Construction and application of provable positive and exact cubature formulas

Jan Glaubitz†

Abstract. Many applications require multi-dimensional numerical integration, often in the form of a cubature formula. These cubature formulas are desired to be positive and exact for certain finite-dimensional function spaces (and weight functions). Although there are several efficient procedures to construct positive and exact cubature formulas for many standard cases, it remains a challenge to do so in a more general setting. Here, we show how the method of least squares can be used to derive provable positive and exact formulas in a general multi-dimensional setting. Thereby, the procedure only makes use of basic linear algebra operations, such as solving a least squares problem. In particular, it is proved that the resulting least squares cubature formulas are ensured to be positive and exact if a sufficiently large number of equidistributed data points is used. We also discuss the application of provable positive and exact least squares cubature formulas to construct nested stable high-order rules and positive interpolatory formulas. Finally, our findings shed new light on some existing methods for multi-variate numerical integration and under which restrictions these are ensured to be successful.

Key words. Multivariate integration, numerical integration, quadrature, cubature, least squares, discrete orthogonal functions, equidistributed sequence, low-discrepancy sequence, Caratheodory–Tchakaloff measure

AMS subject classifications (2020). 65D30, 65D32, 41A55, 41A63, 42C05

DOI. https://doi.org/10.1093/imanum/drac017

1. Introduction. Numerical integration is an omnipresent technique in applied mathematics, engineering, and many other sciences. Prominent examples include numerical differential equations [52, 2], machine learning [68], finance [37], and biology [66].

1.1. Problem Statement. In many cases, the problem can be formulated as follows. Let \( d \geq 2 \) and \( \Omega \subset \mathbb{R}^d \) be bounded with positive volume and boundary of measure zero (in the sense of Lebesgue). Given a function \( f : \Omega \to \mathbb{R} \), we seek to approximate the continuous integral

\[
I[f] := \int_{\Omega} f(x)\omega(x) \, dx
\]

with Riemann integrable weight function \( \omega : \Omega \to \mathbb{R}^+ \) that is assumed to be positive almost everywhere. A prominent approach is to approximate (1.1) by a weighted finite sum over the function values \( f(x_n) \) at some data points \( x_1, \ldots, x_N \), denoted by

\[
C_N[f] := \sum_{n=1}^{N} w_n f(x_n).
\]
Usually, $C_N$ is referred to as an $N$-point cubature formula (CF) and $w_1, \ldots, w_N \in \mathbb{R}$ are called cubature weights. For a 'good' CF, the following properties are often required:

(P1) All data points should lie inside of $\Omega$. That is, $x_n \in \Omega$ for all $n = 1, \ldots, N$.

(P2) The CF should be positive. That is, $w_n > 0$ for all $n = 1, \ldots, N$.\(^1\)

See [31, 24, 28] and references therein. Moreover, in many applications, CFs are desired which are exact for certain finite-dimensional function spaces. Let $F_K(\Omega)$ denote a $K$-dimensional function space spanned by $\varphi_1, \ldots, \varphi_K : \Omega \to \mathbb{R}$. It shall be assumed that all the moments $m_k := I[\varphi_k], k = 1, \ldots, K$, exist. Then, the $N$-point CF (1.2) is said to be $F_K(\Omega)$-exact if

$$C_N[f] = I[f] \quad \forall f \in F_K(\Omega).$$

Usual choices for $F_K(\Omega)$ include the space of all algebraic and trigonometric polynomials up to a certain degree $m$, respectively denoted by $\mathbb{P}_m(\mathbb{R}^d)$ and $\Pi_m(\mathbb{R}^d)$. Another prominent example is approximation spaces of radial basis functions [80, 34, 74, 82].

1.2. Previous Works. The existence of positive and $F_K(\Omega)$-exact CFs with at most $K$ data points is ensured by the following theorem originating from [86]. Also see [26, 4].

**Theorem 1.1 (Tchakaloff, 1957).** Given is (1.1) with nonnegative $\omega$ and a $K$-dimensional function space $F_K(\Omega)$. Then there exist $N$ data points $x_1, \ldots, x_N \in \Omega$ and positive weights $w_1, \ldots, w_N$ with $N \leq K$ such that the corresponding $N$-point CF is $F_K(\Omega)$-exact.

However, it should be noted that Tchakaloff’s original proof is not constructive in nature. A constructive, yet in general not computational practical, prove was provided in [27]. At the same time, it should be pointed out that Theorem 1.1 only provides an upper bound for the number of data points that are needed for a positive and $F_K(\Omega)$-exact CF. Indeed, for standard domains (e.g. $\Omega = [0, 1]^d$) and weight functions (e.g. $\omega \equiv 1$) as well as classic function spaces (e.g. algebraic polynomials), it is possible to construct CFs that use even fewer points [64, 9, 7]. Such CFs are referred to as minimal or near-minimal CFs and usually utilize some kind of symmetry in the domain, weight function, and function space. Unfortunately, they are often limited to algebraic polynomials of low total degrees or specific domains and weight functions. We also mention Smolyak (sparse grid) CFs [79, 15], which rely on a product domain and a separable weight function. Another approach is the use of optimization-based methods [85, 75, 57, 60] to construct near-minimal CFs (sometimes based on heuristic arguments). Unfortunately, the success of these methods often depends on the initial guess for the data points and they can sometimes yield points outside of the computational domain or negative cubature weights. We shall also mention some recent works on constructing nested sampling-based positive CFs [93, 91]. However, it should be noted that these are developed based on an approximate notion of exactness; in the sense that $I[f]$ in the exactness condition (1.3) is replaced by a discrete approximation $I^M[f] = \frac{1}{M} \sum_{m=1}^M f(y_m)$ with $M > N$ and random samples $y_m, m = 1, \ldots, M$. In fact, one might argue that the method proposed in [91] is closer to subsampling [27, 97, 96, 72, 92]. A similar argument can also be made for CFs based on nonnegative least squares (NNLS); see [54, 81, 39]. Finally, a fairly general approach to construct “probable” positive and exact CFs was proposed in [67]. The idea there was to derive

\(^1\)Some authors require the cubature weights to only be nonnegative. However, if $w_n = 0$, then the corresponding data point can—and should—be removed from the CF to avoid an unnecessary loss of efficiency.
randomized CFs based on first approximating the unknown integrand $f$ by a discrete least squares (LS) approximation [21, 23, 47] and to exactly integrate this approximation then. Thereby, the discrete LS approximation was based on the values of $f$ at random samples. The authors proved in [67] that the resulting randomized CFs are positive and exact with a high probability if the number of random samples is sufficiently large. To the best of our knowledge, this result has not been carried over to the deterministic setting (in the sense of the data points not coming from fully-probabilistic random samples) yet. We also mention the two works [69, 50], which we shall address in more detail below. In a nutshell, although many approaches exist to construct provable positive and exact CFs for certain special cases, it remains a challenge to do so in a general deterministic setting.

1.3. Our Contribution. We propose a simple procedure to construct provable positive and exact CFs in a general setting. The procedure is based on the idea of least squares quadrature formulas (LS-QFs). These were introduced in [99, 98] (originally for equidistant points and $\omega \equiv 1$) and generalized in [54, 39] (in one dimension though). Recently, LS formulas have been extended to higher dimensions in [40]. However, this was done under the restriction of the LS-CF being exact for algebraic polynomials up to a certain (total) degree, rather than a general finite-dimensional function space.

The theoretical backbone of the present work is the extension of this result to a general class of finite-dimensional function spaces. We show that LS-CF can be ensured to be positive if a sufficiently large number of equidistributed data points is used. Numerically, we found that the number of data points $N$ has to scale roughly like the squared dimension of the function space $\mathcal{F}_K(\Omega)$. Moreover, we provide an error analysis for the corresponding positive and exact LS-CF, discuss potential applications, and address some connections to other integration methods.

To prove that LS-CFs are positive, we leverage different tools from linear algebra, LS problems, discrete orthonormal functions, and equidistributed sequences from number theory. The sufficient conditions for this result (summarized in Corollary 3.6) to hold, are the following:

(R1) The integration domain $\Omega \subset \mathbb{R}^d$ is bounded with boundary of measure zero.
(R2) The weight function $\omega : \Omega \to \mathbb{R}_+^0$ is Riemann integrable and positive almost everywhere.
(R3) The function space $\mathcal{F}_K(\Omega)$ is spanned by a basis $\{\varphi_k\}_{k=1}^K$ of continuous and bounded functions. Furthermore, $\mathcal{F}_K(\Omega)$ contains constants. In particular, $1 \in \mathcal{F}_K(\Omega)$.

(R1) essentially ensures the existence and simple construction of an equidistributed sequence in $\Omega$. (R2) warrants the existence of a certain continuous inner product (and corresponding orthonormal functions) that can be approximated by a sequence of discrete inner products (and corresponding discrete orthonormal functions). (R3) is needed for technical reasons and is utilized in the proof of the preliminary Lemma 3.4. Well-conditioned computation of the LS-CFs can be ensured if a basis $\{\varphi_k\}_{k=1}^K$ for $\mathcal{F}_K(\Omega)$ of orthonormal functions is available (see Remark 6.1).
1.4. Implications and Potential Applications. Our findings imply a simple\footnote{Indeed, the procedure only utilizes simple operations from linear algebra, such as solving an LS problem.} procedure to construct (deterministic) stable high-order CFs in a general setting. This can be interpreted as an extension of the results on stable high-order randomized CF from \cite{67} to a deterministic setting. Moreover, we discuss the application of the provable positive and exact LS-CF to construct interpolatory CFs (see subsection 5.2). These use a smaller number of $N = K$ data points and can be obtained by several different subsampling strategies. It should also be noted that for a fixed function space $F_K(\Omega)$ and an increasing number $N$ of data points, the LS-CFs discussed here might be interpreted as high-order corrections to Monte Carlo (MC)—for random data points—and quasi-Monte Carlo (QMC)—for low-discrepancy data points—methods. See Remark 5.1 for more details. In particular, this reveals an interesting connection to \cite{69} and potential applications of LS-CFs for variance reduction in MC and QMC methods. We shall also briefly mention the implication of the present work to several other approaches to find positive (interpolatory) CFs. These include NNLS \cite{54, 81, 39} and different optimization strategies \cite{40, 50}. Sometimes, these approaches are justified by Tchakaloff’s theorem (Theorem 1.1), which ensures the existence of a solution to the respective optimization problem if an appropriate set of data points is considered. That said, Tchakaloff’s theorem is providing no information about which data points should be used (it only states their existence and an upper bound for their number). Hence, when applying the above-mentioned procedures without care to an arbitrary set of data points, they cannot always be expected to actually result in a positive interpolatory CF. The present work is providing such an assurance in the sense that Corollary 3.6 combined with the subsampling strategies discussed in subsection 5.2 is telling us that at least some of the positive interpolatory CFs predicted by Tchakaloff’s theorem are supported on a sufficiently large set of equidistributed points in $\Omega$.

1.5. Advantages and Pitfalls. The advantage of the positive and exact CFs discussed in the present work lies in their generality and simple construction. That said, they are neither minimal nor near-minimal. Hence, if the reader is only interested in a certain standard case for which efficient CFs are readily available it is certainly advantageous to use these. The positive and exact CFs presented here, on the other hand, find their greatest utility when a CF is desired for a non-standard domain, weight function, or function space. They might also be of advantage when one is given a fixed and prescribed set of (scattered) data points, which is often the case in applications. Finally, for reasons of computational efficiency, in general, we only recommend using LS-CFs in moderate dimensions (usually $d = 1, 2, 3$). In higher dimensions, they—like many other CFs—might fall victim to the curse of dimensionality \cite{5, 62}. Also see section 4 for more details. Finally, the construction of the positive and exact CFs discussed here relies on knowledge of certain moments, which is discussed in Remark 3.1.

1.6. Outline. The rest of this work is organized as follows. In section 2, we collect some preliminaries, in particular, on unisolvent and equidistributed sequences. Next, section 3 contains our theoretical main results, i.e., exactness and conditional positivity of LS-CFs is proved. In section 4, an error analysis for these CFs is provided. Two specific applications of the provable positive and exact LS-CFs are discussed in section 5. These include the simple construction of stable high-order sequences of CFs (subsection 5.1) and positive interpolatory
CFs (subsection 5.2). Numerical experiments are presented in section 6 and some concluding thoughts are offered in section 7.

2. Preliminaries: Unisolvent and Equidistributed Sequences. Here, we shall provide a few preliminary results. In particular, these address the connection between the exactness of CFs and unisolvent and equidistributed sequences.

2.1. Exactness and Unisolvence. Let \( \{ \varphi_k \}_{k=1}^K \) be a basis of the function space \( \mathcal{F}_K(\Omega) \) and \( m_k = I[\varphi_k], \ k = 1, \ldots, K, \) the corresponding moments. Then the exactness condition (1.3) is equivalent to the data points \( X_N = \{ x_n \}_{n=1}^N \) and weights of a CF solving the nonlinear system

\[
(2.1) \quad \begin{pmatrix} \varphi_1(x_1) & \cdots & \varphi_1(x_N) \\ \vdots & \cdots & \vdots \\ \varphi_K(x_1) & \cdots & \varphi_K(x_N) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_K \end{pmatrix} =: \Phi(X_N)
\]

However, solving (2.1) can be highly nontrivial, and doing so by brute force may result in a CF which points lie outside of the integration domain or with negative weights \[48\]. That said, the situation changes if a fixed set of data points is considered. Then \( \Phi(X_N) = \Phi \), and (2.1) becomes a linear system:

\[
(2.2) \quad \Phi w = m
\]

Note that for \( K < N \) this is an underdetermined linear system. These are well-known to either have no or infinitely many solutions. The latter case arises when we restrict ourselves to unisolvent nodes.

Definition 2.1 (Unisolvent Nodes). The nodes \( X_N = \{ x_n \}_{n=1}^N \subset \mathbb{R}^d \) are called \( \mathcal{F}_K(\Omega) \)-unisolvent if

\[
(2.3) \quad f(x_n) = 0, \ n = 1, \ldots, N \implies f(x) = 0, \ \forall x \in \Omega
\]

holds for all \( f \in \mathcal{F}_K(\Omega) \).

Assuming that the set of data points \( X_N \) is \( \mathcal{F}_K(\Omega) \)-unisolvent the following result follows.

Lemma 2.2. Let \( K < N \) and \( X_N = \{ x_n \}_{n=1}^N \) be \( \mathcal{F}_K(\Omega) \)-unisolvent. Then, the linear system (2.2) induces an \( (N - K) \)-dimensional affine linear subspace of solutions,

\[
(2.4) \quad W := \left\{ w \in \mathbb{R}^N \mid \Phi w = m \right\}
\]

Proof. The case \( \mathcal{F}_K(\Omega) = \mathbb{P}_m(\mathbb{R}^d) \) was shown in [40]. It is easy to verify that the same arguments carry over to the general case discussed here. \(\blacksquare\)

Next, a simple sufficient criterion for \( X_N \subset \Omega \) to be \( \mathcal{F}_K(\Omega) \)-unisolvent is provided. The criterion is based on sequences that are dense in \( \Omega \subset \mathbb{R}^d \). Recall that \( \{ x_n \}_{n \in \mathbb{N}} \subset \mathbb{R}^d \) is called dense in \( \Omega \) if

\[
(2.5) \quad \forall x \in \Omega \ \forall \varepsilon > 0 \ \exists n \in \mathbb{N} : \ |x - x_n| < \varepsilon.
\]
That is, every point in \( \Omega \) can be approximated arbitrarily accurate by an element of \( (x_n)_{n\in\mathbb{N}} \).

**Lemma 2.3.** Let \( (x_n)_{n\in\mathbb{N}} \subset \Omega \) be dense in \( \Omega \) and let \( X_N = \{x_n\}_{n=1}^N \). Moreover, let \( F_K(\Omega) \) be spanned by continuous functions \( \varphi_1, \ldots, \varphi_K : \Omega \to \mathbb{R} \). Then there exists an \( N_0 \in \mathbb{N} \) such that \( X_N \) is \( F_K(\Omega) \)-unisolvent for every \( N \geq N_0 \).

**Proof.** Assume that the assertion is wrong. Then, there exists an \( f \in F_K(\Omega) \) with \( f \neq 0 \) such that \( f(x_n) = 0 \) for all \( n \in \mathbb{N} \). Yet, since \( (x_n)_{n\in\mathbb{N}} \) is dense in \( \Omega \), this either contradicts \( f \neq 0 \) or \( f \) being continuous.

If there exists an \( N_0 \in \mathbb{N} \) such that \( X_N = \{x_n\}_{n=1}^N \) is \( F_K(\Omega) \)-unisolvent for every \( N \geq N_0 \), as in Lemma 2.3, we say that \( (x_n)_{n\in\mathbb{N}} \) is an \( F_K(\Omega) \)-unisolvent sequence. Thus, Lemma 2.3 states that every dense sequence is also \( F_K(\Omega) \)-unisolvent.

**2.2. Equidistributed Sequences.** We just saw that density is a sufficient condition for unisolvence. This will be handy for the subsequent construction of provable positive and exact LS-CFs. Another important property will be for the sequence of data points \( (x_n)_{n\in\mathbb{N}} \subset \Omega \) to satisfy

\[
\lim_{N \to \infty} \frac{|\Omega|}{N} \sum_{n=1}^N g(x_n) = \int_{\Omega} g(x) \, dx
\]

for all measurable bounded functions \( g : \Omega \to \mathbb{R} \) that are continuous almost everywhere (in the sense of Lebesgue). Here, \( |\Omega| \) denotes the \( d \)-dimensional volume of \( \Omega \). Observe that \( (x_n)_{n\in\mathbb{N}} \) being dense in \( \Omega \) is not sufficient for (2.6) to hold.\(^4\) Yet, in [95] it was showed that (2.6) can be connected to \( (x_n)_{n\in\mathbb{N}} \) being equidistributed (also called uniformly distributed). We shall recall that there are many well-known equidistributed sequences for \( d \)-dimensional hypercubes with radius \( R \), denoted by \( C_R^{(d)} = [-R, R]^d \). These include (i) grids of equally spaced points with an appropriate ordering and (ii) low-discrepancy sequences. The latter were developed to minimize the upper bound provided by the famous Koksma–Hlawak inequality [53, 71], used in QMC methods [16, 30, 88]. A special case of low-discrepancy sequences are the Halton points [49], which are a generalization of the one-dimensional van der Corput points, see for example [94, Erste Mitteilung]). To not exceed the scope of this work, we refer to the monograph [61] for more details on equidistributed sequences. That said, we shall demonstrate how equidistributed sequences can be constructed for general bounded domains with a boundary of measure zero.

**Remark 2.4 (Construction of Equidistributed Sequences for General Domains).** Let \( \Omega \subset \mathbb{R}^d \) be contained in the hypercube \( C_R^{(d)} \). Let \( (y_n)_{n\in\mathbb{N}} \) be an equidistributed sequence in \( C_R^{(d)} \), then an equidistributed sequence in \( \Omega \), \( (x_n)_{n\in\mathbb{N}} \), is given by the subsequence of \( (y_n)_{n\in\mathbb{N}} \) for which all elements outside of \( \Omega \) have been removed:

\[
y_n \in (x_n)_{n\in\mathbb{N}} \iff y_n \in \Omega.
\]

\(^3\)The condition (2.5) is independent of the norm since \( \Omega \) is located in a finite-dimensional space.

\(^4\)However, every dense sequence can be rearranged into an (equidistributed) sequence satisfying (2.6).
Let us briefly verify that the resulting subsequence \((x_n)_{n \in \mathbb{N}}\) is an equidistributed sequence in \(\Omega\). To this end, let \(g : \Omega \rightarrow \mathbb{R}\) be a measurable bounded function that is continuous almost everywhere. We can extend \(g\) to the hypercube \(C_R^{(d)}\) by setting \(g\) equal to zero outside of \(\Omega\). This extension, denoted by \(\tilde{g} : C_R^{(d)} \rightarrow \mathbb{R}\), is measurable, bounded, and continuous almost everywhere. The latter follows from \(\Omega\) having a boundary of measure zero. Also observe that

\[
\lim_{M \to \infty} \frac{|\{y_1, \ldots, y_M\} \cap \Omega|}{M} = \frac{|\Omega|}{|C_R^{(d)}|}.
\]

Then, by construction of \((x_n)_{n \in \mathbb{N}}\) and (2.8), we get

\[
\lim_{N \to \infty} \frac{|\Omega|}{N} \sum_{n=1}^{N} g(x_n) = \lim_{M \to \infty} \frac{|C_R^{(d)}|}{M} \sum_{n=1}^{M} \tilde{g}(y_n) = \int_{C_R^{(d)}} \tilde{g}(x) \, dx = \int_{\Omega} g(x) \, dx.
\]

Here, \(M\) is the unique integer such that \(x_N = y_M\), and \(|\{y_1, \ldots, y_M\} \cap \Omega| = N\). Again, we refer to [61] for more details.

Finally, it should be stressed that equidistributed sequences are dense sequences with a specific ordering. We close this section with the following corollary.

**Corollary 2.5.** Let \(\Omega \subset \mathbb{R}^d\) be bounded with a boundary of measure zero. Furthermore, let \((y_n)_{n \in \mathbb{N}}\) be an equidistributed sequence in the hypercube \(C_R^{(d)}\), where \(\Omega \subset C_R^{(d)}\), and let \((x_n)_{n \in \mathbb{N}}\) be the subsequence of \((y_n)_{n \in \mathbb{N}}\) that only contains the elements in \(\Omega\). Then \((x_n)_{n \in \mathbb{N}}\) is equidistributed in \(\Omega\) and \(\mathcal{F}_K(\Omega)\)-unisolvent.

**Proof.** The assertion that \((x_n)_{n \in \mathbb{N}}\) is equidistributed in \(\Omega\) follows from Remark 2.4. Finally, \((x_n)_{n \in \mathbb{N}}\) being \(\mathcal{F}_K(\Omega)\)-unisolvent follows by every equidistributed sequence being dense and Lemma 2.3.

### 3. Provable Positive and Exact Least Squares Cubature Formulas

In this section, it is demonstrated how provable positive and \(\mathcal{F}_K(\Omega)\)-exact LS-CFs can be constructed by using a sufficiently large set of equidistributed data points. This is done by generalizing the LS approach from [99, 98, 54, 38, 39, 40]. Finally, the procedure only relies on determining the LS solution of an underdetermined linear system.

**3.1. Formulation as a Least Squares Problem.** Let \((x_n)_{n \in \mathbb{N}}\) be a \(\mathcal{F}_K(\Omega)\)-unisolvent sequence in \(\Omega\) and let \(X_N = \{x_n\}_{n=1}^N\). As noted before, an \(\mathcal{F}_K(\Omega)\)-exact CF with data points \(X_N\) can be constructed by determining a vector of cubature weights that solves the linear system of exactness conditions (2.2). For \(N > K\), (2.2) induces an \((N-K)\)-dimensional affine linear subspace space of solutions \(W \subset \mathbb{R}^N\). All of these yield an \(\mathcal{F}_K(\Omega)\)-exact CF. The LS approach consists of finding the unique solution \(w \in W\) that minimizes a weighted Euclidean norm:

\[
w^{LS} = \arg \min_{w \in W} \left\| R^{-1/2} w \right\|_2^2.
\]

The elements of \((y_n)_{n \in \mathbb{N}}\) are assumed to be distinct.
where $R^{-1/2}$ is a diagonal weight matrix, given by

$$R^{-1/2} = \text{diag}(1/\sqrt{r_1}, \ldots, 1/\sqrt{r_N}), \quad r_n > 0, \quad n = 1, \ldots, N.$$  

(3.2)

The vector $w^{LS}$ in (3.1) is called the LS solution of $\Phi w = m$. The corresponding $N$-point CF

$$C_N^{LS}[f] = \sum_{n=1}^{N} w^{LS}_n f(x_n)$$  

(3.3)

is called an LS-CF. In subsection 3.4 it is shown that choosing the discrete weights $r_n$ as

$$r_n = \frac{\omega(x_n)|\Omega|}{N}, \quad n = 1, \ldots, N,$$

(3.4)

results in the LS-CFs to be $\mathcal{F}_K(\Omega)$-exact and positive if $N$ is sufficiently large. At least formally, the LS solution is explicitly given by ([20])

$$w^{LS} = R\Phi^T(\Phi R\Phi^T)^{-1}m.$$  

(3.5)

Here, $R\Phi^T(\Phi R\Phi^T)^{-1}$ is the Moore–Penrose pseudoinverse of $R^{-1/2}\Phi$; see [6]. In subsection 3.3, (3.5) will be simplified by utilizing discrete orthonormal bases.

**Remark 3.1 (Computation of the Moments).** The construction of LS-CFs requires the computation of the moments $m_k = I[\varphi_k]$ for a basis $\{\varphi_k\}_{k=1}^K$, which is a requirement for many CFs. Depending on the domain $\Omega$, the weight function $\omega$, and the basis $\{\varphi_k\}_{k=1}^K$, the exact evaluation of $m_k$ might be impractical. However, we can always approximate $m_k$ by another unrelated CF, such as the QMC method, using a larger set of nodes. In some cases, it might also be possible to use certain recurrence relations or differential/difference equations of the basis functions $\varphi_k$ to simplify the computation of the moments.

### 3.2. Continuous and Discrete Orthonormal Bases

Under the restrictions (R1)–(R3),

$$\langle u, v \rangle = \int_{\Omega} u(x)v(x)\omega(x) \, dx, \quad \|u\| = \sqrt{\langle u, u \rangle}$$  

(3.6)

defines an inner product and a corresponding norm on $\mathcal{F}_K(\Omega)$. In particular, (3.6) allows to define an orthonormal basis $\{\pi_k\}_{k=1}^K$ of $\mathcal{F}_K(\Omega)$. That is, the functions $\pi_1, \ldots, \pi_K$ span $\mathcal{F}_K(\Omega)$ and satisfy

$$\langle \pi_k, \pi_l \rangle = \delta_{k,l} := \begin{cases} 1 & ; k = l, \\ 0 & ; k \neq l, \end{cases} \quad k, l = 1, \ldots, K.$$  

(3.7)

Henceforth, we refer to such a basis as a continuous orthonormal basis. Analogously, assuming that $X_N^+ = \{ x_n \mid \omega(x_n) > 0, \ n = 1, \ldots, N \}$ is $\mathcal{F}_K(\Omega)$-unisolvent,

$$[u, v]_N = \sum_{n=1}^{N} r_n u(x_n)v(x_n), \quad \|u\|_N = \sqrt{[u, u]_N}$$  

(3.8)
defines a discrete inner product and a corresponding norm on $\mathcal{F}_K(\Omega)$. Also (3.8) induces an orthonormal basis. This basis satisfies

$$(3.9) \quad [\pi_k, \pi_l]^N = \delta_{k,l}, \quad k, l = 1, \ldots, K,$$

while spanning $\mathcal{F}_K(\Omega)$. It is therefore referred to as a discrete orthonormal basis and its elements are denoted by $\tilde{\pi}^{(N)}_k$. Both bases can be constructed, for instance, by Gram–Schmidt orthonormalization applied to the same initial basis $\{\varphi_k\}_{k=1}^K$ [35, 90]:

$$\tilde{\pi}_k = \varphi_k - \sum_{l=1}^{k-1} \frac{\langle \varphi_k, \pi_l^{(N)} \rangle \pi_l}{\|\tilde{\pi}_k\|}, \quad \pi_k = \frac{\tilde{\pi}_k}{\|\tilde{\pi}_k\|}.$$  

(3.10)

Remark 3.2. Note that we only utilize Gram–Schmidt orthonormalization for theoretical purposes. In our implementation, the LS solution $w^{LS}$ is computed using the Matlab function $\text{lsqminnorm}$, which uses a pivoted QR decomposition of $A = \Phi R^{1/2}$; see [90, 46]. The cost for this is $O(NK^2)$. While we have not tested this in our implementation, it should still be noted that also iterative solvers, such as the preconditioned conjugate gradient (PCG) method, might be used to reduce the costs to $O(NK)$. However, such methods usually rely on sparsity to be efficient and can be more prone to numerical inaccuracies.

3.3. Characterization of the Least Squares Solution. We now collect one more important ingredient to subsequently prove the positivity of LS-CFs. Observe that the matrix product $\Phi R \Phi^T$ in the explicit representation of the LS solution (3.5) can be interpreted as a Gram matrix with respect to the discrete inner product (3.8):

$$\Phi R \Phi^T = \begin{pmatrix} \langle \varphi_1, \varphi_1 \rangle^N & \ldots & \langle \varphi_1, \varphi_K \rangle^N \\ \vdots & \ddots & \vdots \\ \langle \varphi_K, \varphi_1 \rangle^N & \ldots & \langle \varphi_K, \varphi_K \rangle^N \end{pmatrix}.$$  

(3.11)

Thus, if the linear system (2.2) is formulated with respect to the discrete orthonormal basis $\{\varphi_k^{(N)}\}_{k=1}^K$, one gets $\Phi R \Phi^T = I$, where $I$ denotes the $N \times N$ identity matrix. This yields (3.5) to become

$$(3.12) \quad w^{LS} = R \Phi^T m.$$  

In particular, the LS weights are then explicitly given by

$$w^{LS}_n = r_n \sum_{k=1}^K \pi_k^{(N)}(x_n) I[\pi_k^{(N)}], \quad n = 1, \ldots, N,$$

(3.13)

where $I[\pi_k^{(N)}] = \int_{\Omega} \pi_k^{(N)}(x) \, dx$. 

\[\text{PROVABLE POSITIVE AND EXACT CFS}\]
3.4. Positivity of Least Squares Cubature Formulas. We start by presenting two technical lemmas, which will enable us to show that the LS weights are all positive if a sufficiently large number of equidistributed data points is used.

Lemma 3.3. Let \( \Omega \subset \mathbb{R}^d \) be bounded and assume that

\[
\lim_{N \to \infty} [u, v]_N = \langle u, v \rangle \quad \forall u, v \in F_K(\Omega).
\]

Furthermore, let \((u_N)_{N \in \mathbb{N}}, (v_N)_{N \in \mathbb{N}} \subset F_K(\Omega)\) and \(u, v \in F_K(\Omega)\) such that

\[
\lim_{N \to \infty} u_N = u, \quad \lim_{N \to \infty} v_N = v \quad \text{in } (F_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}),
\]

where \(u, v : \Omega \to \mathbb{R}\) are assumed to be bounded. Then,

\[
\lim_{N \to \infty} [u_N, v_N]_N = \langle u, v \rangle.
\]

Recall that \([\cdot, \cdot]_N\) and \(\langle \cdot, \cdot \rangle\) denote the continuous and discrete inner product (3.6) and (3.8), respectively. The corresponding norms are denoted by \(\| \cdot \|\) and \(\| \cdot \|_N\). Moreover, \(\| \cdot \|_{L^\infty(\Omega)}\) is the usual supremum norm with \(\|f\|_{L^\infty(\Omega)} = \sup_{x \in \Omega} |f(x)|\).

Proof. We start by noting that

\[
\left| \langle u, v \rangle - [u_N, v_N]_N \right| \leq \left| \langle u, v \rangle - [u, v]_N \right| + \left| [u, v]_N - [u_N, v_N]_N \right| + \left| [u_N, v_N]_N - [u_N, v_N]_N \right|.
\]

The first term on the right-hand side converges to zero due to (3.14). For the second term, the Cauchy–Schwarz inequality gives

\[
\left| [u, v]_N - [u_N, v_N]_N \right|^2 = \left| [u - u_N, v]_N \right|^2 \leq \| u - u_N \|^2_N \| v \|^2_N.
\]

Furthermore, (3.14) implies \(\| v \|^2_N \to \| v \|^2\) for \(N \to \infty\). Finally, the Hölder inequality and (3.15) yield

\[
\| u - u_N \|^2_N \leq \|1\|^2_N \| u - u_N \|^2_{L^\infty(\Omega)} \to 0, \quad N \to \infty.
\]

Thus, the second term converges to zero as well. A similar argument can be used to show that the third term converges to zero. \(\blacksquare\)

Next, we demonstrate that the discrete orthonormal functions \(\pi_k(N)\) converge uniformly to the corresponding continuous orthonormal functions \(\pi_k\) if the corresponding discrete inner product converges to the continuous one for all elements of \(F_K(\Omega)\).

Lemma 3.4. Let \( \Omega \subset \mathbb{R}^d \) be bounded, let \(\{\varphi_k\}_{k=1}^K\) be a basis of \(F_K(\Omega)\) consisting of continuous and bounded functions, and assume that (3.14) holds. Moreover, let \(\{\pi_k\}_{k=1}^K\) and \(\{\pi_k(N)\}_{k=1}^K\) respectively denote the continuous and discrete orthonormal bases constructed from \(\{\varphi_k\}_{k=1}^K\) by Gram–Schmidt orthonormalization (3.10). Then,

\[
\lim_{N \to \infty} \pi_k(N) = \pi_k \quad \text{in } (F_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}).
\]
Proof. The assertion is proven by induction. For \( k = 1 \), recall that \( \pi_1 = \varphi_1 / \| \varphi_1 \| \) and \( \pi_1^{(N)} = \varphi_1 / \| \varphi_1 \|_N \). Hence, the assertion follows from (3.14) implying that \( \| \varphi_1 \|_N \to \| \varphi_1 \| \) for \( N \to \infty \). Next, it is argued that if the assertion holds for the first \( k - 1 \) orthonormal basis functions, then it also holds for the \( k \)-th one. To this end, let \( l \in \{ 1, 2, \ldots, k - 1 \} \) and assume that

\[
\pi_l^{(N)} \to \pi_l \quad \text{in} \quad (\mathcal{F}_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}), \quad N \to \infty.
\]

Recall that by Gram–Schmidt orthonormalization, the \( k \)-th orthonormal basis functions are given by

\[
(3.10)
\]

Hence, Lemma 3.3 implies

\[
(3.21)
\pi_l^{(N)} \to \pi_l \quad \text{in} \quad (\mathcal{F}_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}), \quad N \to \infty.
\]

Lemma 3.3 yields

\[
(3.22) \quad \varphi_k, \pi_l^{(N)} \to \varphi_k, \pi_l \quad \text{in} \quad (\mathcal{F}_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}), \quad N \to \infty.
\]

This implies

\[
(3.23) \quad \tilde{\pi}_k^{(N)} \to \tilde{\pi}_k \quad \text{in} \quad (\mathcal{F}_K(\Omega), \| \cdot \|_{L^\infty(\Omega)}), \quad N \to \infty.
\]

which completes the proof. 

Lemma 3.3 and Lemma 3.4 now enable us to prove the following theorem.

Theorem 3.5 (The LS-CF is Conditionally Positive). Given is a bounded domain \( \Omega \subset \mathbb{R}^d \), \( \omega: \Omega \to \mathbb{R}_0^+ \), and \( \mathcal{F}_K(\Omega) \subset C(\Omega) \) such that the restrictions (R2) and (R3) are satisfied, i.e.,

(R2) The weight function \( \omega: \Omega \to \mathbb{R}_0^+ \) is Riemann integrable and positive almost everywhere.

(R3) The \( K \)-dimensional vector space \( \mathcal{F}_K(\Omega) \) is spanned by a basis \( \{ \varphi_k \}_{k=1}^K \) of continuous and bounded functions \( \varphi_k: \Omega \to \mathbb{R}, \ k = 1, \ldots, K \). Furthermore, \( \mathcal{F}_K(\Omega) \) contains constants. In particular, \( 1 \in \mathcal{F}_K(\Omega) \).

Moreover, let \( (x_n)_{n \in \mathbb{N}} \subset \Omega \) be \( \mathcal{F}_K(\Omega) \)-unisolvent and \( (r_n)_{n \in \mathbb{N}} \subset \mathbb{R}^+ \) such that

\[
(3.26) \quad \lim_{N \to \infty} \sum_{n=1}^N r_n u(x_n) v(x_n) = \int_{\Omega} \omega(x) u(x) v(x) \, dx \quad \forall u, v \in \mathcal{F}_K(\Omega).
\]

Then, there exists an \( N_0 \in \mathbb{N} \) such that for all \( N \geq N_0 \) the corresponding LS-CF

\[
(3.27) \quad C_N^{LS}[f] = \sum_{n=1}^N u_n^{LS} f(x_n) \quad \text{with} \quad w^{LS} = \arg \min_{w=m} \left\| R^{-1/2} w \right\|_2,
\]

where \( R^{-1/2} = \text{diag}(1/\sqrt{r_1}, \ldots, 1/\sqrt{r_N}) \), is positive and \( \mathcal{F}_K(\Omega) \)-exact.
Proof. First, we note that \( \{x_n\}_{n \in \mathbb{N}} \) being \( F_K(\Omega) \)-unisolvent ensures the existence of a discrete inner product and a discrete orthonormal basis. Let us denote such a basis by \( \{\pi_k^{(N)}\}_{k=1}^K \) and the corresponding continuous orthonormal basis by \( \{\pi_k\}_{k=1}^K \). It can be assumed that both bases are constructed by applying Gram–Schmidt orthonormalization to the same initial basis \( \{\varphi_k\}_{k=1}^K \). Hence, the LS weights are explicitly given by

\[
(3.28) \quad w_{n}^{\text{LS}} = r_n \sum_{k=1}^{K} \pi_k^{(N)}(x_n) I[\pi_k^{(N)}], \quad n = 1, \ldots, N,
\]

for \( N \geq K \), where \( I[\pi_k^{(N)}] = \int_{\Omega} \pi_k^{(N)}(x) \, dx \). Next, let

\[
(3.29) \quad \epsilon_k^{(N)} := [\pi_k^{(N)}, 1]_N - \langle \pi_k^{(N)}, 1 \rangle.
\]

This allows us to rewrite the LS weights as follows:

\[
(3.30) \quad w_{n}^{\text{LS}} = r_n \sum_{k=1}^{K} \pi_k^{(N)}(x_n) \left[ [\pi_k^{(N)}, 1]_N - [\pi_k^{(N)}, 1]_N + \langle \pi_k^{(N)}, 1 \rangle \right]
= r_n \sum_{k=1}^{K} \pi_k^{(N)}(x_n) \left[ [\pi_k^{(N)}, 1]_N - \epsilon_k^{(N)} \right]
\]

for \( n = 1, \ldots, N \). Because of \((R3)\), we can assume that \( \varphi_1 \equiv 1 \) and consequently \( \pi_1^{(N)} = 1/\|1\|_N \). This implies

\[
(3.31) \quad [\pi_k^{(N)}, 1]_N = \|1\|_N [\pi_k^{(N)}, \pi_1^{(N)}]_N = \|1\|_N \delta_{1,k},
\]

since \( \{\pi_k^{(N)}\}_{k=1}^K \) is assumed to be orthonormal with respect to the discrete inner product \([\cdot, \cdot]_N \). Thus, the LS weights can further be rewritten as

\[
(3.32) \quad w_{n}^{\text{LS}} = r_n \sum_{k=1}^{K} \pi_k^{(N)}(x_n) \left( \|1\|_N \delta_{1,k} - \epsilon_k^{(N)} \right)
= r_n \left( \|1\|_N \pi_1^{(N)}(x_n) - \sum_{k=1}^{K} \epsilon_k^{(N)} \pi_k^{(N)}(x_n) \right)
= r_n \left( 1 - \sum_{k=1}^{K} \epsilon_k^{(N)} \pi_k^{(N)}(x_n) \right).
\]

Hence, the assertion \( w_{n}^{\text{LS}} > 0 \) for all \( n = 1, \ldots, N \) is equivalent to

\[
(3.33) \quad \sum_{k=1}^{K} \epsilon_k^{(N)} \pi_k^{(N)}(x_n) < 1, \quad n = 1, \ldots, N.
\]
At the same time, (3.26) and Lemma 3.4 imply that every element of the discrete orthonormal basis converges uniformly to the corresponding element of the continuous orthonormal basis. In particular, for every $k = 1, \ldots, K$, the function sequence $(\pi_k(N))_{N \in \mathbb{N}} \subset \mathcal{F}_K(\Omega)$ is uniformly bounded. Thus, there exists a constant $C > 0$ such that

$$
\sum_{k=1}^{K} \varepsilon_k^{(N)} \pi_k^{(N)}(x_n) \leq C \sum_{k=1}^{K} \left| \varepsilon_k^{(N)} \right|, \quad n = 1, \ldots, N.
$$

Moreover, Lemma 3.3 implies $\lim_{N \to \infty} \varepsilon_k^{(N)} = 0$ for all $k = 1, \ldots, K$. Hence, there exists an $N_0 \geq K$ such that

$$
\left| \varepsilon_k^{(N)} \right| < \frac{1}{KC}, \quad k = 1, \ldots, K,
$$

for all $N \geq N_0$. Finally, this yields (3.33) and therefore the assertion.

A simple consequence of Theorem 3.5 is the subsequent corollary in which the special case of equidistributed data points is considered.

**Corollary 3.6.** Given are $\Omega \subset \mathbb{R}^d$, $\omega : \Omega \to \mathbb{R}_0^+$, and $\mathcal{F}_K(\Omega) \subset C(\Omega)$ such that the restrictions (R2) and (R3) are satisfied, i. e.,

(R2) The weight function $\omega : \Omega \to \mathbb{R}_0^+$ is Riemann integrable and positive almost everywhere.

(R3) The $K$-dimensional vector space $\mathcal{F}_K(\Omega)$ is spanned by a basis $\{\varphi_k\}_{k=1}^K$ of continuous and bounded functions $\varphi_k : \Omega \to \mathbb{R}$, $k = 1, \ldots, K$. Furthermore, $\mathcal{F}_K(\Omega)$ contains constants. In particular, $1 \in \mathcal{F}_K(\Omega)$.

Let $(x_n)_{n \in \mathbb{N}} \subset \Omega$ be a equidistributed sequence with $\omega(x_n) > 0$ for all $n \in \mathbb{N}$. Then, there exists an $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$ and discrete weights

$$
r_n = \frac{\Omega \omega(x_n)}{N}, \quad n = 1, \ldots, N,
$$

the corresponding LS-CF

$$
C_{N}^{\text{LS}}[f] = \sum_{n=1}^{N} u_{\text{LS}}^{f}(x_n) \quad \text{with} \quad w^{\text{LS}} = \arg \min_{\Phi w = m} \left\| R^{-1/2} w \right\|_2^2,
$$

where $R^{-1/2} = \text{diag}(1/\sqrt{r_1}, \ldots, 1/\sqrt{r_N})$, is positive and $\mathcal{F}_K(\Omega)$-exact.

**Proof.** Recall that the equidistributed sequence $(x_n)_{n \in \mathbb{N}} \subset \Omega$ satisfies (2.6) for all measurable bounded functions that are continuous almost everywhere and, by Corollary 2.5, is $\mathcal{F}_K(\Omega)$-unisolvent. In particular, (3.26) holds for $(x_n)_{n \in \mathbb{N}}$ and the discrete weights $(r_n)_{n \in \mathbb{N}}$ defined as in (3.36). In combination with (R2) and (R3), Theorem 3.5 therefore implies the assertion.

---

6If $(y_n)_{n \in \mathbb{N}}$ is a convergent sequence in the normed vector space $(Y, \| \cdot \|)$, then $(y_n)_{n \in \mathbb{N}}$ is bounded. This is because for any $\varepsilon > 0$ we can find an $N \in \mathbb{N}$ such that $\|y - y_n\| \leq \varepsilon$ for all $n > N$, where $y$ denotes the limit of $(y_n)_{n \in \mathbb{N}}$. One can then choose $C = \|y\| + \max\{\|y - y_1\|, \ldots, \|y - y_N\|, \varepsilon\}$ to get $\|y_n\| \leq \|y\| + \|y - y_n\| \leq C$ for all $n$, which shows that $(y_n)_{n \in \mathbb{N}}$ is bounded.
Remark 3.7. (R1) is not necessary for Corollary 3.6 to hold. However, following Remark 2.4, (R1) ensures the existence—and simple construction—of an equidistributed sequence in $\Omega$.

Remark 3.8. Corollary 3.6 implies that if $(x_n)_{n \in \mathbb{N}} \subset \Omega$ is equidistributed, then for sufficiently large $N$, there exists a positive and exact CF on the set $\{x_n\}_{n=1}^N$. As we will discuss in subsection 5.2, thus there also exists a positive interpolatory CF—predicted by Theorem 1.1—on $\{x_n\}_{n=1}^N$. Such sets, on which a positive and interpolatory CF is supported, are referred to as Tchakaloff sets. Hence, Corollary 3.6 also implies that if $(x_n)_{n \in \mathbb{N}} \subset \Omega$ is equidistributed, then for sufficiently large $N$, $\{x_n\}_{n=1}^N$ is a Tchakaloff set. Similar results were obtained in [27] and [96] for everywhere dense sequences. However, Corollary 3.6 does not just provide us with a Tchakaloff set, but also tells us that a positive and exact CF can be obtained in the form of a simple weighted LS solution of the linear system (2.2).

4. Convergence and Error Analysis. Here, we address the convergence of positive and exact LS-CFs.

4.1. The Lebesgue Inequality and Convergence. Assume that the positive LS-CF $C_N$ is exact for all functions from the function space $\mathcal{F}_K(\Omega)$. Moreover, let $f : \Omega \to \mathbb{R}$ be a continuous function, and let us denote a best approximation of $f$ from $\mathcal{F}_K(\Omega)$ with respect to the $L^\infty(\Omega)$-norm by $\hat{s}$. That is,

\[
\hat{s} = \arg \min_{s \in \mathcal{F}_K(\Omega)} \|f - s\|_{L^\infty(\Omega)} \quad \text{with} \quad \|f - s\|_{L^\infty(\Omega)} = \sup_{x \in \Omega} |f(x) - s(x)|.
\]

Then, the following error bound holds:

\[
|C_N[f] - I[f]| \leq \|I\|_\infty \|f - \hat{s}\|_{L^\infty(\Omega)} + \|C_N\|_\infty \|f - \hat{s}\|_{L^\infty(\Omega)}
\]

\[
= \left(\|I\|_\infty + \|C_N\|_\infty\right) \left(\inf_{s \in \mathcal{F}_K(\Omega)} \|f - s\|_{L^\infty(\Omega)}\right).
\]

Inequality (4.2) is commonly known as the Lebesgue inequality; see, e.g., [93] or [12, Theorem 3.1.1]. It is most often encountered in the context of polynomial interpolation [14, 55] but straightforwardly carries over to numerical integration. In this context, the operator norms $\|I\|_\infty$ and $\|C_N\|_\infty$ are respectively given by

\[
\|I\|_\infty = \int_{\Omega} \omega(x) \, dx = I[1], \quad \|C_N\|_\infty = \sum_{n=1}^N |w_n|.
\]

Recall that the CF $C_N$ is positive and exact for constants (we assume that $\mathcal{F}_K(\Omega)$ contains constants). Thus, we have

\[
\|C_N\|_\infty = C_N[1] = I[1] = \|I\|_\infty.
\]

In particular, this implies that the Lebesgue inequality (4.2) simplifies to

\[
|C_N[f] - I[f]| \leq 2\|I\|_\infty \left(\inf_{s \in \mathcal{F}_K(\Omega)} \|f - s\|_{L^\infty(\Omega)}\right).
\]
Based on (4.2), we can note the following: Assume that we are given a sequence of positive CFs \((C_N)_{N \in \mathbb{N}}\) with \(C_N\) being exact for \(F_K(\Omega)\), where \(K = K(N)\). Sequences of CFs are usually referred to as cubature rules (CRs). Let \(F_K(\Omega) \subset F_{K+1}(\Omega)\) for all \(K \in \mathbb{N}\) and let \(\bigcap_{K \in \mathbb{N}} F_K(\Omega)\) be dense in \(C(\Omega)\) with respect to the \(L^\infty(\Omega)\)-norm. If \(K(N) \to \infty\) for \(N \to \infty\), then \((C_N)_{N \in \mathbb{N}}\) converges to the continuous integral, \(I\), for all continuous functions. That is, for all \(f \in C(\Omega)\),

\[
C_N[f] \to I[f], \quad N \to \infty, \quad \text{in } (\mathbb{R}, |\cdot|).
\]

It should be stressed that the particular rate of convergence of \((C_N[f])_{N \in \mathbb{N}}\) to \(I[f]\) depends on the (smoothness) of the function \(f\) as well as the finite-dimensional function spaces \(F_K(\Omega)\). In particular, a more detailed error analysis based on (4.5) relies on some knowledge about the quality of the \(L^\infty(\Omega)\) best approximation from \(F_K(\Omega)\). Results of this flavor are usually referred to as Jackson-type theorems; the subject of constructive function theory [70].

### 4.2. Error Analysis for Analytic Functions

For simplicity, we now restrict ourselves to analytic functions on the \(d\)-dimensional hypercube \(\Omega = [0, 1]^d\). Moreover, we assume that \((C_N)_{N \in \mathbb{N}}\) is a CR with \(C_N\) being positive and exact for all \(d\)-dimensional polynomials up to total degree \(m = m(N)\). That is, \(F_K(\Omega) = \mathbb{P}_m(\mathbb{R}^d)\). In this case, the following result holds for the LS-CFs.

**Lemma 4.1.** Let \(f : [0, 1]^d \to \mathbb{R}\) be analytic in an open set containing \([0, 1]^d\). Then the CR \((C_N)_{N \in \mathbb{N}}\), where \(C_N\) is positive and \(\mathbb{P}_m(\mathbb{R}^d)\)-exact, with \(m = m(N)\), satisfies

\[
|I[f] - C_N[f]| = \mathcal{O}\left(\exp(-cm/\sqrt{d})\right)
\]

for some constant \(c > 0\).

**Proof.** Since \(f\) is analytic in an open set containing \([0, 1]^d\) it can be approximated by a \(d\)-dimensional polynomial of total degree \(m\) as

\[
\inf_{s \in \mathbb{P}_m(\mathbb{R}^d)} \|f - s\|_{L^\infty([0, 1]^d)} = \mathcal{O}\left(\exp(-cm/\sqrt{d})\right);
\]

see [69, Equation 5.8]. Here, \(c\) is a constant depending on the location of the singularity of \(f\) (if there is any) nearest to \([0, 1]^d\) with respect to the radius of the Bernstein ellipse. Finally, combining (4.8) with the Lebesgue inequality (4.5) immediately yields the assertion.

Let us briefly address the relation between the number of data points \(N\) and the maximum total degree \(m\) for which the LS-CF is positive. First, it should be noted that the dimension of \(\mathbb{P}_m(\mathbb{R}^d)\) is \(K = \binom{d+m}{d}\). This implies the asymptotic relation \(\lim_{m \to \infty} K/m^d = 1/d!\); in particular, \(K = \mathcal{O}(m^d)\). Furthermore, in subsection 6.2, we observe the relation between \(K\) and \(N\) to be of the form \(N = \mathcal{O}(K^2)\). Assuming this relation, we get the following version of Lemma 4.1.

---

7The relation between the number of data points \(N\) and the maximum total degree \(m\) remains to be addressed.
Corollary 4.2. Let $f : [0, 1]^d \to \mathbb{R}$ be analytic in an open set containing $[0, 1]^d$ and let $(C_N)_{N \in \mathbb{N}}$ be a CR with $C_N$ being positive and $P_m(\mathbb{R}^d)$-exact. Assume that we have the asymptotic relation $N = O(K^2)$ with $K = \dim P_m(\mathbb{R}^d)$, then

\begin{equation}
|I[f] - C_N[f]| = O\left(\exp\left(-cN^{1/2d}/\sqrt{d}\right)\right)
\end{equation}

for some constant $c > 0$.

Proof. Recall that $g(N) \in O(h(N))$ if and only if there exists a constant $M > 0$ such that $g(N) \leq M h(N)$ for sufficiently large $N$. Henceforth, let $c > 0$ be a generic constant. The assertion follows from noting that there exists a constant $M > 0$ such that

\begin{equation}
|I[f] - C_N[f]| \leq M \exp\left(-cm/\sqrt{d}\right)
\end{equation}

for sufficiently large $N$. Here, the first inequality follows from Lemma 4.1, the second one from the fact that $K = O(m^d)$, and the third one from the assumption that $N = O(K^2)$.

We already point out that for the positive interpolatory CFs discussed in subsection 5.2, we have $N = K$ (rather than just $N = O(K^2)$) and therefore

\begin{equation}
|I[f] - C_N[f]| = O\left(\exp\left(-cN^{1/2d}/\sqrt{d}\right)\right)
\end{equation}

instead of (4.9).

Remark 4.3. The above error analysis can easily be extended whenever a relation analogue to (4.8) is available. See [76, 3] for a discussion of some other classes of functions $f$, domains $\Omega$, and function spaces $F_K(\Omega)$.

Remark 4.4. The 'impossibility' theorem proved in [73] states that any procedure for approximating univariate functions from equally spaced samples that converges exponentially fast must also be exponentially ill-conditioned. Observe that for $d = 1$, (4.11) implies root-exponential convergence for the LS-CF, which allows us to avoid such inherent stability issues.

Remark 4.5. It was argued in some recent works [88, 87, 89]—also see [10]—that for a certain class of functions (analytic in the hypercube with singularities outside), the Euclidean degree should be considered instead of the total or maximum degree. However, we did not observe any advantage in using the Euclidean degree in our numerical tests.

5. Some Applications. We discuss two applications of the provable positive and exact LS-CFs. These address the simple construction of positive high-order CRs (subsection 5.1) and positive interpolatory CFs (subsection 5.2). In both cases, the procedure again only relies on basic linear algebra operations.

5.1. A Simple Procedure To Construct Positive High-Order Cubature Rules. Henceforth, we make the same assumptions as in Corollary 3.6. That is, $\Omega \subset \mathbb{R}^d$ is bounded with a boundary of measure zero and $\omega : \Omega \to \mathbb{R}_0^+$ is Riemann integrable and positive almost
everywhere. Moreover, let \((x_n)_{n \in \mathbb{N}}\) be an equidistributed sequence in \(\Omega\) with \(\omega(x_n) > 0\) for all \(n \in \mathbb{N}\). Given is a sequence of increasing function spaces \((\mathcal{F}_K(\Omega))_{K \in \mathbb{N}}\) with \(1 \in \mathcal{F}_K(\Omega) \subset C(\Omega)\), \(\mathcal{F}_K(\Omega) \subset \mathcal{F}_{K+1}(\Omega)\), and \(\dim \mathcal{F}_K(\Omega) = K\) for all \(K \in \mathbb{N}\). For simplicity, we shall assume that \(\bigcap_{K \in \mathbb{N}} \mathcal{F}_K(\Omega)\) is dense in \(C(\Omega)\) with respect to the \(L^\infty(\Omega)\)-norm. Following the discussion in subsection 4.1, this will ensure convergence of the subsequent CR for all continuous functions.

Under these assumptions, the procedure works as follows: In every step, we increase the dimension \(K\) and find a positive and \(\mathcal{F}_K(\Omega)\)-exact CF by increasing the number of data points until the corresponding LS-CF \(C_N\) is positive. Algorithm 5.1 contains an informal summary of the procedure for fixed \(K\).

**Algorithm 5.1 Constructing Positive High-Order Cubature Formulas**

1: \(N = K, r = 0,\) and \(w_{\text{min}} = 0\)
2: **while** \(r < K\) or \(w_{\text{min}} < 0\) **do**
3: \(X_N = \{x_n\}_{n=1}^N\)
4: Compute the matrix \(\Phi = \Phi(X_N)\)
5: Compute the rank of \(\Phi: r = \text{rank}(\Phi)\)
6: **if** \(r = K\) **then**
7: Compute the LS weights \(w^{\text{LS}}\) as in (3.1)
8: Determine the smallest weight: \(w_{\text{min}} = \min(w^{\text{LS}})\)
9: \(N = 2N\)

Algorithm 5.1 is ensured to terminate due to the theoretical findings presented in section 2 and section 3. In particular, Corollary 2.5 tells us that for a sufficiently large number of (equidistributed) nodes, these will be \(\mathcal{F}_K(\Omega)\)-unisolvent. This is equivalent to the rows of \(\Phi\) to be linearly independent (\(\text{rank}(\Phi) = K\)). Hence, \(r = K\) is ensured for sufficiently large \(N\).

At the same time, Corollary 3.6 implies that the LS weights \(w^{\text{LS}}\) are positive for a sufficiently large \(N\).

**Remark 5.1 (Monte Carlo CFs).** The LS-CFs discussed above can be seen as high-order corrections to QMC methods [16, 30] in the case that low-discrepancy data points are used. Recall that the weights in QMC integration are \(w_n = |\Omega|\omega(x_n)/N\). At the same time, the LS weights are explicitly given by \(w^{\text{LS}}_n = r_n \sum_{k=1}^{K} \pi_k(x_n; r) I[\pi_k(\cdot; r)]\), see (3.13). Here, \(\{\pi_k(\cdot; r)\}_{k=1}^{K}\) is a discrete orthonormal basis. For fixed \(K\) and an increasing number of data points, \(\pi_k(x_n; r) I[\pi_k(\cdot; r)]\) converges to the Kronecker delta \(\delta_{1,k}\). Hence, the difference between the QMC and LS weights converges to zero.

**Remark 5.2 (Exact Integration of Discrete LS Approximations).** The LS-CF \(C_N[f]\) corresponds to exact integration of the following discrete LS approximation of \(f\) from the function space \(\mathcal{F}_K(\Omega) = \text{span}\{\varphi_1, \ldots, \varphi_K\}\) (assuming it is unique):

\[
\hat{f}(x) = \sum_{k=1}^{K} c_k \varphi_k(x) \quad \text{s.t.} \quad \left\| R^{1/2}(\Phi^T c - f) \right\|_2 \text{ is minimized},
\]

where \(c = (c_1, \ldots, c_K)^T\). That is, \(C_N[f] = I[\hat{f}]\). This can be noted by representing \(\hat{f}\) with
respect to a basis of discrete orthonormal functions \( \{ \pi_k(\cdot; r) \}_{k=1}^K \):

\[
\hat{f}(x) = \sum_{k=1}^K [f, \pi_k(\cdot; r)]_N \pi_k(x; r)
\]

Integration therefore yields

\[
I[\hat{f}] = \sum_{k=1}^K [f, \pi_k(\cdot; r)]_N I[\pi_k(\cdot; r)] = \sum_{n=1}^N \left( \sum_{k=1}^K \pi_k(x_n; r) I[\pi_k(\cdot; r)] \right) f(x_n) = C_N[f].
\]

The last equality follows from (3.13). Building upon this connection, in [67] high-order randomized CFs for independent random points were constructed. These were shown to be positive and exact with a high probability if the number of (random) data points is sufficiently larger than \( K \). In particular, it was stated that the proportionality between \( N \) and \( K \) should be at least quadratic. This is in accordance with the results presented here. Also the fundamental work [78] on hyperinterpolation should be mentioned in this context.

**5.2. Constructing Positive Interpolatory Cubature Formulas.** Next, we describe how provable positive and exact LS-CFs can be used to construct interpolatory CFs with the same properties. In contrast to LS-CFs, interpolatory CFs use a smaller subset of \( N = K \) data points, where \( K \) denotes the dimension of the function space \( \mathcal{F}_K(\Omega) \) for which the original LS-CF is exact. In fact, there exist many different approaches to this task. For instance, in [13] a nonlinear optimization procedure is used to downsample formulas one data point at a time (although the work focuses on the one-dimensional case). It is also possible to formulate (3.1) as a basis pursuit problem which can then be solved by linear programming tools [40]. Another option is Caratheodory–Tchakaloff subsampling [72, 11], which may be implemented using linear (or quadratic) programming. Finally, NNLS [63] could be used to recover a sparse nonnegative weight vector (see [54] for the univariate case and [81] for the multivariate case). Another procedure is a method due to Steinitz [83] (also see [27, 97]). While this method might be less efficient than some of the aforementioned approaches, it is fairly simple and—as the construction of the positive LS-CFs—only relies on basic operations from linear algebra. Details on Steinitz’ method can be found in Appendix A.

**6. Numerical Results.** We now come to present several numerical tests to illustrate our theoretical findings as well as to demonstrate the performance of the positive high-order LS-CFs and corresponding interpolatory CFs. All experiments were performed using a 2.6 GHz 6-Core Intel Core i7 processor with 32 GB of random access memory (RAM). The MATLAB code used to generate the subsequent numerical results can be found on GitHub\(^8\).

We consider two different types of data points: (1) Halton points, which are deterministic and from a low-discrepancy sequence\(^9\) [49, 71, 61], and (2) uniformly distributed random points. An illustration for these points is provided by Figure 1.

---

\(^8\)See [https://github.com/jglaubitz/positive_CFs](https://github.com/jglaubitz/positive_CFs)

\(^9\)Recall that low-discrepancy sequences are a subclass of equidistributed sequences.
Remark 6.1. To avoid ill-conditioning caused by a poor choice of $\{\varphi_k\}_{k=1}^K$ we recommend to use continuous orthonormal basis functions satisfying

\[ \int_{\Omega} \varphi_k(x) \varphi_l(x) \, dx = \delta_{kl}, \quad k, l = 1, \ldots, K, \]

to formulate the linear system (2.2) that is solved in the LS sense to obtain the weights of the LS-CFs. Observe that in this case

\[ \lim_{N \to \infty} \frac{1}{N} \Phi \Phi^T = \lim_{N \to \infty} \left( \frac{1}{N} \sum_{n=1}^{N} \varphi_k(x_n) \varphi_l(x_n) \right)_{k,l=1}^{K} = \left( \int_{\Omega} \varphi_k(x) \varphi_l(x) \, dx \right)_{k,l=1}^{K} = I, \]

where $I$ denotes the $(K \times K)$ identity matrix. Then, we further find\(^{10}\)

\[ \lim_{N \to \infty} \kappa(\Phi) = \lim_{N \to \infty} \frac{\lambda_{\text{max}}(\Phi \Phi^T)}{\lambda_{\text{min}}(\Phi \Phi^T)} = \lim_{N \to \infty} \frac{\lambda_{\text{max}}(\frac{1}{N} \Phi \Phi^T)}{\lambda_{\text{min}}(\frac{1}{N} \Phi \Phi^T)} = 1 \]

for the condition number of $\Phi$, where $\lambda_{\text{max}}(A)$ and $\lambda_{\text{min}}(A)$ are the largest and smallest eigenvalue of the matrix $A$.

6.1. Comparison of Different Subsampling Methods. We start by debate the subsampling method used to obtain a positive interpolatory CF from a given positive LS-CF, both exact for the same function space $F_K(\Omega) \subset C(\Omega)$. A few options to construct such positive interpolatory CFs were discussed in subsection 5.2. An exhaustive comparison of all these approaches would exceed the scope of this work, but we at least provide a rudimentary comparison of three especially simple methods: (i) Steinitz’ method, see [83, 27, 97] as well as Appendix A; (ii) NNLS [63] (also see [54] for the univariate case and [81] for the multivariate case); and (iii) basis pursuit formulated as a linear programming problem [40].

\(^{10}\)If $A \in \mathbb{R}^{K \times K}$ has eigenvalues $\lambda_1, \ldots, \lambda_K$ and $p$ is a polynomial, then the matrix $p(A)$ has eigenvalues $p(\lambda_1), \ldots, p(\lambda_K)$. 
Figure 2: A comparison between different subsampling methods to construct a positive interpolatory CF from a positive LS-CF. All tests are performed for the domain $\Omega = [-1,1]^2$ with $\omega \equiv 1$.

Figure 2 provides a comparison of the interpolatory CFs resulting from these three methods. The underlying positive LS-CFs was constructed for $\Omega = [-1,1]^2$ with $\omega \equiv 1$ and to be exact for algebraic polynomials up to a total degree $m \in \{0,1,\ldots,14\}$. For this case all three methods produced positive and interpolatory ($N = K$) CFs. Thus, in Figure 2 we only focus on the efficiency and “exactness” of these methods. Efficiency is measured by the time it took the respective method to produce a positive interpolatory CF from a given positive LS-CF. The normalized times are displayed in Figures 2a and 2b. Based on these results it might be argued that NNLS is the most efficient method, while the Steinitz approach is the least efficient one. On the other hand, “exactness” is measured by the error in the exactness conditions, that is present in the produced interpolatory CF. This is measured by the residual of the moment conditions, $\|\Phi w - m\|_2$. Such errors can be introduced due to computing in finite arithmetics (rounding errors) or a method not arriving at a solution within the performed number of iterations. We observe in Figures 2c and 2d that especially the NNLS method suffers from undesirably high errors. Regarding exactness, we observe that the Steinitz method performs best. We also report that the basis pursuit approach sometimes did not produce
a positive interpolatory CF for total degrees \( m > 14 \). Henceforth, we restrict ourselves to positive interpolatory CFs constructed by the Steinitz method, although we observe it to be inferior in terms of efficiency. This does not take into account that these methods are often used to solve partial differential equations and must be fast. A more detailed investigation and comparison of different subsampling strategies would therefore be of interest.

Remark 6.2. In our numerical tests, we used the Matlab built-routine lsqnonneg. However, it should be pointed out that there are NNLS methods available, which might perform even faster. See [29] and references therein as well as [77].

6.2. Ratio Between \( N \) and \( K \). We now investigate the relation between the number of data points \( N \) and the dimension of the function space \( K \) for which the LS-CF is positive. Recall that Corollary 3.6 stated that the (weighted) LS-CF is ensured to be positive for fixed \( K \) if a sufficiently large number \( N \) of (equidistributed) data points is used. Let \( N(K) \) denote the lowest number of data points for the LS-CF (with discrete weights as in Corollary 3.6) to be positive. Numerically, we found the relation between \( K \) and \( N(K) \) to be of the form \( N(K) = CK^s \) with \( s \) being close to or even below 2 in many different cases.

Figure 3 illustrates this for the hyper-cube \( \Omega = [-1, 1]^2 \) with weight function \( \omega \equiv 1 \). The corresponding LS-CFs were constructed to be exact for algebraic polynomials up to a fixed total degree (Figures 3a and 3d), trigonometric polynomials up to a fixed degree (Figures 3b
and 3e), and cubic polyharmonic RBFs (PHS-RBFs) augmented with a constant (Figures 3c and 3f). See [33] and references therein for more details on RBFs. Figure 3 also reports on the ratio between $K$ and $N(K)$ for the product Legendre rule and the positive interpolatory CF. The positive interpolatory CF was obtained from the LS-CF by subsampling using the Steinitz method.

The numerically observed values for $C$ and $s$ in the assumed relation $N(K) = CK^s$ for some further cases are listed in Table 1. The reported parameters $C$ and $s$ were obtained by performing an LS fit for these based on the values of $K$ and $N = N(K)$ for maximum total degree $m \in \{0, 1, \ldots, 10\}$. We note that the results reported here appear to be in accordance with similar observations made in previous works for certain special cases; see [99, 54, 44, 39, 38] (for univariate LS-QFs) and [40] (for multivariate LS-CFs based on polynomials). Interestingly, similar ratios were also observed in other contexts, such as discrete LS function approximations [21, 23] and stable high-order randomized CFs based on these [67].

6.3. Polynomial Based Cubature Formulas on the Hyper-Cube. We now investigate the accuracy of the positive high-order LS-CFs and corresponding interpolatory CFs. To this end, we consider the hyper-cube $\Omega = [0, 1]^d$ with $\omega \equiv 1$ and the following Genz test functions [36]...
Figure 4: Errors for Genz’ test functions on $\Omega = [0, 1]^2$ with $\omega = 1$. The LS- and interpolatory CF are based on multivariate polynomials of increasing total degree ($F_K(\Omega) = \mathbb{P}_m(\mathbb{R}^2)$).

(also see [93]):

$$g_1(x) = \prod_{i=1}^{d} \left(a_i^{-2} + (x_i - b_i)^2\right)^{-1} \quad \text{(product peak)},$$

$$g_2(x) = \left(1 + \sum_{i=1}^{d} a_i x_i\right)^{-(d+1)} \quad \text{(corner peak)},$$

$$g_3(x) = \exp\left(-\sum_{i=1}^{d} a_i^2(x_i - b_i)^2\right) \quad \text{(Gaussian)}$$

These functions are designed to have different difficult characteristics for numerical integration routines. The vectors $\mathbf{a} = (a_1, \ldots, a_q)^T$ and $\mathbf{b} = (b_1, \ldots, b_q)^T$ respectively contain (randomly chosen) shape and translation parameters. For each case, the experiment was repeated 20 times, with the vectors $\mathbf{a}$ and $\mathbf{b}$ being drawn randomly from $[0, 1]^d$.

Figures 4 and 5 illustrate the errors of different CFs for this test case in two ($d = 2$) and three ($d = 3$) dimensions, respectively. Compared are the positive high-order LS-CF, the corresponding interpolatory CF, the (Q)MC method, and a product Legendre method. In
this context, the term “(Q)MC” refers to a CF of the form

$$C_N[g] = \frac{|\Omega|}{N} \sum_{n=1}^{N} g(x_n) \quad \text{with} \quad g = f \omega.$$  

This corresponds to an MC method if the data points are sampled randomly and to a QMC method if the data points are not fully random but correspond to a (semi- or fully-) deterministic low-discrepancy sequence [71, 16, 30]. Further, the LS-CF was constructed to be exact for all two- and three-dimensional polynomials up to total degree $m \in \{0, 1, \ldots, 20\}$ and $m \in \{0, 1, \ldots, 12\}$, respectively. For this simple test problem, it can be observed from Figures 4 and 5 that the product Legendre rule yields the most accurate results in most cases, followed by the interpolatory and underlying positive high-order LS-CF.

6.4. A Nonconstant Weight Function. We next consider a test case with the nonconstant weight function $\omega(x_1, x_2) = \left(1 - x_1^2\right)^{1/2} \left(1 - x_2^2\right)^{1/2}$ on the hyper-cube $C_2 = [-1, 1]^2$. For this test function and corresponding interpolatory CF are again compared to a (Q)MC method applied to the same data points as the LS-CF and a (transformed) product Legendre rule applied to $f \omega$ as an integrand. We
can note from Figure 6 that the positive interpolatory CF—and for Halton points even the LS-CF—yield more accurate results than the product Legendre rule this time.

### 6.5. A Nonstandard Domain.
As mentioned before, the proposed LS-CF has the advantage of being easily applicable to nonstandard domains, which is demonstrated now. Consider the two-dimensional domain $\Omega$ that is illustrated in Figures 7a and 7b with weight function $\omega \equiv 1$. Figures 7c and 7d report on the errors of the LS-CF and the corresponding interpolatory CF compared to a (Q)MC method. These are again constructed to be exact for algebraic polynomials up to an increasing total degree $m \in \{0, 1, \ldots, 14\}$, and the test function is $f(x_1, x_2) = \exp(x_1^2 + x_2^2)$. It should be noted that even in this case, of a nonstandard domain, the positive high-order LS-CF and the corresponding interpolatory CF display encouraging rates of convergence. At the same time, we were unable to use a simple product Legendre rule for this problem.

### 6.6. A Simple Periodic Problem.
Beside being easily applicable to different domains and weight functions, the LS-CFs can also be constructed to be exact for different function spaces. The specific function space can be chosen based on some prior information about the function $f$ that is integrated. To start with a simple example, consider the periodic function $f(x) = \cos(\pi x)e^{\sin(\pi x)}$ on $\Omega = [-1, 1]$ with weight function $\omega \equiv 1$. In this case, it is of advantage to construct the LS-CF to be exact for trigonometric functions rather than polynomials. The corresponding errors of the LS formulas on equidistant and random points that are exact for polynomials (“LS poly”) and trigonometric functions (“LS trig”) of increasing degree $m \in \{0, 1, \ldots, 30\}$ are illustrated in Figure 8. We also chose this simple problem because it allows us to compare the LS formulas to the (composite) trapezoidal rule when equidistant points are used. In fact, we can observe in Figure 8a that the LS formula which is exact for trigonometric functions coincides with the trapezoidal rule in many cases. In particular, both of them yield highly accurate results for this periodic test problem. Furthermore, the
LS formulas based on trigonometric functions is also able to yield more accurate results than the LS formula based on polynomials for random data points (see Figure 8b).

6.7. Radial Basis Functions. We next consider functions spaces based on positive definite radial basis functions (RBFs) [33]. CFs based on RBFs are a popular tool for scattered data [80, 34, 74, 82]. This is because RBF interpolants can be ensured to uniquely exist for arbitrary point distributions, which is not the case for many other functions (Haar spaces) due to the Mairhuber–Curtis theorem [65, 25].\footnote{The Mairhuber–Curtis theorem tells us that if we want to have a well-posed multivariate scattered data interpolation problem, then the function space needs to depend on the data locations.} The idea behind RBF-CFs is to form an RBF interpolant and to exactly integrate it. However, RBF formulas are not always ensured to be positive [45], which can deteriorate their performance for noisy measurements. Here, we demonstrate that the LS approach can be used to stabilize RBF-CFs. This can be ex-
Figure 8: Errors for $f(x) = \cos(\pi x)e^{\sin(\pi x)}$ on $\Omega = [-1, 1]$ with $\omega \equiv 1$ using stable high-order LS-CFs based on algebraic and trigonometric function spaces as well as the trapezoidal rule.

explained by an LS-CF replacing the (potentially unstable) RBF interpolant by a more stable LS approximation from the same RBF function space. Figure 9 illustrates this for the RBF function space spanned by a constant and the functions $\varphi(\|x - x_k\|)$ on $\Omega = [0, 1]^2$ using a Gaussian kernel $\varphi(r) = e^{(\epsilon r)^2}$ with shape parameter $\epsilon = 0.75$. Further the number of centers was $K \in \{1, 2, \ldots, 60\}$. Figure 9a provides the values of the smallest weight for the CF based on exact integration of RBF interpolants (“RBF”) and the stable LS-CF that is exact for the same RBF function space (“LS-RBF”). Note that the RBF formula is found to feature negative weights, which can result in stability issues, while the LS-RBF formula is having positive weights in all cases. Further, Figure 9b reports on the errors of the RBF and LS-RBF formulas on $N$ random points applied to Genz’ first test function $g_1$ on $\Omega = [0, 1]^2$ with $\omega \equiv 1$. In this example, both formulas perform similarly. In Figures 9c and 9d we repeat this experiment but add uniformly distributed noise of magnitude $10^{-4}$ and $10^{-2}$ to the function values at the data points. Observe that the accuracy of the RBF formulas deteriorates notably stronger than of the LS-RBF formula in the presence of noise, due to the improved stability of the latter. We made the same observation also for other point distributions and Genz test functions.

6.8. Summation-By-Parts Operators for General Function Spaces. Another useful application of the positive and exact LS-CFs presented here is the construction of summation-by-parts (SBP) operators for numerical differentiation, which are able to mimic integration-by-parts on a discrete level—thus the name. For this reason they are popular building blocks for systematically developing stable and high-order accurate numerical methods for time-dependent differential equations, see [84, 32]. SBP operators have been developed based on the idea that the solution is assumed to be well approximated by polynomials up to a certain degree, and the SBP operator should therefore be exact for them. However, polynomials might not provide the best approximation for some problems, and other function spaces should be considered. To illustrate this, consider the boundary layer solution in Figure 10a, which
Figure 9: Smallest weights and errors for Genz’ first test function $g_1$ on $\Omega = [0, 1]^2$ with $\omega \equiv 1$ using an RBF and LS-RBF formula. In all cases random points and a Gaussian kernel $\varphi(r) = e^{(\varepsilon r)^2}$ with shape parameter $\varepsilon = 0.75$ was used.

demonstrates the advantage of using an exponential instead of a polynomial approximation space.

Consequently, in a recent work [43], we developed SBP operators for general functions spaces, referred to as FSBP operators. These allowed us to systematically develop stable and high-order accurate numerical methods that are based on general approximation spaces. The advantage of using such a method is demonstrated in Figure 10b for the inhomogeneous linear advection problem

$$\partial_t u + \partial_x u = 2u, \quad 0 < x < \pi,$$

$$u(x, 0) = 1, \quad 0 \leq x \leq \pi,$$

$$u(0, t) = 1, \quad t \geq 0,$$

with exact steady state solution $u(x) = e^{2x}$. The steady state solution can be expected to
be better approximated using an exponential rather than a polynomial approximation space. However, the construction of an FSBP operator that is based on a function space \( F \) requires that we have a positive and \( (F,F)' \)-exact quadrature, where
\[
(F,F)' = \{ (fg)' \mid f, g \in F \}.
\]
Such quadrature formulas can be constructed using the generalized LS approach introduced in the present work. As an example, consider the following exponential function space on \( \Omega = [0, 1] \):
\[
\mathcal{E}_2 = \text{span}\{1, x, e^x\} \quad \text{with} \quad (\mathcal{E}_2 \mathcal{E}_2)' = \text{span}\{1, x, e^x, xe^x, e^{2x}\}
\]
Using \( N = 5 \) equidistant grid points, we found the LS formula with the following points and weights to be positive and \( (\mathcal{E}_2 \mathcal{E}_2)' \)-exact:
\[
x = \begin{bmatrix} 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \end{bmatrix}^T, \quad w = \begin{bmatrix} 2, 9, 3, 9, 2 \end{bmatrix}^T \begin{bmatrix} \frac{25}{25}, \frac{25}{25}, \frac{25}{25}, \frac{25}{25} \end{bmatrix},
\]
where we have rounded the numbers to the second decimal place. The corresponding FSBP operator \( D \) that was used to generate the numerical results in Figure 10b is
\[
D \approx \begin{bmatrix}
-329 & 859 & -23 & -127 & 49 \\
-9 & 0 & 22 & 29 & -53 \\
7 & -257 & 0 & 129 & -29 \\
53 & -29 & 89 & 0 & 181 \\
-49 & 249 & 12 & -213 & 653 \\
\end{bmatrix}
\]
where we have again rounded the numbers to the second decimal place. See [43] for more details.

The potential advantage of using non-polynomial function spaces to solve time-dependent differential equations has also been pointed out in several other works. For instance, in [58, 59], exponentially fitted schemes were used to solve singular perturbation problems. Discontinuous Galerkin methods based on non-polynomial approximation spaces were considered in [100]. There are also several works on essentially non-oscillatory (ENO) and weighted ENO (WENO) reconstructions based on non-polynomial function approximations [19, 56, 51]. In these cases, it would be of potential advantage to use a LS-CF that is exact for the corresponding function space.

6.9. A Ten-Dimensional Example. For high-dimensional problems it becomes impractical to construct LS-CF that are exact for polynomials of an increasing (total) degree \( m \). This is because the dimension of the corresponding function space \( \mathbb{P}_m(\mathbb{R}^d) \) is given by \( \dim \mathbb{P}_m(\mathbb{R}^d) = \binom{m+d}{m} \), which implies \( \dim \mathbb{P}_m(\mathbb{R}^d) = \mathcal{O}(m^d) \). Thus, in a sense, the LS-CFs based on polynomials fall victim to the curse of dimensionality and cannot be applied directly in high dimensions. However, we now demonstrate how appropriately chosen nonpolynomial function spaces can be used to extend the positive LS-CFs into a high-dimensional setting. Another idea, related to variance reduction in MC and QMC methods, will be discussed in subsection 6.10.

Instead of a polynomial function space, we consider an RBF function space for the ten-dimensional domain \( \Omega = [0, 1]^d \). For sake of simplicity, we use a Gaussian RBF \( \varphi(r) = \exp(-\varepsilon^2 r^2) \) with shape parameter \( \varepsilon = 10^{-2} \sqrt{K} \), where \( K \) is the number of centers used to build the RBF space. We used ten logarithmically spaced values for \( K \) between 1 and \( 10^3 \). The dimension of the RBF space only depends on the number of centers and not on the dimension of the domain \( \Omega \). We make no claim about the optimality of the above choices of a function space of Gaussian kernels and the chosen shape parameters. That said, Figure 11 demonstrates that even in high dimensions LS-CFs are able to yield more accurate results than the MC and QMC method when these are exact for appropriate RBF spaces. Adapting the locations of the centers to some prior knowledge of the function \( f \) (e.g., \( f \) rapidly decays away from the axes) and using, for instance, sparse grids might further improve the results. While this would exceed the scope of this paper, it would also be of interest to combine high-dimensional LS-CF with separable (low-rank) [8, 17] or sparse approximations [18, 22, 1].

6.10. Variance Reduction in MC and QMC Methods. Another possible applications of LS-CFs to high-dimensional problems is related to the idea of variance reduction in MC and QMC methods [69]. In the context of the present paper, the idea is to fix the maximum total degree \( m \), say \( m = 1 \) or \( m = 2 \), and to only increase the number of data points. Henceforth, we refer to such a CF as an LS-QMC formula when random points are used and as an LS-MC formula when a low-discrepancy sequence is used. Figure 12 reports on the errors of these formulas for \( m = 1 \) and \( m = 2 \) compared to a usual (Q)MC formula for the ten-dimensional domain \( \Omega = [0, 1]^d \). We can see that the LS-(Q)MC formulas show the same convergence rate as the (Q)MC formula but with a reduced constant in front of the error. A similar observation was made in [69]. In this context, the present work can be used to ensure positivity of LS-(Q)MC formulas if \( N \) is sufficiently larger than the fixed maximum total degree \( m \). We also refer to [69, Appendix A] for an especially efficient implementation of the LS-(Q)MC formulas.
Figure 11: Errors for Genz’ third test function on $\Omega = [0, 1]^d$ with $\omega \equiv 1$ using a stable LS-CFs based on Gaussian RBFs and a (Q)MC method on the same points.

Figure 12: Errors for Genz’ third test function on $\Omega = [0, 1]^d$ with $\omega \equiv 1$ using a (Q)MC method and the LS-(Q)MC method with $m = 1$ and $m = 2$.

7. Concluding Thoughts. We presented a simple procedure to construct provable positive and exact CFs in a general multi-dimensional setting. It was proved that under relatively mild restrictions such CFs always exist—and can be determined by the method of LS—when a sufficiently large number of equidistributed data points is used. This extends some previous results on LS formulas from one dimension [99, 98, 54, 39] as well as multiple dimensions [40] (but restricted to function spaces of algebraic polynomials). At the same time, our findings can also be seen as an extension of the stable high-order randomized CFs discussed in [67], which are positive and exact with a high probability, into a deterministic framework. Furthermore, similarities with certain methods for variance reduction in the context of MC and
QMC methods [69] should be noted. Our results indicate that such high-order corrections are ensured to yield positive formulas if a sufficiently large number of data points is used. Finally, it is possible to combine the provable positive and exact LS-CFs with subsampling methods, therefore constructing positive interpolatory CFs. While these are already predicted by Tchakaloff’s theorem (see Theorem 1.1), it is often not clear how the data points should be chosen. Our findings indicate that some of the positive interpolatory CFs predicted by Tchakaloff are supported on sets of equidistributed points. This can be considered as an important—yet sometimes missing—justification and design criterion for CFs constructed based on optimization strategies. These include NNLS as well as linear programming approaches.

In forthcoming works, we will address the application of the LS-CF introduced here to derive SBP operators for general function spaces [43] as well as to construct energy-stable RBF methods, a development that we have already initiated in [41, 42].

Appendix A. Steinitz’ Method.

Given is a $K$-dimensional function space $\mathcal{F}_K(\Omega)$ and a positive and $\mathcal{F}_K(\Omega)$-exact CF,

$$C_N[f] = \sum_{n=1}^{N} w_n f(x_n).$$

Here, $N$ denotes the number of data points in a generic sense. If $N > K$, one successively reduces the number of data points until $N \leq K$ by going over to an appropriate subset of data points, while preserving positivity and $\mathcal{F}_K(\Omega)$-exactness. Recall that the vector space $\mathcal{F}_K(\Omega)$ has dimension $K$, and so does its algebraic dual space (the space of all linear functionals defined on $\mathcal{F}_K(\Omega)$). In particular, among the $N$ linear functionals

$$L_n[f] = f(x_n), \quad n = 1, \ldots, N,$$

at most $K$ are linearly independent. That is, if $N > K$, there exist a vector of coefficients $a = (a_1, \ldots, a_M)^T$ such that

$$a_1 L_1[f] + \cdots + a_N L_N[f] = 0 \quad \forall f \in \mathcal{F}_K(\Omega)$$

and $a_n > 0$ for at least one $n$. Let $\sigma = \max_{1 \leq n \leq N} a_n/w_n$. Then, $\sigma > 0$, $\sigma w_n - a_n \geq 0$ for all $n$, and $\sigma w_n - a_n = 0$ for at least one $n$. From (A.3) one therefore has

$$I[f] = \frac{\sigma w_1 - a_1}{\sigma} L_1[f] + \cdots + \frac{\sigma w_N - a_N}{\sigma} L_N[f] \quad \forall f \in \mathcal{F}_K(\Omega).$$

Note that one of the coefficients is zero and, together with the corresponding linear functional (data point), can be removed. Hence, on $\mathcal{F}_K(\Omega)$, the integral $I$ can be expressed as a linear combination of not more than $N-1$ of the linear functionals $L_1, \ldots, L_N$ with positive coefficients. Iterating this process, one finally arrives at a positive interpolatory CF with $N \leq K$ while being exact for all $f \in \mathcal{F}_K(\Omega)$.

An algorithmic description of the Steinitz method is provided in Algorithm A.1. Thereby, $\text{null}(\Phi) = \{a \in \mathbb{R}^N \mid \Phi a = 0\}$ denotes the null space of the matrix $\Phi$. 
Algorithm A.1 The Steinitz Method

1: while $K < N$ do
2:    Compute $\Phi = \Phi(X)$ and $\text{null}(\Phi)$
3:    Determine $a \in \text{null}(\Phi) \setminus \{0\}$ s.t. $a_n > 0$ for at least one $n$ (see Remark A.1)
4:    Compute $\sigma = \max_n a_n/w_n$
5:    Overwrite the cubature weights: $w_n = (\sigma w_n - a_n)/\sigma$
6:    Remove all zero weights as well as the corresponding data points
7:    $N = N - \#\{w_n \mid w_n = 0, n = 1, \ldots, N\}$

Remark A.1. Note that (A.3) is equivalent to $a \in \text{null}(\Phi)$. Essentially every $a \in \text{null}(\Phi) \setminus \{0\}$ can be used (if $a_n \leq 0$ for all $n = 1, \ldots, M$, one can go over to $-a$). Moreover, it was shown Lemma 2.2 that $\text{null}(\Phi)$ has dimension $N - K$. Hence, as long as $N > K$, such a vector of coefficients $a$ can always be found.

REFERENCES

[1] B. Adcock, S. Brugiapaglia, and C. G. Webster, Compressed sensing approaches for polynomial approximation of high-dimensional functions, in Compressed Sensing and its Applications, Springer, 2017, pp. 93–124.
[2] W. F. Ames, Numerical Methods for Partial Differential Equations, Academic press, 2014.
[3] T. Bagby, L. Bos, and N. Levenberg, Multivariate simultaneous approximation, Constructive Approximation, 18 (2002), pp. 569–577.
[4] C. Bayer and J. Teichmann, The proof of Tchakaloff’s theorem, Proceedings of the AMS, 134 (2006), pp. 3035–3040.
[5] R. Bellman, Dynamic programming, Science, 153 (1966), pp. 34–37.
[6] A. Ben-Israel and T. N. Greville, Generalized Inverses: Theory and Applications, vol. 15 of CMS Books in Mathematics, Springer Science & Business Media, 2003.
[7] B. Benouahmane, C. Annie, and I. Yaman, Near-minimal cubature formulae on the disk, IMA Journal of Numerical Analysis, 39 (2019), pp. 297–314.
[8] G. Beylkin, J. Garcke, and M. J. Mohlenkamp, Multivariate regression and machine learning with sums of separable functions, SIAM Journal on Scientific Computing, 31 (2009), pp. 1840–1857.
[9] L. Bos, S. De Marchi, and M. Vianello, Polynomial approximation on Lissajous curves in the $d$-cube, Applied Numerical Mathematics, 116 (2017), pp. 47–56.
[10] L. Bos and N. Levenberg, Bernstein–Walsh theory associated to convex bodies and applications to multivariate approximation theory, Computational Methods and Function Theory, 18 (2018), pp. 361–388.
[11] L. Bos and M. Vianello, CaTchDes: MATLAB codes for Carathéodory–Tchakaloff near-optimal regression designs, SoftwareX, 10 (2019), p. 100349.
[12] H. Brass and K. Petras, Quadrature Theory: The Theory of Numerical Integration on a Compact Interval, no. 178 in Mathematical Surveys and Monographs, AMS, 2011.
[13] J. Bremer, Z. Gimbutas, and V. Rokhlin, A nonlinear optimization procedure for generalized Gaussian quadratures, SIAM Journal on Scientific Computing, 32 (2010), pp. 1761–1788.
[14] L. Brutman, Lebesgue functions for polynomial interpolation—a survey, Annals of Numerical Mathematics, 4 (1996), pp. 111–128.
[15] H.-J. Bungartz and M. Griebel, Sparse grids, Acta Numerica, 13 (2004), pp. 147–269.
[16] R. E. Caflisch, Monte Carlo and quasi-Monte Carlo methods, Acta Numerica, 1998 (1998), pp. 1–49.
[17] M. Chevreuil, R. Lebrun, A. Nouy, and P. Rai, A least-squares method for sparse low rank approximation of multivariate functions, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 897–921.
[18] A. Chkifa, A. Cohen, and C. Schwab, *Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs*, Journal de Mathématiques Pures et Appliquées, 103 (2015), pp. 400–428.

[19] S. N. Christofor, *The study of building blocks for essentially non-oscillatory (ENO) schemes*, Brown University, 1996.

[20] R. Cline and R. J. Plemmons, $\ell_2$-solutions to underdetermined linear systems, SIAM Review, 18 (1976), pp. 92–106.

[21] A. Cohen, M. A. Davenport, and D. Leviatan, *On the stability and accuracy of least squares approximations*, Foundations of Computational Mathematics, 13 (2013), pp. 819–834.

[22] A. Cohen and R. DeVore, *Approximation of high-dimensional parametric pdes*, Acta Numerica, 24 (2015), pp. 1–159.

[23] A. Cohen and G. Migliorati, *Optimal weighted least-squares methods*, The SMAI Journal of Computational Mathematics, 3 (2017), pp. 181–203.

[24] R. Cools, *Constructing cubature formulae: The science behind the art*, Acta Numerica, 6 (1997), pp. 1–54.

[25] P. C. Curtis Jr, *n-parameter families and best approximation.*, Pacific Journal of Mathematics, 9 (1959), pp. 1013–1027.

[26] P. Davis and M. Wilson, *Nonnegative interpolation formulas for uniformly elliptic equations*, Journal of Approximation Theory, 1 (1968), pp. 374–380.

[27] P. J. Davis, *A construction of nonnegative approximate quadratures*, Mathematics of Computation, 21 (1967), pp. 578–582.

[28] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Courier Corporation, 2007.

[29] M. Desole, F. Marcuzzi, and M. Vianello, *Accelerating the Lawson–Hanson NNLS solver for large-scale teckaloff regression designs*, Dolomites Research Notes on Approximation, 13 (2020).

[30] J. Dick, F. Y. Kuo, and I. H. Sloan, *High-dimensional integration: the quasi-Monte Carlo way*, Acta Numerica, 22 (2013), p. 133.

[31] H. Engels, *Numerical Quadrature and Cubature*, Academic Press, 1980.

[32] D. C. D. R. Fernandez, J. E. Hicken, and D. W. Zingg, *Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations*, Computers & Fluids, 95 (2014), pp. 171–196.

[33] B. Fornberg and N. Flyer, *A Primer on Radial Basis Functions With Applications to the Geosciences*, SIAM, 2015.

[34] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright, *Kernel based quadrature on spheres and other homogeneous spaces*, Numerische Mathematik, 127 (2014), pp. 57–92.

[35] W. Gautschi, *Numerical Analysis*, Springer Science & Business Media, 1997.

[36] A. Genz, *Testing multidimensional integration routines*, in Proc. of International Conference on Tools, Methods and Languages for Scientific and Engineering Computation, 1984, pp. 81–94.

[37] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, vol. 53, Springer Science & Business Media, 2013.

[38] J. Glaubitz, *Shock Capturing and High-Order Methods for Hyperbolic Conservation Laws*, Logos Verlag Berlin GmbH, 2020.

[39] J. Glaubitz, *Stable high order quadrature rules for scattered data and general weight functions*, SIAM Journal on Numerical Analysis, 58 (2020), pp. 2144–2164.

[40] J. Glaubitz, *Stable high-order cubature formulas for experimental data*, Journal of Computational Physics, 447 (2021), p. 110693.

[41] J. Glaubitz and A. Gelb, *Stabilizing radial basis function methods for conservation laws using weakly enforced boundary conditions*, Journal of Scientific Computing, 87 (2021), pp. 1–29.

[42] J. Glaubitz, E. Le Meledo, and P. Öffner, *Towards stable radial basis function methods for linear advection problems*, Computers & Mathematics with Applications, 85 (2021), pp. 84–97.

[43] J. Glaubitz, J. Nordström, and P. Öffner, *Summation-by-parts operators for general function spaces*, arXiv preprint arXiv:2108.06375, (2022).

[44] J. Glaubitz and P. Öffner, *Stable discretisations of high-order discontinuous Galerkin methods on equidistant and scattered points*, Applied Numerical Mathematics, 151 (2020), pp. 98–118.
[45] J. Glaubitz and J. Reeger, Towards stability of radial basis function based cubature formulas, arXiv preprint arXiv:2108.06375, (2021).
[46] G. H. Golub and C. F. Van Loan, Matrix Computations, vol. 3, JHU Press, 2012.
[47] L. Guo, A. Narayan, and T. Zhou, Constructing least-squares polynomial approximations, SIAM Review, 62 (2020), pp. 483–508.
[48] S. Haber, Numerical evaluation of multiple integrals, SIAM Review, 12 (1970), pp. 481–526.
[49] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals, Numerische Mathematik, 2 (1960), pp. 84–90.
[50] S. Hayakawa, Monte Carlo cubature construction, Japan Journal of Industrial and Applied Mathematics, 38 (2021), pp. 561–577.
[51] J. S. Hesthaven and F. Mönkeberg, Entropy stable essentially nonoscillatory methods based on RBF reconstruction, ESAIM: Mathematical Modelling and Numerical Analysis, 53 (2019), pp. 925–958.
[52] J. S. Hesthaven and T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, Springer Science & Business Media, 2007.
[53] E. Hlawka, Funktionen von beschränkter Variation in der Theorie der Gleichverteilung, Ann. Mat. Pura Appl., 54 (1961), pp. 325–333.
[54] D. Huybrechs, Stable high-order quadrature rules with equidistant points, Journal of Computational and Applied Mathematics, 231 (2009), pp. 933–947.
[55] B. A. Ibrahimoglu, Lebesgue functions and Lebesgue constants in polynomial interpolation, Journal of Inequalities and Applications, 2016 (2016), pp. 1–15.
[56] A. Iske and T. Sonar, On the structure of function spaces in optimal recovery of point functionals for ENO-schemes by radial basis functions, Numerische Mathematik, 74 (1996), pp. 177–201.
[57] J. D. Jakeman and A. Narayan, Generation and application of multivariate polynomial quadrature rules, Computer Methods in Applied Mechanics and Engineering, 338 (2018), pp. 134–161.
[58] M. Kadalbajoo and K. Patidar, Exponentially fitted spline in compression for the numerical solution of singular perturbation problems, Computers & Mathematics with Applications, 46 (2003), pp. 751–767.
[59] I. Kalashnikova, C. Farhat, and R. Tezaur, A discontinuous enrichment method for the finite element solution of high Péclet advection–diffusion problems, Finite Elements in Analysis and Design, 45 (2009), pp. 238–250.
[60] V. Keshavarzzadeh, R. M. Kirby, and A. Narayan, Numerical integration in multiple dimensions with designed quadrature, SIAM Journal on Scientific Computing, 40 (2018), pp. A2033–A2061.
[61] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, Courier Corporation, 2012.
[62] F. Y. Kuo and I. H. Sloan, Lifting the curse of dimensionality, Notices of the AMS, 52 (2005), pp. 1320–1328.
[63] C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, vol. 15, Siam, 1995.
[64] J. Maeztu, On symmetric cubature formulae for planar regions, IMA Journal of Numerical Analysis, 9 (1989), pp. 167–183.
[65] J. C. Mairhuber, On Haar’s theorem concerning Chebychev approximation problems having unique solutions, Proceedings of the American Mathematical Society, 7 (1956), pp. 609–615.
[66] B. F. Manly, Randomization, Bootstrap and Monte Carlo Methods in Biology, vol. 70, CRC press, 2006.
[67] G. Migliorati and F. Nobile, Stable high-order randomized cubature formulae in arbitrary dimension, Journal of Approximation Theory, (2022), p. 105706, https://doi.org/https://doi.org/10.1016/j.jat.2022.105706.
[68] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT press, 2012.
[69] Y. Nakatsukasa, Approximate and integrate: Variance reduction in Monte Carlo integration via function approximation, arXiv preprint arXiv:1806.05492, (2018).
[70] I. P. Natanson, Constructive Theory of Functions, vol. 1, US Atomic Energy Commission, Office of Technical Information Extension, 1961.
[71] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, 1992.
[72] F. Piazzon, A. Sommariva, and M. Vianello, Carathéodory–Tchakaloff least squares, in International Conference on Sampling Theory and Applications (SampTA), 2017, pp. 672–676.
[73] R. B. Platte, L. N. Trefethen, and A. B. Kuijlaars, Impossibility of fast stable approximation of
analytic functions from equispaced samples, SIAM Review, 53 (2011), pp. 308–318.

[74] J. A. Reeger and B. Fornberg, Numerical quadrature over smooth surfaces with boundaries, Journal of Computational Physics, 355 (2018), pp. 176–190.

[75] E. K. Ryu and S. P. Boyd, Extensions of Gauss quadrature via linear programming, Foundations of Computational Mathematics, 15 (2015), pp. 953–971.

[76] M. H. Schultz, $l^\infty$-multivariate approximation theory, SIAM Journal on Numerical Analysis, 6 (1969), pp. 161–183.

[77] M. Slawski, Non-negative least squares: comparison of algorithms, (2022). https://sites.google.com/site/slawskimartin/code. Accessed January 20, 2022.

[78] I. H. Sloan, Polynomial interpolation and hyperinterpolation over general regions, Journal of Approximation Theory, 83 (1995), pp. 238–254.

[79] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Russian Academy of Sciences, 148 (1963), pp. 1042–1045.

[80] A. Sommariva and M. Vianello, Numerical cubature on scattered data by radial basis functions, Computing, 76 (2006), p. 295.

[81] A. Sommariva and M. Vianello, Compression of multivariate discrete measures and applications, Numerical Functional Analysis and Optimization, 36 (2015), pp. 1198–1223.

[82] A. Sommariva and M. Vianello, RBF moment computation and meshless cubature on general polygonal regions, Applied Mathematics and Computation, 409 (2021), p. 126375.

[83] E. Steinitz, Bedingt konvergente Reihen und konvexe Systeme, Journal für die reine und angewandte Mathematik (Crelle’s Journal), 1913 (1913), pp. 128–176.

[84] M. Svärd and J. Nordström, Review of summation-by-parts schemes for initial–boundary-value problems, Journal of Computational Physics, 268 (2014), pp. 17–38.

[85] M. A. Taylor, B. A. Wingate, and L. P. Bos, A cardinal function algorithm for computing multivariate quadrature points, SIAM Journal on Numerical Analysis, 45 (2007), pp. 193–205.

[86] V. Tchakaloff, Formules de cubatures mécaniques à coefficients non négatifs, Bull. Sci. Math, 81 (1957), pp. 123–134.

[87] L. Trefethen, Multivariate polynomial approximation in the hypercube, Proceedings of the AMS, 145 (2017), pp. 4837–4844.

[88] L. N. Trefethen, Cubature, approximation, and isotropy in the hypercube, SIAM Review, 59 (2017), pp. 469–491.

[89] L. N. Trefethen, Exactness of quadrature formulas, arXiv preprint arXiv:2101.09501, (2021).

[90] L. N. Trefethen and D. Bau III, Numerical Linear Algebra, vol. 50, SIAM, 1997.

[91] L. van den Bos, B. Sanderse, W. Bierbooms, and G. van Bussel, Generating nested quadrature rules with positive weights based on arbitrary sample sets, SIAM/ASA Journal on Uncertainty Quantification, 8 (2020), pp. 139–169.

[92] L. M. van den Bos, B. Koren, and R. P. Dwight, Non-intrusive uncertainty quantification using reduced cubature rules, Journal of Computational Physics, 332 (2017), pp. 418–445.

[93] L. M. van den Bos, B. Sanderse, and W. Bierbooms, Adaptive sampling-based quadrature rules for efficient Bayesian prediction, Journal of Computational Physics, 417 (2020), p. 109537.

[94] J. van der Corput, Verteilungsfunktionen, in Proc. Ned. Akad. v. Wet., vol. 38, 1935, pp. 813–821.

[95] H. Weyl, Über die Gleichverteilung von Zahlen mod. eins, Mathematische Annalen, 77 (1916), pp. 313–352.

[96] D. R. Wilhelmissen, A nearest point algorithm for convex polyhedral cones and applications to positive linear approximation, Mathematics of Computation, 30 (1976), pp. 48–57.

[97] M. W. Wilson, A general algorithm for nonnegative quadrature formulas, Mathematics of Computation, 23 (1969), pp. 253–258.

[98] M. W. Wilson, Discrete least squares and quadrature formulas, Mathematics of Computation, 24 (1970), pp. 271–282.

[99] M. W. Wilson, Necessary and sufficient conditions for equidistant quadrature formula, SIAM Journal on Numerical Analysis, 7 (1970), pp. 134–141.

[100] L. Yuan and C.-W. Shu, Discontinuous Galerkin method based on non-polynomial approximation spaces, Journal of Computational Physics, 218 (2006), pp. 295–323.