



HAL
open science

Approximate D-optimal design and equilibrium measure

Didier Henrion, Jean Bernard Lasserre

► **To cite this version:**

Didier Henrion, Jean Bernard Lasserre. Approximate D-optimal design and equilibrium measure. 2025. hal-04688534v2

HAL Id: hal-04688534

<https://hal.science/hal-04688534v2>

Preprint submitted on 24 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximate D-optimal design and equilibrium measure*

Didier Henrion^{1,2}, Jean Bernard Lasserre^{2,3}

Draft of January 24, 2025

Abstract

We introduce a minor variant of the approximate D-optimal design of experiments with a more general information matrix that takes into account the representation of the design space S . The main motivation (and result) is that if $S \subset \mathbb{R}^d$ is the unit ball, the unit box or the canonical simplex, then remarkably, for every dimension d and every degree n , one obtains an optimal solution in closed form, namely the equilibrium measure of S (in pluripotential theory). Equivalently, for each degree n , the unique optimal solution is the vector of moments (up to degree $2n$) of the equilibrium measure of S . Hence finding an optimal design reduces to finding a cubature for the equilibrium measure, with atoms in S , positive weights, and exact up to degree $2n$. In addition, any resulting sequence of atomic D-optimal measures converges to the equilibrium measure of S for the weak-star topology, as n increases. Links with Fekete sets of points are also discussed. More general compact basic semi-algebraic sets are also considered, and a previously developed two-step design algorithm is easily adapted to this new variant of D-optimal design problem.

1 Introduction

In this paper we consider the approximate D-optimal design problem

$$\max_{\phi \in \mathcal{P}(S)} \log \det \mathbf{M}_n(\phi), \quad (1)$$

¹CNRS; LAAS; Université de Toulouse, France.

²Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia.

³Toulouse School of Economics (TSE), Toulouse, France.

*The second author was supported by the AI Interdisciplinary Institute ANITI funding through the french program "Investing for the Future PISA" under the grant agreement number ANR-19-PI3A-0004. This research is also part of the programme DesCartes and is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

where $S \subset \mathbb{R}^d$ (the design space) is compact, $\mathcal{P}(S)$ is the set of probability measures on S , and with $\mathbf{v}_n(\mathbf{x}) := (\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}_n^d$,

$$\mathbf{M}_n(\phi) := \int_S \mathbf{v}_n(\mathbf{x})\mathbf{v}_n(\mathbf{x})^T d\phi(\mathbf{x})$$

denotes the degree- n moment matrix of ϕ .

The *approximate* qualification is because one considers a (compact) set S and a probability measure ϕ on S rather than a discrete set of points and weights. In their seminal paper on the equivalence theorem, Kiefer & Wolfowitz [22] show that ϕ^* maximizes $\log \det \mathbf{M}_n(\phi)$ over $\mathcal{P}(S)$ if and only if ϕ^* minimizes $\max_{\mathbf{x} \in S} \mathbf{v}_n(\mathbf{x})\mathbf{M}_n(\phi)^{-1}\mathbf{v}_n(\mathbf{x}) =: K_n^\phi(\mathbf{x}, \mathbf{x})$ over \mathcal{C} , attaining the value $s_n := \binom{n+d}{d}$. The polynomial $K_n^\phi(\mathbf{x}, \mathbf{x})$ is a sum of squares and is called the degree- $2n$ Christoffel polynomial (the reciprocal of the Christoffel function) associated with ϕ . Then

$$\max_{\phi \in \mathcal{C}_n} \log \det \mathbf{M}_n(\phi) = \min_{\phi \in \mathcal{C}_n} \max_{\mathbf{x} \in S} K_n^\phi(\mathbf{x}, \mathbf{x}) = \max_{\mathbf{x} \in S} K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) = s_n, \quad (2)$$

and the max is attained at all points of the support of ϕ^* . Moreover, if the support of ϕ^* consists of exactly s_n distinct points, then they all have same weight $1/s_n$ and they are called *Fekete* points of S . However, this latter property happens in rather exceptional cases; see e.g. [4, 5, 6].

In this paper, we restrict ourselves to three specific sets S , namely the Euclidean unit ball, the unit box, and the canonical simplex. We introduce the following slight variant of (1):

$$\max_{\phi \in \mathcal{P}(S)} \log \det \mathbf{M}_n(\phi) + \sum_{g \in G} \log \det \mathbf{M}_{n-d_g}(g\phi) \quad (3)$$

where $g\phi$ is the measure on S which is absolutely continuous with respect to ϕ with density g , $d_g = \lceil \deg(g)/2 \rceil$, and $G \subset \mathbb{R}[\mathbf{x}]$ is an appropriate set of generators of S , i.e. $S = \{\mathbf{x} : g(\mathbf{x}) \geq 0 \forall g \in G\}$. For instance if S is the Euclidean unit ball then G is the singleton $\mathbf{x} \mapsto g(\mathbf{x}) := 1 - \|\mathbf{x}\|^2$, and (3) reads:

$$\max_{\phi \in \mathcal{P}(S)} \log \det \mathbf{M}_n(\phi) + \log \det \mathbf{M}_{n-1}(g\phi).$$

That is, one has replaced the information matrix $\mathbf{M}_n(\phi)$ with the new information matrix

$$\begin{bmatrix} \mathbf{M}_n(\phi) & 0 \\ 0 & \mathbf{M}_{n-1}(g\phi) \end{bmatrix}.$$

Notice that the boundary ∂S of S is now clearly involved via the polynomial g in the measure $g\phi$.

In the univariate case $S = [-1, 1]$ and $x \mapsto g(x) := 1 - x^2$, our framework shares some similarities with the generalization of D-optimal design proposed in [13, 21, 32]. In [13, 32], with a so-called *efficiency* function of the form $\lambda(x) := (1+x)^u(1-x)^v$ with $u, v \in \{0, 1\}$, $\beta_{n-1} = n/(2n+1)$, $\beta_n = n+1/(2n+1)$, one minimizes the criterion

$$\frac{\beta_{n-1} \log \det \mathbf{M}_{n-1}(\lambda\phi)}{n} + \frac{\beta_n \log \det \mathbf{M}_n(\lambda\phi)}{n+1},$$

which differs from (3). That is, in [13, 32] the authors consider the product of information matrices related to $n + 1$ different regressions models and with the *same* efficiency function λ , whereas in (3) one also considers several regression models but each with a specific efficiency function. In fact the matrix $\mathbf{M}_{n-1}(g\phi)$ is associated with a regression model of the form

$$y = v_n(\mathbf{x})^T \theta + \frac{\varepsilon}{\sqrt{g(\mathbf{x})}},$$

i.e., when the noise explodes close to the boundary of S^1 .

Contribution We show that for certain specific domains S , for every degree n and every dimension d , the *equilibrium measure* ϕ^* of S (in pluripotential theory [24]) is an optimal solution of the variant (3) of the D-optimal design problem (1). This variant (3) takes into account explicitly the boundary of S and its criterion tends to favor points in the interior of S (as points in ∂S do not contribute to some of the information matrices $\mathbf{M}_{n-d_g}(g\phi)$). The resulting optimality conditions link the variant (3) with a generalized (polynomial) Pell's equation investigated in [27, 31]. Hence for such design spaces S :

The variant (3) of the D-optimal design problem (1) reduces to finding a cubature for the equilibrium measure ϕ^ of S , with positive weights, atoms in S , and exact up to degree $2n$.*

Appropriate techniques from numerical analysis to compute cubatures [11, 12, 15] could then be combined with the arsenal of existing techniques. Alternatively, the two-step design algorithm provided in [9] to solve (1), reduces only to step-2 as the output of step 1 (moments of the atomic measure supported on an optimal design) is now readily available.

In addition, if ν_n^* is any atomic probability measure, optimal solution of (3) (hence whose support is a degree- n optimal design), then the resulting sequence $(\nu_n^*)_{n \in \mathbb{N}}$ converges weak-star to the equilibrium measure ϕ^* of S , as does any sequence of probability measures equi-supported on so-called *Fekete* points of S . However, as already noted in the literature linking approximation theory with the D-optimal design problem, probability measures supported on Fekete points are rarely optimal solutions to the standard D-optimal design problem (1); see e.g. [4, 5, 6].

Last but not least, for the univariate case $S = [-1, 1]$, the atomic probability measure ν_n^* supported on the $n + 1$ roots of the degree- $(n + 1)$ Chebyshev polynomial of first kind (and with equal weights) is optimal for (3). Hence the support of the celebrated Gauss-Chebyshev quadrature is an optimal solution of the variant (3) of (1). We emphasize that ν_n^* *cannot* be an optimal solution of the standard D-optimal design problem (1). A tensorized version of this result holds for the unit box $[-1, 1]^d$. Interestingly, a design on Chebyshev points of $[-1, 1]$ has been shown to be \mathbf{c} -optimal² in [17]; see also [39] for optimal design on Chebyshev points.

Finally, we extend the variant (3) to the case of arbitrary compact basic semi-algebraic sets $S \subset \mathbb{R}^d$. Of course for such general sets S , an optimal solution $\phi^* \in \mathcal{P}(S)$ of (3) is not available in closed form any more. However, links with the equilibrium measure are still

¹The authors thank Y. de Castro and F. Gamboa for providing such a simple interpretation

²Given a vector \mathbf{c} , a design is \mathbf{c} -optimal if it minimizes $\phi \mapsto \sup_{\mathbf{v}} \mathbf{c}^T \mathbf{v} / \mathbf{v}^T \mathbf{M}_n(\phi) \mathbf{v}$ see e.g. [39].

available asymptotically as the degree n increases. Moreover the two-step design algorithm provided in [9] to solve (1) can be easily adapted. In the convex relaxation defined in step 1 of the algorithm, it suffices to replace the log-det criterion of (1) with that of (3). Then step 2 of the algorithm remains exactly the same.

So in summary, for the three special geometries (unit ball, unit box, and simplex), and for every dimension d and every degree n , the equilibrium measure of S is an optimal solution of the variant (3) of D-optimal design. Moreover, for more general semi-algebraic sets S , this variant also provides (asymptotically) connections with the equilibrium measure of S . It still remains to investigate how this variant compares with the classical version (1) from a statistical point of view.

2 Notation, definitions and preliminaries

2.1 Notation and definitions

Let $\mathbb{R}[\mathbf{x}]_n \subset \mathbb{R}[\mathbf{x}]$ denote the space of polynomials in the variables $\mathbf{x} = (x_1, \dots, x_d)$, of total degree at most n . Let $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$ be the set of sums of squares (SOS) polynomials and let $\Sigma[\mathbf{x}]_n \subset \mathbb{R}[\mathbf{x}]_{2n}$ be its subset of SOS polynomials of total degree at most $2n$. We also denote by $\mathbf{1} \in \mathbb{R}[\mathbf{x}]$, the constant polynomial with value 1. Let \mathbb{N} be the set of natural numbers and $\mathbb{N}_n^d := \{\boldsymbol{\alpha} \in \mathbb{N}^d : |\boldsymbol{\alpha}| (= \sum_i \alpha_i) \leq n\}$, and $s_n := \binom{d+n}{d}$. A polynomial $p \in \mathbb{R}[\mathbf{x}]_n$ reads

$$\mathbf{x} \mapsto p(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_n^d} p_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}},$$

where $\mathbf{p} = (p_{\boldsymbol{\alpha}}) \in \mathbb{R}^{s_n}$ is the vector of coefficients of p in the monomial basis $\mathbf{v}_n(\mathbf{x}) := (\mathbf{x}^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}_n^d}$ with $\mathbf{x}^{\boldsymbol{\alpha}} := \prod_{i=1}^d x_i^{\alpha_i}$. With $S \subset \mathbb{R}^d$ compact, denote by $\mathcal{M}(S)_+$ the convex cone of Borel (positive) measures on S and $\mathcal{P}(S) \subset \mathcal{M}(S)_+$ its subset of probability measures on S . Denote also by $\mathcal{C}(S)$ the space of continuous functions on S .

Given a set of s_n points $\{\mathbf{x}_1, \dots, \mathbf{x}_{s_n}\} \subset S$ denote by $\text{VDM}(\mathbf{x}_1, \dots, \mathbf{x}_{s_n}) \in \mathbb{R}^{s_n \times s_n}$ the Vandermonde matrix associated with $\{\mathbf{x}_1, \dots, \mathbf{x}_{s_n}\}$ and the monomial basis.

Moment and localizing matrix With a given real sequence $\boldsymbol{\phi} = (\phi_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^d}$ (in bold), and $n \in \mathbb{N}$, is associated the Riesz linear functional $\phi \in \mathbb{R}[\mathbf{x}]_n^*$ (not in bold) defined by:

$$p (= \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}) \mapsto \phi(p) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_n^d} p_{\boldsymbol{\alpha}} \phi_{\boldsymbol{\alpha}}, \quad \forall p \in \mathbb{R}[\mathbf{x}]_n.$$

The moment matrix $\mathbf{M}_n(\boldsymbol{\phi})$ (or $\mathbf{M}_n(\phi)$) associated with $\boldsymbol{\phi}$ is the real symmetric matrix with rows and columns indexed by \mathbb{N}_n^d , and with entries

$$\mathbf{M}_n(\boldsymbol{\phi})(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \phi_{\boldsymbol{\alpha} + \boldsymbol{\beta}}, \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_n^d.$$

If $\boldsymbol{\phi}$ has a representing measure ϕ (i.e., $\phi_{\boldsymbol{\alpha}} = \int \mathbf{x}^{\boldsymbol{\alpha}} d\phi$ for all $\boldsymbol{\alpha} \in \mathbb{N}^d$) then $\mathbf{M}_n(\boldsymbol{\phi}) \succeq 0$ for every n . Conversely, $\mathbf{M}_n(\boldsymbol{\phi}) \succeq 0$ for every n , is a necessary but not sufficient condition for $\boldsymbol{\phi}$ to have a representing measure on \mathbb{R}^d .

Given a polynomial $g \in \mathbb{R}[\mathbf{x}]$, $\mathbf{x} \mapsto g(\mathbf{x}) := \sum_{\beta} g_{\beta} \mathbf{x}^{\beta}$, and a real sequence $\phi = (\phi_{\alpha})_{\alpha \in \mathbb{N}^d}$, the localizing matrix $\mathbf{M}_n(g\phi)$ associated with g and ϕ is the moment matrix associated with the Riesz linear functional $g\phi \in \mathbb{R}[\mathbf{x}]_n^*$ defined by:

$$p \mapsto (g\phi)(p) = \phi(gp), \quad \forall p \in \mathbb{R}[\mathbf{x}]_n; \quad (g\phi)_{\alpha} := \sum_{\beta \in \mathbb{N}^d} g_{\beta} \phi_{\beta+\alpha}, \quad \forall \alpha \in \mathbb{N}_n^d.$$

If $g = \mathbf{1}$, the localizing matrix $\mathbf{M}_n(g\phi)$ is simply the moment matrix associated with ϕ , and if ϕ has a representing measure supported on the set $\{\mathbf{x} : g(\mathbf{x}) \geq 0\}$, then $\mathbf{M}_n(g\phi) \succeq 0$ for all n .

Christoffel-Darboux kernel and Christoffel function Given a compact set $S \subset \mathbb{R}^d$ and a Borel measure ϕ on S such that $\mathbf{M}_n(\phi) \succ 0$ for all n , let $(P_{\alpha})_{\alpha \in \mathbb{N}^d} \subset \mathbb{R}[\mathbf{x}]$ be a family of polynomials that are orthonormal w.r.t. ϕ , i.e.:

$$\int_S P_{\alpha} P_{\beta} d\phi = \delta_{\alpha=\beta}, \quad \forall \alpha, \beta \in \mathbb{N}^d,$$

which is guaranteed to exist. Then for every n , the kernel

$$(\mathbf{x}, \mathbf{y}) \mapsto K_n^{\phi}(\mathbf{x}, \mathbf{y}) := \sum_{\alpha \in \mathbb{N}_n^d} P_{\alpha}(\mathbf{x}) P_{\alpha}(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d,$$

is called the Christoffel-Darboux (CD) kernel, the polynomial $\mathbf{x} \mapsto K_n^{\phi}(\mathbf{x}, \mathbf{x})/s_n$ is called the (normalized) CD polynomial, and the rational function

$$\mathbf{x} \mapsto \Lambda_n^{\phi}(\mathbf{x}) = 1/K_n^{\phi}(\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

is called the Christoffel function. Alternatively

$$\Lambda_n^{\phi}(\mathbf{x}) = (\mathbf{v}_n(\mathbf{x})^T \mathbf{M}_n(\phi)^{-1} \mathbf{v}_n(\mathbf{x}))^{-1} \tag{4}$$

$$= \min_{p \in \mathbb{R}[\mathbf{x}]_n} \left\{ \int_S p^2 d\phi : p(\mathbf{x}) = 1 \right\}, \quad \forall \mathbf{x} \in \mathbb{R}^d. \tag{5}$$

Relation (4) is called the ABC theorem in [38] and for (5) see e.g. [29, Theorem 3.1].

Equilibrium measure The notion of equilibrium measure associated with a given set originates from logarithmic potential theory (working with a compact set $E \subset \mathbb{C}$ in the univariate case). It minimizes the energy functional

$$I(\phi) := \int \int \log \frac{1}{|z-t|} d\phi(z) d\phi(t), \tag{6}$$

over all Borel probability measures ϕ supported on E . For instance if E is the interval $[-1, 1] \subset \mathbb{C}$ then the arcsine (or Chebyshev) distribution $\mu = dx/\pi\sqrt{1-x^2}$ is an optimal solution. Generalizations have been obtained in the multivariate case via pluripotential theory in \mathbb{C}^d . In particular, if $E \subset \mathbb{R}^d \subset \mathbb{C}^d$ is compact then its equilibrium measure (let us denote it by μ) is equivalent to the Lebesgue measure on compact subsets of $\text{int}(E)$. It has an

even explicit expression if E is convex and symmetric about the origin; see e.g. [2, Theorems 1.1 and 1.2]. Several examples of sets E with its equilibrium measure given in explicit form can be found in [1]. Importantly, the appropriate approach to define the (intrinsic) equilibrium measure μ of a compact subset of \mathbb{R}^d with $d > 1$, is to consider \mathbb{R}^d as a subset of \mathbb{C}^d and invoke pluripotential theory with its tools from complex analysis (in particular, plurisubharmonic functions (and their regularizations) and the Monge-Ampère operator). For more details on equilibrium measures and pluripotential theory, the interested reader is referred to [1, 2, 24], the discussion in [23, Section 6.8, p. 297] as well as [36, Appendix B], [33], and the references therein. In the sequel, when we speak about the equilibrium measure of a compact subset $E \subset \mathbb{R}^d$, we refer to that in pluripotential theory (i.e., with E considered as a subset of \mathbb{C}^d).

2.2 Background on approximate D-optimal design

The approximate D-optimal design problem is well-known and originates in statistics. Let $S \subset \mathbb{R}^d$ be a compact set with nonempty interior, and let $n \in \mathbb{N}$ be fixed. For a design $\xi = (\xi_1, \dots, \xi_r)$ with positive weights $\gamma = (\gamma_i)_{i \leq r}$ (γ_i is the frequency at which ξ_i is chosen) the matrix $\mathbf{M}_n(\xi) = \sum_{i=1}^r \gamma_i \mathbf{v}_n(\xi_i) \mathbf{v}_n(\xi_i)^T$ is called the information matrix of the design ξ . Among several statistical criteria in parameter estimation, maximizing (over all such atomic designs) the logarithm of the determinant of the information matrix is a popular one. Optimizing the same criterion over *all* probability measures on S yields problem (1), and as one optimizes over $\mathcal{P}(S)$, an optimal solution is called an *approximate* D-optimal design. For a general overview on optimal experimental design the interested reader is referred to the seminal papers [22] and the recent tutorial [19].

If S has nonempty interior then it is shown in [10, Theorem 1] that (1) has an optimal (not necessarily unique) atomic measure supported on m points $\mathbf{x}_j \in S$, $j = 1, \dots, m$, where $s_n \leq m \leq s_{2n}$. That is, there exists a weight vector $0 < \gamma \in \mathbb{R}^m$ such that

$$\phi^* = \sum_{j=1}^m \gamma_j \delta_{\mathbf{x}_j}, \quad \sum_{j=1}^m \gamma_j = 1.$$

Next, with $K_n^{\phi^*}(\mathbf{x}, \mathbf{y})$ being the CD kernel associated with ϕ^* ,

$$s_n - K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in S \tag{7}$$

$$s_n - K_n^{\phi^*}(\mathbf{x}_j, \mathbf{x}_j) = 0, \quad \forall j = 1, \dots, m. \tag{8}$$

See e.g. the equivalence theorem [22, Theorem, p. 364] and [4].

While an optimal atomic measure is not necessarily unique, the resulting optimal moment matrices \mathbf{M}_n are all identical (by strict concavity of the criterion). It turns out that from (7)-(8) one may also link (1) with approximation theory, orthogonal polynomials, Fekete points and Fejér points.

A set of s_n points $\{\mathbf{x}_1^*, \dots, \mathbf{x}_{s_n}^*\} \subset S$ is a Fekete set if it maximizes $\det(\text{VDM}(\mathbf{x}_1, \dots, \mathbf{x}_{s_n}))$ (the Vandermonde determinant) among all sets of s_n points of S . Observe that if μ is the s_n -atomic measure with equal weights $\frac{1}{s_n} \sum_{i=1}^{s_n} \delta_{\mathbf{x}_i}$, then one obtains

$$\mathbf{M}_n(\mu) = \frac{1}{s_n} \text{VDM}(\mathbf{x}_1, \dots, \mathbf{x}_{s_n}) \text{VDM}(\mathbf{x}_1, \dots, \mathbf{x}_{s_n})^T,$$

and $\log \det(\mathbf{M}_n(\mu)) = 2 \log \det(\text{VDM}(\mathbf{x}_1, \dots, \mathbf{x}_{s_n})) - s_n \log s_n$.

Fekete points are in turn connected with the so-called *equilibrium measure* in pluri-potential theory [24]. For instance the sequence of discrete probability measures $(\nu_n)_{n \in \mathbb{N}}$ equi-supported on s_n Fekete points converges to the equilibrium measure of S for the weak-star topology on the space of signed measures on S . See e.g. [3], [30, Theorem 4.5.1], and the references therein.

A set of s_n points $\{\mathbf{x}_1, \dots, \mathbf{x}_{s_n}\} \subset S$ is a Fejér set if $\max_{\mathbf{x} \in S} \sum_{j=1}^{s_n} |\ell_j(\mathbf{x})|^2 = 1$, where $\ell_1, \dots, \ell_{s_n}$ are the Lagrange interpolation polynomials associated with the points $\mathbf{x}_1, \dots, \mathbf{x}_{s_n}$ and the space $\mathbb{R}[\mathbf{x}]_{s_n}$. It turns out that a set of Fejér points is also a set of Fekete points, whereas the reverse is not true in general [5]. Hence it is natural to ask when a set of Fekete points is also the support of an atomic measure ϕ^* in (7)-(8), with equal weights. In fact, an s_n -atomic measure is D-optimal in (1) if and only if it is equally weighted and its support is a Fejér set. See for instance [5, 6]. However, and as noted in [6], such a situation is rather exceptional and not to be expected.

2.3 Convex relaxations

As stated, problem (1) is intractable because one does not know how to optimize over $\mathcal{P}(S)$ for arbitrary sets $S \subset \mathbb{R}^d$. A typical numerical approach is to discretize the design space S . However, if S is the compact basic semi-algebraic set $\{\mathbf{x} \in \mathbb{R}^d : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\}$ with $g_j \in \mathbb{R}[\mathbf{x}]$ for all j , then the authors in [9] have proposed a two-step numerical procedure with good results in a number of cases. In particular, this numerical procedure is mesh-free as it avoids discretizations of the design set S .

Letting $g_0 = \mathbf{1}$, associated with (1) is the following convex optimization problem:

$$\max_{\phi \in \mathbb{R}^{s_{2n}}} \{ \log \det \mathbf{M}_n(\phi) : \phi(\mathbf{1}) = 1; \mathbf{M}_{n-r_j}(g_j \phi) \succeq 0, j = 0, \dots, m \}, \quad (9)$$

where the maximization is over real sequences of pseudo-moments $\phi = (\phi_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$, that is, sequences which do not necessarily have a representing probability measure ϕ on S . For this reason (9) is a *relaxation* of (1). Then in the two-step algorithm proposed in [9] to solve (1):

– Step-1 solves the convex relaxation (9), a convex optimization problem which can be solved efficiently (at least for reasonable dimensions), and one expects to obtain an optimal sequence $\phi^* = (\phi_\alpha^*)_{\alpha \in \mathbb{N}_{2n}^d}$ having a representing measure on S .

– Step-2 of the procedure consists of finding a *flat extension* ψ of the vector $(\phi_\alpha^*)_{\alpha \in \mathbb{N}_{2n}^d}$, which is a vector of pseudo-moments $(\psi_\alpha)_{\alpha \in \mathbb{N}_{2t}^d}$ with $t > n$ such that $\psi_\alpha = \phi_\alpha^*$ for all $\alpha \in \mathbb{N}_{2n}^d$ and $\text{rank } \mathbf{M}_t(\psi) = \text{rank } \mathbf{M}_{t-s}(\psi) =: r$ for some positive integer s . Then from this flat extension, one can use the algorithm of [26, Section 6.1.2] to extract r points of S , which provide a D-optimal design (the support of a D-optimal atomic measure $\nu_n \in \mathcal{P}(S)$). If ϕ^* has indeed a representing measure on S , then a flat extension can be obtained by solving a hierarchy of semidefinite programs for increasing values of $t = n + 1, n + 2, \dots$, by minimizing (with respect to the vector of pseudo-moments $(\psi_\alpha)_{\alpha \in \mathbb{N}_{2t}^d}$) the linear functional $\psi(p)$ for a random strictly positive polynomial p , subject to the constraints $\mathbf{M}_{t-r_j}(\psi) \succeq 0, \forall j = 0, \dots, m$, until the condition $\text{rank } \mathbf{M}_t(\psi) = \text{rank } \mathbf{M}_{t-s}(\psi)$ is satisfied for some positive integer s .

For more details and examples, the interested reader is referred to [10, Section 5].

3 Main result

As we next see, the main result of this paper is obtained from [27, 31] and it is detailed for each of the three cases considered (unit ball, unit box, canonical simplex). What we emphasize here is the significance of such results for the D-optimal design problem. Namely, by using results from [31], one shows that our proposed variant of the D-optimality criterion for optimal design, reveals quite strong links with the equilibrium measure ϕ^* of S . In this variant appears a term which involves the boundary ∂S of the design space S . As a result, for three important sets $S \subset \mathbb{R}^d$, and for all dimensions d and all degrees n , ϕ^* is an optimal solution of (3). Therefore an optimal design can be obtained from any cubature associated with ϕ^* , provided that it has positive weights, atoms in S , and is exact up to degree $2n$. Define :

– the Euclidean unit ball $S^O := \{ \mathbf{x} \in \mathbb{R}^d : 1 - \|\mathbf{x}\|^2 \geq 0 \}$, with associated equilibrium measure

$$\phi^O = d\mathbf{x}/(\pi^d \sqrt{1 - \|\mathbf{x}\|^2}),$$

– the unit box $S^\square := \{ \mathbf{x} \in \mathbb{R}^d : 1 - x_j^2 \geq 0, j = 1, \dots, d \}$, with associated equilibrium measure

$$\phi^\square = d\mathbf{x}/(\pi^d \sqrt{(1 - x_1^2) \cdots (1 - x_d^2)}),$$

– and the canonical simplex $S^\Delta := \{ \mathbf{x} \in \mathbb{R}_+^d : 1 - \sum_{j=1}^d x_j \geq 0 \}$ with associated equilibrium measure

$$\phi^\Delta = d\mathbf{x}/(\pi^d \sqrt{x_1 \cdots x_d \cdot (1 - \sum_{j=1}^d x_j)}).$$

For every integer n , define following subsets of $\mathbb{R}[\mathbf{x}]$:

1. $G_n^O := \{ \mathbf{x} \mapsto 1; \quad \mathbf{x} \mapsto 1 - \|\mathbf{x}\|^2 \}$.
2. $G_n^\square := \{ \mathbf{x} \mapsto \prod_{j=1}^d (1 - x_j^2)^{\varepsilon_j} : \varepsilon \in \{0, 1\}^d; |\varepsilon| \leq n \}$.
3. $G_n^\Delta := \{ \mathbf{x} \mapsto x_1^{\varepsilon_1} \cdots x_d^{\varepsilon_d} (1 - \sum_{j=1}^d x_j)^{\varepsilon_{d+1}} : \varepsilon \in \{0, 1\}^{d+1}; |\varepsilon| \in 2\mathbb{N}; |\varepsilon| \leq n/2 \}$,

associated respectively with S^O , S^\square , and S^Δ . Of course one also has:

$$S^O = \{ \mathbf{x} : g(\mathbf{x}) \geq 0, \forall g \in G_n^O \}; \quad S^\square = \{ \mathbf{x} : g(\mathbf{x}) \geq 0, \forall g \in G_n^\square \}; \quad S^\Delta = \{ \mathbf{x} : g(\mathbf{x}) \geq 0, \forall g \in G_n^\Delta \}$$

Next, fix $n \in \mathbb{N}$, $S^* := S^O, S^\square$, or S^Δ , and instead of (1), consider now the related optimization problem:

$$\rho_n = \max_{\phi \in \mathcal{P}(S^*)} \sum_{g \in G_n^*} \log \det \mathbf{M}_{n-d_g}(g \phi) \tag{10}$$

or equivalently,

$$-\rho_n = \min_{\phi \in \mathcal{P}(S^*)} \sum_{g \in G_n^*} \log \det \mathbf{M}_{n-d_g}(g \phi)^{-1}, \tag{11}$$

where for every $g \in \mathbb{R}[\mathbf{x}]$, $d_g := \lceil \deg(g)/2 \rceil$.

When comparing (10) with (1) observe that we have simply replaced the information matrix $\mathbf{M}_n(\phi)$ with the new block-diagonal information matrix with $|G_n^\star|$ blocks, and where each diagonal block is the localizing matrix $\mathbf{M}_{n-d_g}(g\phi)$, $g \in G_n^\star$. For instance if $\star = O$ (i.e., $S^\star (= S^O)$ is the Euclidean unit ball), the localizing matrix has two diagonal blocks and reads:

$$\begin{bmatrix} \mathbf{M}_n(\phi) & 0 \\ 0 & \mathbf{M}_{n-1}(g\phi) \end{bmatrix} \quad \text{with } g(\mathbf{x}) = 1 - \|\mathbf{x}\|^2, \quad n \in \mathbb{N}.$$

The latter information matrix takes into account the distance to the boundary $\partial S^O = \{\mathbf{x} : \|\mathbf{x}\| = 1\}$ of the design space S^O . Indeed, for a design $\xi = (\xi_1, \dots, \xi_r)$ with weights $\gamma = (\gamma_i)_{i \leq r}$, the information matrix $\mathbf{M}_{n-1}(g\nu_\xi)$ (with $\nu_\xi = \sum_i \gamma_i \delta_{\xi_i}$) reads

$$\mathbf{M}_{n-1}(g\nu_\xi) = \sum_{i=1}^r \gamma_i (1 - \|\xi_i\|^2) \mathbf{v}_{n-1}(\xi_i) \mathbf{v}_{n-1}(\xi_i)^T,$$

and so points $\xi_i \in \partial S^O$ do *not* contribute to $\mathbf{M}_{n-1}(g\nu_\xi)$. So the criterion in (10) put more weight on points in $\text{int}(S^\star)$ than on points in ∂S^\star .

Next, the analogue for (10) of the convex relaxation (9) for (1) reads:

$$\tau_n = \max_{\phi} \left\{ \sum_{g \in G_n^\star} \log \det \mathbf{M}_{n-d_g}(g\phi) : \phi(\mathbf{1}) = 1; \mathbf{M}_{n-d_g}(g\phi) \succeq 0, \forall g \in G_n^\star \right\}, \quad (12)$$

where the maximization is over vectors $\phi = (\phi_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$ of pseudo-moments. We also consider the convex optimization problem :

$$\tau_n^* = \max_{\mathbf{A}_g \succeq 0} \left\{ \sum_{g \in G_n^\star} \log \det \mathbf{A}_g : \sum_{g \in G_n^\star} s_{n-d_g} = \sum_{g \in G_n^\star} \mathbf{v}_{n-d_g}(\mathbf{x})^T \mathbf{A}_g \mathbf{v}_{n-d_g}(\mathbf{x}) g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d \right\}. \quad (13)$$

As proved in [27, Theorem 6, p. 945], (13) is a dual of (12) and also of (11), that is, weak duality $\tau_n^* \leq -\tau_n \leq -\rho_n$ holds, and in fact even strong duality holds³, i.e., $\tau_n^* = -\tau_n = -\rho_n$. So the convex relaxation (12) is in fact exact (i.e. has same optimal value as (11)).

Theorem 1 *Let $n \in \mathbb{N}$ be fixed, arbitrary, and $S^\star = S^O$, S^\square , or S^Δ (fixed). Then*

(i) *The equilibrium measure ϕ^\star of S^\star is an optimal solution of (10), the optimal moment matrices $\mathbf{M}_{n-d_g}(g\phi^\star)$, $g \in G_n^\star$, are unique, and the vector of moments $\phi^\star = (\phi_\alpha^\star)_{\alpha \in \mathbb{N}_{2n}^d}$ is the unique optimal solution of (12).*

(ii) *The unique optimal solution $(\mathbf{A}_g^\star)_{g \in G_n^\star}$ of problem (13) satisfies $\mathbf{A}_g^\star = \mathbf{M}_{n-d_g}(g\phi^\star)^{-1}$, and therefore*

$$\sum_{g \in G_n^\star} s_{n-g} = \sum_{g \in G_n^\star} \Lambda_{n-d_g}^{g\phi^\star}(\mathbf{x})^{-1} g(\mathbf{x}) = \sum_{g \in G_n^\star} K_{n-d_g}^{g\phi^\star}(\mathbf{x}, \mathbf{x}) g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d. \quad (14)$$

³In the present context, the condition $\mathbf{1} \in Q_n(G)$ in [27, Theorem 6] is satisfied as soon as the constraint in (13) is satisfied for some matrices $\mathbf{A}_g \succ 0$, for every $g \in G_n^\star$.

(iii) There are $s_n \leq m \leq s_{2n}$ distinct points $\mathbf{x}_j^* \in S^*$, and a positive vector $0 < \boldsymbol{\gamma} \in \mathbb{R}^m$ such that the atomic probability measure $\nu_n^* := \sum_{j=1}^m \gamma_j \delta_{\mathbf{x}_j^*}$ has the same moments as ϕ^* , up to degree $2n$. The normalized CD polynomial

$$\mathbf{x} \mapsto K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) \quad (15)$$

attains its maximum $\sum_{g \in G_n^*} s_{n-d_g}$ at all points of the boundary ∂S^* if S^* is the Euclidean unit ball, only at the 2^d vertices of ∂S^* , if S^* is the unit box, and only at the $d+1$ vertices of S^* if S^* is the canonical simplex.

(iv) With ν_n^* as in (iii), the sequence $(\nu_n^*)_{n \in \mathbb{N}}$ converges to ϕ^* for the weak-star topology of $\mathcal{M}(S^*)$, i.e.,

$$\lim_{n \rightarrow \infty} \int_S f d\nu_n^* = \int_S f d\phi^*, \quad \forall f \in \mathcal{C}(S^*).$$

Proof: For (i) and (ii) see [27] and [31, Theorem 4.1, 4.3, 4.5]. For the first statement in (iii) see e.g. [10, Theorem 1]. For the second statement in (iii), observe that by (14) one obtains

$$\sum_{g \in G_n^*} s_{n-d_g} - K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) = \sum_{1 \neq g \in G_n^*} K_{n-d_g}^{g\phi^*}(\mathbf{x}, \mathbf{x}) g(\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in S^*,$$

as $g \geq 0$ on S^* for all $g \in G_n^*$, and so $K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) \leq \sum_{g \in G_n^*} s_{n-d_g}$ for all $\mathbf{x} \in S^*$.

Next, recall that if $g = \mathbf{1}$ then $K_{n-d_g}^{g\phi^*}(\mathbf{x}, \mathbf{x}) = K_n^{\phi^*}(\mathbf{x}, \mathbf{x})$. Hence if $S^* = S^O$, then $K_n^{\phi^*}(\mathbf{x}, \mathbf{x}) = s_n + s_{n-1}$ for every $\mathbf{x} \in \partial S^O$ as $\|\mathbf{x}\| = 1$ on ∂S^O . On the other hand, if $S^* = S^\square$, observe that only at the 2^d points $\mathbf{u} = (\pm 1, \pm 1, \dots, \pm 1)$ $g(\mathbf{u}) = 0$ for every $g \in G_n^\square$ with $g \neq \mathbf{1}$. Similarly, if $S^* = S^\Delta$ then only at the vertices \mathbf{u} of S^* , $g(\mathbf{u}) = 0$ for every $g \in G_n^\Delta$ with $g \neq \mathbf{1}$. This yields the desired conclusion.

Finally for (iv), as S^* is compact, there exists a subsequence $(n_k)_{k \in \mathbb{N}}$ and a probability measure ψ on S such that $\lim_{k \rightarrow \infty} \int_S f d\nu_{n_k}^* = \int_{S^*} f d\psi$ for all $f \in \mathcal{C}(S^*)$. In particular, by definition of ν_n^* ,

$$\phi_\alpha^* = \lim_{k \rightarrow \infty} \int_{S^*} \mathbf{x}^\alpha d\nu_{n_k}^* = \int_{S^*} \mathbf{x}^\alpha d\psi, \quad \forall \alpha \in \mathbb{N}^d,$$

and so as S^* is compact, $\psi = \phi^*$. But this also implies that the whole sequence $(\nu_n^*)_{n \in \mathbb{N}}$ converges to ϕ^* . \square

Hence when S^ is the unit ball, the unit box, or the simplex, its equilibrium measure ϕ^* is an optimal solution of (10) for all degrees n . Therefore the support of any available cubature for ϕ^* , with atoms in S^* , positive weights, and exact up to degree $2n$, provides an optimal design.*

Concerning the construction of such atomic-measures in Theorem 1(iii), the interested reader is referred to [16] where several cubature formula are provided for the unit ball and unit sphere for various weight functions (including the Chebyshev weight $1/\sqrt{1-\|\mathbf{x}\|^2}$). For numerical computation of cubatures the interested reader is referred to e.g. [11, 12, 15].

Incidentally, by Theorem 1, with $s := \sum_{g \in G_n^*} s_{n-d_g}$, the polynomials $(g(\mathbf{x}) K_{n-d_g}^{g\phi^*}(\mathbf{x}, \mathbf{x})/s)_{g \in G_n^*}$ are nonnegative on S^* , and provide S^* with a partition of unity.

Example 1 Let $S^* = S^\circ$ (the Euclidean unit ball). In the case $d = 2$ (disk) and with $n = 2$ (degree 4), let us compare the D -optimal designs obtained with the approach in [10] briefly described in Section 2.3, with the approach described in this paper, i.e., solving variant problem (3). In solving (9) at Step-1 of the two-step moment-SOS algorithm [10] (with S

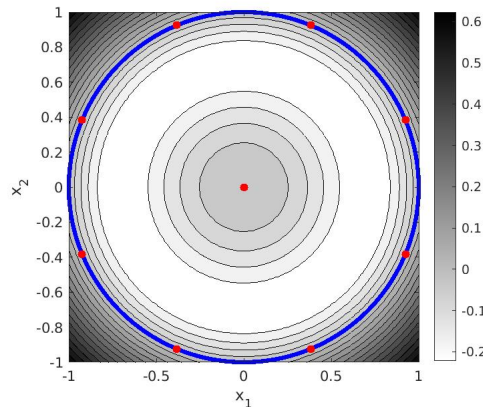


Figure 1: Degree 4 D -optimal design on the disk solving problem (1). The optimal points (red) are located on the CD polynomial unit level set (blue).

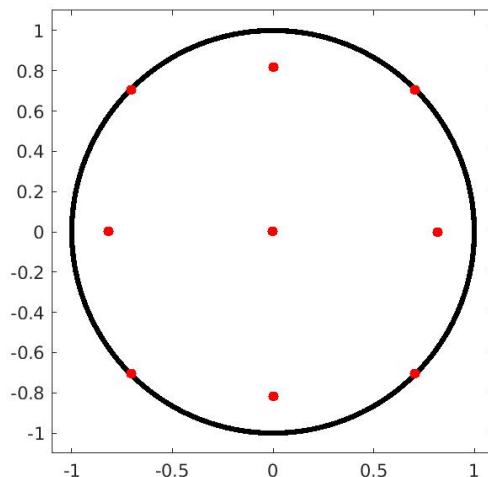


Figure 2: Degree 4 D -optimal design on the disk (black) with points (red) solving variant problem (3).

described as $\{\mathbf{x} : 1 - \|\mathbf{x}\|^2 \geq 0\}$ with one generator),, a moment vector⁴

$$\boldsymbol{\mu}_4^* := \int_{S^*} \mathbf{v}_4(\mathbf{x}) d\mu(\mathbf{x}) \in \mathbb{R}^{15}$$

⁴The unique solution $\boldsymbol{\mu}_4^*$ is not guaranteed to come from a measure $\mu^* \in \mathcal{P}(S^*)$, but in our numerical experiments in [10], it does.

is computed numerically. Its non-zero entries are $\mu_{00}^* = 1$, $\mu_{20}^* = \mu_{02}^* = 0.4167$, $\mu_{40}^* = \mu_{04}^* = 0.3125$, $\mu_{22}^* = 0.1042$ (to 4 significant digits). In step-2, an atomic measure supported at 9 points is computed at relaxation order 5, i.e. by extending the moment vector up to degree 10. The points are displayed in red on Figure 1. They are located on the unit level set (blue) of the CD polynomial. Other level sets (gray) are represented in logarithmic scale.

In solving variant problem (3), the moment vector $\phi_4^* \in \mathbb{R}^{15}$ of the equilibrium measure ϕ^* of S^* is computed, either in closed form or numerically as above in step-1 (but with the new log det criterion). Its non-zero entries read $\phi_{00}^* = 1$, $\phi_{20}^* = \phi_{02}^* = 1/3$, $\phi_{40}^* = \phi_{04}^* = 1/5$, $\phi_{22}^* = 1/15$. In step-2, an atomic measure supported on 9 points of S^* is computed at relaxation order 5, i.e. by extending the moment vector ϕ_4^* up to degree 10. The points are displayed in Figure 2. As expected, we observe that 5 out of 9 points are in the interior of the disk, whereas only a single point (the origin) is in interior of the disk for original problem (1).

Example 2 Let $S^* = S^\Delta$ (the simplex). With $d = 2$ and $2n = 4$, in solving the standard D -optimal design (1) via the two-step algorithm in [9] described in Section 2.3 (with S described as $\{\mathbf{x} : x_1 \geq 0, x_2 \geq 0, 1 - x_1 - x_2 \geq 0\}$ with 3 generators), one obtains 6 points (the three vertices and the mid-points on each facet). On the other hand in solving (3) one obtains a cubature for the equilibrium measure, supported on 8 points on the boundary and one point inside.

The univariate case $S^\square = [-1, 1]$

Let $\phi^* = dx/\pi\sqrt{1-x^2}$ and $x \mapsto g(x) = 1 - x^2$. Then (14) reads:

$$s_n + s_{n-1} = \Lambda_n^{\phi^*}(x)^{-1} + (1 - x^2) \Lambda_{n-1}^{g \cdot \phi^*}(x)^{-1}, \quad \forall x \in \mathbb{R}, \quad (16)$$

or, equivalently:

$$s_n + s_{n-1} = K_n^{\phi^*}(x, x) + (1 - x^2) K_{n-1}^{g \cdot \phi^*}(x, x), \quad \forall x \in \mathbb{R}. \quad (17)$$

Next, the Gauss-Chebyshev quadrature for the equilibrium measure ϕ^* is supported on the zeros of the degree- $(n+1)$ Chebyshev polynomial of first kind, and reads $\nu_n = \frac{1}{n+1} \sum_{i=1}^{n+1} \delta_{x_i}$ with

$$x_i = \cos\left(\frac{2i-1}{2(n+1)}\pi\right) \in (-1, 1), \quad i = 1, \dots, n+1, \quad (18)$$

and is exact up to degree $2n+2-1 = 2n+1$, i.e.,

$$\int_{-1}^1 p d\nu_n = \int_{-1}^1 p d\phi^*, \quad \forall p \in \mathbb{R}[x]_{2n+1}.$$

Moreover it turns out that

$$K_n^{\phi^*}(x_i, x_i) = \Lambda_n^{\phi^*}(x_i)^{-1} = \Lambda_n^{\nu_n}(x_i)^{-1} = n+1 = s_n,$$

and so by (16), $(1 - x_i^2) \Lambda_{n-1}^{g \cdot \phi^*}(x_i)^{-1} = s_{n-1}$, for all $i = 1, \dots, n+1$.

So the distinguished equidistributed atomic probability measure ν_n is an optimal solution of (10). Moreover

$$\frac{\text{VDM}(x_1, \dots, x_{n+1})\text{VDM}(x_1, \dots, x_{n+1})^T}{n+1} = \mathbf{M}_n(\nu_n).$$

However, its support (zeros of the degree- $(n+1)$ Chebyshev polynomial) is *not* a set of Fekete points because

$$\phi^* \neq \arg \max_{\mu \in \mathcal{P}(S)} \det(\mathbf{M}_n(\mu)),$$

(i.e., ϕ^* does not solve the D-optimal design⁵ (1)). Indeed $\max_{x \in S} K_n^{\phi^*}(x, x) > s_n$ because by (16)

$$K_n^{\phi^*}(x, x) (= s_n + s_{n-1}) > s_n \quad \text{for } x = \pm 1.$$

So the points $(x_i)_{i \leq n+1}$ have the remarkable property that the equidistributed measure ν_n (supported on $\text{int}(S)$) is an optimal solution of (10) for every n . And of course, again, $(\nu_n)_{n \in \mathbb{N}}$ converges weakly to the equilibrium measure ϕ^* as does any sequence of probability measures equi-distributed on Fekete points (see e.g. [3, 30]).

However, even for S being the unit ball or the unit sphere, cases (d, n) where a D-optimal measure $\mu \in \mathcal{P}(S)$ is equi-supported on Fekete points are exceptional; see e.g. Bos [5, p. 134]. The univariate atomic measure ν_n in (18) has an immediate extension to the multivariate $S = [-1, 1]^d$ case.

Corollary 1 *Let ν_n be the Gauss-Chebyshev atomic (univariate) measure in (18). Then the multivariate atomic product measure $\nu_n^* := \underbrace{\nu_n \otimes \dots \otimes \nu_n}_{d \text{ times}}$ on $S = [-1, 1]^d$ is an optimal solution of (10). Moreover ν_n^* converges to ϕ^* for the weak-star topology of $\mathcal{M}(S^*)$.*

Proof: As $\phi^* = \prod_{j=1}^d dx_j / (\pi \sqrt{1-x_j^2})$ is an optimal solution of (10), ν_n^* provides with a cubature for ϕ^* , exact up to maximum degree $2n+1$ in *each* variable, and hence also exact for total degree $2n$. The final statement on weak-star convergence is due to the product structure of ϕ^* and ν_n^* . \square

Example 3 *In the case $d = 1$ (interval) and $n = 8$ (degree 16) let us compare the D-optimal designs obtained with the approach in [9] described in Section 2.3 for solving original problem (1), and with the approach described in this paper for solving variant problem (3).*

In solving original problem (1), in Step-1 of the moment-SOS algorithm one obtains a CD polynomial of degree $2n = 16$ (black curve) whose local maxima (located on the unit level set) are the optimal 9 points (red dots), see Figure 3. These points are also optimal Fekete points and roots of $x \mapsto (1-x^2)p'_n(x)$ where p_n is the degree- n Legendre polynomial [6, 18]. Also represented (blue dots) are the roots of the degree 9 Chebyshev polynomial, which cannot be optimal Fekete points, as discussed above.

In solving variant problem (3), as $\phi^ = dx / (\pi \sqrt{1-x^2})$, the roots of the degree 9 Chebyshev polynomial provide the support of an optimal atomic probability measure. (Step-1 of*

⁵An approximate D-optimal design for (1) is equally supported on the zeros of $(1-x^2)P'(x)$ where P_n is the degree- n Legendre polynomial; see [18]. It is also a set of Fekete points [6].

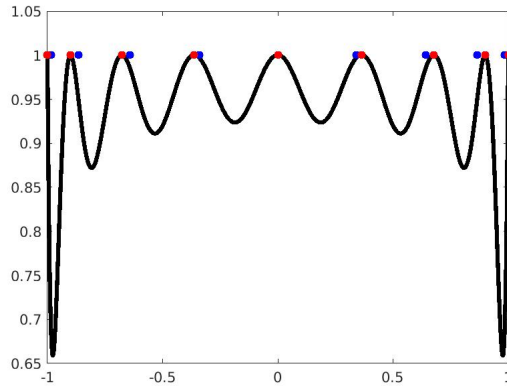


Figure 3: Degree 16 D-optimal design on unit interval: CD polynomial unit level set points (red) solve problem (1), whereas roots of the degree 9 Chebyshev polynomials (blue) solve variant problem (3).

our moment-SOS algorithm [10] also provides the moment vector ϕ_{18}^ of ϕ^* , rounded to 4 significant digits.)*

3.1 Comparison with the original formulation

An optimal atomic solution μ^* of (1) is such that the degree- $2n$ SOS normalized CD polynomial $\mathbf{x} \mapsto K_n^{\mu^*}(\mathbf{x}, \mathbf{x})/s_n$ satisfies

$$0 \leq K_n^{\mu^*}(\mathbf{x}, \mathbf{x})/s_n \leq 1, \quad \forall \mathbf{x} \in S,$$

and μ^* is supported on r points $\mathbf{z}_i \in S$, $i = 1, \dots, r$, with $s_n \leq r \leq s_{2n}$, and $K_n^{\mu^*}(\mathbf{z}_i, \mathbf{z}_i) = s_n$ for all $i = 1, \dots, r$. If $r = s_n$ then the points $\{z_1, \dots, z_r\}$ form a set of Fejér points and are also Fekete points. However this case is not to be expected in general and even in special cases as the ones considered in this paper; see e.g. [5, 6]. Notice that an optimal measure $\mu^* \in \mathcal{P}(S)$ of (1) is not provided explicitly and in view of Theorem 1(iii), μ^* cannot be the equilibrium measure ϕ^* of S . Only the optimal moment matrix $\mathbf{M}_n(\mu^*)$ is obtained and is unique (while μ^* is not unique).

On the other hand, if one considers the proposed variant (10) of the D-optimal design problem, then remarkably for S^* being the unit ball, the unit box, or the simplex:

- For every degree n and every dimension d , the associated equilibrium measure ϕ^* of S^* is an optimal solution.
- Every cubature of ϕ^* with positive weights, atoms in S^* , and exact up to degree $2n$, provides with a D-optimal atomic measure ν_n^* whose support is an optimal design.
- The sequence $(\nu_n^*)_{n \in \mathbb{N}}$ converges to the equilibrium measure ϕ^* of S^* for the weak-star topology of $\mathcal{M}(S^*)$.

3.2 Computing an optimal atomic measure

As we have seen, it reduces to that of searching for *any* degree- $2n$ cubature with positive weights and atoms in S^* , for the (known) equilibrium measure ϕ^* of S^* . For the unit box $[-1, 1]^n$, the (tensorized) Gauss-Chebyshev cubature already provides a D-optimal design with no computation; see Corollary 1. For the Euclidean ball and the simplex, some cubatures for ϕ^* are also known in some cases. For numerical procedures for computing cubatures, the interested reader is referred to [11, 12, 15]. An alternative is to adapt the two-step procedure described in [9] and sketched in Section 2.3, to solve the variant (10).

4 More general semi-algebraic sets

For the special geometries S^* of Section 3, we have seen that the infinite-dimensional problem (10) is also equivalent to solving the single finite-dimensional convex relaxation (12); see Theorem 1.

So while the equilibrium measure ϕ^* is not the unique optimal solution of (10) (any probability measure on \mathbb{R}^d with same degree- $2n$ moments as ϕ^* is also optimal), its vector ϕ_{2n}^* of moments up to degree $2n$, is the unique optimal solution of (12).

Basic semi-algebraic sets S Of course, for the special geometries S^* (unit ball, unit box or simplex), we know ϕ^* (and so ϕ_{2n}^* as well) and there is no need to solve (12). But for other geometries, e.g., if $S \subset \mathbb{R}^d$ is the compact basic semi-algebraic set

$$S = \{ \mathbf{x} \in \mathbb{R}^d : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m \}, \quad (19)$$

for some $g_j \in \mathbb{R}[\mathbf{x}]$, $j = 1, \dots, m$, then inspired by (10), we propose the following analogue of the D-optimal variant (3):

$$\max_{\phi \in \mathcal{P}(S)} \sum_{j=0}^m \log \det \mathbf{M}_{n-r_j}(g_j \cdot \phi), \quad (20)$$

where $g_0 = \mathbf{1}$, and $r_j = \lceil \deg(g_j)/2 \rceil$, $j = 0, \dots, m$. The set

$$Q_n(g_1, \dots, g_m) := \{ \mathbf{x} \mapsto \sum_{j=0}^m \sigma_j(\mathbf{x}) g_j(\mathbf{x}) : \sigma_j \in \Sigma[\mathbf{x}]; \deg(\sigma_j g_j) \leq 2n \} \quad (21)$$

is called the truncated quadratic module associated with the polynomials g_1, \dots, g_m , a central object in the Moment-SOS hierarchy for polynomial optimization; see e.g. [26, 28]. As S is compact, $S \subset \{ \mathbf{x} : M - \|\mathbf{x}\|^2 \geq 0 \}$ for some $M > 0$, and therefore, we may and will also assume that $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$ (if not we include this redundant quadratic constraint in the definition of S without changing S).

Theorem 2 Let $S \subset \mathbb{R}^d$ in (19) be compact with nonempty interior. Then every optimal solution $\mu \in \mathcal{P}(S)$ of (20) satisfies

$$\sum_{j=0}^m s_{n-r_j} - \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \mu}(\mathbf{x}, \mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in S \quad (22)$$

$$\sum_{j=0}^m s_{n-r_j} - \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \mu}(\mathbf{x}, \mathbf{x}) = 0, \quad \mu\text{-a.e.} \quad (23)$$

In addition, if $\mu = f d\mathbf{x}$ with $f > 0$ on S , is an optimal solution then

$$\sum_{j=0}^m s_{n-r_j} = \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \mu}(\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d. \quad (24)$$

Proof: Introduce the two (finite-dimensional) convex cones

$$\mathcal{C}_n := \left\{ \phi = (\phi_\alpha)_{\alpha \in \mathbb{N}_{2n}^d} : (\phi_\alpha) = \left(\int_S \mathbf{x}^\alpha d\phi \right) \text{ for some } \phi \in \mathcal{M}(S)_+ \right\}$$

$$\mathcal{P}_n := \{ p \in \mathbb{R}[\mathbf{x}]_{2n} : p \geq 0 \text{ on } S \},$$

and let \mathcal{C}_n^* and \mathcal{P}_n^* denote their respective duals. From the Riesz-Haviland Theorem (see e.g. [26, Theorem 2.34]) and since S is compact with nonempty interior, it turns out that $\mathcal{C}_n = \mathcal{P}_n^*$ and $\mathcal{C}_n^* = \mathcal{P}_n$. For every $j \leq m$, let the real symmetric matrices \mathbf{B}_α^j of size $n - r_j$ be defined by

$$\sum_{j=0}^m g_j(\mathbf{x}) \mathbf{v}_{n-r_j}(\mathbf{x}) \mathbf{v}_{n-r_j}(\mathbf{x})^T = \sum_{\alpha \in \mathbb{N}_{2n}^d} \mathbf{B}_\alpha^j \mathbf{x}^\alpha, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

Note that by construction it holds

$$\mathbf{M}_{n-r_j}(g_j \phi) = \sum_{\alpha \in \mathbb{N}_{2n}^d} \mathbf{B}_\alpha^j \phi_\alpha.$$

Then (20) also reads:

$$\max_{\phi \in \mathcal{C}_n} \left\{ \sum_{j=0}^m \log \det \mathbf{M}_{n-r_j}(g_j \phi) : \phi_0 = 1 \right\}, \quad (25)$$

which is also a convex program. Let μ^* be an optimal solution of (20) and let $\boldsymbol{\mu}_{2n}^* \in \mathcal{C}_n$ be its vector of moments up to degree $2n$ (hence an optimal solution of (25)). Recalling that \mathbf{X}^{-1} is the gradient of $\log \det \mathbf{X}$, the necessary KKT-optimality conditions imply that there exists some scalar λ^* and some polynomial $q^* \in \mathcal{P}_n$ such that

$$\lambda^* \mathbf{1}_{\alpha=0} - \sum_{j=0}^m \langle \mathbf{M}_{n-r_j}(g_j \boldsymbol{\mu}_{2n}^*)^{-1}, \mathbf{B}_\alpha^j \rangle = q_\alpha^*, \quad \forall \alpha \in \mathbb{N}_{2n}^d; \quad \langle \boldsymbol{\mu}_{2n}^*, q^* \rangle = 0 \quad [\text{complementarity}].$$

Multiplying by $(\mu_{2n}^*)_\alpha$ and summing up, yields $\lambda^* = \sum_{j=0}^m s_{n-r_j}$, and multiplying by \mathbf{x}^α and summing up yields the desired result (22)-(23). To get (24) observe that with $f > 0$ on S , and $q^* \in \mathcal{P}_n$

$$0 = \langle \boldsymbol{\mu}_{2n}^*, q^* \rangle = \int_S q^* d\mu^* = \int_S q^* f d\mathbf{x} \quad \Rightarrow \quad q^* = 0, \text{ a.e. on } S,$$

and since S has nonempty interior, $q^* = 0$ for all $\mathbf{x} \in \mathbb{R}^d$. \square Of course (20) (or, equivalently (25)) is not solvable in general because no tractable description of \mathcal{C}_n is available. Therefore we also introduce its convex relaxation (the analogue of (9))

$$\max_{\boldsymbol{\mu}=(\mu_\alpha)} \left\{ \sum_{j=0}^m \log \det \mathbf{M}_{n-r_j}(g_j \boldsymbol{\mu}) : \mu_0 = 1; \quad \mathbf{M}_{n-r_j}(g_j \boldsymbol{\mu}) \succeq 0, j = 0, \dots, m \right\}, \quad (26)$$

where $\boldsymbol{\mu} = (\mu_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$ is a vector of pseudo-moments up to degree $2n$.

In Step-1 of the algorithm described in [9, 10] and in Section 2.3 to solve (1), one solves almost the same problem as (26) except that the criterion is simply $\log \det \mathbf{M}_n(\boldsymbol{\mu})$. So it is straightforward to adapt its Step-1 for solving (26). Then Step-2 (whose input is the output $\boldsymbol{\mu}_{2n}^*$ of Step-1), remains exactly the same. Its goal is to extract the support of an atomic probability measure on S with same moments as $\boldsymbol{\mu}_{2n}^*$, that is, an optimal design. Equivalently, (26) reduces to

$$\rho_n = \min_{\boldsymbol{\mu}=(\mu_\alpha)} \left\{ - \sum_{j=0}^m \log \det \mathbf{M}_{n-r_j}(g_j \boldsymbol{\mu}) : \phi_0 = 1; \quad \mathbf{M}_{n-r_j}(g_j \boldsymbol{\mu}) \succeq 0, \forall j \leq m \right\}, \quad (27)$$

whose optimal value ρ_n is minus that of (26). Consider the convex optimization problem:

$$\rho_n^* = \max_{\mathbf{A}_j \succeq 0} \left\{ \sum_{j=0}^m \log \det \mathbf{A}_j : \sum_{j=0}^m s_{n-r_j} = \sum_{j=0}^m g_j(\mathbf{x}) \mathbf{v}_{n-r_j}(\mathbf{x})^T \mathbf{A}_j \mathbf{v}_{n-r_j}(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d \right\}. \quad (28)$$

As shown in [27] and [31, Theorem 3.2], if $\mathbf{1} \in \text{int}(Q_n(g_1, \dots, g_m))^6$ then (28) is a dual of (27), and strong duality holds, i.e., $\rho_n = \rho_n^*$. Moreover (27) (resp. (28)) has a unique optimal solution $\boldsymbol{\mu}_{2n}^*$ (resp. (\mathbf{A}_j^*)), and $\mathbf{A}_j^* = \mathbf{M}_{n-r_j}(g_j \cdot \boldsymbol{\mu}_{2n}^*)^{-1}$ for all $j = 0, \dots, m$. Therefore

$$\sum_{j=0}^m s_{n-r_j} = \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \boldsymbol{\mu}_{2n}^*}(\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d. \quad (29)$$

The difference with the three special cases of S in Section 3, is that now the linear functional $\boldsymbol{\mu}_{2n}^* \in \mathbb{R}[\mathbf{x}]_{2n}^*$ is not guaranteed to have a representing measure μ^* on S (Step-1 of the algorithm in [10] and its adaptation (26) proposed for (3) assumes that it is the case), let alone μ^* being the equilibrium measure of S . However as shown in [27, Theorem 10], each accumulation point of the sequence $(\boldsymbol{\mu}_{2n}^*)_{n \in \mathbb{N}}$ has a representing probability measure μ^* on S .

Remark 1 Recall that if (20) has an optimal solution $\mu \in \mathcal{P}(S)$ such that $\mu = f d\mathbf{x}$ with $f > 0$ on S , and if S has nonempty interior, then by Theorem 2

$$\sum_{j=0}^m s_{n-r_j} = \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \mu}(\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

and its vector $\boldsymbol{\mu}_{2n} := (\mu_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$ of moments up to degree $2n$ is an optimal solution of (25) of which (26) is a relaxation. In other words, the relaxation (26) (or (27)) is exact. That is, the vector of pseudo moments $\boldsymbol{\mu}_{2n}^*$, unique optimal solution of (26), is in fact the vector of moments (up to degree $2n$) of μ .

⁶If $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$ then the condition $\mathbf{1} \in \text{int}(Q_n(g_1, \dots, g_m))$ is satisfied; see [27].

Link between μ_{2n}^* and the equilibrium measure of S

With ϕ^* being the equilibrium measure of S , consider the polynomial $p_n^* \in \mathbb{R}[\mathbf{x}]_{2n}$ defined by:

$$\mathbf{x} \mapsto \left(\sum_{j=0}^m s_{n-r_j} \right) p_n^*(\mathbf{x}) := \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \phi^*}(\mathbf{x}, \mathbf{x}), \quad (30)$$

for all $\mathbf{x} \in \mathbb{R}^d$. Under some regularity conditions on (S, ϕ^*) :

- The sequence of measures $(p_n^* \phi^*)_{n \in \mathbb{N}}$ weak star converges to ϕ^* ,
- and under an additional condition, one obtains the stronger result, $\lim_{n \rightarrow \infty} p_n^*(\mathbf{x}) = 1$, uniformly on compact subsets of $\text{int}(S)$, which shows that somehow (30) is close to (29) (at least when $\mathbf{x} \in \text{int}(S)$).

Indeed fix $j \leq m$ arbitrary, and consider the measure $\nu_j := g_j \phi^*$ on S and its associated Christoffel polynomial $K_n^{\nu_j}(\mathbf{x}, \mathbf{x})$. We also assume that $g_j > 0$ on $\text{int}(S)$. If (S, ν_j) satisfies the Bernstein-Markov property (see [25, (1.3) p. 603]) then by [30, Theorem 4.4.4],

$$\lim_{n \rightarrow \infty} \int_S f(\mathbf{x}) \frac{K_n^{g_j \cdot \phi^*}(\mathbf{x}, \mathbf{x})}{s_n} g_j(\mathbf{x}) d\phi^* = \lim_{n \rightarrow \infty} \int_S f(\mathbf{x}) \frac{K_n^{\nu_j}(\mathbf{x}, \mathbf{x})}{s_n} d\nu_j = \int_S f d\phi^*, \quad \forall f \in \mathcal{C}(S).$$

Hence as j was arbitrary, it follows that

$$\lim_{n \rightarrow \infty} \int_S f(\mathbf{x}) p_n^*(\mathbf{x}) d\phi^*(\mathbf{x}) = \int_S f d\phi^*, \quad \forall f \in \mathcal{C}(S), \quad (31)$$

i.e., $p_n^* \phi^* \Rightarrow \phi^*$ as $n \rightarrow \infty$ (for the weak convergence of probability measures).

In addition, if $\lim_{n \rightarrow \infty} K_n^{\phi^*}(\mathbf{x}, \mathbf{x})/s_n = 1$, uniformly on compact subsets of $\text{int}(S)$, then

$$\lim_{n \rightarrow \infty} K_n^{\nu_j}(\mathbf{x}, \mathbf{x})/s_n = 1/g_j(\mathbf{x}), \quad \text{uniformly on compact subsets of } \text{int}(S).$$

See e.g. [30, Theorem 4.4.1] (which is a consequence of (1.5) in [25, Theorem 1.1] and [25, Remark(b), p.63]). Hence, as j was arbitrary,

$$\lim_{n \rightarrow \infty} g_j(\mathbf{x}) K_n^{\nu_j}(\mathbf{x}, \mathbf{x})/s_n = 1 \quad \text{uniformly on compact subsets of } \text{int}(S),$$

for every $0 \leq j \leq m$, which in turn implies

$$\lim_{n \rightarrow \infty} p_n^*(\mathbf{x}) = 1, \quad \text{uniformly on compact subsets of } \text{int}(S). \quad (32)$$

Comparing (29) with (30)-(32) show that even if the linear functional μ_{2n}^* , unique optimal solution of (26), does not have a representing measure, asymptotically it is strongly connected to the equilibrium measure ϕ^* of S . Indeed, for sufficiently large degree n and for every $\mathbf{x} \in \text{int}(S)$:

$$\sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \phi^*}(\mathbf{x}, \mathbf{x}) \approx \sum_{j=0}^m s_{n-r_j} = \sum_{j=0}^m g_j(\mathbf{x}) K_{n-r_j}^{g_j \cdot \mu_{2n}^*}(\mathbf{x}, \mathbf{x}).$$

Hence this asymptotic property supports our claim that the variant (20) of the D-optimal design problem still has a strong connection with the equilibrium measure of S .

5 Conclusion

We have introduced a variant of the D-optimal design problem with a more general information matrix which takes into account the geometry of the design space $S \subset \mathbb{R}^d$. One main reason to introduce such a variant is that remarkably, for the three cases where S is the Euclidean ball, the unit box and the canonical simplex (in any dimension), the equilibrium measure ϕ^* of S (in pluripotential theory) provides an optimal solution for every degree n . Therefore finding a D-optimal design (for this variant) reduces to finding a cubature for ϕ^* , with positive weights, atoms in S , and exact up to degree $2n$. Moreover and trivially, the associated resulting sequence of atomic probability measures $(\nu_n^*)_{n \in \mathbb{N}}$ converges to ϕ^* for the weak-star topology.

If the link between statistics and approximation theory has been largely invoked in the literature since the pioneer works [22] and [20], this new variant makes it even stronger and striking, at least for the three specific cases of S . Hence such a remarkable property suggests that one might use the log-det criterion of this new information matrix, even for more general compact basic semi-algebraic sets $S \subset \mathbb{R}^d$. In particular, the two-step algorithm proposed in [10] is easily adapted to this new variant of the D-optimal design problem.

References

- [1] M. Baran, Complex equilibrium measure and Bernstein type theorems for compact sets in \mathbb{R}^n , *Proc. Amer. Math. Soc.* **123** (1995), No. 2, p. 485-494.
- [2] E. Bedford, B. A. Taylor, The complex equilibrium measure of a symmetric convex set in \mathbb{R}^n , *Trans. Amer. Math. Soc.* **294** (1986), No. 2, p. 705-717.
- [3] R. J. Berman, S. Boucksom, D. W. Nyström, Fekete points and convergence towards equilibrium measures on complex manifolds, *Acta. Math.* **207** (2011), No. 1, p. 1-27.
- [4] T. Bloom, L. Bos, N. Levenberg, S. Waldron, On the convergence of optimal measures, *Constr. Approx.* **32** (2010), p. 150-179.
- [5] L. Bos, Some remarks on the Fejér Problem for Lagrange interpolation in several variables, *J. Approx. Theory* **60** (1990), p. 133-140.
- [6] L. Bos, On Fekete points for a real Simplex, *Indag. Math.* **34** (2023), No. 2, p. 274-293.
- [7] L. Bos, N. Levenberg, S. Waldron, On the spacing of Fekete points for a sphere, ball or simplex, *Indag. Math. (N.S.)* **19** (2008), No. 2, p. 163-176.
- [8] M. Briani, A. Sommariva, M. Vianello, Computing Fekete and Lebesgue points: simplex, square, disk., *J. Comput. Appl. Math.* **236** (2012), p. 2477-2486.
- [9] Y. de Castro, F. Gamboa, D. Henrion, R. Hess, J. B. Lasserre, Approximate optimal designs for multivariate polynomial regression, *Ann. Stat* **47** (2019), No. 1, p. 127-155.
- [10] Y. D. Castro, F. Gamboa, D. Henrion, J. B. Lasserre, Dual optimal design and the Christoffel-Darboux polynomial, *Optim. Letters* **15** (2021), No. 1, p. 3-8.

- [11] R. Cools, An encyclopaedia of cubature formulas, *J. Complexity* **19** (2003), No. 3, p. 445-453.
- [12] R. Cools, A. Haegemans, Algorithm 824: CUBPACK: A package for automatic cubature; framework description, *Trans. Math. Software* **29** (2003), No. 3, p. 287-296.
- [13] H. Dette, A generalization of D - and D_1 -optimal designs in polynomial regression, *Ann. Stat.* **18** (1990), No. 4, p. 1784-1804.
- [14] C. Dunkl, Y. Xu, *Orthogonal polynomials of several variables*, Encyclopedia of Mathematics and its Applications, vol. 155, Cambridge Univ. Press, 2014, 2nd edition.
- [15] W. Gautschi, Orthogonal Polynomials: Applications and Computations, *Acta Numerica* **5** (1996), p. 45-119.
- [16] S. Heo, Y. Xu, Constructing cubature formula for spheres and balls, *J. Approx. Theory* **112** (1999), No. 1-2, p. 95-119.
- [17] P. G. Hoel, A. Levine, Optimal spacing and weighing in polynomial prediction, *Ann. Math. Statist.* **35** (1964), p. 1553-1560.
- [18] P. J. Hoel, Efficiency in polynomial estimation, *Ann. Math. Statist.* **29** (1958), p. 1134-1145.
- [19] X. Huan, J. Jagalur, Y. Marzouk, Optimal experimental design: Formulations and computations, *Acta Num.* **33** (2024), p. 1-129.
- [20] S. Karlin, W. J. Studden, *Tchebycheff Systems: With Applications in Analysis and Statistics*, Interscience, New York, 1966.
- [21] J. Kiefer, Optimum designs in regression problems II, *Ann. Math. Statist.* **32** (1961), p. 298-325.
- [22] J. Kiefer, J. Wolfowitz, The equivalence of two extremum problems, *Canad. J. Math.* **12** (1960), p. 363-366.
- [23] S. Kirsch, Transfinite Diameter, Chebyshev Constant and Capacity, in *Handbook of Complex Analysis: Geometric Function Theory*, vol. 2, Elsevier B.V., 2005, p. 243-305.
- [24] M. Klimek, *Pluripotential Theory*, Clarendon Press, UK, 1992.
- [25] A. Kroó, D. S. Lubinsky, Christoffel functions and universality in the bulk for multi-variable orthogonal polynomials, *Canad. J. Math.* **65** (2012), No. 3, p. 600-620.
- [26] J. B. Lasserre, *An Introduction to Polynomial and Semi-Algebraic Optimization*, Cambridge University Press, Cambridge, UK, 2015.
- [27] J. B. Lasserre, Pell's equation, sum-of-squares and equilibrium measures on a compact set, *Comptes Rendus Mathématique* **361** (2023), p. 935-952.
- [28] J. B. Lasserre, The Moment-SOS hierarchy: Applications and related topics, *Acta Numerica* **33** (2024), p. 841-908.

- [29] J. B. Lasserre, E. Pauwels, The empirical Christoffel function with applications in data analysis, *Adv. Comp. Math.* **45** (2019), p. 1439-1468.
- [30] J. B. Lasserre, E. Pauwels, M. Putinar, *The Christoffel-Darboux Kernel for Data Analysis*, Cambridge University Press, Cambridge, UK, 2022.
- [31] J. B. Lasserre, Y. Xu, A generalized Pell's equation for a class of multivariate orthogonal polynomials, *Trans. Amer. Math. Soc.* (2024).
- [32] E. Lauter, Experimental design in a class of models, *Math. Oper. Statist.* **5** (1974), p. 379-398.
- [33] N. Levenberg, Approximation in \mathbb{C}^N , *Surveys in Approximation Theory* **2** (2006), p. 92-140.
- [34] F. Piazzon, Pluripotential numerics, *Constr. Approx* **49** (2019), p. 227-263.
- [35] H. Richter, Parameterfreie Abschatzung und Realisierung von Erwartungswerten, *Bl. der Deutsch. Ges. Versicherungsmath* **3** (1957), p. 147-161.
- [36] E. B. Saff, V. Totik, *Logarithmic Potentials with External Fields*, Springer, Berlin, Heidelberg, 2010.
- [37] K. Schmudgen, *The Moment Problem*, Graduate Texts in Mathematics, Springer International Publishing AG, 2017.
- [38] B. Simon, The Christoffel-Darboux kernel, in *Perspective in Partial Ddifferential Erations, Harmonic Analysis and Applications: Proc. Symp. Pure Math.*, vol. 79, American Mathematical Society, Providence, RI, 2008, p. 295-335.
- [39] W. J. Studden, Optimal designs on Tchebycheff points, *Ann. Math. Stat.* **39** (1968), No. 5, p. 1435-1447.