))^{n-1}], \end{aligned}$$

where

$$Z_2(\mathbf{x}) = \mu_d(B_{X,\rho(X,\mathbf{x})} \cup B_{X',\rho(X',\mathbf{x})})$$

Table 2. Numerical Results of Solving (3.4) for $n = 4(t + 1)^2$

Trial 1				
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	2.303655E-15	2.390542E-15	3.097006s	20
32	3.185510E-14	4.099474E-14	1.089420s	24
64	1.969465E-13	2.341838E-13	56.693948s	43
128	2.962170E-13	3.981251E-13	299.824356s	52
256	3.819095E-13	7.150612E-13	14576.090757s	150
512	4.066202E-03	5.510995E-02	$\approx 59\text{h}$	201
Trial 2				
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	3.500963E-15	7.052921E-15	0.709574s	21
32	2.767770E-14	2.929239E-14	1.089420s	22
64	1.647513E-13	1.777890E-13	25.140223s	33
128	3.785107E-13	6.031171E-13	1088.606990s	92
256	4.167524E-13	8.449096E-13	2856.551934s	61
512	1.036127E-03	1.675591E-02	$\approx 46\text{h}$	201
Trial 3				
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	2.401973E-15	3.126254E-15	0.590745s	19
32	3.123671E-14	3.260830E-14	1.988796s	31
64	1.889524E-13	2.290627E-13	76.218829s	51
128	2.553135E-13	3.624777E-13	169.328444s	40
256	4.349658E-13	1.763420E-12	4003.917936s	72
512	5.023403E-13	7.931863E-13	$\approx 22\text{h}$	169
Trial 4				
t	$\mathcal{L}$	$\ \nabla \mathcal{L}\ _\infty$	Time	Iteration
16	1.744179E-15	2.584496E-15	0.614434s	19
32	6.004930E-14	6.404889E-14	2.000308s	34
64	1.606202E-13	2.159116E-13	14.824280s	34
128	3.207168E-13	3.248754E-13	172.495088s	52
256	4.041109E-13	6.541325E-13	1881.816718s	81
512	2.329341E-03	3.467551E-02	$\approx 59\text{h}$	201

with X and X' being independent and identically distributed with respect to μ_d . Note that X and X' are also independent of $X_1, \dots, X_n$. Then, we have

$$n^2 \mathbb{E}[M_n(\mathbf{x})^2] = n^2 \mathbb{E}[(1 - Z_2(\mathbf{x}))^{n-1}] = n^2 \int_0^1 \mathbb{P}(\{(1 - Z_2(\mathbf{x}))^{n-1} \geq s\}) \, ds.$$

By a change of variables with $z = 1 - s^{\frac{1}{n-1}}$, we have

$$\int_0^1 \mathbb{P}(\{(1 - Z_2(\mathbf{x}))^{n-1} \geq s\}) ds = (n-1) \int_0^1 \mathbb{P}(\{Z_2(\mathbf{x}) \leq z\})(1-z)^{n-2} dz$$

Note that for a uniformly distributed Y , we have

$$Z(\mathbf{x}) = \mu_d(B_{Y,\rho(Y,\mathbf{x})}) = \mu_d(B_{\mathbf{x},\rho(Y,\mathbf{x})})$$

and that $\mu_d(B_{\mathbf{x},\rho(Y,\mathbf{x})})$ is uniformly distributed on $[0, 1]$. Now, for the integration $\int_0^1 \mathbb{P}(\{Z_2(\mathbf{x}) \leq z\})(1-z)^{n-2} dz$, by the independence of X and X' , we have

$$\begin{aligned} & \int_0^1 \mathbb{P}(\{Z_2(\mathbf{x}) \leq z\})(1-z)^{n-2} dz \\ & \leq \int_0^1 \mathbb{P}(\{\mu_d(B_{X,\rho(X,\mathbf{x})}) \leq z\} \cap \{\mu_d(B_{X',\rho(X',\mathbf{x})}) \leq z\})(1-z)^{n-2} dz \\ & = \int_0^1 \mathbb{P}(\{\mu_d(B_{\mathbf{x},\rho(X,\mathbf{x})}) \leq z\} \cap \{\mu_d(B_{\mathbf{x},\rho(X',\mathbf{x})}) \leq z\})(1-z)^{n-2} dz \\ & = \int_0^1 \mathbb{P}(\{\mu_d(B_{\mathbf{x},\rho(X,\mathbf{x})}) \leq z\})^2 (1-z)^{n-2} dz \\ & = \int_0^1 z^2 (1-z)^{n-2} dz = \frac{2}{n(n-1)(n+1)}. \end{aligned}$$

Consequently, we obtain

$$n^2 \mathbb{E}[M_n(\mathbf{x})^2] = n^2(n-1) \int_0^1 \mathbb{P}(\{Z_2(\mathbf{x}) \leq z\})(1-z)^{n-2} dz \leq \frac{2}{1+\frac{1}{n}} < 2.$$

Integrating both sides of $n^2 \mathbb{E}[M_n(\mathbf{x})] < 2$ with respect to X_1 gives

$$n^2 \mathbb{E}[M_n(X_1)^2] < 2.$$

The result

$$n^2 \text{Var}[M_n(X_i)] < 1$$

follows directly from $\text{Var}[M_n(X_i)] = \mathbb{E}[M_n(X_i)^2] - \mathbb{E}[M_n(X_i)]^2$. For the cases of $X_2, \dots, X_n$, they can be proved analogously. $\square$

Lemma 4.2. *For any $c > 1$, with probability at least $1 - \frac{1}{c^2}$, we have*

$$\max_i |M_n(X_i) - \mathbb{E}[M_n(X_i)]| = \max_i \left| M_n(X_i) - \frac{1}{n} \right| < \frac{c}{\sqrt{n}}. \quad (4.1)$$

Proof. Denote $\sigma := \sqrt{\text{Var}(M_n(X_i))}$, then $\sigma^2 < \frac{1}{n^2}$ by Lemma 4.1. Using the Chebyshev inequality

$$\mathbb{P}(|X - \mathbb{E}X| \geq s\sigma) \leq \frac{1}{s^2}, \quad s > 0,$$

16

and letting $s\sigma = c/\sqrt{n}$, we have

$$\mathbb{P}\left(\left\{\left|M_n(X_i) - \frac{1}{n}\right| \geq \frac{c}{\sqrt{n}}\right\}\right) \leq \frac{n\sigma^2}{c^2} < \frac{1}{nc^2}$$

for all $i \in \{1, \dots, n\}$, which implies

$$\begin{aligned} \mathbb{P}\left(\left\{\max_i \left|M_n(X_i) - \frac{1}{n}\right| \geq \frac{c}{\sqrt{n}}\right\}\right) &= \mathbb{P}\left(\bigcup_{i=1}^n \left\{\left|M_n(X_i) - \frac{1}{n}\right| \geq \frac{c}{\sqrt{n}}\right\}\right) \\ &\leq \sum_{i=1}^n \mathbb{P}\left(\left\{\left|M_n(X_i) - \frac{1}{n}\right| \geq \frac{c}{\sqrt{n}}\right\}\right) \\ &< \frac{1}{c^2}. \end{aligned}$$

This completes the proof. $\square$

4.2. The diameters of Voronoi cells

Recall that $D_n(\mathbf{x})$ denote the diameter of the Voronoi cell associated with $\mathbf{x}$ with respect to the Voronoi tessellation of $\mathbf{x}, X_2, \dots, X_n$. Thus, the probability of the events of $D_n(\mathbf{x})$ involves the joint distribution of $X_2, \dots, X_n$ while for $D_n(X_1)$, the probability of the events of $D_n(X_1)$ involves the joint distribution of $X_1, X_2, \dots, X_n$. The diameters $D_n(X_2), \dots, D_n(X_n)$ are similarly defined.

Below we need to use the so-called “geodesic cones” on $\mathbb{S}^d$. Let $\vec{u}$ be a unit-length vector in $T_{\mathbf{x}}\mathbb{S}^d$, the tangent space of $\mathbf{x}$. For each $\mathbf{y}$, choose a unit-speed geodesic α that connects $\mathbf{x}$ and $\mathbf{y}$ and satisfies $\alpha(0) = \mathbf{x}$. Thus we have a velocity vector $\vec{v} := \dot{\alpha}(0)$. Define the geodesic cone $\mathfrak{C}(\mathbf{x}, \vec{u})$ to be the set of points in $\mathbb{S}^d$ such that the angle θ between $\vec{u}$ and $\vec{v}$ satisfies $\theta \leq \frac{\pi}{8}$. Such cones $\mathfrak{C}$ at $\mathbf{x}$ are considered to have an “angle” $\frac{\pi}{4}$. Take $\mathbf{x}$ as the north pole in $\mathbb{S}^2$ for example, each area consecutively covering longitude of 45 degree is a cone with angle $\frac{\pi}{4}$.

We next provide the proof of Theorem 2.2 on the diameter of the Voronoi cells.

Proof. [Proof of Theorem 2.2] We first show that

$$\mathbb{P}(\{D_n(X_i) > \delta\}) \leq c_{1,d}(1 - c_{2,d,\delta})^{n-1}, \quad i \in \{1, \dots, n\}.$$

As in the proof of Lemma 4.1, we fix $\mathbf{x} \in \mathbb{S}^d$. Let γ_d be the minimal numbers of geodesic cones $\mathfrak{C}_1, \dots, \mathfrak{C}_{\gamma_d}$ of angle $\frac{\pi}{4}$ at $\mathbf{x}$ such that their union covers $\mathbb{S}^d$. For each $j = 1, \dots, \gamma_d$, let $R_{n,j}$ be the distance between $\mathbf{x}$ and the nearest neighbor among $X_2, \dots, X_n$ belonging to $\mathfrak{C}_j$. Define $R_{n,j} = \infty$ if no such point exists.

We bound the diameter of the Voronoi cell associated with $\mathbf{x}$ by observing that

$$D_n(\mathbf{x}) \leq \max_{j=1, \dots, \gamma_d} \frac{1}{\sqrt{2}} R_{n,j}$$

To see this, consider an arbitrary point $\mathbf{y} \in \mathcal{C}_{1,n}$ and let $\mathfrak{C}_j$ be a cone (among the covering cones) containing $\mathbf{y}$. If ξ_j is the nearest neighbor of $\mathbf{x}$ among $X_2, \dots, X_n$

belonging to $\mathfrak{C}_j$, then $\rho(\mathbf{x}, \mathbf{y}) \leq \rho(\mathbf{y}, \boldsymbol{\xi}_j)$. We obtain a geodesic triangle by connecting $\mathbf{x}, \mathbf{y}$ and $\boldsymbol{\xi}_j$ using minimal geodesics. Let s_1, s_2, s_3 denote the sides of the geodesic triangle that connects $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x}$ and $\boldsymbol{\xi}_j$, $\mathbf{y}$ and $\boldsymbol{\xi}_j$, respectively. Let l_1, l_2, l_3 denote the length of s_1, s_2, s_3 . Note that the angle α between s_1 and s_2 is bounded by $\frac{\pi}{4}$ and we have

$$l_1 = \rho(\mathbf{x}, \mathbf{y}), \quad l_2 = \rho(\mathbf{x}, \boldsymbol{\xi}_j) = R_{n,j}, \quad l_3 = \rho(\mathbf{y}, \boldsymbol{\xi}_j).$$

By the Toponogov's theorem (cf. [32, Chapter 11] or [5, Theorem 2.1]), we have the Law of Cosines on $\mathbb{S}^d$:

$$l_3^2 \leq l_1^2 + l_2^2 - 2l_1l_2 \cos \alpha.$$

Hence we have $2l_1l_2 \cos \alpha \leq l_1^2 - l_3^2 + l_2^2$. By that $l_1 \leq l_3$, we deduce that $2l_1l_2 \cos \alpha \leq l_2^2$. Consequently, from

$$\rho(\mathbf{x}, \mathbf{y}) = l_1 \leq \frac{l_2}{2 \cos \alpha} \leq \frac{R_{n,j}}{\sqrt{2}},$$

we conclude that $D_n(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{C}_{1,n}} \rho(\mathbf{x}, \mathbf{y}) \leq \max_{j=1, \dots, \gamma_d} \frac{1}{\sqrt{2}} R_{n,j}$.

Now for any $\delta > 0$, we have

$$\begin{aligned} \{D_n(\mathbf{x}) > \delta\} &\subset \left\{ \max_{j=1, \dots, \gamma_d} R_{n,j} > \sqrt{2}\delta \right\} \\ &\subset \bigcup_{j=1}^{\gamma_d} \left\{ \mathfrak{C}_j \cap B_{\mathbf{x}, \sqrt{2}\delta} \text{ has no point among } X_2, \dots, X_n \right\}. \end{aligned}$$

Let P_j be the event $\left\{ \mathfrak{C}_j \cap B_{\mathbf{x}, \sqrt{2}\delta} \text{ has no point among } X_2, \dots, X_n \right\}$. Then, we have $\mathbb{P}(P_j) = \left(1 - \mu_d\left(\mathfrak{C}_j \cap B_{\mathbf{x}, \sqrt{2}\delta}\right)\right)^{n-1}$. Hence, we deduce that

$$\mathbb{P}(\{D_n(\mathbf{x}) > \delta\}) \leq \sum_{j=1}^{\gamma_d} \mathbb{P}(P_j) = \sum_{j=1}^{\gamma_d} \left(1 - \mu_d\left(\mathfrak{C}_j \cap B_{\mathbf{x}, \sqrt{2}\delta}\right)\right)^{n-1}.$$

Let $c_{1,d} := \gamma_d$ and $c_{2,d,\delta} := \mu_d\left(\mathfrak{C}_1 \cap B_{\mathbf{x}, \sqrt{2}\delta}\right)$. Noting that $c_{1,d}, c_{2,d,\delta}$ are independent of $\mathbf{x}$, we obtain

$$\mathbb{P}(\{D_n(\mathbf{x}) > \delta\}) \leq c_{1,d} (1 - c_{2,d,\delta})^{n-1} \quad \forall \mathbf{x} \in \mathbb{S}^d.$$

Integrating on both sides with respect to X_1 gives

$$\mathbb{P}(\{D_n(X_1) > \delta\}) \leq c_{1,d} (1 - c_{2,d,\delta})^{n-1}$$

For $X_2, \dots, X_n$, this can be similarly proved.

Now, the result

$$\mathbb{P}(\{\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta\}) \leq nc_{1,d} (1 - c_{2,d,\delta})^{n-1}$$

obviously holds. In fact,

$$\begin{aligned} \mathbb{P}(\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta) &= \mathbb{P}\left(\left\{\max_i D_n(X_i) > \delta\right\}\right) \\ &= \mathbb{P}\left(\bigcup_{i=1}^n \{D_n(X_i) > \delta\}\right) \\ &\leq \sum_{i=1}^n \mathbb{P}(\{D_n(X_i) > \delta\}) \leq nc_{1,d}(1 - c_{2,d,\delta})^{n-1}. \end{aligned}$$

This completes the proof. $\square$

4.3. Marcinkiewicz–Zygmund inequalities and the existence of quadrature rules

Sufficient conditions for the existence of quadrature rules appear in several works [12, 28, 29, 31]. The existence is guaranteed by satisfying the Marcinkiewicz–Zygmund (M–Z) inequalities, which in turn is implied by the conditions of the partition norm. Therefore, we organize the lemmas by first stating the existence of quadrature rules from the implication of the M–Z inequalities to, and then the M–Z inequalities from the bounded partition norm. Moreover, we particularly select conditions with explicitly defined constants, which serve our purpose of providing the concrete probability estimates. Therefore, to obtain a summary of related results in different works, we follow [28, 29, 31] and provide the following lemma, in which the settings are slightly generalized so as to facilitate the discussions related to [10].

Lemma 4.3 *Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be n points in $\mathbb{S}^d$ and $w_1, \dots, w_n$ be n positive real numbers.*

(a) *For $\eta \in (0, 1)$, if the M–Z inequality*

$$(1-\eta) \int_{\mathbb{S}^d} |P(\mathbf{x})| d\mu_d(\mathbf{x}) \leq \sum_{i=1}^n |P(\mathbf{x}_i)| w_i \leq (1+\eta) \int_{\mathbb{S}^d} |P(\mathbf{x})| d\mu_d(\mathbf{x}) \quad (4.2)$$

holds for all $P \in \Pi_t^d$, then there exist quadrature weights $a_1, \dots, a_n \in \mathbb{R}$ such that $((a_i, \mathbf{x}_i))_{i=1}^n$ is a t -exact quadrature rule for $\mathbb{S}^d$ satisfying

$$|a_i| \leq \frac{w_i}{1-\eta}, \quad i \in \{1, \dots, n\}. \quad (4.3)$$

(b) *For $\eta \in (0, \frac{1}{2})$, if*

$$\frac{1-2\eta}{1-\eta} \sum_{i=1}^n P(\mathbf{x}_i) w_i \leq \int_{\mathbb{S}^d} P(\mathbf{x}) d\mu_d(\mathbf{x}) \leq \frac{1}{1-\eta} \sum_{i=1}^n P(\mathbf{x}_i) w_i. \quad (4.4)$$

holds for all $P \in \{P \mid P \in \Pi_t^d, P(\mathbf{x}_i) \geq 0, i = 1, \dots, n\}$, then there exist positive quadrature weights $b_1, \dots, b_n \in \mathbb{R}$ such that $((b_i, \mathbf{x}_i))_{i=1}^n$ is a t -exact quadrature rule for $\mathbb{S}^d$ satisfying

$$\frac{1-2\eta}{1-\eta} w_i \leq b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}, \quad i \in \{1, \dots, n\}. \quad (4.5)$$

Proof. Define a linear map

$$F : \Pi_t^d \rightarrow \mathbb{R}^n, \quad F(P) := (P(\mathbf{x}_1), \dots, P(\mathbf{x}_n)).$$

Let the spaces Π_t^d and $\mathbb{R}^n$ be equipped with the norms

$$\int_{\mathbb{S}^d} |P(\mathbf{x})| d\mu_d(\mathbf{x}) \text{ for } P \in \Pi_t^d \text{ and } \sum_{i=1}^n |v_i| w_i \text{ for } \mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n, \quad (4.6)$$

respectively. Let $V \subseteq \mathbb{R}^n$ denote the range of F . By (4.2), F is injective. Thus, the inverse $F^{-1} : V \rightarrow \Pi_t^d$ exists. Moreover, the operator norm $\|F^{-1}\|$ satisfies $\|F^{-1}\| \leq (1 - \eta)^{-1}$. Let f denote the linear functional on Π_t^d by

$$f(P) := \int_{\mathbb{S}^d} P(\mathbf{x}) d\mu_d(\mathbf{x}), \quad \forall P \in \Pi_t^d.$$

By the Hölder's inequality, we have $\|f\|_{(\Pi_t^d)^*} \leq 1$. Let $\tilde{f} := f(F^{-1}(\mathbf{v}))$, $\mathbf{v} \in V$ be a linear functional on V . We have $\|\tilde{f}\|_{V^*} \leq (1 - \eta)^{-1}$. By the Hahn-Banach theorem, $\tilde{f}$ has a norm-preserving extension to a linear function $\tilde{f}_{ext}$ on $\mathbb{R}^n$. Therefore, there exists $(a_1, \dots, a_n) \in \mathbb{R}^n$ such that for any $\mathbf{v} := (v_1, \dots, v_n) \in \mathbb{R}^n$, we have

$$\tilde{f}_{ext}(\mathbf{v}) = \sum_{i=1}^n a_i v_i, \quad \|a_i\|_{(\mathbb{R}^n)^*} \leq (1 - \eta)^{-1}.$$

Note that $\|\cdot\|_{(\mathbb{R}^n)^*}$ is induced by the norm in (4.6). Thus,

$$\|a_i\|_{(\mathbb{R}^n)^*} \leq (1 - \eta)^{-1} = \max_i |a_i| w_i^{-1} \leq (1 - \eta)^{-1}.$$

By construction, $f(P) = \tilde{f}(\mathbf{v}) = \tilde{f}_{ext}(\mathbf{v})$, where $\mathbf{v} = F(P)$. This shows that $((a_i, \mathbf{x}_i))_{i=1}^n$ is a t -exact quadrature rule of Π_t^d with weights satisfying (4.3).

When $\eta \in (0, \frac{1}{2})$, let $\mathbb{R}_+^n$ denote the positive cone $\{\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n \mid v_i \geq 0\}$ and define the linear functional $\tilde{g}$ on V

$$\tilde{g}(\mathbf{v}) = \tilde{f}(\mathbf{v}) - \frac{1 - 2\eta}{1 - \eta} \sum_{i=1}^n v_i w_i.$$

For $\mathbf{v} \in V \cap \mathbb{R}_+^n$, by (4.4), $\tilde{g}$ is positive on $V \cap \mathbb{R}_+^n$. Note that $(1, \dots, 1) \in V$ since the constant polynomial is in Π_t^d , and it is an interior point of $\mathbb{R}_+^n$. By the Krein-Rutman theorem [17], there exists a positive linear functional $\tilde{g}_{ext}$ that extends $\tilde{g}$ on $\mathbb{R}^n$. Hence, there exist weights $c_i \geq 0, i \in \{1, \dots, n\}$ such that $\tilde{g}_{ext}(\mathbf{v}) = \sum_{i=1}^n c_i v_i$. Therefore, we have positive weights

$$b_i := c_i + \frac{1 - 2\eta}{1 - \eta} w_i, \quad b_i \geq \frac{1 - 2\eta}{1 - \eta} w_i, \quad i \in \{1, \dots, n\}$$

such that $((\mathbf{x}_i, b_i))_{i=1}^n$ is an exact quadrature rule of Π_t^d . To derive an upper bound for $b_i, i \in \{1, \dots, n\}$, we fix i and set

$$P(\mathbf{x}) := \sum_{j=1}^{N_{\lfloor \frac{t}{2} \rfloor}} Y_j(\mathbf{x}) Y_j(\mathbf{x}_i),$$

20

where Y_i are the spherical harmonics in $\Pi_{\lfloor \frac{t}{2} \rfloor}^d$. Due to the orthogonality of Y_i and that $P^2 \in \Pi_t^d$, we have

$$\int_{\mathbb{S}^d} |P(\mathbf{x})|^2 d\mu_d(\mathbf{x}) = P(\mathbf{x}_i) = \sum_{j=1}^n b_j P^2(\mathbf{x}_j) \geq b_i P^2(\mathbf{x}_i).$$

Note that $P(\mathbf{x}_i) = N_{\lfloor \frac{t}{2} \rfloor}$ (cf. [37, (1)]). Thus, $b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}$. $\square$

The following lemma, which is originally described for general compatible partitions, are adapted to the special case of the Voronoi tessellation.

Lemma 4.4. *Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be n distinct points in $\mathbb{S}^d$ with $d \in \mathbb{N}$. Let $\tau_d := 6(3^{\frac{d}{2}}\pi + 2d + 3)$.*

- (a) *For $\eta \in (0, 1)$, if $\|\mathcal{R}(\mathbf{m}_{n,d})\| \leq \frac{\eta}{\tau_d(t+d^2)}$, then the M-Z inequality in (4.2) holds for all $P \in \Pi_t^d$ with $w_i = \mu_d(\mathcal{C}_{i,n})$, $i \in \{1, \dots, n\}$.*
- (b) *For $\eta \in (0, 1/2)$, if $\|\mathcal{R}(\mathbf{m}_{n,d})\| \leq \frac{\eta}{\tau_d(t+d^2)}$, then the inequality (4.4) holds for all $P \in \{P \mid P \in \Pi_t^d, P(\mathbf{x}_i) \geq 0, i = 1, \dots, n\}$ with $w_i = \mu_d(\mathcal{C}_{i,n})$, $i \in \{1, \dots, n\}$.*

Proof. (a) directly comes from [10, Theorem 4.1]. To prove (b), it is sufficient to prove

$$\sum_{i=1}^n \int_{\mathcal{C}_{i,n}} |P(\mathbf{x}) - P(\mathbf{x}_i)| d\mu_d(\mathbf{x}) \leq \eta \int_{\mathbb{S}^d} |P(\mathbf{x})| d\mu_d(\mathbf{x}), \quad \forall P \in \Pi_t^d, \quad (4.7)$$

since by noting that (a) is true in this case and that by assumption $P(\mathbf{x}_i) \geq 0, i = 1, \dots, n$, we have

$$\begin{aligned} \left| \int_{\mathbb{S}^d} P(\mathbf{x}) d\mu_d(\mathbf{x}) - \sum_{i=1}^n P(\mathbf{x}_i) w_i \right| &\leq \sum_{i=1}^n \int_{\mathcal{C}_{i,n}} |P(\mathbf{x}) - P(\mathbf{x}_i)| d\mu_d(\mathbf{x}) \\ &\leq \eta \int_{\mathbb{S}^d} |P(\mathbf{x})| d\mu_d(\mathbf{x}) \leq \frac{\eta}{1-\eta} \sum_{i=1}^n P(\mathbf{x}_i) w_i, \end{aligned}$$

which implies (4.4). In fact, (4.7) is an intermediate result shown in the proof of [10, Theorem 4.1] when $p = 1$. $\square$

Now we can provide the proof of Theorem 2.4.

Proof. [Proof of Theorem 2.4]. The result in [10, Theorem 5.1] and its proof show that when satisfying (2.4), with probability at least $1 - \frac{1}{n}$, the sampled points satisfy the M-Z inequality (4.2) with $w_i \leq 4 \frac{\ln n}{n}$. Therefore, by Lemma 4.3 (a), we have the existence of the quadrature weights with the same probability lower bound stated in (i) of Theorem 2.4.

As for (ii) of Theorem 2.4, note that the proof of [10, Theorem 5.1] is based on the event such that each interior of the cells of an equal-area partition contains

at least one sampled point. We associate each interior with one sampled point. Moreover, we assign each associated sampled point a weight equal to the measure of the cell. For the rest of the sampled points, we assign them a zero weight. (ii) of Theorem 2.4 can be deduced by using similar arguments as in (b) of Lemma 4.4 and combining it with (b) of Lemma 4.3. Note that the lower bound of the weights from (b) of Lemma 4.3 is discarded due to the possibility of having points assigned with zero weight. $\square$

Remark 4.1. As pointed out in [29], the Krein-Rutman extensions of non-negative functionals are not guaranteed to be norm-preserving, and thus one should not directly apply the statements related to quadrature rules with positive weights in [28]. On the other hand, in [31, Section 4.2], theorems related to the existence of quadrature rules with positive weights are given, in which the upper bounds of weights are derived using other arguments. In Lemma 4.3, we follow [31] instead of [28] to obtain the bounds of non-negative quadrature weights.

Remark 4.2. Our definition of the M–Z inequality (4.2) corresponds to case $p = 1$ in which L_1 and (weighted) l_1 norms are involved. The general definition of the M–Z inequality replaces the integrations with L_p norms and the discrete sum with (weighted) l_p norms, $p \in [1, \infty]$ (cf. [10, 28]). Assuming that $\mathbf{x}_1, \dots, \mathbf{x}_n$ satisfy the M–Z with $p > 1$, by following [28], one can obtain bounds for the weights similar to those of a_i in Lemma 4.3, but depicted with (weighted) l_p norms. However, as mentioned in [10, 11], the applications of the Riesz-Thorin interpolation in [19, 28] are not fully justified, and thus one has to be cautious about the M–Z inequalities with $p \in (1, \infty)$. We refer to [10], which provides an alternative proof of the M–Z inequalities with rigorous applications of the Riesz-Thorin interpolation.

Besides [10, Theorem 4.1], by [12, Theorem 4.2, Remark 4.3], the upper bound for $\|\mathcal{R}(\mathbf{m}_{n,d})\|$ in Lemma 4.4 can be replaced by

$$\|\mathcal{R}(\mathbf{m}_{n,d})\| \leq \frac{\eta}{\tilde{\tau}_d t},$$

$$\tilde{\tau} := \begin{cases} (2\sqrt{3})^d(5d+1), & d \geq 3, \\ 111, & d = 2. \end{cases}$$

For $d = 2$, the original upper bound is approximately

$$\frac{\eta}{98.55(t+4)}.$$

Compared with the alternative bound $\frac{\eta}{111t}$, the original bound gives a smaller denominator for sufficiently large t . For $d > 2$, the original bound also seems better since the order of d in the form $6(3^{\frac{d}{2}}\pi + 2d + 3)t$ is smaller in contrast to the order in $(2\sqrt{3})^d(5d+1)t$. However, the addendum $6(3^{\frac{d}{2}}\pi + 2d + 3)d^2$ in the denominator should also be considered.

4.4. Quadrature weights for uniform sampled points

Note that in Lemma 4.3 the points are assumed to be distinct. On the other hand, it is possible that $X_i = X_j, i \neq j$. Therefore, when combining the statements from Sections 4.1, 4.2, and 4.3, it should be understood that the results are first established for the distinct values of the random vectors, i.e. the set formed by the distinct values of $X_1, X_2, \dots, X_n$. Then, for $X_i, X_j, X_k, \dots$ corresponding to the same Voronoi cell, we can simply set one point to have non-zero weight and other points' weights to be zero, or an average weight for each point. It is obvious that choosing either way will not affect the validity of Theorem 2.3.

Now we are ready to present the proof of Theorem 2.3. This immediately follows from Lemma 4.3, Lemma 4.4, Theorem 2.1 and Theorem 2.2. We provide details below.

Proof. [Proof of Theorem 2.3] We first prove Item (i). For any $c > 1$ and $\eta \in (0, 1)$, we have $\delta = \frac{\eta}{\tau_d(t+d^2)}$. Denote the events

$$P_1 := \{\|\mathcal{R}(\mathbf{m}_{n,d})\| \leq \delta\} \text{ and } P_2 := \{\max_i |M_n(X_i) - \mathbb{E}[M_n(X_i)]| < \frac{c}{\sqrt{n}}\}.$$

Their complements are denoted as

$$\bar{P}_1 := \{\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta\} \text{ and } \bar{P}_2 := \{\max_i |M_n(X_i) - \mathbb{E}[M_n(X_i)]| \geq \frac{c}{\sqrt{n}}\}.$$

Then, by Theorem 2.1 and Theorem 2.2, we have

$$\mathbb{P}(\bar{P}_1 \cup \bar{P}_2) \leq \mathbb{P}(\bar{P}_1) + \mathbb{P}(\bar{P}_2) \leq nc_{1,d}(1 - c_{2,d,\delta})^{n-1} + c^{-2},$$

which implies that

$$\mathbb{P}(P_1 \cap P_2) = 1 - \mathbb{P}(\bar{P}_1 \cup \bar{P}_2) \geq 1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1} - c^{-2}.$$

Hence, with probability at least $1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1} - c^{-2}$, both events P_1 and P_2 hold.

Now, from the event P_1 and Lemma 4.4, we see that the M-Z inequality in (4.2) holds for any $P \in \Pi_t^d$ and thus by Part (a) of Lemma 4.3, there exist quadrature weights $a_1, \dots, a_n \in \mathbb{R}$ such that $((a_i, \mathbf{x}_i))_{i=1}^n$ is a t -exact quadrature rule for $\mathbb{S}^d$ satisfying (4.3). Moreover, from the event P_2 , we see that $\mu_d(\mathcal{C}_{i,n}) = M_n(X_i) \leq \max_i M_n(X_i) < \frac{1}{n} + \frac{c}{\sqrt{n}}$, which leads to the result in (2.2).

Next, we prove Item (ii). For the result in (2.3), we need to use the result of Part (b) in Lemma 4.3. To this end, for any $c > 1$ and $\eta \in (0, \frac{1}{2})$, we still set δ to be $\delta := \frac{\eta}{\tau_d(t+d^2)}$. The event $P_3 := \{\|\mathcal{R}(\mathbf{m}_{n,d})\| > \delta\}$ is defined with respect to such a δ . By Theorem 2.2, with probability at least $1 - nc_{1,d}(1 - c_{2,d,\delta})^{n-1}$, the event P_3 holds. Then, by Lemma 4.4 with $\eta \in (0, \frac{1}{2})$, the M-Z inequality in (4.2) holds for any $P \in \Pi_t^d$. Hence, by Part (b) of Lemma 4.3, we see that there exist quadrature weights $b_1, \dots, b_n \in \mathbb{R}$ such that $\{(b_i, \mathbf{x}_i)\}_{i=1}^n$ is a t -exact quadrature rule for $\mathbb{S}^d$ satisfying (4.5), which directly implies $0 \leq b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}$ for all $i \in \{1, \dots, n\}$. This completes the proof of Item (ii) and the proof of the theorem. $\square$

Remark 4.3. As for the non-negative weights $(b_i)_{i=1}^n$, we consider only the event P_3 since Lemma 4.3, Lemma 4.4, Theorem 2.2 together implies the existence of quadrature rule and also (2.3), that is, each b_i satisfies

$$\frac{1-2\eta}{1-\eta}\mu_d(\mathcal{C}_{i,n}) \leq b_i \leq \frac{1}{N_{\lfloor \frac{t}{2} \rfloor}}.$$

The event P_2 in this case is not considered since it can only imply a lower bound for b_i which, by Theorem 2.1, is given by

$$b_i \geq \frac{1-2\eta}{1-\eta}\mu_d(\mathcal{C}_{i,n}) \geq \frac{1-2\eta}{1-\eta} \cdot \frac{1-c\sqrt{n}}{n}.$$

It is negative under our assumptions and is therefore discarded. On the other hand, as mentioned earlier it is possible that some of the sample points X_i are identical and thus the weights are either be divided or set to zero. If a positive lower bound for b_i is given, then either way of adjusting the weights will result in having weights smaller than the given bound. Therefore, we do not consider providing such positive lower bounds. However, if we assume that the values of X_i are distinct, then it is possible to derive such positive lower bounds by applying the Cantelli's inequality which provides one-sided tail bounds.

5. Conclusion and Final Remarks

We have provided concrete probability lower bounds on obtaining exact quadrature rules with bounded weights given n points uniformly sampled on the sphere. The estimates are either directly obtained from the established results in the literature or derived from the probability quantities of the Voronoi cells, which are part of our main contributions. These results in turn provide upper bounds on the number of points to obtain a t -exact quadrature rule on spheres with high probability. Nonetheless, the lower bounds are pessimistic compared with the numerical findings. In contrast, finding possible relations between the number of points n and the degree t might provide more practical significance, for it allows tuning the specific constants in the relations instead of directly adjusting the number n . Our estimates based on the Voronoi tessellations provide a general probability lower bound that allows a unified way to test the validity of such relations.

In particular, we investigated the problem of setting the number of points n to be $n = C_d t^d$, where C_d is independent of t . Simple qualitative analysis based on our probability lower bound suggests that the aforementioned problem is possibly intractable for large t in practice, and our experiments did verify this conjecture. Note that the results are affected by the choices of optimization settings and numerical algorithms, for which we might further investigate using high-performance computing resources to conduct a more thorough test. On the other hand, our lower bounds depict the requirement to satisfy certain sufficient conditions that imply the existence of the t -exact quadrature rules. The failure to satisfy the sufficient conditions does not preclude the existence of the rules, as it has not been established

that these conditions are necessary. Interestingly, the results related to the relation between n and t conform to the conditions, at least empirically, to a certain extent.

Finally, we also showed that by adding logarithmic factors to the relation $n = C_d t^d$, the new relations hold with high probability. Despite being consistent with the results derived from the literature, we did not provide better relations. Nonetheless, our results give a different perspective, which provides some insights into the possibility of having better relations. In detail, we are interested in whether we could replace the logarithmic factors with functions of smaller order. From Remark 2.4 and 2.5, we see that the logarithmic factors come from the first n of $nc_{1,d}(1 - c_{2,d,\delta})^{n-1}$. This n appears due to trivially handling the events of n Voronoi cells. Similar comments are also mentioned in Remark 2.6. Indeed, the n Voronoi cells are not independent, and we might reduce the first n to n^α , $0 < \alpha < 1$ while keeping the exponential tails of the Voronoi cells by exploiting geometric relations among the cells. However, we still can only obtain logarithmic factors since α is absorbed into the constants. Vaguely speaking, it seems not likely to have better relations. Nonetheless, this remains an interesting question to explore.

Acknowledgments

The authors thank the anonymous reviewers for their constructive comments and valuable suggestions that greatly help the improvement of the quality of this article. This work was supported in part by the Research Grants Council of Hong Kong (Project no. CityU 11309122, CityU 11302023, CityU 11301224, and CityU 11300825) and the National Natural Science Foundation of China (NSFC 12471400).

References

- [1] E. Bannai and E. Bannai, A survey on spherical designs and algebraic combinatorics on spheres, *European Journal of Combinatorics* **30**(6) (2009) 1392–1425.
- [2] A. Bondarenko, D. Radchenko and M. Viazovska, Optimal asymptotic bounds for spherical designs, *Annals of Mathematics* (2013) 443–452.
- [3] M. A. Branch, T. F. Coleman and Y. Li, A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems, *SIAM Journal on Scientific Computing* **21**(1) (1999) 1–23.
- [4] R. H. Byrd, R. B. Schnabel and G. A. Shultz, Approximate solution of the trust region problem by minimization over two-dimensional subspaces, *Mathematical Programming* **40**(1) (1988) 247–263.
- [5] J. da Cruz Neto, L. De Lima and P. Oliveira, Geodesic algorithms in Riemannian geometry, *Balkan Journal of Geometry and Its Applications* **3**(2) (1998) 89–100.
- [6] F. Dai and Y. Xu, *Approximation Theory and Harmonic Analysis on Spheres and Balls* (Springer, 2013).
- [7] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, 1992).
- [8] L. Devroye, L. Györfi, G. Lugosi and H. Walk, On the measure of Voronoi cells, *Journal of Applied Probability* **54**(2) (2017) 394–408.
- [9] H. Feng, S. Huang and D.-X. Zhou, Generalization analysis of CNNs for classification on spheres, *IEEE Transactions on Neural Networks and Learning Systems* **34**(9) (2021) 6200–6213.

- [10] F. Filbir, R. Hielscher, T. Jahn and T. Ullrich, Marcinkiewicz–Zygmund inequalities for scattered and random data on the q -sphere, *Applied and Computational Harmonic Analysis* **71** (2024) p. 101651.
- [11] F. Filbir and H. N. Mhaskar, Marcinkiewicz–Zygmund measures on manifolds, *Journal of Complexity* **27**(6) (2011) 568–596.
- [12] F. Filbir and W. Themistoclakis, Polynomial approximation on the sphere using scattered data, *Mathematische Nachrichten* **281**(5) (2008) 650–668.
- [13] Q. L. Gia, I. H. Sloan and H. Wendland, Multiscale analysis in Sobolev spaces on the sphere, *SIAM Journal on Numerical Analysis* **48**(6) (2010) 2065–2090.
- [14] M. Gräf, S. Kunis and D. Potts, On the computation of nonnegative quadrature weights on the sphere, *Applied and Computational Harmonic Analysis* **27**(1) (2009) 124–132.
- [15] B. Han, *Framelets and Wavelets: Algorithms, Analysis, and Applications*, 1st edn. (Birkhäuser Basel, 2018).
- [16] K. Hesse and R. S. Womersley, Numerical integration with polynomial exactness over a spherical cap, *Advances in Computational Mathematics* **36** (2012) 451–483.
- [17] R. B. Holmes, *Geometric Functional Analysis and its Applications*, volume 24 (Springer Science & Business Media, 2012).
- [18] Z. Kabluchko and C. Thäle, The typical cell of a Voronoi tessellation on the sphere, *Discrete & Computational Geometry* **66** (2021) 1330–1350.
- [19] J. Keiner, S. Kunis and D. Potts, Efficient reconstruction of functions on the sphere from scattered data, *Journal of Fourier Analysis and Applications* **13** (2007) 435–458.
- [20] J. Keiner and D. Potts, Fast evaluation of quadrature formulae on the sphere, *Mathematics of Computation* **77**(261) (2008) 397–419.
- [21] S. Kunis and D. Potts, Fast spherical Fourier algorithms, *Journal of Computational and Applied Mathematics* **161**(1) (2003) 75–98.
- [22] Q. T. Le Gia and H. N. Mhaskar, Localized linear polynomial operators and quadrature formulas on the sphere, *SIAM Journal on Numerical Analysis* (2008) 440–466.
- [23] J. Li, H. Feng and X. Zhuang, Convolutional neural networks for spherical signal processing via area-regular spherical Haar tight framelets, *IEEE Transactions on Neural Networks and Learning Systems* **35**(4) (2022) 4400–4410.
- [24] S. Li, Concise formulas for the area and volume of a hyperspherical cap, *Asian Journal of Mathematics and Statistics* **4**(1) (2011) 66–70.
- [25] S.-B. Lin, D. Wang and D.-X. Zhou, Sketching with spherical designs for noisy data fitting on spheres, *SIAM Journal on Scientific Computing* **46**(1) (2024) A313–A337.
- [26] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way* (Academic Press, 2008).
- [27] T. Mao, Z. Shi and D.-X. Zhou, Approximating functions with multi-features by deep convolutional neural networks, *Analysis and Applications* **21**(01) (2023) 93–125.
- [28] H. N. Mhaskar, F. J. Narcowich and J. D. Ward, Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature, *Mathematics of Computation* **70**(235) (2001) 1113–1130.
- [29] H. N. Mhaskar, F. J. Narcowich and J. D. Ward, Corrigendum to “spherical Marcinkiewicz–Zygmund inequalities and positive quadrature”, *Mathematics of Computation* **71**(237) (2002) 453–454.
- [30] C. Müller, *Spherical Harmonics*, volume 17 (Springer, 2006).
- [31] F. J. Narcowich, P. Petrushev and J. D. Ward, Localized tight frames on spheres, *SIAM Journal on Mathematical Analysis* **38**(2) (2006) 574–594.
- [32] P. Petersen, *Riemannian Geometry*, volume 171 (Springer, 2006).
- [33] A. Reznikov and E. B. Saff, The covering radius of randomly distributed points on a

- manifold, *International Mathematics Research Notices* **2016**(19) (2016) 6065–6094.
- [34] I. H. Sloan, Polynomial interpolation and hyperinterpolation over general regions, *Journal of Approximation Theory* **83**(2) (1995) 238–254.
- [35] Y. G. Wang and X. Zhuang, Tight framelets and fast framelet filter bank transforms on manifolds, *Applied and Computational Harmonic Analysis* **48**(1) (2020) 64–95.
- [36] Y. Xiao and X. Zhuang, Spherical framelets from spherical designs, *SIAM Journal on Imaging Sciences* **16**(4) (2023) 2072–2104.
- [37] R. Zheng and X. Zhuang, On the existence and estimates of nested spherical designs, *Applied and Computational Harmonic Analysis* (2024) p. 101672.