Exact discretization, tight frames
and recovery via D-optimal designs

Felix Bartel, Lutz Kämmerer, Kateryna Pozharska,
Martin Schäfer & Tino Ullrich

Exact discretization, tight frames and recovery via D-optimal designs

Felix Bartel [1]
Lutz Kämmerer [2]
Kateryna Pozharska [3]
Martin Schäfer [4]
Tino Ullrich [5]

[1] Mathematisch-Geographische Fakultät, KU Eichstätt-Ingolstadt, 85270 Eichstätt, Germany.
E-mail address: felix.bartel@ku.de
[2] Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany.
E-mail address: kaemmerer@math.tu-chemnitz.de
[3] Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
and Institute of Mathematics of NAS of Ukraine.
E-mail address: pozharska.k@gmail.com
[4] Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany.
E-mail address: martin.schaefer@math.tu-chemnitz.de
[5] Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany.
E-mail address: tino.ullrich@math.tu-chemnitz.de.

Abstract. D-optimal designs originate in statistics literature as an approach for optimal experimental designs. In numerical analysis points and weights resulting from maximal determinants turned out to be useful for quadrature and interpolation. Also recently, two of the present authors and coauthors investigated a connection to the discretization problem for the uniform norm. Here we use this approach of maximizing the determinant of a certain Gramian matrix with respect to points and weights for the construction of tight frames and exact Marcinkiewicz–Zygmund inequalities in L_2. We present a direct and constructive approach resulting in a discrete measure with at most $N \leq n^2 + 1$ atoms, which discretely and accurately subsamples the L_2-norm of complex-valued functions contained in a given n-dimensional subspace. This approach can as well be used for the reconstruction of functions from general RKHS in L_2 where one only has access to the most important eigenfunctions. We verifiably and deterministically construct points and weights for a weighted least squares recovery procedure and pay in the rate of convergence compared to earlier optimal, however probabilistic approaches. The general results apply to the d-sphere or multivariate trigonometric polynomials on $\mathbb{T}^d$ spectrally supported on arbitrary finite index sets $I \subset \mathbb{Z}^d$. They can be discretized using at most $|I|^2 - |I| + 1$ points and weights. Numerical experiments indicate the sharpness of this result. As a negative result we prove that, in general, it is not possible to control the number of points in a reconstructing lattice rule only in the cardinality $|I|$ without additional condition on the structure of I. We support our findings with numerical experiments.

2020 Mathematics Subject Classification. 41A25, 65D32, 94A20.

Keywords. Integral norm discretization, exact quadrature, sampling recovery, D-optimal designs.

F. BARTEL, L. KÄMMERER, *ET AL.*

1. **Introduction**

In this paper we investigate D-optimal designs, which are the result of optimizing the determinant of rescaled Gramian matrices. This approach is known from optimal experimental designs in statistics literature [22] or in point constructions for quadrature and interpolation, like Fekete points where one maximizes the determinant of the Vandermonde matrix [5, 7, 8, 9, 36]. We take a frame-theoretical viewpoint and obtain novel results for quadrature and frame analysis. These two perspectives are connected in the following way: Let $\boldsymbol{\varphi}(x) = (\varphi_1(x), \ldots, \varphi_n(x))^\top \in \mathbb{C}^n$ for $\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$ be linearly independent continuous functions on a compact domain Ω. Then

$$V_n = \mathrm{span}_{\mathbb{C}}\{\varphi_1, \ldots, \varphi_n\} \subset C(\Omega)$$

represents an n-dimensional space of complex-valued functions on Ω. Further, there exist probability measures μ on Ω such that the family $(\boldsymbol{\varphi}(x))_{x \in \Omega}$ becomes a continuous frame in $\mathbb{C}^n$ with respect to μ. The notion of continuous frame in a separable Hilbert space $\mathcal{H}$ can be recalled, e.g. in [28, Def. 1.2]. For $\mathcal{H} = \mathbb{C}^n$ it is essentially characterized by the property

$$A\|\boldsymbol{a}\|_2^2 \leq \int_\Omega |\langle \boldsymbol{a}, \boldsymbol{\varphi}(x)\rangle|^2 \, \mathrm{d}\mu(x) \leq B\|\boldsymbol{a}\|_2^2 \quad \text{for all} \quad \boldsymbol{a} \in \mathbb{C}^n$$

and fixed constants $0 < A, B < \infty$, the so-called frame bounds. If $A = B = 1$ the frame is called Parseval.

While, in the described setup, there always exist probability measures on Ω such that $(\boldsymbol{\varphi}(x))_{x \in \Omega}$ is a frame, the existence of such measures that turn $(\boldsymbol{\varphi}(x))_{x \in \Omega}$ into a Parseval frame is not always guaranteed. The following properties are equivalent, see e.g. [13, Thm. 2.2] for a discrete version:

- The functions $\varphi_1, \ldots, \varphi_n$ are orthonormal with respect to μ.

- The family $(\boldsymbol{\varphi}(x))_{x \in \Omega} \subset \mathbb{C}^n$ is a continuous Parseval frame with respect to μ.

In the first central result of this paper, Theorem 3.1, we present four elementary equivalent conditions under which the system $\varphi_1, \ldots, \varphi_n$ turns out to be discretely orthonormal with respect to a suitably chosen probability measure ϱ^*. Or, in other words, we give equivalent criteria when the family $(\boldsymbol{\varphi}(x))_{x \in \Omega}$ can be turned into a Parseval frame in $\mathbb{C}^n$. The result is constructive in the sense that the measure ϱ^* is the optimizer of a certain D-optimal design procedure. As such it is a discrete probability measure $\varrho^* = \sum_{i=1}^{M_\varrho} \varrho_i \delta_{x_i}$ which maximizes a specific Gramian determinant, i.e.,

$$\varrho^* = \arg\max_\varrho \det\left(\sum_{i=1}^{M_\varrho} \varrho_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^*\right). \tag{1.1}$$

A variation of this design technique is utilized in Section 3.1, where we use it to produce discrete equal norm tight frames from $(\boldsymbol{\varphi}(x))_{x \in \Omega}$ with certain additional properties, see Algorithm 3 and Theorem 3.5.

1.1. **Parseval subframes and exact discretization**

As a direct application of Theorem 3.1, in Section 4, we seek points $x_1, \ldots, x_N \in \Omega$ and weights $\varrho_1, \ldots, \varrho_N \geq 0$, summing up to 1, forming an exact L_2-Marcinkiewicz–Zygmund (MZ) inequality [42] for V_n, i.e.,

$$\int_\Omega |f(x)|^2 \, \mathrm{d}\mu(x) = \sum_{i=1}^N \varrho_i |f(x_i)|^2 \quad \text{for all} \quad f \in V_n. \tag{1.2}$$

In terms of the language of finite frames (1.2) can be rephrased as finding a weighted Parseval subframe $(\sqrt{\varrho_i}\boldsymbol{\varphi}(x_i)))_{i=1}^N$ such that

$$\|\boldsymbol{a}\|_2^2 = \int_\Omega |\langle \boldsymbol{a}, \boldsymbol{\varphi}(x)\rangle|^2 \, \mathrm{d}\mu(x) = \sum_{i=1}^N \varrho_i |\langle \boldsymbol{a}, \boldsymbol{\varphi}(x_i)\rangle|^2 \quad \text{for all} \quad \boldsymbol{a} \in \mathbb{C}^n .$$

The problem of the discretization of continuous frames is deeply studied in literature. Here we refer to [27, 28] and the references therein. Based on Theorem 3.1, we present an approach using specific D-optimal designs to obtain (1.2) with at most

$$N = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\varphi_k \overline{\varphi_l}\}_{k,l=1}^n) + 1 \leq n^2 + 1 \tag{1.3}$$

points and weights. In addition we give numerical evidence on the sharpness of this result. We further include L_p-MZ inequalities for even $p > 2$ which supports the findings in [27]. We also contribute to Open Problem 10 in [18], i.e., we give an algorithmic method for constructing an exact L_p-MZ inequality with positive weights, cf. Corollary 4.4. To obtain the bound (1.3) on the required number N of points, see the proof of Theorem 3.1, we use a version of Carathéodory's theorem, Proposition 2.7, which allows for subsampling convex (conic) combinations of (complex) Hermitian matrices. Carathéodory subsampling has been considered before by several authors, see [9, 38, 49]. The latter reference also comments on its implementation.

Instead of aiming for (1.2), one may ask for an exact discretization of the measure μ in the space V_n. Earlier works in this direction rely on the classical 1957 Tchakaloff theorem [56], a corner stone for the theory of exact quadrature formulas for spaces of polynomials, see also Shapiro [52, Thm. 3.1.1], Novak [47], Putinar [50], as well as [44, Thm. 1] and [18, Thm. 4.1] for more general versions. From our results we are able to obtain an extended version of the classical Tchakaloff theorem for complex-valued functions, Theorem 4.3. Finding good quadrature formulas becomes a challenging task with growing dimension, see e.g. [17, 23] for a history overview and [44] for results on randomized quadrature for real-valued functions defined on domains in $\mathbb{R}^d$ with arbitrary dimension d.

In Section 5, we apply the general results in two specific contexts. First, we state a result for the d-sphere $\mathbb{S}^d$ in Theorem 5.5 for even exponents $p \geq 2$. As a consequence, we obtain in Theorem 5.6 that there exist points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^N \in \mathbb{S}^d$ forming an exact (weighted) integration rule for the polynomials Π_m up to degree m on the d-sphere, i.e.,

$$\int_{\mathbb{S}^d} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) = \sum_{i=1}^N w_i f(\boldsymbol{x}^i) \quad \text{for all} \quad f \in \Pi_m$$

with

$$N \leq \dim_{\mathbb{R}} \Pi_m \leq \left(\frac{9m}{d}\right)^d .$$

Quadrature formulas of the above type with equal weights are called t-designs (or m-design according to our notation, where m is the degree of the polynomial) introduced in [20].

The equal weight condition makes this problem much harder and only a limited number of constructions of spherical designs are known. However, there are also approaches to obtain spherical designs being exact up to machine precision numerically, cf. [29, 60]. In general the existence of spherical designs is known with the optimal asymptotic rate $N \geq C_d m^d$, cf. [6, Thm. 1]. This matches the number of points in our weighted result.

In Corollary 5.1 we state that (1.2) holds for trigonometric polynomials with frequencies in any arbitrary finite index set $I \subset \mathbb{Z}^d$. For

$$D(I) := \{\boldsymbol{k} - \boldsymbol{\ell} : \boldsymbol{k}, \boldsymbol{\ell} \in I\}$$

the difference set, we find $N \leq |D(I)| \leq |I|^2 - |I| + 1$ points and positive weights such that the exact MZ inequality (1.2) holds for all $f \in V(I) := \mathrm{span}_{\mathbb{C}}\{\exp(2\pi i \langle \boldsymbol{k}, \cdot \rangle) : \boldsymbol{k} \in I\}$. In general, these points

$(x_i)_{i=1}^N$ should have a very irregular structure. In fact, we show in Corollary 5.3 that at least for lattice rules it is not possible to control the number of points only in the cardinality of the index set $|I|$ (without additional conditions on I as in [34]). This contradicts [57, Thm. 4.1] and contributes to Open Problem 2 in [18].

Also here the question arises whether all of this can be done with equal weights $w_i = 1/N$. In case of trigonometric polynomials such a discretization is always possible with a large enough N depending on the "largest" frequency in I. However, it is still not known, whether the equal-weighted (1.2) can be arranged for any index set I with a number of points N only depending on the cardinality of $|I|$. We leave this as an open problem.

1.2. **Numerics**

Our approach to determine the points $(x_i)_i$ and weights $(\varrho_i)_i$ such that (1.2) holds true is constructive in the sense that these point-weight pairs can be obtained by a maximization, respectively minimization, procedure, see Theorem 3.1. Based on this insight, we propose two algorithms, Algorithm 1 and Algorithm 2, whereby the first utilizes a D-design principle requiring the maximization of a determinant and the second rests on the minimization of a Frobenius norm. To our knowledge, the D-design method is new, whereas the norm minimization in Algorithm 2 is an already known principle, suggested and analyzed in [18, §2.5] in the case of real weights.

Several experiments connected to these two algorithms are conducted in Section 6, the numerics section. Since our primary interest are the resulting point-weight pairs, we only implemented a version of Algorithm 2. According to Theorem 3.1 both algorithms should produce the same results. As the experiments show, the outcomes align with our theoretical findings:

- In Experiment 1 we indicate the numerical applicability of our approach as we recover known lattice point constructions. Furthermore, we compute exact MZ points in settings, where we know that all lattice rules of that size fail as discussed above.

- In Experiment 2 we investigate a dimensional dependence and observe that for a fixed number of random frequencies fewer points suffice with increasing dimension. This suggests a certain "blessing of dimension" in this particular setting. A similar effect was considered in [29] for the sphere $\mathbb{S}^d$.

- Experiment 3 indicates the necessity of the oversampling $N = |D(I)|$ in general as we do not find exact MZ inequalities with fewer points for this particular choice of frequencies in I. We do not have a rigorous proof. Theoretical lower bounds indicating at least the quadratic scaling in the number of samples are given in [18, Thm. 3.3, 3.4].

It should be remarked that implementation of Algorithms 1 or 2 requires the solution of a non-convex optimization problem, for which general efficient methods are lacking. The search for such methods lies beyond the scope of the conducted research. For the implementation of Algorithm 2 we used the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, see Section 6 for more details.

1.3. **Equal norm tight frames**

In the well-known paper by Kiefer and Wolfowitz [36] the authors prove the equivalence of two different optimization problems. The D-optimal design for an n-dimensional subspace of real-valued functions $V_n \subset C(\Omega)$ on the one hand and the min-max problem for the corresponding Christoffel function $\eta_n(x) = \sum_{k=1}^n |\varphi_k(x)|^2$ on the other hand. This is the problem of finding a Borel probability measure μ on Ω which minimizes $\sup_{f \in V_n \setminus \{0\}} \|f\|_\infty^2 / \|f\|_{L_2(\mu)}^2$. The latter is sometimes called G-optimal design, see

Bos [8]. The equivalence in [36] is useful in a different context. In the recent paper [38] two of the authors with coauthors extended the technique in [36] to complex-valued functions (where the continuity is even lacking) and the optimal measure turns out to be discrete (Tchakaloff's theorem [50, 56] is not required). Based on the proof in [38] we give, in Section 3.1, a D-optimal design procedure to generate an equal norm tight frame (ENTF) from a given arbitrary continuous frame $(\boldsymbol{\varphi}(x))_x \subset \mathbb{C}^n$. ENTFs play an important role in the frame community, see Remark 3.8. The corresponding algorithm is formulated in Algorithm 3. It computes, see Theorem 3.5, a discrete probability measure $\lambda = \sum_{i=1}^{N} \lambda_i \delta_{x_i}$ with $N \leq n^2 + 1$ and a related linear transformation $\boldsymbol{A}_\lambda$ with $\boldsymbol{\psi} = \boldsymbol{A}_\lambda^{-1/2} \cdot \boldsymbol{\varphi}$ such that

$$\|\boldsymbol{a}\|_2^2 = \sum_{i=1}^{N} \lambda_i |\langle \boldsymbol{a}, \boldsymbol{\psi}(x_i) \rangle|^2 \quad \text{for all} \quad \boldsymbol{a} \in \mathbb{C}^n,$$

with the additional feature that $\|\boldsymbol{\psi}(x)\|_2 \leq \sqrt{n}$ for all $x \in \Omega$ and $\|\boldsymbol{\psi}(x)\|_2 = \sqrt{n}$ on supp λ. Within the scope of this article, we did not test Algorithm 3 numerically, however.

1.4. **Guaranteed and verifiable recovery of functions**

In the last section, Section 7, we consider a recovery problem, which has been first addressed by Wasilkowski and Woźniakowski [59] in 2001 and drew a lot of attention in the past 6 years [3, 24, 39, 45]. We aim for the problem to practically find stable recovery algorithms based on suitable sample points and weights which recover verifiably any function from the unit ball of a reproducing kernel Hilbert space (RKHS) $H(k)$ with a prescribed accuracy in $L_2(\mu)$. Here we consider a further restriction which is certainly motivated from practical computation. We only have access to the kernel function and the n most important (largest) singular values and corresponding eigenfunctions of the corresponding integral operator. The main result reads as follows. We construct a linear sampling recovery operator $S_{n,N}^{k,\mu}$ depending on the kernel $k(\cdot,\cdot) : \Omega \times \Omega \to \mathbb{C}$ and the target measure μ which uses $N \leq n^2 + 1$ many sample points such that

$$\sup_{\|f\|_{H(k)} \leq 1} \|f - S_{n,N}^{k,\mu} f\|_{L_2(\mu)}^2 \leq 3 \sum_{j \geq n+1} \sigma_j^2 \leq 3\, c_n(H(k), C(\Omega))^2 \,, \tag{1.4}$$

where $c_n(H(k), C(\Omega))$ denotes the n-Gelfand number of the embedding $I : H(k) \to C(\Omega)$

$$c_n(H(k), C(\Omega)) = \inf_{\substack{\theta:\, \mathbb{C}^n \to C(\Omega) \\ L \in \mathcal{L}(H(k), \mathbb{C}^n)}} \sup_{\|f\|_{H(k)} \leq 1} \|f - \theta \circ L(f)\|_{C(\Omega)} \,.$$

Our attempt will be to construct an exact discretization for the space of the most important eigenfunctions. In special cases like periodic functions this is done using, e.g., rank-1 lattices for the given frequency sets in $\mathbb{Z}^d$, like hyperbolic crosses, in [34]. As this specific approach does not work in general, we rely on optimal D-designs working in the most general context.

Notation. As usual, $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$ denote the natural (without zero), integer, real, and complex numbers. If not indicated otherwise $\log(\cdot)$ denotes the natural logarithm. $\mathbb{C}^n$ shall denote the complex n-space and $\mathbb{C}^{m \times n}$ the set of complex $m \times n$-matrices. Vectors and matrices are usually typeset boldface. We use $\boldsymbol{y}^* := \overline{\boldsymbol{y}}^\top$. In general, the adjoint of a matrix $\boldsymbol{L} \in \mathbb{C}^{m \times n}$ is denoted by $\boldsymbol{L}^*$. For the spectral norm we use $\|\boldsymbol{L}\|_{2 \to 2}$, whereas the Frobenius norm is denoted with $\|\boldsymbol{L}\|_F$. For a compact topological space Ω we use $C(\Omega)$ for the space of complex-valued continuous functions on Ω.

2. **Auxiliary tools**

Equalities as in the following proposition have a long history and are related to a variety of topics. Our specific result is a generalization of a classical identity (for $M = m$), see e.g. [55, (2.1.10), p. 24],

and it turns out to be also a special case of the recent result [21, Lem. 2.3] relating the expectation of the determinant of a random matrix and the determinant of the expectation of a matrix. As things simplify in our special case we give an elementary proof for the convenience of the reader.

Proposition 2.1. *Let* $\psi_1, \dots, \psi_m \colon \Omega \to \mathbb{C}$ *be orthonormal with respect to a measure* μ *on* Ω*. Let further* $\boldsymbol{\psi}(x) := (\psi_1(x), \dots, \psi_m(x))^\top \in \mathbb{C}^m$*. Then for fixed* $M \in \mathbb{N}$ *with* $M \geq m$ *we have*

$$\int_\Omega \cdots \int_\Omega \det\left(\frac{1}{M} \sum_{i=1}^M \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right) d\mu(x_1) \dots d\mu(x_M) = \frac{M}{M} \frac{M-1}{M} \cdots \frac{M-m+1}{M}.$$

Proof. By Leibniz' formula for the determinant we have

$$\det\left(\frac{1}{M} \sum_{i=1}^M \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right)$$

$$= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{k=1}^m \frac{1}{M} \sum_{i=1}^M \psi_k(x_i) \overline{\psi_{\sigma(k)}(x_i)}$$

$$= M^{-m} \sum_{i_1,\dots,i_m=1}^M \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot \psi_1(x_{i_1}) \overline{\psi_{\sigma(1)}(x_{i_1})} \dots \psi_m(x_{i_m}) \overline{\psi_{\sigma(m)}(x_{i_m})}. \qquad (2.1)$$

We now show that the inner sum evaluates to zero whenever some $i_l = i_k$ for some $1 \leq l \neq k \leq m$. Without loss of generality assume $i_1 = i_2 = j$. We split S_m into the permutations with positive and negative sign by using $\operatorname{sgn}(\sigma \circ (1\ 2)) = -\operatorname{sgn}(\sigma)$, i.e.,

$$S_m = \{\sigma \in S_m : \operatorname{sgn}(\sigma) = 1\} \dot\cup \{\sigma \in S_m : \operatorname{sgn}(\sigma) = -1\}$$

$$= \{\sigma \in S_m : \operatorname{sgn}(\sigma) = 1\} \dot\cup \{\sigma \circ (1\ 2) \in S_m : \operatorname{sgn}(\sigma) = 1\}.$$

For the inner sum of (2.1) it follows in case $i_1 = i_2 = j$

$$\sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot \psi_1(x_j) \overline{\psi_{\sigma(1)}(x_j)} \cdot \psi_2(x_j) \overline{\psi_{\sigma(2)}(x_j)} \dots \psi_m(x_{i_m}) \overline{\psi_{\sigma(m)}(x_{i_m})}$$

$$= \sum_{\substack{\sigma \in S_m \\ \operatorname{sgn}(\sigma)=1}} \psi_1(x_j) \overline{\psi_{\sigma(1)}(x_j)} \cdot \psi_2(x_j) \overline{\psi_{\sigma(2)}(x_j)} \dots \psi_m(x_{i_m}) \overline{\psi_{\sigma(m)}(x_{i_m})}$$

$$\qquad - \psi_1(x_j) \overline{\psi_{\sigma(2)}(x_j)} \cdot \psi_2(x_j) \overline{\psi_{\sigma(1)}(x_j)} \dots \psi_m(x_{i_m}) \overline{\psi_{\sigma(m)}(x_{i_m})}$$

$$= 0.$$

Thus, in (2.1) we only need to sum over pairwise different indices $i_l \neq i_k$ for $1 \leq l, k \leq m$. It remains to apply the integrals. Using Fubini's theorem gives

$$\int_\Omega \cdots \int_\Omega \det\left(\frac{1}{M} \sum_{i=1}^M \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right) d\mu(x_1) \dots d\mu(x_M)$$

$$= M^{-m} \sum_{\substack{i_1,\dots,i_m=1 \\ i_l \neq i_k}}^M \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot \int_\Omega \psi_1(x_{i_1}) \overline{\psi_{\sigma(1)}(x_{i_1})} \, d\mu(x_{i_1}) \cdots \int_\Omega \psi_m(x_{i_m}) \overline{\psi_{\sigma(m)}(x_{i_m})} \, d\mu(x_{i_m})$$

$$= M^{-m} \sum_{\substack{i_1,\dots,i_m=1 \\ i_l \neq i_k}}^M \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot \delta_{1,\sigma(1)} \dots \delta_{m,\sigma(m)}$$

$$= M^{-m} \sum_{\substack{i_1,\dots,i_m=1 \\ i_l \neq i_k}}^M 1.$$

The number of elements in $\{i_1, \ldots, i_m = 1, \ldots, M : i_l \neq i_k\}$ evaluates to $M(M-1) \ldots (M-m+1)$, which shows the assertion. ∎

Corollary 2.2. *Let $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$ be orthonormal with respect to a probability measure μ and $\boldsymbol{\psi}(x) = (\psi_1(x), \ldots, \psi_m(x))^\top$. Then for all $\varepsilon > 0$ there exists an $M \in \mathbb{N}$ and $x_1, \ldots, x_M \in \Omega$ such that*

$$\det\left(\frac{1}{M}\sum_{i=1}^{M} \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right) \geq 1 - \varepsilon .$$

Proof. In Proposition 2.1 we computed the mean of the determinant for m orthonormal functions and M points to be $M(M-1) \ldots (M-m+1)/M^m$. Since μ is assumed to be a probability measure here, there must therefore exist M points $x_1, \ldots, x_M \in \Omega$ such that this value is actually attained or exceeded. Consequently, as this value tends to 1 for a growing number of points M, the assertion follows. ∎

Proposition 2.3. *Let $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$ be functions. As above we put $\boldsymbol{\psi}(x) := (\psi_1(x), \ldots, \psi_m(x))^\top \in \mathbb{C}^m$. Further, let $M \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_M \geq 0$ with $\alpha_1 + \cdots + \alpha_M = 1$, and $x_1, \ldots, x_M \in \Omega$. Then*

$$\sqrt[m]{\det\left(\sum_{i=1}^{M}\alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right)} \leq \frac{1}{m}\operatorname{trace}\left(\sum_{i=1}^{M}\alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right) \leq \frac{1}{m}\sup_{x\in\Omega}\sum_{k=1}^{m}|\psi_k(x)|^2 .$$

Proof. The first inequality is a consequence of the inequality between arithmetic and geometric mean. For a positive semi-definite Hermitian matrix $\boldsymbol{A} \in \mathbb{C}^{m\times m}$ with eigenvalues $\lambda_1(\boldsymbol{A}), \ldots, \lambda_m(\boldsymbol{A}) \geq 0$, we have

$$\sqrt[m]{\det(\boldsymbol{A})} = \left(\prod_{k=1}^{m}\lambda_k(\boldsymbol{A})\right)^{1/m} \leq \frac{1}{m}\sum_{k=1}^{m}\lambda_k(\boldsymbol{A}) = \frac{\operatorname{trace}(\boldsymbol{A})}{m} .$$

For the second inequality, we compute the trace of the given matrix via its diagonal entries

$$\operatorname{trace}\left(\sum_{i=1}^{M}\alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*\right) = \sum_{k=1}^{m}\sum_{i=1}^{M}\alpha_i[\boldsymbol{\psi}(x_i)\cdot\boldsymbol{\psi}(x_i)^*]_{k,k}$$
$$= \sum_{i=1}^{M}\alpha_i\sum_{k=1}^{m}|\psi_k(x_i)|^2$$
$$\leq \left(\sum_{i=1}^{M}\alpha_i\right)\sup_{x\in\Omega}\sum_{k=1}^{m}|\psi_k(x)|^2 .$$

Taking $\alpha_1 + \cdots + \alpha_M = 1$ into account, the proof is complete. ∎

Lemma 2.4 (A version of Lemma 10 from [38]). *Let $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$ be functions and denote $\boldsymbol{\psi}(x) = (\psi_1(x), \ldots, \psi_m(x))^\top$. Then it holds*

$$\dim_{\mathbb{C}}(\operatorname{span}_{\mathbb{C}}\{\psi_1, \ldots, \psi_m\}) = \dim_{\mathbb{C}}(\operatorname{span}_{\mathbb{C}}\{\boldsymbol{\psi}(x) \colon x \in \Omega\}) ,$$

where we consider a linear space of functions on the left-hand side. On the right-hand side a subspace of $\mathbb{C}^m$ is considered.

Lemma 2.5. *Let* $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$ *be functions and denote* $\boldsymbol{\psi}(x) = (\psi_1(x), \ldots, \psi_m(x))^\top$. *For the set*

$$\mathcal{M} := \{\boldsymbol{\psi}(x) \cdot \boldsymbol{\psi}(x)^* \colon \ x \in \Omega\} \tag{2.2}$$

of complex-valued Hermitian matrices in $\mathbb{C}^{m \times m}$ *it holds*

$$\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^m) = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}} \mathcal{M}) = \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}} \mathcal{M}) \le m^2 \,.$$

Proof. We consider the set (2.2) of matrices with the entries $\psi_k(x)\overline{\psi_l(x)}$, $k, l = 1, \ldots, m$. These matrices can also be interpreted as vectors $\widetilde{\boldsymbol{\psi}}(x) := (f_1(x), \ldots, f_{m^2}(x))^\top \in \mathbb{C}^{m^2}$ with component functions $f_j : \Omega \to \mathbb{C}$, $j = 1, \ldots, m^2$. Further, $\widetilde{\mathcal{M}} := \{\widetilde{\boldsymbol{\psi}}(x) \colon \ x \in \Omega\}$ can be identified with $\mathcal{M}$. With Lemma 2.4 we then obtain

$$\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^m) = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{f_j\}_{j=1}^{m^2}) = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}} \widetilde{\mathcal{M}}) = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}} \mathcal{M}) \,,$$

where $\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{f_j\}_{j=1}^{m^2}) \le m^2$ is obvious. Further, the space $\mathrm{span}_{\mathbb{C}} \mathcal{M}$ can be decomposed as a direct sum of $\mathrm{span}_{\mathbb{R}} \mathcal{M}$ and $\mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M}$, i.e.,

$$\mathrm{span}_{\mathbb{C}} \mathcal{M} = \mathrm{span}_{\mathbb{R}} \mathcal{M} \oplus_{\mathbb{R}} \mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M} \,.$$

For this, note that $\mathrm{span}_{\mathbb{R}} \mathcal{M}$ is an $\mathbb{R}$-linear space of Hermitian matrices, whereas $\mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M}$ is an $\mathbb{R}$-linear space consisting of skew-Hermitian matrices. Therefore $\mathrm{span}_{\mathbb{R}} \mathcal{M} \cap (\mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M}) = \{0\}$, and for the dimensions we deduce $\dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}} \mathcal{M}) = \dim_{\mathbb{R}}(\mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M})$ and

$$\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}} \mathcal{M}) = \frac{1}{2} \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{C}} \mathcal{M}) = \frac{1}{2} \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}} \mathcal{M} \oplus_{\mathbb{R}} \mathrm{i} \cdot \mathrm{span}_{\mathbb{R}} \mathcal{M}) = \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}} \mathcal{M}) \,.$$

The proof is finished. ∎

Remark 2.6. Note that in general for complex-valued functions $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$, in contrast to the set $\mathcal{M}$ of Hermitian matrices, we have

$$\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^m) \ne \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^m).$$

The poof of Lemma 2.5 heavily depends on the structure of sets in the underlying spans. A simple example is the system $\psi_1 = 1$, $\psi_2 = \mathrm{i}$, for which one gets

$$\dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^2) = 1 \ne 2 = \dim_{\mathbb{R}}(\mathrm{span}_{\mathbb{R}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^2).$$

Proposition 2.7 (Carathéodory subsampling for Hermitian matrices). *Let* $\psi_1, \ldots, \psi_m \colon \Omega \to \mathbb{C}$ *be functions and denote* $\boldsymbol{\psi}(x) = (\psi_1(x), \ldots, \psi_m(x))^\top$. *Further, let* $M \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_M \ge 0$ *and* $x_1, \ldots, x_M \in \Omega$ *such that*

$$\boldsymbol{A} = \sum_{i=1}^{M} \alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^* \,.$$

Let further $N = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\psi_k \overline{\psi_l}\}_{k,l=1}^m) \le m^2$. *Then the following holds.*

(i) *There exists a subset* $\{y_1, \ldots, y_N\} \subset \{x_1, \ldots, x_M\} \subset \Omega$ *of the initial points and weights* $\beta_1, \ldots, \beta_N \ge 0$ *such that* $\boldsymbol{A}$ *can be represented as*

$$\boldsymbol{A} = \sum_{i=1}^{N} \beta_i \boldsymbol{\psi}(y_i) \cdot \boldsymbol{\psi}(y_i)^* \,.$$

(ii) *If additionally* $\alpha_1 + \cdots + \alpha_M = 1$ *then we have a subset* $\{z_1, \ldots, z_{N+1}\} \subset \{x_1, \ldots, x_M\} \subset \Omega$ *and weights* $\gamma_1, \ldots, \gamma_{N+1} \ge 0$ *satisfying* $\gamma_1 + \cdots + \gamma_{N+1} = 1$ *with*

$$\boldsymbol{A} = \sum_{i=1}^{N+1} \gamma_i \boldsymbol{\psi}(z_i) \cdot \boldsymbol{\psi}(z_i)^* \,.$$

Proof. The convex hull of the set (2.2) of matrices $\mathcal{M}$ is defined via

$$\operatorname{conv}\mathcal{M} := \left\{ \sum_{i=1}^{M} \alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^* \colon M \in \mathbb{N},\ \alpha_i \geq 0,\ \sum_{i=1}^{M} \alpha_i = 1,\ x_i \in \Omega \right\}$$

and clearly lies in $\operatorname{span}_{\mathbb{R}} \mathcal{M}$. Consequently, it can be considered a convex subset of $\mathbb{R}^N$ (see Lemma 2.5). By Carathéodory's theorem, see [51, Thm. 1.1.4], we have that each element of a convex hull, which is a subset of $\mathbb{R}^N$, can be represented as a convex combination of $N + 1$ elements. By the above identification this reduction transfers to $\operatorname{conv}\mathcal{M}$ and proves (ii).

Considering instead of $\operatorname{conv}\mathcal{M}$ the conic hull and, respectively, dropping the condition that the coefficients α_i sum up to one, a similar reduction can be shown with a straight forward modification of the proof. Without the convexity assumption we may even reduce to N summands (instead of $N+1$). This gives (i). ■

3. Discrete orthogonality and tight frames via D-optimal designs

We are now heading for conditions under which a given system of linearly independent, bounded and continuous functions $\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$ turns out to be (discretely) orthonormal with respect to some probability measure μ on Ω. We will prove a characterization first, which is interesting on its own.

In case the characteristic function $\mathbb{1}_\Omega$ belongs to $V_n := \operatorname{span}_{\mathbb{C}}\{\varphi_1(\cdot), \ldots, \varphi_n(\cdot)\}$ we may define

$$\eta_n(x) := \frac{\sum_{k=1}^{n} |\varphi_k(x)|^2}{n},$$

which is a continuous function on Ω bounded away from zero (otherwise the $\mathbb{1}_\Omega$-function would have zeros). In case the $\mathbb{1}_\Omega$-function is not contained in V_n we augment the space to

$$V_{n+1} := \operatorname{span}_{\mathbb{C}}\{\varphi_1(\cdot), \ldots, \varphi_n(\cdot), \varphi_{n+1} := \mathbb{1}_\Omega\}.$$

In this case we obtain a basis $\{\varphi_1(\cdot), \ldots, \varphi_n(\cdot), \mathbb{1}_\Omega\}$ and the corresponding Christoffel-type function

$$\eta_{n+1}(x) := \frac{1 + \sum_{k=1}^{n} |\varphi_k(x)|^2}{n+1},$$

which is again bounded away from zero and continuous at any $x \in \Omega$.

Next, consider a discrete (atomic) measure on Ω

$$\varrho := \sum_{i=1}^{M_\varrho} \varrho_i \delta_{x_i}, \tag{3.1}$$

where the non-negative weights $(\varrho_i)_i$ sum up to one and the points $(x_i)_i \subset \Omega$ are fixed. Further, define

$$\langle f, g \rangle_\varrho := \sum_{i=1}^{M_\varrho} \varrho_i f(x_i) \overline{g(x_i)}$$

and the $m \times m$ Gramian

$$\boldsymbol{A}_\varrho := \left(\left\langle \frac{\varphi_k}{\sqrt{\eta_m}}, \frac{\varphi_l}{\sqrt{\eta_m}} \right\rangle_\varrho \right)_{k,l=1}^{m}, \tag{3.2}$$

where either $m := n$ or $m := n + 1$ depending on the cases above. The matrix $\boldsymbol{A}_\varrho$ can then also be represented in the form $\boldsymbol{A}_\varrho = \sum_{i=1}^{M_\varrho} \varrho_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^*$ with $\boldsymbol{\psi}(x) := (\psi_1(x), \ldots, \psi_m(x))^\top$ and

$$\psi_k(x) := \varphi_k(x)/\sqrt{\eta_m(x)}, \quad k = 1, \ldots, m. \tag{3.3}$$

This transition from the functions φ_k to the functions ψ_k performs the "change of measure".

Let us finally consider the corresponding set $\mathcal{M}$ from (2.2) in the proof of Proposition 2.7. This is clearly a compact set of rank-1 matrices $\boldsymbol{\psi}(x) \cdot \boldsymbol{\psi}(x)^* \in \mathbb{C}^{m \times m}$ if Ω is assumed to be compact due to the continuity of the mapping $x \mapsto \boldsymbol{\psi}(x) \cdot \boldsymbol{\psi}(x)^*$. Besides, each element of $\mathcal{M}$ is Hermitian and positive semi-definite with real determinant, i.e., $\det(\boldsymbol{\psi}(x) \cdot \boldsymbol{\psi}(x)^*) \geq 0$ for all $x \in \Omega$. Furthermore, by Proposition 2.7 the convex hull remains closed and bounded and hence compact.

The determinant is a continuous mapping on $\mathbb{C}^{m \times m}$, which implies the existence of a maximizer $\boldsymbol{A}_\alpha = \sum_{i=1}^{M_\alpha} \alpha_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^* \in \operatorname{conv} \mathcal{M}$ which satisfies (for weights ϱ as in (3.1) and $\boldsymbol{A}_\varrho$ as in (3.2))

$$\det(\boldsymbol{A}_\alpha) = \max_\varrho \det(\boldsymbol{A}_\varrho) = \sup_{\boldsymbol{M} \in \operatorname{conv} \mathcal{M}} \det(\boldsymbol{M}) \,. \tag{3.4}$$

From Proposition 2.3 we obtain that $\det(\boldsymbol{A}_\alpha) \leq 1$ since $\sum_{k=1}^m |\psi_k(x)|^2 = m$.

The subsequent theorem characterizes the property that a set of given continuous functions turns out to be orthonormal with respect to a suitably chosen Borel probability measure μ. This characterization is central for our investigations and might be of independent interest. Note, that in case the constant function belongs to their span, i.e., $\sum_{i=1}^n \gamma_i \varphi_i = \mathbb{1}_\Omega$, the condition $\sum_{i=1}^n |\gamma_i|^2 = 1$ is required. However, this can be guaranteed by a simple rescaling of the basis functions.

Theorem 3.1. *Let Ω be a compact topological space and $\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$ linearly independent continuous functions on Ω. Let further $N = \dim_\mathbb{C}(\operatorname{span}_\mathbb{C}\{\varphi_k \overline{\varphi_l}\}_{k,l=1}^n)$. The following assertions are equivalent.*

(i) *We have*

$$\boldsymbol{I}_n \in \operatorname{conv}\left\{ (\varphi_k(x)\overline{\varphi_l(x)})_{k,l=1}^n \colon x \in \Omega \right\} \,.$$

(ii) *There exist $N+1$ points $x_1, \ldots, x_{N+1} \in \Omega$ and non-negative weights with $\mu_1 + \cdots + \mu_{N+1} = 1$ such that*

(ii a) *the system $(\varphi_l)_{l=1}^n$ is orthonormal with respect to discrete measure $\varrho = \sum_{i=1}^{N+1} \mu_i \delta_{x_i}$, i.e.,*

$$\sum_{i=1}^{N+1} \mu_i \varphi_k(x_i)\overline{\varphi_l(x_i)} = \delta_{k,l} \,, \quad k, l = 1, \ldots, n;$$

(ii b) *we have*

$$\min_{(x_i)_i, (\mu_i)_i} \left\| \sum_{i=1}^{N+1} \mu_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^* - \boldsymbol{I}_n \right\|_F = 0 \tag{3.5}$$

for the minimization problem with respect to the Frobenius norm $\|\cdot\|_F$. Here $\boldsymbol{\varphi}(x) = (\varphi_1(x), \ldots, \varphi_n(x))^T$.

(iii) *There exists a Borel probability measure μ on Ω such that*

$$\int_\Omega \varphi_k(x)\overline{\varphi_l(x)} \, \mathrm{d}\mu(x) = \delta_{k,l} \,, \quad k, l = 1, \ldots, n,$$

i.e., the system $(\varphi_l)_{l=1}^n$ is orthonormal with respect to the measure μ.

(iv) *The representation $\mathbb{1}_\Omega = \sum_{i=1}^m \gamma_i \varphi_i$ holds true with $\sum_{i=1}^m |\gamma_i|^2 = 1$ (Recall that $\varphi_{n+1} := \mathbb{1}_\Omega$ in case $m = n+1$.) and the maximizer $\boldsymbol{A}_\alpha$ in (3.4) satisfies one of the following two (equivalent) conditions:*

(iv a)

$$\det(\boldsymbol{A}_\alpha) = 1 \,,$$

(iv b)

$$\boldsymbol{A}_\alpha = \boldsymbol{I}_m \,.$$

Proof.

Step 1. Using Proposition 2.7 yields (i) $\implies$ (ii). The implications (ii a) $\iff$ (ii b) $\implies$ (iii) are immediate.

Step 2. We prove (iii) $\implies$ (iv). First, note that in case $m = n$ indeed

$$1 = \int_\Omega \mathbb{1}_\Omega(x) \, \mathrm{d}\mu(x) = \sum_{k,l=1}^m \gamma_k \overline{\gamma_l} \int_\Omega \varphi_k \overline{\varphi_l(x)} \, \mathrm{d}\mu(x) = \sum_{k,l=1}^m \gamma_k \overline{\gamma_l} \delta_{k,l} = \sum_{k=1}^m |\gamma_k|^2 \,.$$

In case $m = n + 1$ we have $\sum_{i=1}^m |\gamma_i|^2 = 1$ due to $\gamma_m = 1$ and $\gamma_1 = \cdots = \gamma_n = 0$. Second, we already know that the maximizer $\boldsymbol{A}_\alpha$ in (3.4) exists and satisfies $\det(\boldsymbol{A}_\alpha) \leq 1$. From Corollary 2.2 applied to the functions ψ_k, $k = 1, \ldots, m$, with the corresponding measure $\eta_m(\cdot)\mathrm{d}\mu(\cdot)$ we obtain that $\det(\boldsymbol{A}_\alpha) \geq 1 - \varepsilon$ for all $\varepsilon > 0$ and hence $\det(\boldsymbol{A}_\alpha) \geq 1$, which proves (iv a).
Further, from Proposition 2.3, it follows that $\frac{1}{m} \sum_{k=1}^m \lambda_k(\boldsymbol{A}_\alpha) = 1$ due to $\sum_{k=1}^m |\psi_k(x)|^2 = m$. Therefore, the geometric mean of the eigenvalues of $\boldsymbol{A}_\varrho$ equals the arithmetic mean (see Proposition 2.3) which is only possible for equal eigenvalues. This implies $\lambda_1 = \cdots = \lambda_m = 1$ and hence $\boldsymbol{A}_\alpha = \boldsymbol{I}_m$ since $\boldsymbol{I}_m$ is the only Hermitian matrix that has all eigenvalues 1, i.e., the statement (iv b).

Step 3. It remains to prove the implication (iv) $\implies$ (i). First note that if (iv a) is fulfilled for a maximizer $\boldsymbol{A}_\alpha$ also (iv b) is fullfilled by the arguments in Step 2. Now we distinguish the cases $m = n + 1$ and $m = n$. In case $m = n + 1$ we obtain from $\varphi_{n+1} \equiv 1$,

$$1 = \left\langle \frac{\varphi_m}{\sqrt{\eta_m}}, \frac{\varphi_m}{\sqrt{\eta_m}} \right\rangle_\alpha = \sum_{i=1}^{M_\alpha} \frac{\alpha_i}{\eta_m(x_i)} \,. \tag{3.6}$$

In addition, we obtain for $1 \leq k, l \leq n$

$$\delta_{k,l} = \left\langle \frac{\varphi_k}{\sqrt{\eta_m}}, \frac{\varphi_l}{\sqrt{\eta_m}} \right\rangle_\alpha = \sum_{i=1}^{M_\alpha} \frac{\alpha_i}{\eta_m(x_i)} \varphi_k(x_i) \overline{\varphi_l(x_i)} \,. \tag{3.7}$$

Together with (3.6) this gives (i). The case $m = n$ is a bit more involved, since in this case we may not have $\varphi_k \equiv 1$ for any $k = 1, \ldots, n$. Clearly, (3.7) is still satisfied. To obtain an analogon to (3.6), we now exploit the condition $\mathbb{1}_\Omega = \sum_{i=1}^n \gamma_i \varphi_i$ with $\sum_{i=1}^n |\gamma_i|^2 = 1$. Indeed, then

$$1 = \sum_{k,l=1}^n \gamma_k \overline{\gamma_l} \delta_{k,l} = \sum_{k,l=1}^n \sum_{i=1}^{M_\alpha} \gamma_k \overline{\gamma_l} \frac{\alpha_i}{\eta_n(x_i)} \varphi_k(x_i) \overline{\varphi_l(x_i)} = \sum_{i=1}^{M_\alpha} \frac{\alpha_i}{\eta_n(x_i)} \,.$$

The proof is finished. ∎

Remark 3.2. If Ω is a finite set and $\varphi_1, \ldots, \varphi_n : \Omega \to \mathbb{C}$ linearly independent vectors in $\mathbb{C}^{|\Omega|}$ then the characterization in Theorem 3.1 remains valid.

Remark 3.3. If we have $\mathbb{1}_\Omega \in V_n$ from the start (and therefore $m = n$ in (3.2)), the condition (i) in Theorem 3.1, namely $\boldsymbol{I}_n \in \mathrm{conv}\{(\varphi_k(x)\overline{\varphi_l(x)})_{k,l=1}^n : x \in \Omega\}$, is equivalent to

$$\boldsymbol{I}_n \in \mathrm{cone}\left\{(\varphi_k(x)\overline{\varphi_l(x)})_{k,l=1}^n : x \in \Omega\right\}, \tag{i$'$}$$

i.e., we do not need to require the summability to 1 of the coefficients explicitly. In this case, the number of points needed in (ii) can be reduced to N.

Remark 3.4.

(i) In the language of the frame community the result in Theorem 3.1 essentially tells us that if we start with a continuous Parseval frame $(\boldsymbol{\varphi}(x))_{x\in\Omega}$ in the sense that for some Borel probability measure μ it holds (see also Definition 1.2 in [28])

$$\|\boldsymbol{a}\|_2^2 = \int_\Omega |\langle \boldsymbol{a}, \boldsymbol{\varphi}(x)\rangle|^2 \, \mathrm{d}\mu(x) \,, \qquad \boldsymbol{a} \in \mathbb{C}^n \,, \tag{3.8}$$

then there are non-negative weights μ_i, which sum up to one, and points x_i such that

$$\|\boldsymbol{a}\|_2^2 = \sum_{i=1}^{N+1} \mu_i |\langle \boldsymbol{a}, \boldsymbol{\varphi}(x_i)\rangle|^2 \,, \qquad \boldsymbol{a} \in \mathbb{C}^n \,. \tag{3.9}$$

In other words, the frame is still Parseval with respect to a discrete probability measure with at most $N+1$ atoms.

(ii) Note, that in the classical frame setting the system of vectors

$$\boldsymbol{\varphi}_\mu(x_i) = \sqrt{\mu_i}(\varphi_1(x_i), \dots, \varphi_n(x_i))^\top \in \mathbb{C}^n \,, \qquad i = 1, \dots, N+1, \tag{3.10}$$

constructed in Theorem 3.1, constitutes a discrete Parseval frame in $\mathbb{C}^n$. Hence, there is a relation to the notion of "scalable frames" as introduced in [11, 12, 41]. Let us emphasize that the property in (3.9) is stronger, since we require $\sum_{i=1}^{N+1} \mu_i = 1$.

Besides ensuring the existence of the points x_i and weights μ_i in Remark 3.4, Theorem 3.1 also provides in (ii b) and (iv) a constructive way of how to retrieve those. Corresponding algorithms are formulated below in Algorithm 1 and Algorithm 2. The first algorithm is based on characterization (iv) and a specific D-design method. The second algorithm aims for the minimization of a Frobenius norm as in (ii b). The latter idea is not new and has e.g. already been proposed in [18, §2.5] in a similar context.

Some numerical tests are conducted in Section 6, whereby our focus is on the structure of the computed outcomes and not on the computational performance of the algorithms themselves. For this we implemented Algorithm 2, where we used the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm for the optimization procedure. In both of our algorithmic formulations, the optimization procedures, which are difficult non-convex problems, are left as a black box. Here, further research is needed to find efficient practical routines.

3.1. **Equal norm tight frames with respect to a probability measure**

Assume that $\varphi_1, \dots, \varphi_n : \Omega \to \mathbb{C}$ are orthonormal functions on a domain Ω with respect to some probability measure μ on Ω or, in other words, that the corresponding system of vectors $\boldsymbol{\varphi}(x) = (\varphi_1(x), \dots, \varphi_n(x))^\top \in \mathbb{C}^n$ represents a Parseval frame in $\mathbb{C}^n$. In this setting condition (iii) of Theorem 3.1 is fulfilled and from the previous consideration we learned that, as long as the set $\mathcal{M}$ in (2.2) is a compact subset of $\mathbb{C}^{n\times n}$, we always find a discrete probability measure $\varrho = \sum_{i=1}^{M_\varrho} \varrho_i \delta_{x_i}$ via special D-optimal designs preserving the orthogonality. These are essentially the statements of Theorem 3.1 and Remark 3.4.

We now again start with a system of functions $\varphi_1, \dots, \varphi_n : \Omega \to \mathbb{C}$ and, in what follows, we again require the compactness of the set $\mathcal{M}$ from (2.2). This would be the case, for example, in the considered settings of Theorem 3.1 and Remark 3.2, where either Ω is finite or Ω is a compact topological space and $\varphi_i : \Omega \to \mathbb{C}$, $i = 1, \dots, n$, represent continuous functions. However, this time we do not demand

---

Algorithm 1 `Construction of a discrete Parseval frame`

Input: Ω compact topological space and Borel probability measure μ

          $\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$ mutually orthonormal and continuous functions on Ω w.r.t. μ

Output: $(x_i)_i$ and $(\mu_i)_i$ satisfying (3.9) and μ_i summing up to 1, $i = 1, \ldots, N+1$

1: put $N \coloneqq \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^n)$
2: **if** $\mathbb{1}_\Omega \in \mathrm{span}_{\mathbb{C}}\{\varphi_1, \ldots, \varphi_n\}$ **then**
3:     put $m \coloneqq n$
4: **else**
5:     put $\varphi_{n+1} \coloneqq \mathbb{1}_\Omega$ and $m \coloneqq n+1$
6: **end if**
7: put $\widetilde{N} \coloneqq \dim_{\mathbb{C}}(\mathrm{span}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^m)$
8: put $\eta_m(x) \coloneqq \frac{1}{m}\sum_{k=1}^m |\varphi_k(x)|^2$ and $\boldsymbol{\varphi}(x) \coloneqq (\varphi_1(x), \ldots, \varphi_m(x))^\top$
9: put $\mathcal{M} \coloneqq \left\{ \frac{\boldsymbol{\varphi}(x)\cdot\boldsymbol{\varphi}(x)^*}{\eta_m(x)} : x \in \Omega \right\}$
10: find $x_1, \ldots, x_{\tilde{N}+1} \in \Omega$ and $\alpha_1, \ldots, \alpha_{\tilde{N}+1} \geq 0$ with $\alpha_1 + \cdots + \alpha_{\tilde{N}+1} = 1$ such that

$$\det\left(\sum_{i=1}^{\tilde{N}+1} \alpha_i \frac{\boldsymbol{\varphi}(x_i)\cdot\boldsymbol{\varphi}(x_i)^*}{\eta_m(x_i)} \right) = \max_{\boldsymbol{M}\in\mathrm{conv}\,\mathcal{M}} \det(\boldsymbol{M}) = 1$$

11: **if** $N = \widetilde{N}$ **then**
12:     put $\mu_i \coloneqq \alpha_i/\eta_m(x_i)$ for $i = 1, \ldots, N+1$
13: **else**
14:     use Carathéodory subsampling to reduce the number of points to $N+1$ and calculate corresponding weights μ_i summing up to 1
15: **end if**
16: **return** $(x_i)_i$ and $(\mu_i)_i$

Algorithm 2 `Construction of a discrete Parseval frame`

Input: Ω compact topological space and Borel probability measure μ

          $\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$ mutually orthonormal and continuous functions on Ω w.r.t. μ

Output: $(x_i)_i$ and $(\mu_i)_i$ satisfying (3.9) and μ_i summing up to 1, $i = 1, \ldots, N+1$

1: $\boldsymbol{\varphi}(x) \coloneqq (\varphi_1(x), \ldots, \varphi_n(x))^\top$
2: put $N \coloneqq \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^n)$
3: find $x_1, \ldots, x_{N+1} \in \Omega$ and $\mu_1, \ldots, \mu_{N+1} \geq 0$ with $\mu_1 + \cdots + \mu_{N+1} = 1$ such that

$$\left\| \sum_{i=1}^{N+1} \mu_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^* - \boldsymbol{I}_n \right\|_F^2 = 0$$

4: **return** $(x_i)_i$ and $(\mu_i)_i$

---

orthonormality of the functions and thus do not need a measure μ on Ω in the beginning. Instead, we are satisfied with linear independence. The functions then define an n-dimensional linear function space V_n.

Interestingly, it turned out in [36, 38] that D-optimal designs are actually useful to additionally control the norms $\|f\|_\infty$ of the functions $f \in V_n$. In fact, as we will see below, we always find a discrete

probability measure $\lambda = \sum_{i=1}^{N} \lambda_i \delta_{x_i}$ on Ω with $N \leq n^2 + 1$ such that

$$\sup_{x \in \Omega} |f(x)| \leq \sqrt{n} \left(\sum_{i=1}^{N} \lambda_i |f(x_i)|^2 \right)^{1/2} \qquad \text{for all } f \in V_n \,. \tag{3.11}$$

As above, denote with $\varrho = \sum_{i=1}^{M_\varrho} \varrho_i \delta_{x_i}$ a discrete probability measure on Ω. Putting $\boldsymbol{A}_\varrho = \sum_{i=1}^{M_\varrho} \varrho_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^*$, due to the compactness of $\mathcal{M}$, we then also have a maximizer

$$\det(\boldsymbol{A}_\lambda) = \max_{\varrho} \det(\boldsymbol{A}_\varrho) = \sup_{\boldsymbol{M} \in \text{conv}\,\mathcal{M}} \det(\boldsymbol{M}) \,. \tag{3.12}$$

This time a rescaling by the Christoffel function as in (3.3) is not necessary. Taking Carathéodory's theorem into account, the measure λ gives non-negative weights $(\lambda_i)_{i=1}^{N}$ and points $(x_i)_{i=1}^{N}$ with $N = \dim_{\mathbb{C}}(\text{span}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^{n}) + 1$.

The following theorem states our result in the case of a compact topological space Ω and continuous functions $\varphi_i : \Omega \to \mathbb{C}$, $i = 1, \ldots, n$, i.e., in the setting corresponding to Theorem 3.1. An essential part of it is actually proved in [38].

Theorem 3.5. *Let Ω denote a compact topological space and $\boldsymbol{\varphi}(x) = (\varphi_1(x), \ldots, \varphi_n(x))^\top$ a vector of n linearly independent continuous functions on Ω. Further, let $N = \dim_{\mathbb{C}}(\text{span}_{\mathbb{C}}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^{n})$. Algorithm 3 then results in a discrete probability measure $\lambda = \sum_{i=1}^{N+1} \lambda_i \delta_{x_i}$ and a corresponding strictly positive definite Hermitian matrix $\boldsymbol{A}_\lambda = \sum_{i=1}^{N+1} \lambda_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^*$ such that the new system $\boldsymbol{\psi}(x) := \boldsymbol{A}_\lambda^{-1/2} \cdot \boldsymbol{\varphi}(x)$ satisfies*

$$\sum_{i=1}^{N+1} \lambda_i |\langle \boldsymbol{a}, \boldsymbol{\psi}(x_i) \rangle|^2 = \|\boldsymbol{a}\|_2^2 \tag{3.13}$$

for all $\boldsymbol{a} \in \mathbb{C}^n$. In addition, we have $\|\boldsymbol{\psi}(x)\|_2 \leq \sqrt{n}$ for any $x \in \Omega$ and $\|\boldsymbol{\psi}(x)\|_2 = \sqrt{n}$ on $\text{supp}\,\lambda$.

Proof. The orthogonality relation in (3.13) with respect to the induced inner product coming from the discrete measure λ is straightforward and a consequence of $\boldsymbol{\psi} = \boldsymbol{A}_\lambda^{-1/2} \cdot \boldsymbol{\varphi}$. At this point it is not relevant that λ is the maximizing measure in (3.12). In fact, note that for any full rank matrix $\boldsymbol{B} \in \mathbb{C}^{n \times (N+1)}$ and $\boldsymbol{A} := \boldsymbol{B} \cdot \boldsymbol{B}^*$ one has

$$(\boldsymbol{A}^{-1/2} \cdot \boldsymbol{B}) \cdot (\boldsymbol{A}^{-1/2} \cdot \boldsymbol{B})^* = \boldsymbol{I}_n \,.$$

Applying this identity to $\boldsymbol{B} = (\sqrt{\lambda_j} \varphi_i(x_j))_{i=1,j=1}^{n,N+1}$ yields (3.13).

The property that λ is the maximizing measure becomes relevant to prove $\|\boldsymbol{\psi}(x)\|_2 \leq \sqrt{n}$ for all $x \in \Omega$. For this part we refer to [38, Prop. 9], which yields this result (with $\varepsilon = 0$ in our context). It remains to argue that $\|\boldsymbol{\psi}(x)\|_2 = \sqrt{n}$ on $\text{supp}\,\lambda$. First note that (3.13) can be rewritten in the form

$$\boldsymbol{a}^* \left(\sum_{i=1}^{N+1} \lambda_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^* \right) \boldsymbol{a} = \|\boldsymbol{a}\|_2^2 \,.$$

As a consequence, the sum in the brackets is the identity matrix which has trace n. We thus get

$$n = \text{trace} \left(\sum_{i=1}^{N+1} \lambda_i \boldsymbol{\psi}(x_i) \cdot \boldsymbol{\psi}(x_i)^* \right) = \sum_{i=1}^{N+1} \lambda_i \|\boldsymbol{\psi}(x_i)\|_2^2 \,.$$

Since $\sum_{i=1}^{N+1} \lambda_i = 1$, $\lambda_i > 0$ for $x_i \in \text{supp}\,\lambda$, and $\|\boldsymbol{\psi}(x)\|_2 \leq \sqrt{n}$ for all $x \in \Omega$, we obtain that $\|\boldsymbol{\psi}(x_i)\|_2 = \sqrt{n}$ for $x_i \in \text{supp}\,\lambda$. ∎

Remark 3.6. Theorem 3.5 can be adapted to the setting of Remark 3.2, i.e., linearly independent functions $\varphi_1, \ldots, \varphi_n : \Omega \to \mathbb{C}$ on a finite set Ω, in a straight-forward manner.

Algorithm 3 `Construction of an equal norm tight frame`

Input:	Ω	compact topological space
	$\varphi_1, \ldots, \varphi_n \colon \Omega \to \mathbb{C}$	linearly independent continuous functions on Ω

Output:	$(x_i)_i$ and $(\lambda_i)_i$	satisfying (3.13) and $\lambda_i > 0$
	$\boldsymbol{\psi}(x)$	such that $\|\boldsymbol{\psi}(x_i)\|_2 = \sqrt{n}$ and $\|\boldsymbol{\psi}(x)\|_2 \le \sqrt{n}$ for any $x \in \Omega$

1: $\boldsymbol{\varphi}(x) := (\varphi_1(x), \ldots, \varphi_n(x))^\top$
2: put $\mathcal{M} := \{\boldsymbol{\varphi}(x) \cdot \boldsymbol{\varphi}(x)^* : x \in \Omega\}$
3: put $N := \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\varphi_k \overline{\varphi}_l\}_{k,l=1}^n)$
4: find $x_1, \ldots, x_{N+1} \in \Omega$ and $\lambda_1, \ldots, \lambda_{N+1} \ge 0$ with $\lambda_1 + \cdots + \lambda_{N+1} = 1$ such that

$$\det\left(\sum_{i=1}^{N+1} \lambda_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^*\right) = \max_{M \in \mathrm{conv}\,\mathcal{M}} \det(\boldsymbol{M})$$

5: put $\boldsymbol{A}_\lambda = \sum_{i=1}^{N+1} \lambda_i \boldsymbol{\varphi}(x_i) \cdot \boldsymbol{\varphi}(x_i)^*$
6: put $\boldsymbol{\psi}(x) := \boldsymbol{A}_\lambda^{-1/2} \cdot \boldsymbol{\varphi}(x)$ for $x \in \Omega$
7: filter out the indices i with $\lambda_i > 0$; reindex accordingly
8: **return** $\boldsymbol{\psi}(x)$; $(x_i)_i$ and $(\lambda_i)_i$ for indices i with $\lambda_i > 0$

Remark 3.7. The property $\|\boldsymbol{\psi}(x)\|_2 \le \sqrt{n}$ yields (3.11). Indeed, for any $f \in V_n$ with $f(x) = \langle \boldsymbol{a}, \boldsymbol{\psi}(x)\rangle$ we have the estimate

$$|f(x)|^2 = |\langle \boldsymbol{a}, \boldsymbol{\psi}(x)\rangle|^2 \le \|\boldsymbol{a}\|_2^2 \|\boldsymbol{\psi}(x)\|_2^2 \le n \sum_{i=1}^{N+1} \lambda_i |\langle \boldsymbol{a}, \boldsymbol{\psi}(x_i)\rangle|^2 = n \sum_{i=1}^{N+1} \lambda_i |f(x_i)|^2 \,.$$

Remark 3.8. The procedure described in Theorem 3.5 reminds on the way of making a given frame $\boldsymbol{\varphi}$ tight by multiplying with $\boldsymbol{A}^{-1/2}$, where $\boldsymbol{A} = \boldsymbol{T}_{\boldsymbol{\varphi}}^* \cdot \boldsymbol{T}_{\boldsymbol{\varphi}}$ denotes the frame matrix of $\boldsymbol{\varphi}$ and $\boldsymbol{T}_{\boldsymbol{\varphi}}$ its analysis operator. However, the transformed frame $\boldsymbol{\psi} = \boldsymbol{A}^{-1/2} \cdot \boldsymbol{\varphi}$ then usually does not consist of equal norm frame elements. A subsequent normalization and corresponding reweighting would yield an equal norm tight frame of type (3.13) or (3.8) with respect to a discrete probability measure. However, the separate normalization step would typically change the coefficient space such that $\mathrm{Im}(\boldsymbol{T}_{\boldsymbol{\psi}}) \neq \mathrm{Im}(\boldsymbol{T}_{\boldsymbol{\varphi}})$. In Theorem 3.5 the weights and points are constructed simultaneously and result in a discrete measure and a corresponding equal norm tight frame $\boldsymbol{\psi}$ where $\mathrm{Im}(\boldsymbol{T}_{\boldsymbol{\psi}}) = \mathrm{Im}(\boldsymbol{T}_{\boldsymbol{\varphi}})$. This construction thus maintains a stronger connection to the original frame. Let us emphasize that there is significant interest in the frame community for the construction of unit/equal norm tight frames (ENTF,UNTF), see [14, 15].

4. Application: Exact Marcinkiewicz–Zygmund inequalities and quadrature

Frames have a related structure to L_2-MZ inequalities and tightness in frames means exactness in L_2-MZ inequalities. With that, we immediately obtain a method for constructing points and weights fulfilling the identity in the following theorem.

Theorem 4.1. *Let μ be a Borel probability measure on a compact topological space Ω and V_n be an n-dimensional subspace of $C(\Omega)$, the space of continuous complex-valued functions on Ω. Then there exists an $N \le \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{f \cdot \overline{g} : f, g \in V_n\}) \le n^2$ and $N + 1$ points $x_1, \ldots, x_{N+1} \in \Omega$ together with*

non-negative weights satisfying $\sum_{i=1}^{N+1} \mu_i = 1$ such that

$$\int_\Omega |f(x)|^2 \, \mathrm{d}\mu(x) = \sum_{i=1}^{N+1} \mu_i |f(x_i)|^2 \quad \text{for all} \quad f \in V_n \,. \tag{4.1}$$

Proof. Choose an orthonormal basis in V_n. By Theorem 3.1 this basis is also discretely orthonormal. The statement is obtained by Parseval's identity. ∎

The respective points and weights (the discrete measure) are determined by finding a maximizer through (3.4) taking Carathéodory subsampling into account. Exact L_2-MZ inequalities have a useful property. They also discretize (subsample) the corresponding inner product as the following corollary shows.

Corollary 4.2. *Let μ be a Borel probability measure on a compact topological space Ω and V_n be an n-dimensional subspace of the complex-valued $C(\Omega)$. Then there exists an $N \le n^2$ and $N+1$ points $x_1, \ldots, x_{N+1} \in \Omega$ together with non-negative weights satisfying $\sum_{i=1}^{N+1} \mu_i = 1$ such that for all $f, g \in V_n$ it holds*

$$\int_\Omega f(x)\overline{g(x)} \, \mathrm{d}\mu(x) = \sum_{i=1}^{N+1} \mu_i f(x_i)\overline{g(x_i)} \,.$$

Proof. Due to Theorem 4.1, for any $f, g \in V_n$ the relations (4.1) hold. Hence, they also hold for $f+g$, $f-g$, $f+ig$ and $f-ig$ that belong to the subspace V_n. Using the polarization identity

$$\int_\Omega f(x)\overline{g(x)} \, \mathrm{d}\mu(x) = \frac{1}{4}\left(\int_\Omega |f(x)+g(x)|^2 \, \mathrm{d}\mu(x) - \int_\Omega |f(x)-g(x)|^2 \, \mathrm{d}\mu(x) \right)$$
$$+ \frac{i}{4}\left(\int_\Omega |f(x)+ig(x)|^2 \, \mathrm{d}\mu(x) - \int_\Omega |f(x)-ig(x)|^2 \, \mathrm{d}\mu(x) \right),$$

we then obtain

$$\int_\Omega f(x)\overline{g(x)} \, \mathrm{d}\mu(x) = \frac{1}{4}\left(\sum_{i=1}^{N+1} \mu_i |f(x_i)+g(x_i)|^2 - \sum_{i=1}^{N+1} \mu_i |f(x_i)-g(x_i)|^2 \right)$$
$$+ \frac{i}{4}\left(\sum_{i=1}^{N+1} \mu_i |f(x_i)+ig(x_i)|^2 - \sum_{i=1}^{N+1} \mu_i |f(x_i)-ig(x_i)|^2 \right)$$
$$= \sum_{i=1}^{N+1} \mu_i f(x_i)\overline{g(x_i)} \,. \qquad \blacksquare$$

We obtain general quadrature formulas in the sense of Tchakaloff using our approach. See Tchakaloff [56] and also [50], [44, Thm. 1], [18, Thm. 4.1], which handle the case of real functions.

Theorem 4.3 (Tchakaloff for complex-valued functions). *Let μ be a Borel probability measure on a compact topological space Ω and $V_n \subset C(\Omega)$ (complex-valued). Then there exists an $N \le 2n$ and $x_1, \ldots, x_{N+1} \in \Omega$ such that with non-negative weights satisfying $\mu_i + \cdots + \mu_{N+1} = 1$ it holds*

$$\int_\Omega f(x) \, \mathrm{d}\mu(x) = \sum_{i=1}^{N+1} \mu_i f(x_i) \quad \text{for all} \quad f \in V_n \,.$$

Proof.

Step 1. Applying Corollary 4.2 to $V_n + \mathrm{span}_C(\mathbb{1}_\Omega)$ and setting $g = \mathbb{1}_\Omega$ gives the stated quadrature condition with $M \le (n+1)^2$ points $y_1, \ldots, y_{M+1}$ and weights $\beta_1, \ldots, \beta_{M+1}$.

Step 2. In order to obtain the stated number of points, we need to apply a second Carathéodory step. For $\varphi_1, \ldots, \varphi_n$ a basis of V_N, the points and weights from Step 1 fulfill $\boldsymbol{A}(\beta_1 \ \ldots \ \beta_{M+1})^\top = \boldsymbol{b}$ with

$$\boldsymbol{A} = \begin{pmatrix} \mathrm{Re}(\varphi_1(y_1)) & \ldots & \mathrm{Re}(\varphi_1(y_{M+1})) \\ \mathrm{Im}(\varphi_1(y_1)) & \ldots & \mathrm{Im}(\varphi_1(y_{M+1})) \\ \vdots & \ddots & \vdots \\ \mathrm{Re}(\varphi_n(y_1)) & \ldots & \mathrm{Re}(\varphi_n(y_{M+1})) \\ \mathrm{Im}(\varphi_n(y_1)) & \ldots & \mathrm{Im}(\varphi_n(y_{M+1})) \end{pmatrix} \in \mathbb{R}^{2n \times M+1} \quad \text{and} \quad \boldsymbol{b} = \begin{pmatrix} \mathrm{Re}(\int_\Omega \varphi_1 \, \mathrm{d}\mu) \\ \mathrm{Im}(\int_\Omega \varphi_1 \, \mathrm{d}\mu) \\ \vdots \\ \mathrm{Re}(\int_\Omega \varphi_n \, \mathrm{d}\mu) \\ \mathrm{Im}(\int_\Omega \varphi_n \, \mathrm{d}\mu) \end{pmatrix} \in \mathbb{R}^{2n}.$$

We interpret this as $\boldsymbol{b}$ being represented by a convex combination of the column vectors of $\boldsymbol{A}$. Applying Carathéodory here gives $N \leq 2n$, non-negative real weights $\mu_1, \ldots, \mu_{N+1}$, and a subset of points $\{x_1, \ldots, x_N\} \subset \{y_1, \ldots, y_M\}$ such that we have in $\mathbb{C}^n$

$$\sum_{i=1}^{N+1} \mu_i (\varphi_1(x_1), \ldots, \varphi_n(x_n))^\top = \left(\int_\Omega \varphi_1 \, \mathrm{d}\mu, \ldots, \int_\Omega \varphi_n \, \mathrm{d}\mu \right)^\top.$$

The statement in Theorem 4.3 follows by linearity since any f can be represented as complex linear combination in the functions $\varphi_1, \ldots, \varphi_n$. ∎

Next, we show a method for constructing exact L_p-MZ inequalities with even p by using similar techniques as in [18, Thm. 3.1] but utilizing the exact L_2 result in Theorem 4.1 instead of Tchakaloff's Theorem on exact quadrature.

Corollary 4.4. *Let μ be a Borel probability measure on a compact topological space Ω and $V_n \subset C(\Omega)$ (complex-valued) and p even. Then there exists an $N \leq 2\binom{n+p-1}{p} \sim n^p$ and $x_1, \ldots, x_{N+1} \in \Omega$ with non-negative weights satisfying $\mu_i + \cdots + \mu_{N+1} = 1$ such that*

$$\int_\Omega |f(x)|^p \, \mathrm{d}\mu(x) = \sum_{i=1}^{N+1} \mu_i |f(x_i)|^p \quad \text{for all} \quad f \in V_n. \tag{4.2}$$

Proof. Using a basis of $V_n = \mathrm{span}_\mathbb{C}\{\varphi_1, \ldots, \varphi_n\}$ we define the space

$$W_n = \mathrm{span}_\mathbb{C}\left\{ \varphi_1^{k_1} \ldots \varphi_n^{k_n} : k_1, \ldots, k_n \in \mathbb{N}_0, \ k_1 + \cdots + k_n = \frac{p}{2} \right\}.$$

We have

$$\dim_\mathbb{C}(\mathrm{span}_\mathbb{C}\{f \cdot g : f, g \in W_n\}) = \dim_\mathbb{C}\left(\mathrm{span}_\mathbb{C}\left\{ \varphi_1^{k_1} \ldots \varphi_n^{k_n} : k_1, \ldots, k_n \in \mathbb{N}_0, \ k_1 + \cdots + k_n = p \right\} \right)$$

$$\leq \left| \{k_1, \ldots, k_n \in \mathbb{N}_0 : k_1 + \cdots + k_n = p\} \right|$$

$$= \binom{n+p-1}{p},$$

where the last equality is known as the number of weak compositions of the integer p into n parts. Applying Theorem 4.1 to W_n gives an $N \leq \binom{n+p-1}{p}$ with points $x_1, \ldots, x_{N+1} \in \Omega$ and non-negative weights satisfying $\mu_i + \cdots + \mu_{N+1} = 1$ such that

$$\int_\Omega |f(x)|^2 \, \mathrm{d}\mu(x) = \sum_{i=1}^{N+1} \mu_i |f(x_i)|^2 \quad \text{for all} \quad f \in W_n.$$

In particular, it holds for all $f^{p/2}$ with $f \in V_n$, which is the assertion. ∎

Corollary 4.4 contributes to [18, Open problem 10] in the sense that it gives a constructive method to obtain an L_p-MZ inequality for even p. Note, constructiveness here is meant in the sense of [18,

§2.5] where a greedy algorithm is said to be constructive. Since we have positive weights, it goes even further to make [18, Cor. 4.2] constructive.

Note, in general the number of points is optimal, cf. [18, Thm. 3.2]. The condition of p being an even number is necessary as it has been shown that otherwise such inequalities do not exist in general, cf. [18, Prop. 3.3].

5. Specific examples

5.1. The d-torus

Let $\mathbb{T} = [0, 1)$ be a torus where the endpoints of the interval are identified. By $\mathbb{T}^d$ we denote a d-dimensional torus and equip it with the normalized Lebesgue measure $\mathrm{d}\boldsymbol{x}$. Let further $L_2 := L_2(\mathbb{T}^d, \mathrm{d}\boldsymbol{x})$ and $V = V(I) := \operatorname{span}_{\mathbb{C}}\{\exp(2\pi\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{x} \rangle) \colon \boldsymbol{k} \in I \subset \mathbb{Z}^d, \ \boldsymbol{x} \in \mathbb{T}^d\}$. In what follows, we assume that $|I|$ is finite. Let us denote with $\mathcal{D}(I) := \{\boldsymbol{k} - \boldsymbol{\ell} \ : \ \boldsymbol{k}, \boldsymbol{\ell} \in I\}$.

Corollary 5.1. *There exist points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^N \in \mathbb{T}^d$ and non-negative weights $w_1, \ldots, w_N$ satisfying $w_1 + \cdots + w_N = 1$ with $N = |\mathcal{D}(I)| \leq |I|^2 - |I| + 1$ such that for all $f \in V$ it holds*

$$\int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} = \sum_{i=1}^{N} w_i |f(\boldsymbol{x}^i)|^2 \,. \tag{5.1}$$

Proof. The statement with $N + 1$ instead of N follows immediately from Theorem 4.1 by the fact, that $\dim_{\mathbb{C}}(\operatorname{span}_{\mathbb{C}}\{f \cdot \overline{g} : f, g \in V\}) = |\mathcal{D}(I)| \leq |I|^2 - |I| + 1$. In order to reduce further to $N = |\mathcal{D}(\mathcal{I})|$ points we return to the characterization in Theorem 3.1. Recall, that for the reduction to $N + 1$ we used Carathéodory's theorem in Proposition 2.7. This time we interpret the convex combination of the identity matrix $\boldsymbol{I}_n \in \operatorname{conv} \mathcal{M}$ as a conical combination (just non-negative coefficients which do not necessarily sum up to one). With Proposition 2.7 (i) we obtain (5.1) initially with N non-negative weights w_i. However, since the space $V(I)$ contains complex monomials with modulus one everywhere we obtain $w_1 + \cdots + w_N = 1$ by testing (5.1) with such a monomial. ∎

Next, we compare Corollary 5.1 to existing results in the torus setting. Let us emphasize that existing exact MZ-discretization results need additional restrictions on the index set. This is not the case in Corollary 5.1. However, existing results use equal weights. This problem turns out to be much harder.

Equidistant points

It is known that if $\sqrt[d]{n} \in \mathbb{N}$ and $I \subset \{-\sqrt[d]{n}/2, \ldots, \sqrt[d]{n}/2 - 1\}^d$, then the grid $\boldsymbol{G}$ of equidistant points $\boldsymbol{X} := \{\boldsymbol{k}/\sqrt[d]{n} \colon \boldsymbol{k} \in \{1, \ldots, \sqrt[d]{n}\}^d\}$ with equal weights $1/n$ satisfies the exact Marcinkiewicz–Zygmund inequality, i.e., for all $f \in V$ it holds

$$\int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} = \frac{1}{n} \sum_{\boldsymbol{x} \in \boldsymbol{X}} |f(\boldsymbol{x})|^2 \,. \tag{5.2}$$

Equidistant points are an example of an exact L_2-MZ inequality with equal weights. If I is small compared to the frequency cube $\{-\sqrt[d]{n}/2, \ldots, \sqrt[d]{n}/2 - 1\}^d$ this setting is not very useful in high dimensions d. However, if I denotes the entire frequency cube, no oversampling is needed, i.e., $|I| = |\boldsymbol{G}| = n$.

Rank-1 lattices [34] *and quadratic oversampling*

For a set $I \subset \mathbb{Z}^d$,

$$\max_{j=1,\ldots,d} \max_{\boldsymbol{h},\boldsymbol{k} \in I} |h_j - k_j| < |I|, \tag{5.3}$$

there is some $M \in \mathbb{N}$ satisfying $|I| \leq M \leq |\mathcal{D}(I)| \leq |I|^2 - |I| + 1$ such that one can construct a rank-1 lattice $\boldsymbol{X} := \{\frac{i}{M}\boldsymbol{z} \bmod \mathbb{1} : i = 0, \ldots, M-1\}$ such that for all $f \in V(I)$ (5.2) holds. Moreover, such a rank-1 lattice can be efficiently determined using a component-by-component approach, cf. [40]. It is worth mentioning that rank-1 lattices have a similar quadratic oversampling as in Corollary 5.1. In [57] it is claimed that the additional restriction (5.3) on the index set can be dropped. In fact, it is claimed that a Korobov lattice of size $M \leq 2d|\mathcal{D}(I)|$ always exists, such that (5.2) is true for all $f \in V(I)$. A closer inspection of the proof in [57, §4], however, shows that an additional assumption similar to (5.3) is used in the proof.

Theorem 5.2. *Let* $M \in \mathbb{N}$, $\boldsymbol{R}_K = \{0, 1/K, \ldots, (K-1)/K\}$ *the full one-dimensional grid with* K *equidistant points and* $\boldsymbol{R}^{M,d} = \left(\bigcup_{K=1}^{M} \boldsymbol{R}_K\right)^d$ *a d-dimensional rectilinear grid. Then there always exist index sets* $I \subset \mathbb{Z}^d$ *with two elements such that for all* $\boldsymbol{X} \subset \boldsymbol{R}^{M,d}$ *the exact MZ inequality* (5.1) *is violated.*

Proof. We simply specify a range of sets that fulfills the assertion. Let

$$I_{\boldsymbol{a},\boldsymbol{b}} := \left\{\boldsymbol{a}, \boldsymbol{a} + \left(\prod_{k=1}^{M} k\right)\boldsymbol{b}\right\}, \qquad \boldsymbol{a} \in \mathbb{Z}^d, \boldsymbol{b} \in \mathbb{Z}^d \setminus \{\boldsymbol{0}\},$$

Clearly, for any $\boldsymbol{x} \in \boldsymbol{R}^{M,d}$ and $j = 1, \ldots, d$, there exist $K_j \in \mathbb{N} \cap [0, M]$ and $l_j \in \mathbb{N} \cap [0, K_j - 1]$ such that $x_j = l_j/K_j$, which yields

$$\exp(2\pi\mathrm{i}\langle\boldsymbol{a},\boldsymbol{x}\rangle) = \prod_{j=1}^{d} \exp(2\pi\mathrm{i}\,a_j x_j) = \prod_{j=1}^{d} \exp\left(2\pi\mathrm{i}\,a_j \frac{l_j}{K_j}\right).$$

Due to the fact that $\frac{l_j}{K_j}\left(\prod_{k=1}^{M} k\right)b_j \in \mathbb{N}_0$ for each j, we observe

$$\exp(2\pi\mathrm{i}\langle\boldsymbol{a},\boldsymbol{x}\rangle) = \prod_{j=1}^{d} \left(\exp\left(2\pi\mathrm{i}\,a_j \frac{l_j}{K_j}\right)\exp\left(2\pi\mathrm{i}\,\frac{l_j}{K_j}\left(\prod_{k=1}^{M} k\right)b_j\right)\right)$$

$$= \exp\left(2\pi\mathrm{i}\left\langle\boldsymbol{a} + \left(\prod_{k=1}^{M} k\right)\boldsymbol{b}, \boldsymbol{x}\right\rangle\right).$$

As a consequence, for the function $f: \mathbb{T}^d \to \mathbb{C}$, $\boldsymbol{x} \mapsto \exp(2\pi\mathrm{i}\langle\boldsymbol{a},\boldsymbol{x}\rangle) - \exp(2\pi\mathrm{i}\langle\boldsymbol{a}',\boldsymbol{x}\rangle)$, $\boldsymbol{a}' = \boldsymbol{a} + \left(\prod_{k=1}^{M} k\right)\boldsymbol{b}$, we observe $f \in V(I_{\boldsymbol{a},\boldsymbol{b}})$ and that the right hand side of (5.1) is zero independent of $\boldsymbol{X} \subset \boldsymbol{R}^{M,d}$, while the left hand side is not. ∎

Corollary 5.3. *For any* $n \in \mathbb{N}$, $n \geq 2$, *and* $M \in \mathbb{N}$ *there are index sets* $I \subset \mathbb{Z}^d$, $|I| = n$, *such that* (5.1) *is violated for any lattice rule with not more than M elements.*

Proof. First we assume the lattice rule being a rank-m lattice rule with $M' \leq M$ sampling nodes $\boldsymbol{X}$. Moreover, we assume that an expression of the lattice rule is given in canonical form, cf. [54, Thm. 4.5], which implies that each used sampling node $\boldsymbol{x} \in \boldsymbol{X}$ can be written as

$$\boldsymbol{x} = \frac{j_1 \boldsymbol{z}_1}{M'_1} + \cdots + \frac{j_m \boldsymbol{z}_m}{M'_m} \bmod \mathbb{1},$$

where $j_1, \ldots, j_m \in \mathbb{N}_0$, $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m \in \mathbb{Z}^d$, and $M' = \prod_{j=1}^{m} M'_j$. Obviously, each component of each sampling node $\boldsymbol{x} \in \boldsymbol{X} \subset [0,1)^d$ can be written as a rational with denominator at most M', which

means that $\boldsymbol{X} \subset \boldsymbol{R}^{M',d} \subset \boldsymbol{R}^{M,d}$ and we can apply Theorem 5.2 and subsequently add arbitrary indices from $\mathbb{Z}^d$ to I until $|I| = n$ is reached. ∎

Remark 5.4. The index sets and fooling functions constructed in the proofs of Theorem 5.2 and Corollary 5.3 can also be applied to (4.2) with $\Omega = \mathbb{T}^d$, $V_n = V(I)$, and $d\mu(x)$ the normalized Lebesgue measure. Consequently, sampling sets consisting of rational sampling nodes with bounded denominators do not provide exact L_p-MZ discretizations for arbitrary spans $V(I)$, $|I| < n$, of trigonometric monomials. In particular, lattice rules of bounded cardinality cannot fulfill this characteristic either.

Random points [58] *and logarithmic oversampling*

For any $\varepsilon > 0$ there exist $M \leq \frac{6}{\varepsilon^2}|I|\log|I|$ uniformly drawn points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M$ such that with high probability it holds for all $f \in V$ the tight bound

$$(1 - \varepsilon) \int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} \leq \frac{1}{M} \sum_{i=1}^{M} |f(\boldsymbol{x}^i)|^2 \leq (1 + \varepsilon) \int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} \,. \tag{5.4}$$

Random points show the existence of L_2-MZ inequalities with logarithmic instead of quadratic oversampling but loose the exactness condition.

Constant oversampling [3, 4, 57]

Let $b > 1$ and $M = \lceil b|I| \rceil$. Then there exist subsampled random points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M \in \mathbb{T}^d$ and positive weights $w_1, \ldots, w_M$ such that it holds for all $f \in V$ that

$$\int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} \leq \sum_{i=1}^{M} w_i |f(\boldsymbol{x}^i)|^2 \leq \left(\frac{\sqrt{b} + 1}{\sqrt{b} - 1} \right)^2 \int_{\mathbb{T}^d} |f(\boldsymbol{x})|^2 \, d\boldsymbol{x} \,.$$

Here the question arises if constant weights can be used here. Based on the groundbreaking works by Marcus, Spielman, Srivastava [43] and Nitzan, Olevskii, Ulanovskii [46] one should mention the following recent result by Kosov [37, Cor. 1.3]. If $\varepsilon > 0$ then there is a set of M points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M \in \mathbb{T}^d$ with $M \leq 10^5 \varepsilon^{-2}|I|$ such that (5.4) holds. All these results show discretization inequalities for the square norm with merely linear oversampling. The last one even gives a tight bound with ε-distortion.

5.2. The d-sphere

Let $\mathbb{S}^d = \{\boldsymbol{x} \in \mathbb{R}^{d+1} : \|\boldsymbol{x}\|_2 = 1\}$ be the unit sphere with dimension d and let $\mu(\boldsymbol{x})$ be the normalized surface measure. We have a look at the space of polynomials of degree at most m restricted to the sphere

$$\Pi_m := \operatorname{span}_{\mathbb{R}} \left\{ f : \mathbb{S}^d \to \mathbb{R} \ : \ f(\boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{k}} = x_1^{k_1} \ldots x_d^{k_d}, \boldsymbol{k} \in \mathbb{N}_0^d, \|\boldsymbol{k}\|_1 \leq m \right\} \,.$$

These are commonly used in approximation on the sphere as they can be represented by spherical harmonics allowing for applying analytic tools and fast algorithms, cf. [19]. For the dimension of this space, we have by [19, Cor. 1.1.5]

$$\dim_{\mathbb{R}} \Pi_m = \binom{m + d}{d} + \binom{m + d - 1}{d} \leq \left(\frac{9m}{d} \right)^d \,,$$

where $d \leq m$ and $\binom{n}{k} \leq (n \cdot e/k)^k$ was used in the last inequality.

Theorem 5.5. *Let $m \geq d$, $p \in \mathbb{N}$ even. There exist points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^N$ in $\mathbb{S}^d$ and positive weights $w_1 + \cdots + w_N = 1$ with $N \leq \dim_{\mathbb{R}}(\Pi_{pm}) \leq (9pm/d)^d$ such that*

$$\int_{\mathbb{S}^d} |f(\boldsymbol{x})|^p \, \mathrm{d}\mu(\boldsymbol{x}) = \sum_{i=1}^N w_i |f(\boldsymbol{x}^i)|^p \quad for\ all \quad f \in \Pi_m \, .$$

Proof. The proof works analogous to Corollary 4.4: Using a basis of $\Pi_m = \operatorname{span}_{\mathbb{R}}\{\varphi_1, \ldots, \varphi_{\dim_{\mathbb{R}}(\Pi_m)}\}$ we define the space

$$W_m = \operatorname{span}_{\mathbb{R}}\left\{\varphi_1^{k_1} \cdots \varphi_{\dim_{\mathbb{R}}(\Pi_m)}^{k_n} : k_1, \ldots, k_n \in \mathbb{N}_0, \ k_1 + \cdots + k_n = \frac{p}{2}\right\}.$$

Because of the polynomial structure, we have $W_m = \Pi_{pm/2}$. It is left to apply Theorem 4.1 to $\Pi_{pm/2}$ and use $\dim_{\mathbb{R}}(\{f \cdot \overline{g} : f, g \in \Pi_{pm/2}\}) = \dim_{\mathbb{R}}(\Pi_{pm}) \leq (9pm/d)^d$.

Because $\mathbb{1}_{\mathbb{S}^d} \in \Pi_m$ for all m, the number of points can be reduced to $\dim_{\mathbb{R}}(\Pi_{pm})$ instead of $\dim_{\mathbb{R}}(\Pi_{pm}) + 1$ by the same reasoning as in Corollary 5.1. ∎

Next, we compare Theorem 5.5 to existing results.

- For $\mathbb{S}^2$ the assertion above gives the existence of an exact L_2-MZ inequality for $N \leq (2m+1)^2$ points, which matches the lower bound of $N \geq m(m+1)/2$ in terms of the rate from [18].

- Similar to the torus, one can draw uniform points $\{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\} \subset \mathbb{S}^d$ with $M \geq C \dim_{\mathbb{R}}(\Pi_m) \log(\dim_{\mathbb{R}} \Pi_m)\varepsilon^{-2}$ at random to obtain an equal-weight, non-exact L_2-MZ inequality, i.e.,

$$(1-\varepsilon)\|f\|_{L_2(\mu)}^2 \leq \frac{1}{M} \sum_{i=1}^M |f(\boldsymbol{x}^i)|^2 \leq (1+\varepsilon)\|f\|_{L_2(\mu)}^2 \quad for\ all \quad f \in \Pi_m \, .$$

  This result is stated in [26, Thm. 3.2].

- Furthermore, in [26, Thm. 4.1] the existence of M points with $M \leq C \dim_{\mathbb{R}}(\Pi_m)\varepsilon^{-d}$ was shown which simultaneously fulfill an equal-weight, non-exact L_p-MZ inequality for $p \in [1, \infty]$.

Moving on to exact quadrature, we immediately obtain the following result.

Theorem 5.6. *Let $m \geq d$. There exist points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^N$ in $\mathbb{S}^d$ and positive weights $w_1 + \cdots + w_N = 1$ with $N \leq \dim_{\mathbb{R}}(\Pi_m) \leq (9m/d)^d$ such that*

$$\int_{\mathbb{S}^d} f(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) = \sum_{i=1}^N w_i f(\boldsymbol{x}^i) \quad for\ all \quad f \in \Pi_m \, .$$

Proof. The proof works analogous to Theorem 4.3. But since we have real-valued basis functions and $\mathbb{1}_{\mathbb{S}^d} \in \Pi_m$ for all m, we use the conical variant of Carathéodory's Theorem (as in Corollary 5.1) in order to obtain $\dim_{\mathbb{R}}(\Pi_m)$ points instead of $2 \dim_{\mathbb{R}}(\Pi_m) + 1$. ∎

Quadrature formulas of the above type with equal weights are called *t*-designs introduced in [20] (or *m*-design according to our notation, where m is the degree of the polynomial). The equal weight condition makes this problem much harder and only a limited number of constructions of spherical designs are known. However, there are also approaches to obtain spherical designs being exact up to machine precision numerically, cf. [29, 60]. Let us also mention [1], where the authors use a similar technique of maximizing a certain Gram determinant (for the case $d = 2$) in order to computationally construct well-conditioned spherical design with a number of $M \geq (m+1)^2$ points. In general the existence of spherical designs is known with the optimal asymptotic rate $M \geq C_d m^d$, cf. [6, Thm. 1]. This matches the number of points of our weighted result in Theorem 5.6.

6. Numerical experiments

In this section we test Algorithm 2 numerically in order to obtain points and weights forming an exact L_2-MZ inequality. For that let us assume a product type basis $\varphi_{\boldsymbol{k}}(\boldsymbol{x}) = \varphi_{k_1}(x_1) \ldots \varphi_{k_d}(x_d)$ and, as before, $\boldsymbol{\varphi}(x) = (\varphi_{\boldsymbol{k}}(\boldsymbol{x}))_{\boldsymbol{k} \in I}^{\top}$ for some multi-index set I. For the minimization process we use the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, which is an iterative quasi-Newton optimizer for unconstrained non-linear optimization problems, in order to update the points and weights in an alternating fashion. It uses the objective function

$$f(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^{N+1}, \alpha_1, \ldots, \alpha_{N+1}) = \left\| \sum_{i=1}^{N+1} \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m \right\|_F^2 = \sum_{\boldsymbol{k},\boldsymbol{l} \in I} \left| \sum_{i=1}^{N+1} \alpha_i \varphi_{\boldsymbol{k}}(\boldsymbol{x}^i) \overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^i)} - \delta_{\boldsymbol{k}\boldsymbol{l}} \right|^2 .$$

and the partial derivatives: noting that for any complex numbers z_1 and z_2 it holds $\mathrm{Re}(z_1)\mathrm{Re}(z_2) + \mathrm{Im}(z_1)\mathrm{Im}(z_2) = \mathrm{Re}(z_1\overline{z_2})$, they evaluate to

$$\frac{\partial f}{\partial x_j^{i'}} = \sum_{\boldsymbol{k},\boldsymbol{l} \in I} 2\,\mathrm{Re}\left(\sum_{i=1}^{N+1} \alpha_i \varphi_{\boldsymbol{k}}(\boldsymbol{x}^i)\overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^i)} - \delta_{\boldsymbol{k},\boldsymbol{l}} \right) \mathrm{Re}\left(\alpha_{i'}\varphi_{\boldsymbol{k}}(\boldsymbol{x}^{i'})\left(\frac{\partial}{\partial x_j^{i'}} \overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^{i'})} \right) + \alpha_{i'}\left(\frac{\partial}{\partial x_j^{i'}} \varphi_{\boldsymbol{k}}(\boldsymbol{x}^{i'}) \right)\overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^{i'})} \right)$$

$$+ 2\,\mathrm{Im}\left(\sum_{i=1}^{N+1} \alpha_i \varphi_{\boldsymbol{k}}(\boldsymbol{x}^i)\overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^i)} - \delta_{\boldsymbol{k},\boldsymbol{l}} \right) \mathrm{Im}\left(\alpha_{i'}\varphi_{\boldsymbol{k}}(\boldsymbol{x}^{i'})\left(\frac{\partial}{\partial x_j^{i'}} \overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^{i'})} \right) + \alpha_{i'}\left(\frac{\partial}{\partial x_j^{i'}} \varphi_{\boldsymbol{k}}(\boldsymbol{x}^{i'}) \right)\overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^{i'})} \right)$$

$$= 4\alpha_{i'}\,\mathrm{Re}\left(\sum_{\boldsymbol{k},\boldsymbol{l} \in I} \overline{\varphi_{\boldsymbol{k}}(\boldsymbol{x}^{i'})}\left(\sum_{i=1}^{N+1} \alpha_i \varphi_{\boldsymbol{k}}(\boldsymbol{x}^i)\overline{\varphi_{\boldsymbol{l}}(\boldsymbol{x}^i)} - \delta_{\boldsymbol{k}\boldsymbol{l}} \right)\left(\frac{\partial}{\partial x_j^{i'}} \varphi_{\boldsymbol{l}}(\boldsymbol{x}^{i'}) \right) \right)$$

$$= 4\alpha_{i'}\,\mathrm{Re}\left(\boldsymbol{\varphi}(\boldsymbol{x}^{i'})^*\left(\sum_{i=1}^{N+1} \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m \right) \frac{\partial}{\partial x_j^{i'}} \boldsymbol{\varphi}(\boldsymbol{x}^{i'}) \right)$$

and

$$\frac{\partial f}{\partial \alpha_{i'}} = 2\,\mathrm{Re}\left(\boldsymbol{\varphi}(\boldsymbol{x}^{i'})^*\left(\sum_{i=1}^{N+1} \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m \right) \boldsymbol{\varphi}(\boldsymbol{x}^{i'}) \right).$$

In our experiments, we use trigonometric polynomials $\varphi_{\boldsymbol{k}}(\boldsymbol{x}) = \exp(2\pi\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{x} \rangle)$ on $\mathbb{T}^d = [0,1)^d$ for different dimensions d and frequency index sets $I \subset \mathbb{Z}^d$, $|I| = m$. The initial draw of points is made randomly and the weights are set to be equal. The optimization procedure is done several times for different initial draws to counteract a bad draw of points. According to Corollary 5.1, we have here $N = \dim_{\mathbb{C}}(\mathrm{span}_{\mathbb{C}}\{\varphi_{\boldsymbol{k}}\overline{\varphi}_{\boldsymbol{l}}\}_{\boldsymbol{k},\boldsymbol{l} \in I}) \leq m^2 - m + 1$ for the theoretical bound N for the number of needed points. The actual number of points n may be smaller.

Experiment 1. For the first experiments we use dimension $d = 2$ and three different frequency index sets: An ℓ_1-ball with $m = |I_1| = 41$ frequencies, a hyperbolic cross with $m = |I_2| = 33$ frequencies, i.e.,

$$I_1 = \left\{ \boldsymbol{k} \in \mathbb{Z}^2 : \|\boldsymbol{k}\|_1 \leq 4 \right\} \quad \text{and} \quad I_2 = \left\{ \boldsymbol{k} \in \mathbb{Z}^2 : \prod_{j=1}^{2}(|k_j| + 1) \leq 6 \right\}$$

and the third choice is based on the argumentations leading to Corollary 5.3 with $m = |I_3| = 10$, where we know that the minimal lattice size is 113 for exact reconstruction and our theoretical bound
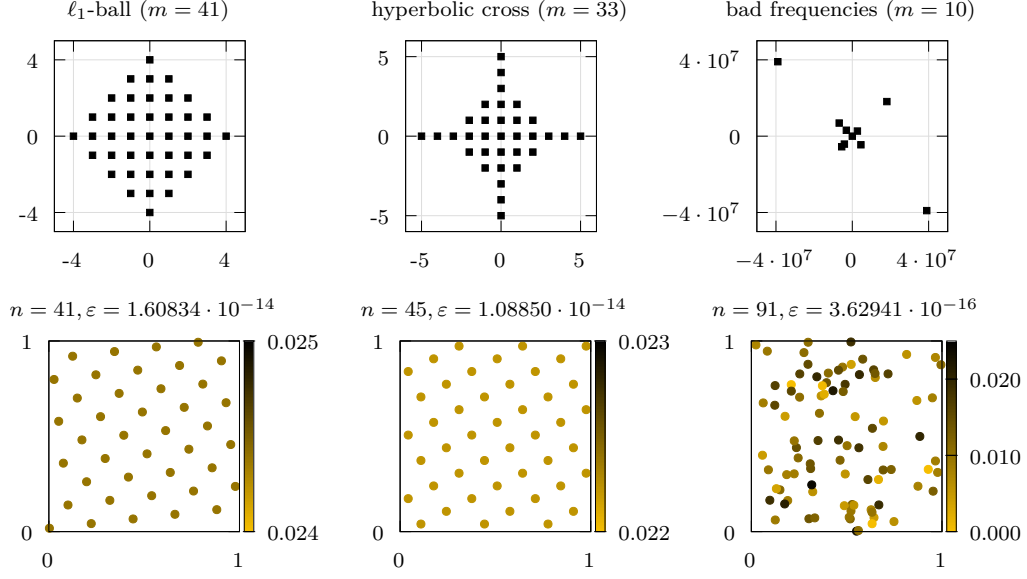
FIGURE 6.1. Experiment 1: points and weights forming an exact L_2-MZ inequality for frequencies in an ℓ_1-ball (left), a hyperbolic cross (middle), and 2-dimensional bad frequencies (right).

is $N = \dim_{\mathbb{C}}\{V \cdot \overline{V}\} = 91$:

$$
I_3 = \left\{
\begin{array}{l}
\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2\,671\,704 \\ 2\,671\,704 \end{pmatrix}, \begin{pmatrix} -3\,111\,990 \\ 3\,111\,990 \end{pmatrix}, \begin{pmatrix} -4\,145\,974 \\ -4\,145\,974 \end{pmatrix}, \begin{pmatrix} 4\,520\,742 \\ -4\,520\,742 \end{pmatrix}, \\[2mm]
\begin{pmatrix} -5\,553\,600 \\ -5\,553\,600 \end{pmatrix}, \begin{pmatrix} -6\,867\,835 \\ 6\,867\,835 \end{pmatrix}, \begin{pmatrix} 18\,119\,640 \\ 18\,119\,640 \end{pmatrix} \begin{pmatrix} 39\,011\,940 \\ -39\,011\,940 \end{pmatrix}, \begin{pmatrix} -39\,021\,892 \\ 39\,021\,892 \end{pmatrix}
\end{array}
\right\}.
$$

As for the number of points n, we seeked the smallest number such that we still have exact reconstruction up to a certain threshold, i.e., an L_2-MZ constant of $\|\sum_{i=1}^{n} \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m\|_{2\to 2} = \varepsilon < 10^{-13}$.

The outcome is depicted in Figure 6.1. For the ℓ_1-ball and the hyperbolic cross we obtain very regular sets of sampling nodes with almost equal weights, reminiscent of shifted rank-1 lattices. In particular, we achieve an oversampling factor of one for the ℓ_1-ball, i.e., $n = m = |I_1| = 41$, where also a reconstructing lattice is known. The third frequency set I_3 appears to be more challenging as we were not able to find points and weights for $n < 91$, which suggests the sharpness of our results. It is also an example of non-equal weights. We further tried several strategies to obtain an equal-weight exact L_2-MZ inequality by only optimizing the points but did not succeed with our algorithm.

Experiment 2. Now we want to explore a possible dimension-dependence of the needed number of points for an exact L_2-MZ inequality. We choose 20 random frequencies in $\{-100, \dots, 100\}^d$ for $d = 1, \dots, 7$ and empirically minimize $\|\sum_{i=1}^{n} \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m\|_F$ for different predetermined numbers n of points, ranging in $\{20, \dots, 300\}$. For each n, we started with 50 different random initial draws of n points and used the best obtained minimization as approximant for the corresponding L_2-MZ constant. The weights were set equal to $\alpha_i = 1/n$ as the observed outcome did not differ compared to optimizing the weights as well.
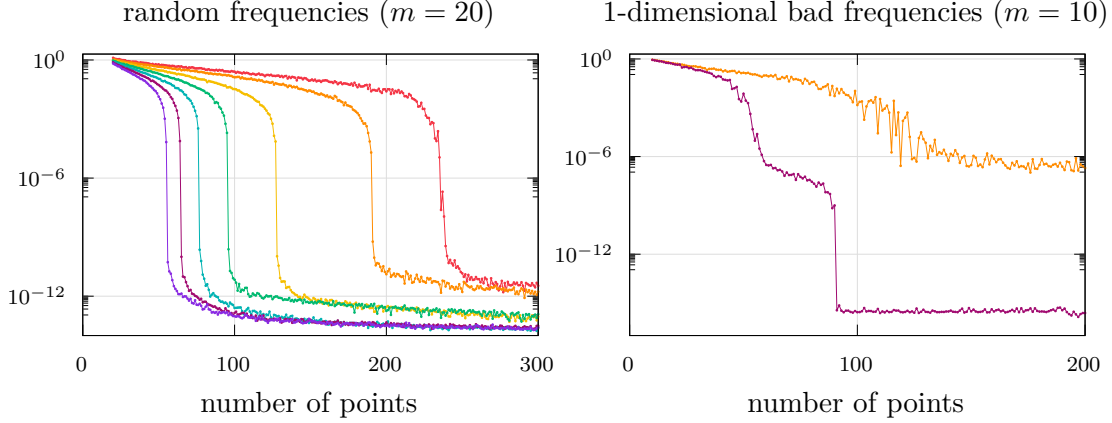
random frequencies ($m = 20$)　　　　1-dimensional bad frequencies ($m = 10$)

number of points　　　　　number of points

FIGURE 6.2. Depiction of the L_2-MZ constant $\varepsilon = \|\sum_{i=1}^n \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m\|_{2\to2}$ for: Experiment 2 in $d = 1$, $d = 2$, $d = 3$, $d = 4$, $d = 5$, $d = 6$, and $d = 7$ (left) and Experiment 3 for $d = 1$ with weights and without weights (right).

The results are depicted in the left of Figure 6.2. One can see that the threshold of finding a good L_2-MZ inequality is distinctly indicated by a sharp drop in the L_2-MZ constant $\varepsilon = \|\sum_{i=1}^n \alpha_i \boldsymbol{\varphi}(\boldsymbol{x}^i) \cdot \boldsymbol{\varphi}(\boldsymbol{x}^i)^* - \boldsymbol{I}_m\|_{2\to2}$. This happens around in accordance of our bound with $N = \dim_{\mathbb{C}}\{V \cdot \overline{V}\} \leq 381$. Note, that for lower dimension this number may be smaller because of the fewer possibilities of differences. In particular, we have $\dim_{\mathbb{C}}\{V \cdot \overline{V}\} \approx 220$ for $d = 1$, $\dim_{\mathbb{C}}\{V \cdot \overline{V}\} \approx 375$ for $d = 2$. Furthermore, the results suggest a "blessing of dimensionality" as a good L_2-MZ inequality can be found with less points for higher dimensions. In particular, the below table shows the first $n^\star(d)$ such that $\varepsilon < 10^{-10}$ with respect to the dimension d.

d	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$n^\star(d)$	241	192	129	97	78	66	56	49	44	39	36	33	31	29	27	25	24	23	20	20
$\lfloor 400/d \rfloor$	400	200	133	100	80	66	57	50	44	40	36	33	30	28	26	25	23	22	21	20

We added the values $\lfloor 400/d \rfloor$ which are close to $n^\star(d)$ suggesting an inverse linear dependence on the dimension. A similar effect for the sphere $\mathbb{S}^d$ is also discussed in [29], where it is argued that with higher dimension there are more degrees of freedom to be utilized.

Experiment 3. The third experiment aims to indicate the sharpness of our result. For that we choose $d = 1$ and $I = \{0, 107\,062, 124\,928, 1\,033\,760, 1\,414\,818, 2\,142\,995, 2\,820\,145, 4\,210\,229, 4\,645\,143, 5\,264\,579\}$. This is based on Corollary 5.3, where we know that the minimal lattice size is 103 for exact reconstruction and our theoretical bound is $N = \dim_{\mathbb{C}}\{V \cdot \overline{V}\} = 91$.

Analogous to the second experiment, we computed the L_2-MZ constant varying a fixed number of points $n \in \{20, \ldots, 200\}$. For each n, we started with $1\,000$ different random initial draws of n points and used the best result. We repeated this experiment with and without optimizing the weights.

The results are seen in the right of Figure 6.2. The first observation is the improvement gained from utilizing non-equal weights. It seems it makes the optimization procedure more stable as we know of the existence of an exact L_2-MZ inequality for $n = 103$ which the equal-weighted algorithm did not find. The weights in this case are non-equal. The second observation is the visible drop to finding an exact L_2-MZ inequality at $n = 91$, which is the same as our proposed bound for the needed number of points $\dim_{\mathbb{C}}\{V \cdot \overline{V}\}$. While this does not show the non-existence of possibly even equal-weighted exact L_2-MZ inequalities for $n \leq 90$, it is a numerical indication of the sharpness of our result.

7. Guaranteed and verifiable recovery in $L_2(\mu)$

Let us consider a recovery problem, which has been first addressed by Wasilkowski and Woźniakowski [59] in 2001 and drew a lot of attention in the past 6 years [3, 24, 39, 45]. In this section we will comment on the problem how to practically find stable recovery algorithms and suitable points which are guaranteed to recover any function from the unit ball in a RKHS $H(k)$ with a prescribed accuracy in $L_2(\mu)$. In certain special case like Besov–Sobolev-spaces with dominating mixed smoothness, sparse grid and Smolyak type algorithms serve as a practical near optimal way to recover such functions, see [25, Ch. 5]. Other approaches include rank-1-lattice algorithms, see, e.g., [34] or [40, §6], and versions thereof [2]. The question arises whether it is possible to mimic such constructive approaches in the general setting of a reproducing kernel Hilbert space where we have access to the most important eigenfunctions of the embedding operator into $L_2(\mu)$. Our attempt will be to construct an exact L_2-discretization on the subspace spanned by these eigenfunctions via D-optimal designs. This is supposed to mimic the construction of a reproducing rank-1-lattice on spaces of trigonometric polynomials for given frequency sets in $\mathbb{Z}^d$, like hyperbolic crosses, see [32, 33, 40].

Given a probability measure μ on a compact topological space Ω and a bounded, continuous Mercer kernel $k(\cdot, \cdot) : \Omega \times \Omega \to \mathbb{C}$ we will construct a direct recovery algorithm for this specific situation using D-optimal designs. We start with the Mercer decomposition of the kernel $k(\cdot, \cdot)$ which is

$$k(x, y) = \sum_{j=1}^{\infty} \sigma_j^2 \psi_j(x)\overline{\psi_j(y)}. \tag{7.1}$$

Here the system $(\psi_j(\cdot))_j$ represents the system of orthonormal (w.r.t. measure μ) eigenfunctions of the corresponding integral operator with respect to the kernel $k(\cdot, \cdot)$ mapping from $L_2(\mu)$ to $L_2(\mu)$. The sequence $(\sigma_j^2)_{j \in \mathbb{N}}$ represents the corresponding non-increasing sequence of non-negative eigenvalues satisfying

$$\int_{\Omega} k(x, x)\, \mathrm{d}\mu(x) = \sum_{j=1}^{\infty} \sigma_j^2 < \infty\,.$$

Let us fix $n < \dim_{\mathbb{C}}(H(k))$ and assume that we have access to the first n (largest) singular values σ_j, the corresponding functions $\psi_j(\cdot)$ and the truncated trace

$$k_n(x, x) := \sum_{j=1}^{n} \sigma_j^2 |\psi_j(x)|^2\,.$$

We construct the sampling points and the linear sampling recovery operator $S_{n,N}^{k,\mu} : H(k) \to L_2(\mu)$ and start with the density function

$$\omega_n(x) := \frac{1}{2} + \frac{k(x, x) - k_n(x, x)}{2\int_{\Omega}(k(x, x) - k_n(x, x))\, \mathrm{d}\mu(x)} = \frac{1}{2} + \frac{\sum_{j=n+1}^{\infty} \sigma_j^2 |\psi_j(x)|^2}{2\sum_{j=n+1}^{\infty} \sigma_j^2}\,. \tag{7.2}$$

Clearly, $\int_{\Omega} \omega_n(x)\, \mathrm{d}\mu(x) = 1$. We continue applying Theorem 3.1 to the modified system

$$\varphi_j(\cdot) := \frac{\psi_j(\cdot)}{\sqrt{\omega_n(\cdot)}}\,, \qquad j = 1, \ldots, n\,.$$

In particular, we apply Algorithm 2 to the system $(\varphi_j(\cdot))_{j=1}^n$ and obtain points $(x_i)_{i=1}^N$ with $N \leq n^2 + 1$, and weights $(\lambda_i)_{i=1}^N$ summing up to 1 such that the system matrix $\boldsymbol{A} = \boldsymbol{D}_{\boldsymbol{\lambda}}^{1/2} \cdot \boldsymbol{D}_{\boldsymbol{\gamma}}^{-1/2} \cdot \boldsymbol{\Psi}$ for the linear

system

$$\begin{pmatrix} \sqrt{\frac{\lambda_1}{\omega_n(x^1)}}\psi_1(x^1) & \cdots & \sqrt{\frac{\lambda_1}{\omega_n(x^1)}}\psi_n(x^1) \\ \vdots & \ddots & \vdots \\ \sqrt{\frac{\lambda_N}{\omega_n(x^N)}}\psi_1(x^N) & \cdots & \sqrt{\frac{\lambda_N}{\omega_n(x^N)}}\psi_n(x^N) \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{\lambda_1}{\omega_n(x^1)}}f(x^1) \\ \vdots \\ \sqrt{\frac{\lambda_N}{\omega_n(x^N)}}f(x^N) \end{pmatrix}$$

has orthonormal columns. Introducing the matrices $\boldsymbol{\Psi} := (\psi_k(x^j))_{j,k}$, $\boldsymbol{D_\lambda} := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $\boldsymbol{D_\omega} := \mathrm{diag}(\omega_n(x^1), \ldots, \omega_n(x^N))$ we find

$$\boldsymbol{c} = (\boldsymbol{D}_\lambda^{1/2} \cdot \boldsymbol{D}_\omega^{-1/2} \cdot \boldsymbol{\Psi})^* \cdot \boldsymbol{D}_\lambda^{1/2} \cdot \boldsymbol{D}_\omega^{-1/2} \cdot \boldsymbol{f} = \boldsymbol{\Psi}^* \cdot \boldsymbol{D}_\lambda \cdot \boldsymbol{D}_\omega^{-1} \cdot \boldsymbol{f},$$

where $\boldsymbol{f} = (f(x^1), \ldots, f(x^N))^T = N(f)$, $\boldsymbol{c} = (c_1, \ldots, c_n)^T$. The operator $S_{n,N}^{k,\mu} : H(K) \to L_2(\mu)$ is finally given by

$$S_{n,N}^{k,\mu} := E_n \circ \boldsymbol{A}^* \circ \boldsymbol{D}_\lambda^{1/2} \circ \boldsymbol{D}_\omega^{-1/2} \circ N$$

with $N : f \mapsto (f(x^1), \ldots, f(x^N))^T$ and $E_n : \boldsymbol{c} \mapsto \sum_{j=1}^n c_j \psi_j$. This implies that the computation of $S_{n,N}^{k,\mu} f$ is possible in $O(n \cdot N)$ arithmetic operations. Note, that $N \le n^2 + 1$, see Theorem 3.1.

In [24] the authors improve on a result in [3] and [45] and give an asymptotically sharp result in terms of the number of samples n in cases where the singular values decay fast enough. It is shown that there is an algorithm A_n which uses $N = 43200 \cdot 866 \cdot n$ points such that for all $\|f\|_{H(k)} \le 1$

$$\|f - A_n(f)\|_{L_2(\mu)}^2 \le \frac{1}{n} \sum_{j \ge n} \sigma_j^2.$$

Despite its sharpness, the results in [24, 45] have two main drawbacks. Although the number of samples N scales linearly in the dimension n of the subspace, the oversampling constant may be huge. In addition, the point set is highly non-constructive and only its existence is proved. Also the non-optimal results in [3, 35, 39] propose points which are the result of a random draw that can not be "verified".

The following theorem is a first "verifiable" attempt which guarantees the proposed accuracy. We refine an approach by Gröchenig [31], where error estimates for function classes are obtained from given Marcinkiewicz–Zygmund families. Here we use the exact ones from Section 3 which lead to direct algorithms.

Theorem 7.1. *Let Ω be a compact topological measurable space with a Borel probability measure μ and $k(\cdot, \cdot) : \Omega \times \Omega \to \mathbb{C}$ a Mercer kernel. Then the sampling recovery algorithm $S_{n,N}^{k,\mu} : H(k) \to L_2(\mu)$ defined above yields the recovery bound*

$$\sup_{\|f\|_{H(k)} \le 1} \|f - S_{n,N}^{k,\mu} f\|_{L_2(\mu)}^2 \le 3 \sum_{j \ge n+1} \sigma_j^2 \le 3\, c_n(H(k), C(\Omega))^2, \tag{7.3}$$

where $c_n(H(k), C(\Omega))$ denotes the nth Gelfand number of the identity operator from $H(k)$ into $C(\Omega)$. In addition, $S_{n,N}^{k,\mu} f$ uses $N \le n^2 + 1$ many function samples and a value $S_{n,N}^{k,\mu} f(x)$ can be directly and accurately computed with less than $c \cdot n^3$ arithmetic operations, where $c > 0$ is an absolute constant.

Proof. Let the projection operator $A_n : H(k) \to L_2(\mu)$ be given as

$$A_n f := \sum_{j=1}^n \langle f, \sigma_j \psi_j \rangle_{H(k)} \sigma_j \psi_j(\cdot).$$

By a straight-forward computation we get for all $x \in \Omega$

$$|f(x) - A_n f(x)|^2 \le \|f\|_{H(k)}^2 \cdot \sum_{j=n+1}^\infty \sigma_j^2 |\psi_j(x)|^2 \tag{7.4}$$

but also $\|f - A_n f\|_{L_2(\mu)}^2 \leq \sigma_{n+1}^2 \|f\|_{H(k)}^2$ which will be used in the following estimation. Fix $f \in H(k)$ with $\|f\|_{H(k)} \leq 1$. By $L_2(\mu)$-orthogonality we have

$$
\begin{aligned}
\|f - S_{n,N}^{k,\mu} f\|_{L_2(\mu)}^2 &= \|f - A_n f\|_{L_2(\mu)}^2 + \|A_n f - S_{n,N}^{k,\mu} f\|_{L_2(\mu)}^2 \\
&= \sigma_{n+1}^2 + \|S_{n,N}^{k,\mu}(A_n f - f)\|_{L_2(\mu)}^2 \\
&= \sigma_{n+1}^2 + \|E_n \circ \boldsymbol{A}^* \circ \boldsymbol{D}_{\boldsymbol{\lambda}}^{1/2} \circ \boldsymbol{D}_{\boldsymbol{\omega}}^{-1/2} \circ N(A_n f - f)\|_{L_2(\mu)}^2 \\
&= \sigma_{n+1}^2 + \left\| \boldsymbol{A}^* \circ \boldsymbol{D}_{\boldsymbol{\lambda}}^{1/2} \circ \boldsymbol{D}_{\boldsymbol{\omega}}^{-1/2} \circ N(A_n f - f) \right\|_{\ell_2^n}^2 \qquad (7.5) \\
&\leq \sigma_{n+1}^2 + \left\| \boldsymbol{D}_{\boldsymbol{\lambda}}^{1/2} \circ \boldsymbol{D}_{\boldsymbol{\omega}}^{-1/2} \circ N(A_n f - f) \right\|_{\ell_2^N}^2 \\
&\leq \sigma_{n+1}^2 + \sum_{i=1}^N \lambda_i \frac{|f(x^i) - A_n f(x^i)|^2}{\omega_n(x^i)}
\end{aligned}
$$

In the last but one estimate we used that $\|\boldsymbol{A}^*\|_{2\to 2} = \|\boldsymbol{A}\|_{2\to 2} = 1$. Taking (7.2) and (7.4) into account yields

$$
\|f - S_{n,N}^{k,\mu} f\|_{L_2(\mu)}^2 \leq \sigma_{n+1}^2 + 2 \sum_{j=n+1}^\infty \sigma_j^2 \sum_{i=1}^N \lambda_i \leq 3 \sum_{j=n+1}^\infty \sigma_j^2.
$$

The second inequality in (7.3) follows from [16, Lem. 3.3] and the fact that Gelfand and approximation numbers coincide in our situation, see [48, Thm. 4.8]. ∎

Remark 7.2 (Point cloud constructions)**.** Recent approaches for obtaining point clouds involve an initial random draw, see [3, 24, 39, 45]. Our approach characterizes the point set via a maximum search for determinants / Frobenius norms. In special cases, i.e., rank-1-lattices, one may prove better results [2, 10]. In addition, so-called sparse grids yield another constructive approach for suitable point clouds which work particularly well in the special case of mixed smoothness RKHS, see [25] and [53]. We may further BSS-subsample [3] the points obtained in the optimization process to get a similar bound with fewer points like in [38]. The right-hand side in (7.3) keeps valid up to constants, whereas the number of point samples shrinks to $N = O(n)$. The subsampling is constructive.

Remark 7.3 (Direct and stable recovery algorithms)**.** Using points from an exact L_2-MZ inequality gives a direct and stable method to obtain the approximation, i.e., only a multiplication with the perfectly conditioned adjoint system matrix $\boldsymbol{A}$ is needed rather than a matrix inversion. With the naive matrix-vector multiplication, this yields a computational complexity of $\mathcal{O}(n \cdot N) = \mathcal{O}(n^3)$ when using n basis functions. In contrast to that, current subsampling techniques yield $N = O(n)$ many points with a near-optimal error behavior with respect to the number of points [2, 3]. These points do not fulfill an exact L_2-MZ inequality and one usually computes the approximation by (weighted) least squares algorithms. With r iterations, this gives a computational complexity of $\mathcal{O}(r \cdot n \cdot N) = \mathcal{O}(r \cdot n^2)$. For an exact solution $r = n$ iterations could be necessary which yields also a computational complexity of $\mathcal{O}(n^3)$. However, the error decay is exponential in r and often only 20 iterations are sufficient, cf. [30, Thm. 3.1.1], which would yield a computational complexity of $\mathcal{O}(n^2)$.

References

[1] Congpei An, Xiaojun Chen, Ian H. Sloan, and Robert S. Womersley. Well Conditioned Spherical Designs for Integration and Interpolation on the Two-Sphere. *SIAM J. Numer. Anal.*, 48(6):2135–2157, 2010.

[2] Felix Bartel, Lutz Kämmerer, Daniel Potts, and Tino Ullrich. On the reconstruction of functions from values at subsampled quadrature points. *Math. Comput.*, 93(346):785–809, 2024.

[3] Felix Bartel, Martin Schäfer, and Tino Ullrich. Constructive subsampling of finite frames with applications in optimal function recovery. *Appl. Comput. Harmon. Anal.*, 65:209–248, 2023.

[4] Joshua Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-Ramanujan sparsifiers. *SIAM J. Comput.*, 41(6):1704–1721, 2012.

[5] Thomas Bloom and Norman Levenberg. Transfinite diameter notions in $\mathbb{C}^N$ and integrals of Vandermonde determinants. *Ark. Mat.*, 48(1):17–40, 2010.

[6] Andriy Bondarenko, Danylo Radchenko, and Maryna Viazovska. Optimal asymptotic bounds for spherical designs. *Ann. Math.*, 178(2):443–452, 2013.

[7] Len Bos. On optimal designs for a d-cube. *Dolomites Res. Notes Approx.*, 15(4):20–34, 2022.

[8] Len Bos. On Fekete points for a real simplex. *Indag. Math.*, 34(2):274–293, 2023. special issue on the occasion of Jaap Korevaar's 100-th anniversary.

[9] Len Bos, Federico Piazzon, and Marco Vianello. Near G-optimal Tchakaloff designs. *Comput. Stat.*, 35(2):803–819, 2020.

[10] Glenn Byrenheid, Lutz Kämmerer, Tino Ullrich, and Toni Volkmer. Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness. *Numer. Math.*, 136(4):993–1034, 2017.

[11] Jameson Cahill and Xuemei Chen. A note on scalable frames. In *Conference proceedings of SampTA*, pages 93–96. Zenodo, 2013.

[12] Peter G. Casazza and Xuemei Chen. Frame scalings: A condition number approach. *Linear Algebra Appl.*, 523:152–168, 2017.

[13] Peter G. Casazza, Laura De Carli, and Tin T. Tran. Remarks on scalable frames. *Oper. Matrices*, 17(2):327–342, 2023.

[14] Peter G. Casazza, Matthew Fickus, and Dustin G. Mixon. Auto-tuning unit norm frames. *Appl. Comput. Harmon. Anal.*, 32(1):1–15, 2012.

[15] Peter G. Casazza and Jelena Kovačević. Equal-norm tight frames with erasures. *Adv. Comput. Math.*, 18(2/4):387–430, 2003.

[16] Fernando Cobos, Thomas Kühn, and Winfried Sickel. Optimal approximation of multivariate periodic Sobolev functions in the sup-norm. *J. Funct. Anal.*, 270(11):4196–4212, 2016.

[17] Ronald Cools. Constructing cubature formulae: the science behind the art. In *Acta numerica, 1997*, volume 6 of *Acta Numerica*, pages 1–54. Cambridge University Press, 1997.

[18] Feng Dai, Andriy Prymak, Vladimir Temlyakov, and Sergey Yu. Tikhonov. Integral norm discretization and related problems. *Russ. Math. Surv.*, 74(4):579–630, 2019.

[19] Feng Dai and Yuan Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer Monographs in Mathematics. Springer, 2013.

[20] Philippe Delsarte, Jean-Marie Goethals, and Johan J. Seidel. Spherical codes and designs. *Geom. Dedicata*, 6(3):363–388, 1977.

[21] Michał Dereziński, Manfred K. Warmuth, and Daniel Hsu. Unbiased estimators for random design regression. *J. Mach. Learn. Res.*, 23: article no. 167 (46 pages), 2022.

[22] Holger Dette and William J. Studden. *The theory of canonical moments with applications in statistics, probability, and analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, 1997.

[23] Josef Dick, Frances Y. Kuo, and Ian H. Sloan. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer.*, 22:133–288, 2013.

[24] Matthieu Dolbeault, David Krieg, and Mario Ullrich. A sharp upper bound for sampling numbers in L_2. *Appl. Comput. Harmon. Anal.*, 63:113–134, 2023.

[25] Dinh Dũng, Vladimir Temlyakov, and Tino Ullrich. *Hyperbolic cross approximation.* Advanced Courses in Mathematics – CRM Barcelona. Birkhäuser/Springer, 2018. edited and with a foreword by Sergey Tikhonov.

[26] Frank Filbir, Ralf Hielscher, Thomas Jahn, and Tino Ullrich. Marcinkiewicz-Zygmund inequalities for scattered and random data on the q-sphere. *Appl. Comput. Harmon. Anal.*, 71: article no. 101651 (18 pages), 2024.

[27] Daniel Freeman and Dorsa Ghoreishi. Discretizing L_p norms and frame theory. *J. Math. Anal. Appl.*, 519(2): article no. 126846, 2023.

[28] Daniel Freeman and Darrin Speegle. The discretization problem for continuous frames. *Adv. Math.*, 345:784–813, 2019.

[29] Manuel Gräf and Daniel Potts. On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms. *Numer. Math.*, 119(4):699–724, 2011.

[30] Anne Greenbaum. *Iterative methods for solving linear systems*, volume 17 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics, 1997.

[31] Karlheinz Gröchenig. Sampling, Marcinkiewicz–Zygmund inequalities, approximation, and quadrature rules. *J. Approx. Theory*, 257: article no. 105455, 2020.

[32] Lutz Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along rank-1 lattices. *SIAM J. Numer. Anal.*, 51:2773–2796, 2013.

[33] Lutz Kämmerer. Reconstructing Multivariate Trigonometric Polynomials from Samples Along Rank-1 Lattices. In Gregory E. Fasshauer and Larry L. Schumaker, editors, *Approximation Theory XIV: San Antonio 2013*, Springer Proceedings in Mathematics & Statistics, pages 255–271. Springer, 2014.

[34] Lutz Kämmerer, Daniel Potts, and Toni Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on sampling along rank-1 lattice with generating vector of Korobov form. *J. Complexity*, 31(3):424–456, 2015.

[35] Lutz Kämmerer, Tino Ullrich, and Toni Volkmer. Worst-case recovery guarantees for least squares approximation using random samples. *Constr. Approx.*, 54(2):295–352, 2021.

[36] Jack Kiefer and Jacob Wolfowitz. The equivalence of two extremum problems. *Can. J. Math.*, 12:363–366, 1960.

[37] Egor D. Kosov. Remarks on sampling discretization of integral norms of functions. *Tr. Mat. Inst. Steklova*, 319:202–212, 2022.

[38] David Krieg, Kateryna Pozharska, Mario Ullrich, and Tino Ullrich. Sampling projections in the uniform norm. https://arxiv.org/abs/math/2401.02220, 2024.

[39] David Krieg and Mario Ullrich. Function values are enough for L_2-approximation. *Found. Comput. Math.*, 21(4):1141–1151, 2021.

[40] Frances Y. Kuo, Giovanni Migliorati, Fabio Nobile, and Dirk Nuyens. Function integration, reconstruction and approximation using rank-1 lattices. *Math. Comput.*, 90(330):1861–1897, 2021.

[41] Gitta Kutyniok, Kasso A. Okoudjou, and Friedrich Philipp. Scalable frames and convex geometry. In *Operator methods in wavelets, tilings, and frames. AMS special session on harmonic analysis of frames, wavelets, and tilings*, volume 626 of *Contemporary Mathematics*, pages 19–32. Springer, 2014.

[42] J. Marcinkiewicz and A. Zygmund. Sur les fonctions indépendantes. *Fundam. Math.*, 29(1):60–90, 1937.

[43] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. Math. (2)*, 182(1):327–350, 2015.

[44] Giovanni Migliorati and Fabio Nobile. Stable high-order randomized cubature formulae in arbitrary dimension. *J. Approx. Theory*, 275: article no. 105706, 2022.

[45] Nicolas Nagel, Martin Schäfer, and Tino Ullrich. A new upper bound for sampling numbers. *Found. Comput. Math.*, 22(2):445–468, 2022.

[46] Shahaf Nitzan, Alexander Olevskii, and Alexander Ulanovskii. Exponential frames on unbounded sets. *Proc. Am. Math. Soc.*, 144(1):109–118, 2016.

[47] Erich Novak. Quadrature and widths. *J. Approx. Theory*, 47(3):195–202, 1986.

[48] Erich Novak and Henryk Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society, 2008.

[49] Federico Piazzon, Alvise Sommariva, and Marco Vianello. Caratheodory-Tchakaloff subsampling. *Dolomites Res. Notes Approx.*, 10(1):5–14, 2017.

[50] Mihai Putinar. A note on Tchakaloff's theorem. *Proc. Am. Math. Soc.*, 125(8):2409–2414, 1997.

[51] Rolf Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 151 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, expanded edition, 2014.

[52] Harold S. Shapiro. *Topics in approximation theory*, volume 187 of *Lecture Notes in Mathematics*. Springer, 1971. with appendices by Jan Boman and Torbjörn Hedberg.

[53] Winfried Sickel and Tino Ullrich. The Smolyak algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness. *East J. Approx.*, 13(4):387–425, 2007.

[54] Ian H. Sloan and James N. Lyness. The Representation of Lattice Quadrature Rules as Multiple Sums. *Math. Comput.*, 52(185):81–94, 1989.

[55] Gábor Szegö. *Orthogonal Polynomials*. American Mathematical Society, 1975.

[56] Vladimir Tchakaloff. Formules de cubatures mécaniques à coefficients non négatifs. *Bull. Sci. Math. (2)*, 81:123–134, 1957.

[57] Vladimir Temlyakov. The Marcinkiewicz-type discretization theorems for the hyperbolic cross polynomials. *Jaen J. Approx.*, 9(1):37–63, 2017.

[58] Joel A. Tropp. User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.*, 12(4):389–434, 2012.

[59] Grzegorz W. Wasilkowski and Henryk Woźniakowski. On the power of standard information for weighted approximation. *Found. Comput. Math.*, 1(4):417–434, 2001.

[60] Robert S. Womersley. *Efficient Spherical Designs with Good Geometric Properties*, pages 1243–1285. Springer, 2018.