arXiv:2303.01460v1 [math.NA] 2 Mar 2023

球的并集上的压缩的 QMC 体积与曲面积分

Giacomo Elefante[a]、Alvise Sommariva[a]、Marco Vianello[a]

[a] *University of Padova, Italy*

摘要

本文讨论一种算法，用于对球的并集（多气泡）上的准蒙特卡罗（QMC）体积／曲面积分进行 Tchakaloff 型压缩。关键工具包括 Davis-Wilhelmsen 定理，该定理针对多项式空间上正线性泛函的所谓“Tchakaloff 集”；以及用于非负最小二乘（NNLS）的 Lawson-Hanson算法。我们提供了相应的 Matlab 软件包，并给出若干示例。

关键词： 多气泡、球的并集、准蒙特卡罗公式、体积分、曲面积分、Tchakaloff 集、Davis-Wilhelmsen 定理、求积压缩、非负最小二乘。*2020* 年数学学科分类： 65D32。

1. 引言

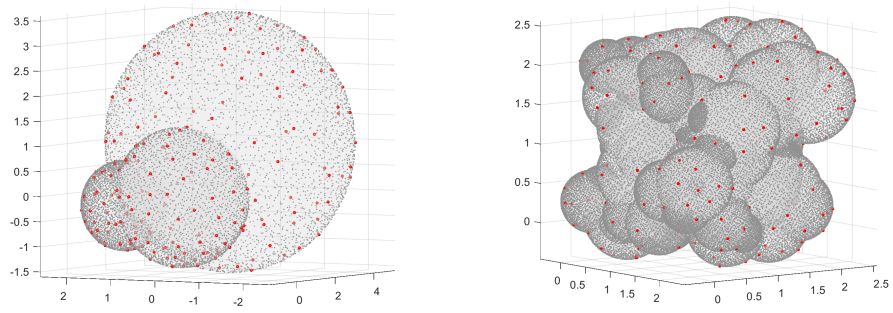
由有限个圆盘、球和球面构成的数值建模，与多个应用领域息息相关。例如，分子建模、计算几何、计算光学和无线网络分析中都会涉及这类几何对象的交集、并集与差集问题；参见 [1, 2, 12, 16, 19, 22] 及其中的参考文献。一个基本问题是计算这类集合的面积和体积，随后更具挑战的任务则是通过合适的求积公式计算其上的体积积分和曲面积分。

事实上，近期已有研究处理平面圆盘交集与并集上的数值求积问题， [28, 30], 给出了具有正权重和内部节点的低基数代数公式。另一方面，尽管已有一些文献（主要涉及分子建模领域）研究任意球的并集（多气泡）的体积和表面积计算，但据我们所知，目前尚无针对这类域的专用数值积分程序。

本文沿着 [13]的思路，开始填补这一空白，为多气泡上的体积和曲面积分提供压缩的准蒙特卡罗（QMC）公式。这类公式能够保留 *QMC* 的逼近能力，使误差达到给定次数下被积函数的最佳一致多项式逼近误差，同时大幅减少采样点的数量；图1展示了两个多气泡及其 QMC 采样压缩的示例。关键工具包括针对正线性泛函的所谓“切卡洛夫集”的 Davis-Wilhelmsen 定理，以及用于非负最小二乘的 Lawson-Hanson算法；后者能够从庞大的低差异序列中提取一组经重新加权的“等效”节点。

我们强调，与 [13], 不同，本文方法不仅能够压缩 QMC 体积积分，还能压缩代数曲面紧子集上的 QMC 积分（特别是多气泡的曲面，它是球体并集的一个子集）。请注意，曲面积分实例中的主要难点之一，是调整压缩算法，使其适用于限制在代数簇上的多项式空间，并找到合适的多项式基。据我们所知，本文是在 QMC 框架内朝这一方向进行的首次尝试。

*通讯作者：Alvise Sommariva*Email addresses:* giacomo.elefante@unipd.it (Giacomo Elefante), alvise@math.unipd.it (Alvise Sommariva),marcov@math.unipd.it (MarcoVianello)



(a) (b)

图1：在球并集曲面上，次数为 $n = 9$ 时，从低差异点（灰色）中提取的压缩 QMC 点（红色）。左：从约 8200 个点中提取 200 个点（3个球），压缩比为 43；右：从约 69000 个点中提取 220 个点（100个球），压缩比超过 300。

本文结构如下。第2节讨论 $\mathbb{R}^3$中体积积分和曲面积分的 QMC 压缩所涉及的理论与计算问题。第3节介绍我们在三维多气泡上的实现，并展示若干数值测试。开源代码可在 [14]免费获取。

2. 压缩准蒙特卡罗公式

QMC 公式的压缩不过是离散测度压缩的一个特例。这一主题在过去十年的文献中日益受到关注，无论是在概率框架还是确定性框架下都是如此。事实上，已有多篇论文和一些软件致力于从高基数离散测度的支撑集中提取较小的一组重新加权的质量点，同时确保其矩保持到给定的多项式次数；例如参见 [15, 20, 23, 27, 33] 及其中的参考文献。

从求积的角度来看，这一主题与著名的 Tchakaloff 定理[32] 密切相关。该定理证明了存在具有正权重的低基数公式。另一方面，Tchakaloff 定理本身包含在 Wilhelmsen [34]提出的一项更深刻、却在某种程度上被忽视的结果之中；该结果涉及有限维函数空间上正线性泛函的离散表示（推广了 Davis [5]先前的一项结果）。事实上，直到最近，这一定理才作为通过自适应非负最小二乘矩匹配实现正权积分类的一项基本工具而重新受到关注，参见 [13, 18, 29, 31]。

定理1. *(Davis，1967年 - Wilhelmsen，1976)* 设 Ψ 是定义在紧集$\Omega \subset \mathbb{R}^d$上的连续实值线性无关函数 $\{\phi_j\}_{j=1,\dots,N}$ 的线性张成空间。假设 Ψ 满足克雷因条件（即存在至少一个在Ω上不为零的 $f \in \Psi$ ），且 L 是 Ψ上的正线性泛函，即对于任意在Ω上不处处为零的 $f \in \Psi$、 $f \geq 0$ ，均有 $L(f) > 0$ 。

若 $\{P_i\}_{i=1}^{\infty}$ 是Ω中的处处稠密子集，则当 m足够大时，集合$X_m = \{P_i\}_{i=1,\dots,m}$是一个切卡洛夫集，也就是说，存在权重 $w_k > 0$、 $k = 1$、.. 、 ν以及节点$\{\mathcal{T}_k\}_{k=1,\dots,\nu} \subset X_m \subset \Omega$，满足$\nu = \mathit{card}(\{\mathcal{T}_k\}) \leq N$，使得

$$L(f) = \sum_{k=1}^{\nu} w_k f(\mathcal{T}_k) \ , \quad \forall f \in \Psi \ . \tag{1}$$

由此可立即得到以下结论

推论1. 设 λ 是定义在 Ω 上的正测度，且 $\mathrm{supp}(\lambda)$ 是 $\mathbb{P}_n^d(\Omega)$ 的决定集；这里 $\mathbb{P}_n^d(\Omega)$ 是次数不超过 n、限制在Ω 上的全次数多项式空间（即，在 Ω 上消失的 $\mathbb{P}_n^d(\Omega)$ 中的多项式必处处为零）。那么，定理*1*的结论对 $L(f) = \int_\Omega f \, d\lambda$ 成立。

事实上，非负且不恒为零的多项式 $f \in \mathbb{P}_n^d(\Omega)$ 的积分必为正（否则 f 会在 $supp(\lambda)$ 上消失）。注意，Tchakaloff 定理的经典版本对应于

$$L(f) = L_{\text{INT}}(f) = \int_\Omega f(P)\, dP \,,$$

与 $\Psi = \mathbb{P}_n^d(\Omega)$ 和

$$N = N_n^d = dim(\mathbb{P}_n^d(\Omega)) \,. \tag{2}$$

从现在起，我们将重点讨论三维情形 $(d = 3)$，尽管大多数讨论都可以推广到一般维数。请注意，Davis-Wilhelmsen 定理的表述足够一般，能够涵盖体积分，即 Ω 是某个有界开集的闭包，且 $N = dim(\mathbb{P}_n^d(\mathbb{R}^3)) = \binom{n+3}{3} = (n+1)(n+2)(n+3)/6$；也涵盖代数簇紧子集上的曲面积分（此时，曲面测度为 $dP = d\sigma$）。在后一种情形下，多项式空间的维数可能会退化；例如，当 $\Omega = S^2 \subset \mathbb{R}^3$ 时，我们有 $N = (n+1)^2 < \binom{n+3}{3} = (n+1)(n+2)(n+3)/6$。

在另一面 ，Wilhelmsen 定理也可应用于准蒙特卡罗公式这类离散泛函 应用于 $f \in C(\Omega)$

$$L(f) = L_{\text{QMC}}(f) = \frac{\mu(\Omega)}{M} \sum_{i=1}^{M} f(P_i) \approx \int_\Omega f(P)\, dP \,, \tag{3}$$

其中

$$X_M = \{P_i\}_{i=1,\dots,M} \,, \quad M > N \,,$$

是在 Ω 上的 低差异序列 ，而 $\mu(\Omega)$ 可以是体积，也可以是表面积。通常，在包围 Ω 的包围盒或包围曲面 $\mathcal{B} \supset$上生成一个基数为 M_0 的低差异序列，再通过合适的域内算法从中提取 Ω 上的低差异序列。我们注意到，如果 $\mu(\Omega)$ 未知或难以计算（如多气泡的情况），则可将其近似为 $\mu(\Omega) \approx \mu(\mathcal{B})M/M_0$。

当集合 X_M 是 $\mathbb{P}_n^3(\Omega)$-决定集时，泛函对于 $f \in \Psi = \mathbb{P}_n^3(\Omega)$ 的正性即可得到保证；也就是说，在该集合上消失的多项式在 Ω 上处处消失，等价地， $dim(\mathbb{P}_n^3(X_M)) = N = dim(\mathbb{P}_n^3(\Omega))$，甚至

$$rank(V_M) = N \,, \tag{4}$$

其中

$$V_M = V^{(n)}(X_M) = [\phi_j(P_i)] \in \mathbb{R}^{M \times N} \tag{5}$$

是相应的矩形类Vandermonde矩阵。注意，由于 X_M 是一个序列，对于每个$k \le M$，我们都有

$$V_k = V^{(n)}(X_k) = [(V_M)_{ij}] \,, \ 1 \le i \le k \,, \ 1 \le j \le N \,. \tag{6}$$

实际上，对 V_M 的满秩要求并不苛刻。事实上，对于均匀分布点， $det(V_N) = 0$成立的概率为零，因为前一个方程定义了Ω^N中一个多项式的零点集，其乘积测度为零（关于这一点的更完整讨论，参见例如 [21, §§3-4] ）。

根据定理1，当 $M \gg N$ 时，我们可以尝试寻找一个切卡洛夫集 X_m，其中 $N \le m < M$，使得存在一个稀疏非负解向量 u ，满足欠定的矩匹配系统

$$V_m^t u = p = V_M^t e \,, \quad e = \frac{\mu(\Omega)}{M} (1, \dots, 1)^t \,. \tag{7}$$

在实 际应用中，我们通过 Lawson-Hanson 主动集方法 [17] 求解（7），该方法用于求解非负最小二乘问题

$$\min_{u \ge 0} \|V_m^t u - p\|_2 \,, \tag{8}$$

当残差大小较小时接受该解，例如

$$\|V_m^t u - p\|_2 < \varepsilon \tag{9}$$

其中， ε 是给定的容差。随后， u 的非零分量确定从 X_m中提取的压缩 QMC 公式的节点和权重，即 $\{w_k\} = \{u_i : \ u_i > 0\}$ 和 $\{\mathcal{T}_k\} = \{P_i : \ u_i > 0\}$，得到

$$L^*_{\mathrm{QMC}}(f) = \sum_{k=1}^{\nu} w_k f(\mathcal{T}_k) \ , \ \ \nu \le N \ll M \ \ , \tag{10}$$

其中，对于每个 $f \in \mathbb{P}^3_n(\Omega)$， $L^*{}_{\mathrm{QMC}}(f) = L_{\mathrm{QMC}}(f)$。

请注意，根据有限维锥组合的 Carathéodory 定理，将该定理应用于 V_M^t 的各列，即可保证 $m = M$ 存在类似于 (10) 的表示（参见 [23] ，其中对离散测度压缩的一般框架下这一点进行了全面讨论）。然而，这样一来，我们就必须处理一个大得多的矩阵，也就是说，必须直接求解

$$\min_{u \ge 0} \|V_M^t u - p\|_2 \ . \tag{11}$$

相反，沿着一系列规模递增的小问题 $m := m_1, \ m_2, \ m_3, \ \ldots\ldots$求解 (8)，并令$m_1 < m_2 < m_3 < \cdots \le M$，

$$\min_{u \ge 0} \|V_{m_j}^t u - p\|_2 \ , \ \ j = 1, 2, 3, \ldots \ , \ m_1 \ge N \ , \tag{12}$$

这些问题对应于越来越稠密的子集 $X_{m_1} \subset X_{m_2} \subset \cdots \subseteq X_M$（可以称为“自底向上”），持续进行直至残差足够小，这样做可以显著降低计算成本。事实上，如 [13], 所示，若适当选择序列 $\{m_j\}$ ，经过少数几次迭代后，残差便会变得极小，最终提取的基数也远小于 M。

关于 QMC 压缩的逼近能力，按照 [13], 很容易推导出以下误差估计

$$\begin{aligned} |L^*_{\mathrm{QMC}}(f) - L_{\mathrm{INT}}(f)| &\le \mathcal{E}_{\mathrm{QMC}}(f) + 2\,\mu(\Omega)\,E_n(f;X) \\ &\le \mathcal{E}_{\mathrm{QMC}}(f) + 2\,\mu(\Omega)\,E_n(f;\Omega) \ , \end{aligned} \tag{13}$$

对每个 $f \in C(\Omega)$ 均成立，其中 $\mathcal{E}_{\mathrm{QMC}}(f) = |L_{\mathrm{QMC}}(f) - L_{\mathrm{INT}}(f)|$ 且定义 $E_n(f;\ K) = \inf_{\phi \in \mathbb{P}^3_n(K)} \|f - \phi\|_{\infty,\ K}$其中 K 为离散或连续紧集。

(13) 的含义是，压缩后的 QMC 泛函 $L^*{}_{\mathrm{QMC}}$ 保留了原始准蒙特卡罗公式的逼近能力，误差至多相差一个与 f在 X 上一致范数下的最佳多项式逼近误差成正比的量（因此，根据包含关系，它也与在 Ω 上一致范数下的最佳多项式逼近误差成正比）。我们回顾一下，后者可根据 f 的正则性，通过多元 Jackson 型定理进行估计；例如，参见 [24] ，其中讨论的是 Ω 为有界开集闭包时的体积分。

另一方面，这里不深入探讨 QMC 收敛与误差估计这一内容丰富且已有大量研究的主题，仅回顾一点：（粗略而言）对于光滑函数, QMC 误差 $\mathcal{E}_{\mathrm{QMC}}(f)$ 接近于 $\mathcal{O}(1/M)$，可与蒙特卡罗的 $\mathcal{O}(1/\sqrt{M})$ 误差相比较。关于 QMC 理论的基本概念与结果，例如差异度、星差异度、Hardy-Krause 变差、Erd¨os-Tur´an-Koksma 不等式和 Koksma-Hlawka 不等式，读者可参阅专门的综述，例如 [11]。

注 1. QMC 压缩算法可以轻松扩展到 Ω（无论是体积还是曲面）为若干不相交子集的有限并的情形，例如 $\Omega = \cup_{\ell=1}^{L} \Omega_\ell$，且已知包围集 $\mathcal{B}_\ell \supset \Omega_\ell$ 上的低差异点序列。在这种情况下，整体QMC点为 $X = \cup_{\ell=1}^{L} Y_\ell$，其中 $Y_\ell = \{P_{\ell,i}\}_{i=1}^{M_\ell}$ 和 $M = card(X) = \sum_{\ell=1}^{L} M_\ell$，而 Y_ℓ 是位于 Ω_ℓ 中的 $\mathcal{B}_\ell$ 的低差异点。我们强调，必须交替选择低差异点，以便在整个 Ω 上构造均匀分布序列 X_M : 先选取每个 Ω_ℓ 中的第一个点，再选取每个 Ω_ℓ 中的第二个点，依此类推，即序列 $\{P_{1,1}, P_{2,1}, \ldots, P_{L,1}, P_{1,2}, P_{2,2}, \ldots, P_{L,2}, \ldots\}$。

此外，利用积分的可加性，QMC泛函变为

$$L_{QMC}(f) = \sum_{\ell=1}^{L} \sum_{i=1}^{M_\ell} w_{\ell,i} f(P_{\ell,i}) \approx \sum_{\ell=1}^{L} \int_{\Omega_\ell} f(P)\,dP = \int_{\Omega} f(P)\,dP \ , \tag{14}$$

其中 $w_{\ell i} = \mu(\Omega_\ell)/M_\ell$, $i = 1, \ldots, M_\ell$, 因此必须使用这样的权重来计算式 (7) 中的 QMC 矩。

3. 实现与数值测试

为了展示上一节所述自底向上压缩过程的有效性，我们简要勾勒一种可能的实现方式，并介绍一些针对任意球并集（多气泡）的体积和曲面积分数值测试。

事实上，我们将通用软件包 $dCATCH$ [10], 中实现的多元离散测度“Caratheodory-Tchakaloff”压缩方法与自底向上方法进行比较。所有测试均在配备 48 GB 内存的 AMD Ryzen 5 3600 CPU 上运行 Matlab R2022a 完成。Matlab 代码和演示程序汇集于名为 $Qbubble$ 的软件包中，可从 [14] 免费获取。

下面首先介绍已实现的多气泡算法的主要特点，具体如下:

- 对于多气泡体积分，我们只需取较小包围盒内的 Halton 点

$$[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] \supset \Omega$$

并选取其中属于 Ω 的点；对于多气泡曲面积分，我们按照注 1 中概述的流程，在每个球面上通过保面积变换（见下文第 3.2 节中的 (19)）取 $\mathcal{B}_\ell$ 低差异映射 Halton 点，然后选取其中属于曲面的点；

- 鉴于标准单项式基的条件数极高，我们从 Ω 较小包围盒上的乘积型切比雪夫全次数基出发（无论是体积还是曲面），即

$$p_j(x, y, z) = T_{\alpha_1(j)}(\sigma_1(x))\, T_{\alpha_2(j)}(\sigma_2(y))\, T_{\alpha_3(j)}(\sigma_3(z)) \ , \quad j = 1, \ldots, J \ ,$$

其中 $J = (n+1)(n+2)(n+3)/6$, $\sigma_i(t) = \frac{2t - b_i - a_i}{b_i - a_i}, \quad i = 1, 2, 3$, 且 $j \mapsto \alpha(j)$ 对应于三元组 $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ 的分次字典序, $0 \le \alpha_1 + \alpha_2 + \alpha_3 \le n$;

- 对于曲面积分，我们通过计算秩，并在必要时对三元切比雪夫-范德蒙德矩阵进行列主元 QR 分解以选择列，从而确定合适的多项式基；

- 为应对范德蒙德类矩阵 V_{m_j} （其条件数会随次数增加而增大）的病态性，我们进行一次带列主元的 QR 分解 $V_{m_j} = Q_{m_j} R_{m_j}$，以构造关于离散内积 $\langle f, g \rangle_{X_{m_j}} = \sum_{i=1}^{m_j} f(P_i) g(P_i)$ 的正交多项式基，并在式（12）中用 Q_{m_j} 替换 V_{m_j}；因此，式（7）中的 QMC 矩 p 必须通过高斯消元修改为 $(R_{m_j}^{-1})^t p$ ；

- 采用最新实现的 Lawson–Hanson 主动集方法 LHDM 求解（经修改的）自底向上非负最小二乘问题 (12)。该方法基于“偏差最大化”而非“列主元选取”来进行底层 QR 分解；实验表明，与标准 Matlab 函数 lsqnonneg 相比，它的加速比至少为 2（参见 [6, 8, 9]）。

在接下来的小节中，我们将介绍若干数值测试，以展示自底向上方法对多气泡进行体积和曲面 QMC 压缩的有效性。

3.1. 多气泡上的体积分

In 本小节考虑球的并集（实心多气泡）上的体积分，记为 ly

$$\Omega = \bigcup_{j=1}^{s} B(C_j, r_j) \tag{15}$$

其中 $B(C_j, r_j) \subset \mathbb{R}^3$ 是以 C_j 为球心、半径为 r_j的闭三维球。这里，我们在 Ω 的最小笛卡尔包围盒中生成一列 Halton点，然后通过检查 $\|P_i - C_j\|_2 \le r_j$ 是否对某个 j成立，选出属于该并集的点，记为 $X = \{P_i\}$。

更确切地说，我们考虑以下内容

- 第一个示例：由球心为 $C_1 =$(0，0，0)、 $C_2 =$(0，1．3，-0.2)、 $C_3 =$(2．5，0，1) 且半径分别为 $r_1 = 1.4$、 $r_2 = 0.9$、 $r_3 = 1$ 的 3个球的并集；
- 第二个示例：由随机选取后固定的球心和 $[0, 2]^3$ 中的半径构成的100个球的并集$[0, 2, 0.6]$。

关于自底向上方法应用的结果汇总于表1–2中。我们对包含一百多万个Halton点的QMC体积积分进行压缩，同时保留次数最高为3、6、9、12、15的多项式矩（这些矩对应于球并集的最小笛卡尔包围盒上的乘积切比雪夫基）。

我们从包围盒中的2,400,000个Halton点出发，并设定 $m_1 = 2N$ 和 $m_{j+1} = 2m_j$、 $j \geq 1$。残差容差为 $\varepsilon = 10^{-10}$。我们将当前的自底向上压缩算法（简称 Q_c^{bu}）与一种作用于完整Halton序列 X_M的全局压缩算法进行比较；后者是 $dCATCH$ 中开发的通用离散测度压缩器 [10], ，其基本上通过 [27, 23]提出的Carathéodory-Tchakaloff子采样直接求解式（11）。

具体而言，我们展示了基数和压缩比、构造低差异序列（CPU Halton 序列）所需的 CPU 时间，以及计算压缩求积公式所需的时间。在当前次数范围内，新算法的加速比约为 6 至超过 24，并且最多经过 3 次迭代即可确保矩残差始终低于所需容差。值得强调的是，在 [13], 中已观察到一种现象： Q_c^{dCATCH}可能会失败，在某些情况下产生的残差远大于 Q_c^{bu}。

为检验 QMC 压缩求积公式的多项式精确性，我们在图2–3中展示了 100 次试验中多项式的 QMC 相对压缩误差及其对数平均值（即误差对数之和除以试验次数）：

$$g(P) = (ax + by + cz + d)^n \ , \quad P = (x, y, z) \tag{16}$$

其中 a，b，c，d是 $[0, 1]$中的均匀随机变量。此外，表3–4展示了三个正则性不同的测试函数的积分相对误差，具体为

$$\begin{aligned} f_1(P) &= |P - P_0|^5 \\ f_2(P) &= \cos(x + y + z) \\ f_3(P) &= \exp(-|P - P_0|^2) \end{aligned} \tag{17}$$

其中 $P_0 =$(0，0，0) $\in \Omega$，第一个属于 C^4 类，具有不连续的五阶导数，而第二个和第三个是解析函数。积分的参考值是从包围盒中的 10^8 Halton点出发，使用准蒙特卡罗公式计算得到的。

我们看到，基于超过一百万个点的压缩公式具有量级相近的误差；正如估计式（13）所预期的那样，随着多项式压缩次数增加，误差逐渐减小，直至接近 QMC 误差（但请注意，在表3中，次数为 $n = 15$时，只有自底向上算法的误差已达到 QMC 误差的量级）。

3.2. 多气泡上的曲面积分

现在我们转向曲面积分，考虑域 Ω，它是任意球并集的边界，即

$$\Omega = \partial \bigcup_{j=1}^{s} B(C_j, r_j) = \bigcup_{j=1}^{s} \partial B(C_j, r_j) \backslash \bigcup_{j=1}^{s} \overset{\circ}{B}(C_j, r_j) \ , \tag{18}$$

也就是说，该集合包含所有位于某个球面 ∂B(${C_j}$，${r_j}$)、 $j = 1$，.....，s 上，但不在任何一个球内部的点B(${C_k}$，${r_k}$)、 $k \neq j$。我们给出两个例子，它们对应于上文体积分中考虑的相同球心和半径，即第3.1节中3个球和100个球的并集的表面。注意，Ω是一个代数曲面的子集，也就是相应球面的并集。虽然可以通过代数几何方法从理论上计算多项式空间的维数（参见例如 [4]），但我们不深入

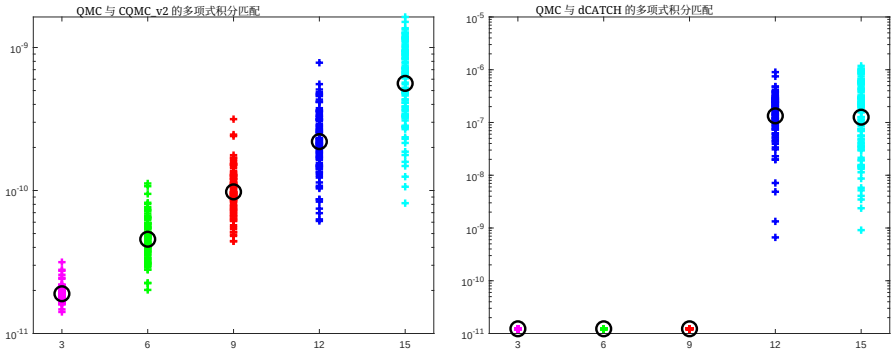


图 2：在 3 个球的并集上，对随机多项式进行 100 次试验，比较自底向上算法（左）和 $dCATCH$ （右）的 QMC 相对压缩误差及其对数平均值（圆圈）。请注意，左右图的刻度不同。

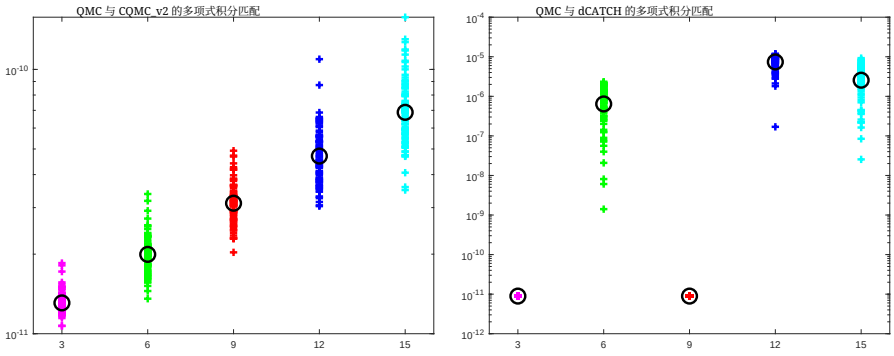


图 3：在 100 个球的并集上，对随机多项式进行 100 次试验，比较自底向上算法（左）和 $dCATCH$（右）的 QMC 相对压缩误差及其对数平均值（圆圈）。请注意，左右图的刻度不同。

deg	3	6	9	12	15
基数 QMC	$M = 1{,}128{,}709$				
基数 Q_c^{dCATCH}	20	84	220	452	806
基数 Q_c^{bu}	20	84	220	455	816
压缩比	5.6e+04	1.3e+04	5.1e+03	2.5e+03	1.4e+03
CPU Halton 序列	9.0e-01s				
CPU Q_c^{dCATCH}	3.4e+00秒	1.9e+01秒	4.9e+01秒	1.4e+02秒	3.1e+02秒
CPU Q_c^{bu}	2.2e-01秒	9.0e-01秒	2.4e+00秒	5.7e+00秒	2.6e+01秒
加速比	15.4	21.1	20.5	24.4	11.9
矩残差 Q_c^{dCATCH}	8.9e-12	8.9e-12	8.9e-12	⋆ 5.1e-06	⋆ 1.1e-05
矩残差 Q_c^{bu}					
迭代 1	4.55e-16	1.51e-02	1.63e-01	3.81e-01	7.12e-01
迭代 2		1.12e-15	1.85e-15	3.62e-15	8.06e-15

Ta表 1：3 个球的并集示例，位于包含 2,400,000 个低差异点的包围盒中。

deg	3	6	9	12	15
基数 QMC	$M = 1{,}195{,}806$				
基数 Q_c^{dCATCH}	20	83	220	450	795
基数 Q_c^{bu}	20	84	220	455	816
压缩比	5.6e+04	1.3e+04	5.1e+03	2.8e+03	1.5e+03
CPU Halton 序列	1.3e+00s				
CPU Q_c^{dCATCH}	3.4e+00s	2.3e+01s	6.5e+01s	1.5e+02s	3.7e+02s
CPU Q_c^{bu}	2.5e-01s	8.7e-01s	2.6e+00s	9.5e+00s	6.7e+01s
加速比	13.8	26.6	25.0	15.7	5.6
矩残差 Q_c^{dCATCH}	1.1e-11	⋆ 1.2e-05	1.1e-11	⋆ 5.6e-05	⋆ 7.3e-05
矩残差 Q_c^{bu}					
迭代 1	2.08e-16	9.41e-02	4.99e-01	1.51e+00	1.78e+00
迭代 2		1.32e-15	2.20e-15	4.72e-15	8.30e-02
迭代 3					7.32e-15

表2：使用 2,400,000 个 Halton点，在一个包围盒中计算 100 个球的并集的示例。

这里的细微问题在于，算法通过对一个类Vandermonde矩阵采用揭示秩的方法，数值计算出这一维数。

在这种情况下，我们采用了注 1 中讨论的扩展方法：在每个球面 $\partial B(C_j,\ r_j)$ 上取大量低差异点，并从中选出属于对并集表面有贡献的球面部分（即不位于任何其他球体内部的部分），从而在整个 Ω 上构造均匀分布序列 X_M。具体而言，我们在每个球面上通过保面积变换，将矩形 $[-1,1]\times[0,2\pi]$ 中的Halton点映射得到的点

$$(t,\phi)\mapsto C_j + r_j(\sqrt{1-t^2}\,\cos(\phi),\sqrt{1-t^2}\,\sin(\phi),t)\ , \tag{19}$$

该变换也保留了低差异性。最后，通过在并集表面的每个有效部分交替选取一个点，并为每个点附加一个局部权重，对这些点进行排序。图1展示了压缩点的示例，这些点由每个球面上 4000 个映射后的 Halton 点提取得到。

在表5–6中，我们针对这些曲面积分示例报告与表1–2中体积分所列相同的量。我们再次使用 [9] 中的 $dCATCH$ code来压缩整个 X_M上的QMC公式，因为该算法的设计初衷也是处理可能受限于代数曲面的多项式空间。在这里，对于3个球的示例，我们从每个球面上映射得到的500,000个Halton点开始；对于100个球的示例，则从60,000个点开始，从而在相应的球并集曲面上得到约一百万个低差异点的序列。和之前一样，我们设置 $m_{j+1}=2m_j$、$j\ge 1$，其中 $m_1=2N$ 和 $\varepsilon=10^{-10}$。

我们再次获得了令人瞩目的压缩比，加速比从约 5 到 16 以上不等。此外，自底向上算法始终能使残差低于给定容差，而 $dCATCH$

deg	3	6	9	12	15
$E^{QMC}(f_1)$	3.5e-04				
$E^{dCATCH}(f_1)$	1.3e-01	3.4e-04	3.5e-04	3.5e-04	3.5e-04
$E^{bu}(f_1)$	2.3e-03	3.2e-04	3.5e-04	3.5e-04	3.5e-04
$E^{QMC}(f_2)$	7.3e-04				
$E^{dCATCH}(f_2)$	2.4e+00	7.0e-02	4.3e-03	7.3e-04	7.3e-04
$E^{bu}(f_2)$	7.5e-01	3.7e-03	4.8e-04	7.4e-04	7.3e-04
$E^{QMC}(f_3)$	8.7e-05				
$E^{dCATCH}(f_3)$	7.1e-01	1.4e-01	9.4e-03	2.1e-03	1.1e-04
$E^{bu}(f_3)$	5.8e-01	2.8e-02	1.5e-02	9.5e-04	2.5e-05

表3：示例 e 对于 3个球（参考值通过从 10^8 Halton点开始进行 QMC 计算得出，位于包围盒）。

deg	3	6	9	12	15
$E^{QMC}(f_1)$	1.1e-04				
$E^{dCATCH}(f_1)$	8.3e-02	8.8e-05	1.1e-04	1.1e-04	1.1e-04
$E^{bu}(f_1)$	1.7e-03	9.8e-05	1.1e-04	1.1e-04	1.1e-04
$E^{QMC}(f_2)$	1.7e-04				
$E^{dCATCH}(f_2)$	2.9e-01	8.7e-04	1.6e-04	1.7e-04	1.7e-04
$E^{bu}(f_2)$	5.6e-02	1.5e-04	1.7e-04	1.7e-04	1.7e-04
$E^{QMC}(f_3)$	2.2e-04				
$E^{dCATCH}(f_3)$	2.3e-01	2.3e-03	8.4e-04	2.3e-04	2.2e-04
$E^{bu}(f_3)$	6.1e-03	3.6e-03	1.2e-04	2.3e-04	2.2e-04

表4：包含100个球的示例（参考值通过从包围盒中的 10^8 Halton点开始进行 QMC 计算得到）。

事实证明，其更容易失败（请参见包含3个球的示例中次数为 $n = 15$ 的情形，以及包含100个球的示例中次数为 $n = 9$, 15的情形下的残差）。

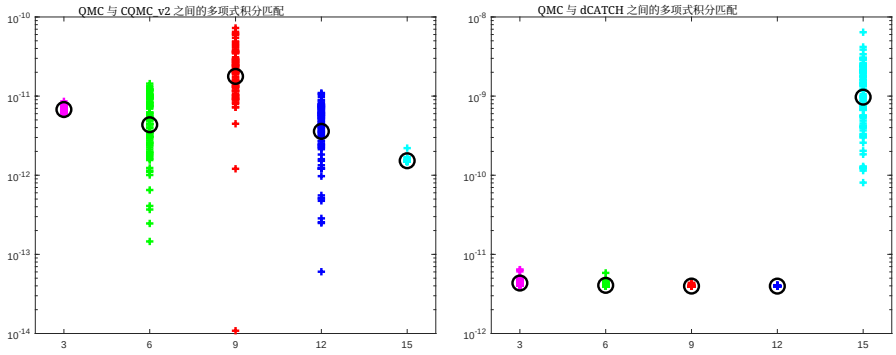
随机多项式（16）在球并集边界上的曲面积分的对数平均误差绘于图4–5中。在表7–8中，我们展示了式（17）中三个测试函数的曲面积分误差，其中 P_0 是球并集曲面上适当选取的一个点。我们再次看到，在超过一百万个点上，压缩公式的误差具有相近的数量级；正如估计式（13）所预期的那样，随着多项式压缩次数的增加，误差逐渐减小，直至达到接近 QMC 误差的大小。

4. 致谢

本研究部分由帕多瓦大学的DOR基金及两年期项目BIRD 192932资助，并得到INdAM-GNCS 2022项目“多变量积分模型的方法与软件”的资助。本研究是在RITA“意大利近似研究网络”、UMI TAA小组“逼近理论与应用”（G. Elefante、A. Sommariva）以及SIMAI ANA&A活动组（A. Sommariva、M. Vianello）的框架内完成的。

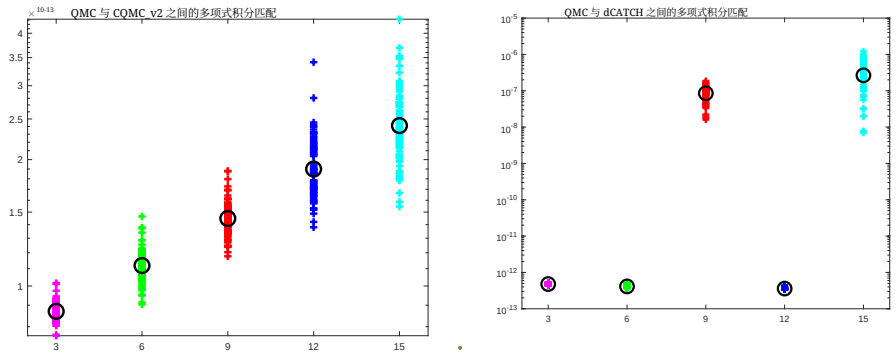
参考文献

[1] D. Avis、B.K. Bhattacharya、H. Imai，《球体并集的体积计算》，The Visual Computer 3（1988）323–328。[2] B. Bauman、H. Xiao，《非圆形光瞳和视场及宽波长范围的光学设计高斯求积》，Proc. SPIE 7652，2010年。[3] L. Bittante、S. De Marchi、G. Elefante，《一种基于非负最小二乘和近似费克特点的新型准蒙特卡罗技术》，Numer. Math. Theory Methods Appl. 9（2016），640–663。



(a) (b)

图 4：自底向上算法（左）和 $dCATCH$ （右）对3个球并集表面上的随机多项式进行 100 次试验所得的 QMC 相对压缩误差及其对数平均值（圆圈）。注意，左右图的刻度不同。



(a) (b)

图 5：在 100 个球并集的表面上，对随机多项式采用自底向上算法（左图）和 $dCATCH$（右图）进行 100 次试验所得的 QMC 相对压缩误差及其对数平均值（圆圈）。请注意，左右图的刻度不同。

deg	3	6	9	12	15
基数 QMC	$M = 1{,}024{,}179$				
基数 Q_c^{dCATCH}	20	83	200	371	572
基数 Q_c^{bu}	20	83	200	371	596
压缩比	5.1e+04	1.2e+04	5.1e+03	2.8e+03	1.7e+03
CPU Halton 序列	8.8e-01s				
CPU Q_c^{dCATCH}	2.8e+00s	1.7e+01s	5.0e+01s	1.4e+02s	3.2e+02s
CPU Q_c^{bu}	3.1e-01秒	1.1e+00秒	2.7e+00秒	5.8e+00秒	6.5e+01秒
加速比	9.0	15.9	18.2	24.0	4.9
矩残差 Q_c^{dCATCH}	8.6e-12	8.9e-12	8.9e-12	8.9e-12	⋆ 1.6e-06
矩残差 Q_c^{bu}					
迭代 1	7.2e-01	1.4e-15	2.8e-15	4.2e-15	2.6e-01
迭代 2	3.7e-16				1.3e-01
迭代 3					3.3e-12

表5：从每个球面上的 500,000 个低差异点开始，对 3 个球的并集进行曲面QMC积分压缩。

deg	3	6	9	12	15
基数 QMC	$M = 1{,}032{,}718$				
基数 Q_c^{dCATCH}	20	84	219	455	807
基数 Q_c^{bu}	20	84	220	455	816
压缩比	5.2e+04	1.2e+04	4.7e+03	2.3e+03	1.3e+03
CPU Halton 序列	1.5e+01秒				
CPU Q_c^{dCATCH}	2.8e+00秒	1.6e+01秒	4.3e+01秒	1.1e+02s	2.4e+02s
CPU Q_c^{bu}	3.2e-01s	1.1e+00s	3.0e+00s	6.8e+00s	2.4e+01s
加速比	8.7	14.5	14.3	16.2	9.9
矩残差 Q_c^{dCATCH}	9.0e-13	9.1e-13	⋆ 3.2e-06	9.3e-13	⋆ 1.8e-05
矩残差 Q_c^{bu}					
迭代 1	2.09e+00	1.22e+00	5.53e-01	6.18e-01	1.41e-01
迭代 2	7.49e-16	1.30e-15	2.52e-15	5.16e-15	1.18e-14

表6：从每个球面上的 60,000 个低差异点开始，对 100 个球的并集进行曲面QMC积分压缩。

[4] D.A. Cox、J. Little、D. O'Shea，《理想、簇与算法》，第4版，Springer，2015年。[5] P.J. Davis，非负近似求积公式的构造，《数学计算》21 (1967年) 578–582。[6] M. Dell'Orto、M. Dessole、F. Marcuzzi，《带偏差最大化的 Lawson-Hanson 算法：有限收敛与稀疏恢复》，《数值线性代数及其应用》，2023年1月13日在线发表。[7] S. De Marchi、G. Elefante，利用映射低差异点和贪心最小黎曼 s-能量点进行流形上的准蒙特卡罗积分，《应用数值数学》127 (2018年) 110–124。[8] M. Dessole、F. Marcuzzi，秩揭示QR分解的偏差最大化，《数值算法》91 (2022年) 1047-1079。[9] M. Dessole、F. Marcuzzi、M. Vianello，利用快速非负最小二乘加速大规模Tchakaloff回归设计，《多洛米蒂近似研究札记》DRNA 13 (2020年)，20–29。[10] M. Dessole、F. Marcuzzi、M. Vianello，dCATCH：通过快速非负最小二乘实现 d元近似G最优Tchakaloff回归的数值软件包，MDPI-Mathematics 8(7) (2020年) - 特刊"数值方法"。[11] J. Dick 和 F. Pillichshammer，《数字网与序列：差异理论与准蒙特卡罗积分》，剑桥大学出版社，剑桥，2010年。[12] X. Duan、C. Quan、B. Stamm，d维球的基于边界划分的Voronoi图：定义、性质与应用，《计算数学进展》46 (2020年)。[13] G. Elefante、A. Sommariva、M. Vianello，CQMC：用于低维压缩准蒙特卡罗求积的改进代码，《多洛米蒂近似研究札记》DRNA 15 (2022年) 92–100。[14] G. Elefante、A. Sommariva、M. Vianello，Qbubble：多气泡上的压缩QMC体积与曲面积分（使用Matlab），`https://www.math.unipd.it/~alvise/software.html`。[15] S. Hayakawa，蒙特卡罗求积构造，《日本工业与应用数学杂志》38 (2021年) 561-577。[16] I.J. Kim、H. Na，一种计算共有溶剂可及体积的高效算法，PLoS ONE 17(3): e0265614 (2022年)。[17] C.L. Lawson、R.J. Hanson，《求解最小二乘问题》。应用数学经典丛书15，SIAM，费城，1995年。

deg	3	6	9	12	15
$E^{QMC}(f_1)$	3.9e-06				
$E^{dCATCH}(f_1)$	2.3e-04	8.3e-06	4.0e-06	3.9e-06	3.9e-06
$E^{bu}(f_1)$	3.7e-04	3.6e-06	4.0e-06	3.9e-06	3.9e-06
$E^{QMC}(f_2)$	8.6e-05				
$E^{dCATCH}(f_2)$	3.5e-01	3.2e-02	8.3e-04	8.4e-05	8.6e-05
$E^{bu}(f_2)$	1.3e+00	2.2e-02	8.3e-06	8.5e-05	8.6e-05
$E^{QMC}(f_3)$	5.8e-06				
$E^{dCATCH}(f_3)$	3.9e-01	5.8e-03	6.9e-04	5.8e-05	8.9e-06
$E^{bu}(f_3)$	2.5e-02	3.8e-03	4.7e-06	9.1e-05	6.0e-06

表7：三个球并集上的曲面QMC积分压缩（参考值通过从每个球面上的 10^6 个点开始进行 QMC 计算得到）。

deg	3	6	9	12	15
$E^{QMC}(f_1)$	4.0e-05				
$E^{dCATCH}(f_1)$	2.9e-03	2.8e-05	3.9e-05	4.0e-05	4.0e-05
$E^{bu}(f_1)$	3.2e-02	3.7e-05	3.9e-05	4.0e-05	4.0e-05
$E^{QMC}(f_2)$	2.0e-04				
$E^{dCATCH}(f_2)$	1.4e-01	6.4e-04	1.7e-04	2.0e-04	2.0e-04
$E^{bu}(f_2)$	1.4e-01	6.2e-05	1.9e-04	2.0e-04	2.0e-04
$E^{QMC}(f_3)$	1.6e-04				
$E^{dCATCH}(f_3)$	2.9e-02	1.7e-02	4.3e-04	1.5e-04	1.6e-04
$E^{bu}(f_3)$	1.6e-02	1.7e-03	2.6e-04	1.5e-04	1.6e-04

表8：100个球并集上的曲面QMC积分压缩（参考值通过从每个球面上的 10^6 个点开始进行 QMC 计算得到）。

[18] G. Legrain，《虚拟区域方法的非负矩拟合求积公式》，Comput. Math. Appl. 99（2021年）270–291。[19] F. Librino、M. Levorato、M. Zorzi，《计算圆交面积的算法解及其在无线通信中的应用》，Wireless Communications and Mobile Co mputing 14（2014）1672–1690。[20] C. Litterer、T. Lyons，《高阶重组及其在Wiener空间求积中的应用》，Ann. Appl. Probab. 22（2012）1301–1327。[21] E. Pauwels、M. Putinar、J.-B. Lasserre，《基于经验矩和Christoffel函数的数据分析》，Found. Comput. Math. 21（2021年）243–273。[22] M. Petitjean，《球的并集与交集及其在分子建模中的若干应用》，载于：A. Mucherino等编，《距离几何：理论、方法与应用》，Springer New York，2013，61–83页。[23] F. Piazzon、A. Sommariva、M. Vianello，《Carathéodory-Tchakaloff子采样》，《多洛米蒂近似研究札记》DRNA 10（2017）5–14。[24] W. Pl´esniak，《多元Jackson不等式》，J. Comput. Appl. Math. 233（2009）815–820。[25] D.L. Ragozin，《球面与射影空间上的构造性多项式逼近》，Trans. Amer. Math. Soc. 162（1971）157–170。[26] M. Slawski，《非负最小二乘：算法比较》，`https://sites.google.com/site/slawskimartin`。[27] A. Sommariva、M. Vianello，《多元离散测度压缩及其应用》，Numer. Funct. Anal. Optim. 36（2015）1198–1223。[28] A. Sommariva、M. Vianello，《平面圆盘交集上的数值求积》，FILOMAT 31（2017）4105–4115。[29] A. Sommariva、M. Vianello，《样条曲线多边形上的Tchakaloff型求积公式计算》，《多洛米蒂近似研究札记》DRNA 14（2021年）1–11。[30] A. Sommariva、M. Vianello，《圆盘并集上的正权重求积公式》，《多洛米蒂近似研究札记》DRNA 15（2022年）73–81（纪念S. De Marchi 60岁特刊）。[31] A. Sommariva、M. Vianello，《NURBS形状区域上的低基数正内部求积》，BIT Numer. Math.，待发表。[32] V. Tchakaloff，《具有非负系数的机械求积公式》，´（法语），`，Bull. Sci. Math. 81（1957）123–134。[33] M. Tchernychova，《Carathéodory求积测度》。数学博士论文（导师：T. Lyons），牛津大学，2015年。[34] D.R. Wilhelmsen，《凸多面体锥的最近点算法及其在正线性逼近中的应用》，Math. Comp. 30（1976）48–57。