QUADRATURE ERRORS, DISCREPANCIES AND THEIR RELATIONS TO HALFTONING ON THE TORUS AND THE SPHERE

MANUEL GRÄF*, DANIEL POTTS‡, AND GABRIELE STEIDL§

Abstract. This paper deals with a special halftoning process, also called stippling, which aims to create the illusion of a gray-value image by appropriately distributing black dots. Recently a framework for this task was proposed by minimizing an attraction-repulsion functional consisting of the difference of two continuous, convex functions. One of them describes attracting forces caused by the image gray values, the other one enforces repulsion between dots. In this paper, we generalize this approach by considering quadrature error functionals on reproducing kernel Hilbert spaces (RKHSs) with respect to the quadrature nodes, where we ask for optimal distributions of these nodes. For special reproducing kernels these quadrature error functionals coincide with discrepancy functionals. It turns out that the attraction-repulsion functional appears for a special RKHS of functions on $\mathbb{R}^2$. Moreover, our more general framework enables us to consider optimal point distributions not only in $\mathbb{R}^2$ but also on the torus $\mathbb{T}^2$ and the sphere $\mathbb{S}^2$. For a large number of points the computation of such point distributions is a serious challenge and requires fast algorithms. To this end, we work in RKHSs of bandlimited functions on $\mathbb{T}^2$ and $\mathbb{S}^2$. Then the quadrature error functional can be rewritten as a least squares functional. We propose a nonlinear conjugate gradient method to compute a minimizer of this functional and show that each iteration step can be computed in an efficient way by fast Fourier transforms at nonequispaced nodes on the torus and the sphere. Numerical examples demonstrate the good halftoning results obtained by our method.

Math Subject Classifications. 65T40, 65K10 53B21 49M15, 33C55.
Keywords and Phrases. Halftoning, dithering, stippling, point distributions, quadrature rules, discrepancies, optimization methods on Riemannian manifolds, CG method, nonequispaced fast Fourier transform, spherical Fourier transform.

1. Introduction. Halftoning is a method for creating the illusion of a continuous tone image having only a small number of tones available. In this paper, we focus just on two tones, black and white and ask for appropriate distributions of the black 'dots'. Applications of halftoning include printing and geometry processing [46] as well as sampling problems occurring in rendering [50], re-lighting [31] or object placement and artistic non-photorealistic image visualization [4, 41]. Halftoning has been an active field of research for many years. *Dithering* methods which place the black dots at image grid points include, e.g., ordered dithering [7, 39], error diffusion [10, 18, 27, 44, 45] global or direct binary search [1, 6] and structure-aware halftoning [37]. In this paper we consider *stippling* algorithms which use the continuous domain as possible black dot positions instead of the image grid. More precisely, consider a gray-value image $u : \mathcal{G} \to [0,1]$ on a (squared) grid $\mathcal{G} := G \times G$, where $G := \{\frac{1}{2n} + \frac{i}{n} : i = 0, \dots, n-1\}$. Since 'black' is 0 and 'white' 1, we will later use the corresponding weight distribution $w := 1 - u$. Now one intends to find the positions $p_i \in [0,1]^2$, $i = 1, \dots, M$, of M black dots which create the illusion of the original gray-value image u, see Figure 1.1 for illustration. One prominent method for placing such points was proposed by Secord [41]. It is based on weighted centroidal Voronoi tessellations and Lloyd's iterative algorithm [23, 34]. A capacity-constrained variant of Lloyd's algorithm was introduced by Balzer et al. [4]. Recently, a novel halftoning framework was proposed in [47], where the vector $\boldsymbol{p} := \big(p_i\big)_{i=1}^M \in \mathbb{R}^{2M}$ of the black dot positions was determined to be a minimizer of the functional

$$E(\boldsymbol{p}) := \sum_{i=1}^{M} \sum_{x \in \mathcal{G}} w(x) \|p_i - x\|_2 - \frac{\lambda}{2} \sum_{i,j=1}^{M} \|p_i - p_j\|_2. \tag{1.1}$$

Here $\lambda := \frac{1}{M} \sum_{x \in \mathcal{G}} w(x)$ is an equilibration parameter between the 'opposite' functionals. The intention for considering minimizers of this functional as 'good' black dot positions comes originally from electrostatic principles used in [40] for halftoning. The black points are considered as small

*Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz, Germany, m.graef@mathematik.tu-chemnitz.de
‡Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz, Germany, potts@mathematik.tu-chemnitz.de
§University of Mannheim, Dept. of Mathematics and Computer Science, A5, 68131 Mannheim, Germany, steidl@kiwi.math.uni-mannheim.de

particles of equal size moving in an environment, e.g., a glass pane above the image w. The particles are *attracted* by the image forces $w(x)$ at the points $x \in \mathcal{G}$. On the other hand, there is a force of *repulsion* between the particles modeled by the negative sign of the second sum which becomes minimal if the distances between the particles are maximized.

FIG. 1.1. *Left: Original* 256×256 *image. Right: Stippling result by minimizing* (1.1) *with* $m = 30150$ *points using the technique from [47].*

In this paper, we deal with the continuous version of the above attraction-repulsion functional

$$E(\boldsymbol{p}) := \sum_{i=1}^{M} \int_{[0,1]^2} w(x) \|p_i - x\|_2 \, \mathrm{d}x - \frac{\lambda}{2} \sum_{i,j=1}^{M} \|p_i - p_j\|_2, \tag{1.2}$$

where $w : [0,1]^2 \to [0,1]$ is defined on the whole cube $[0,1]^2$ and $\lambda := \frac{1}{M} \int_{[0,1]^2} w(x) \, \mathrm{d}x$. We can also consider (1.2) with more general functions $\varphi : [0,\infty) \to \mathbb{R}$:

$$E_\varphi(\boldsymbol{p}) := \frac{\lambda}{2} \sum_{i,j=1}^{M} \varphi(\|p_i - p_j\|_2) - \sum_{i=1}^{M} \int_{[0,1]^2} w(x) \varphi(\|p_i - x\|_2) \, \mathrm{d}x. \tag{1.3}$$

In (1.2) the function $\varphi(r) = -r$ was used. In [40] the function $\varphi(r) = -\log(r)$ with a modification near zero was applied. Further, the authors in [47] also mentioned $\varphi(r) = -r^\tau$, $0 < \tau < 2$ and $\varphi(r) = r^{-\tau}$, $\tau > 0$ with a modification near zero as possible choices.

In this paper we look at the halftoning problem from different points of view. Our framework arises primarily from approximation theory but touches many different areas in mathematics as well. The proposed setting is quite general and enables us to consider in some sense optimal point distributions not only in $\mathbb{R}^2$ but also on the torus $\mathbb{T}^2$ and the sphere $\mathbb{S}^2$. Let us remark that even in the seemingly easiest case with $w \equiv 1$ the search for optimal point configurations is a very tough problem, at least in more than one dimension. For example on the sphere with the Coulomb potential $\varphi(r) = r^{-1}$ we are confronted with the Thomson problem [48], which asks for the ground states of a given number of electrons on the sphere. This famous problem originated lots of publications concerning the computations [53], asymptotics [32] and characteristics [9] of optimal distributions on the sphere, to name but a few. Another interesting application of our halftoning procedure on the sphere may be found in methods for solving partial differential equations arising in geoscience [19].

In the following, we consider worst case quadrature errors on RKHSs in dependence on the quadrature nodes and ask for optimal node distributions. In the literature this was mainly done for constant weights $w \equiv 1$. A weighted setting appears in connection with the so-called 'importance sampling', see [36] and the references therein. It turns out that the attraction-repulsion functional (1.2) leads to the same optimal point distributions as the quadrature error functional for a certain RKHS of functions on $\mathbb{R}^2$ with the Euclidean distance kernel. For special reproducing kernels we show that the quadrature error functionals coincide with discrepancy functionals. This adds another interesting point of view which is closely related to the notation of 'capacity constraints' in halftoning, see [3, 4]. As already mentioned before, the main challenge for computing optimal point distributions is the design of fast algorithms. Here, we present an algorithm which works on RKHSs of bandlimited functions on $\mathbb{T}^2$ and $\mathbb{S}^2$. We show that the quadrature error functional can be rewritten as a least squares functional. Then we propose a nonlinear conjugate gradient (CG) method for computing a minimizer. Indeed, on $\mathbb{S}^2$ we apply the CG method on manifolds, see [16, 43]. This method was also successfully used for the approximation of spherical designs in [22]. We show how each step within the CG method can be realized in an efficient way by fast Fourier transforms at nonequispaced nodes on the torus (NFFT) and the sphere (NFSFT), respectively. Finally, we provide proof-of-concept numerical examples based on the NFFT library [29].

Our paper is organized as follows: In Section 2, we introduce the worst case quadrature errors on RKHSs in dependence on the quadrature nodes and show that the attraction-repulsion functional (1.2) appears as a special case. The relation to discrepancy functionals is proved in Section 3. Furthermore, we provide discrepancy kernels on $\mathbb{S}^1$, $\mathbb{T}^2$ and $\mathbb{S}^2$ and compare them numerically with the kernels $-\|x-y\|_2$. Section 4 deals with the efficient computation of optimal point distributions. In Subsection 4.1, the functionals are considered on RKHSs of bandlimited functions on $\mathbb{S}^1$, $\mathbb{T}^2$ and $\mathbb{S}^2$, where they can be rewritten as least squares functionals. We show that the evaluation of these functionals as well as the computation of their gradients and vector multiplications with their Hessians are realized in a fast way by using NFFTs/NFSFTs. Subsection 4.2 provides the CG algorithms with respect to our setting. Finally, we present stippling examples on $\mathbb{T}^2$ and $\mathbb{S}^2$ in Section 5.

2. Quadrature Errors in RKHSs and Halftoning. In this section, we consider worst case quadrature errors in RKHSs. We show that for special RKHSs the minimizers of the corresponding error functional coincide with those of the halftoning functional (1.3). Of course our more general setting can be used as starting point for various applications. In this paper, we will use it to design halftoning procedures on submanifolds of $\mathbb{R}^d$, more precisely on the torus $\mathbb{T}^2$ and the sphere $\mathbb{S}^2$.

2.1. Quadrature Error in RKHSs. Let $\mathcal{X} \in \{\mathbb{R}^d, [0,1]^d, \mathbb{S}^1, \mathbb{T}^2, \mathbb{S}^2\}$. A symmetric function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is said to be *positive semi-definite* if for any $M \in \mathbb{N}$ points $x_1, \dots, x_M \in \mathcal{X}$ and any $a = (a_1, \dots, a_M)^{\mathrm{T}} \neq 0$ the relation

$$a^{\mathrm{T}} \left(K(x_i, x_j)\right)_{i,j=1}^{M} a \geq 0, \tag{2.1}$$

holds true and *positive definite* if we have strict inequality in (2.1). A (real) *reproducing kernel Hilbert space* (RKHS) is a Hilbert space having a *reproducing kernel*, i.e., a function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ which fulfills

$$\begin{aligned} K_x := K(\cdot, x) \in H_K \quad & \forall\, x \in \mathcal{X}, \\ f(x) = \langle f, K(\cdot, x) \rangle_{H_K} \quad & \forall x \in \mathcal{X} \quad \text{and} \quad \forall f \in H_K. \end{aligned} \tag{2.2}$$

An equivalent definition of a RKHS says that it is a Hilbert space on which the point evaluation functionals are continuous. To every RKHS there corresponds a unique positive semi-definite kernel and conversely given a positive semi-definite function K there exists a unique RKHS of real-valued function having K as its reproducing kernel, see [52, Theorem 1.1.1].
If K is in addition continuous and in $L_2(\mathcal{X} \times \mathcal{X})$ it can be expanded into an absolutely and

uniformly convergent series

$$K(x,y) := \sum_{l=1}^{\infty} \lambda_l \psi_l(x)\overline{\psi_l(y)} = \sum_{l=1}^{\infty} \lambda_l \overline{\psi_l(x)}\psi_l(y)$$

of orthonormal eigenfunctions $\psi_l \in L_2(\mathcal{X})$ and associated eigenvalues $\lambda_l > 0$ of the integral operator T_K associated with the kernel K

$$T_K f(x) := \int_{\mathcal{X}} K(x,y) f(y) \,\mathrm{d}y.$$

For more information on RKHSs we refer to [2].
In the following, let $w : \mathcal{X} \to \mathbb{R}_{\geq 0}$ be a nontrivial, continuous function. In the case $\mathcal{X} = \mathbb{R}^d$, which is only of interest in Subsection 2.2, we suppose that w has compact support. Furthermore, we suppose that

$$h_w(x) := \int_{\mathcal{X}} w(y) K(x,y) \,\mathrm{d}y \ \in H_K, \tag{2.3}$$

i.e.,

$$\begin{aligned} \|h_w\|_{H_K}^2 = \langle \int_{\mathcal{X}} w(y) K(\cdot,y) \,\mathrm{d}y, h_w \rangle_{H_K} \ &= \ \int_{\mathcal{X}} w(y) \langle K(\cdot,y), h_w \rangle_{H_K} \,\mathrm{d}y \\ = \int_{\mathcal{X}} w(y) h_w(y) \,\mathrm{d}y \ &= \ \int_{\mathcal{X}} \int_{\mathcal{X}} w(x) w(y) K(x,y) \,\mathrm{d}x\mathrm{d}y < \infty. \end{aligned} \tag{2.4}$$

Set $\boldsymbol{p} := (p_1, \ldots, p_M) \in \mathcal{X}^M$. We are interested in approximating the integrals

$$I_w(f) := \int_{\mathcal{X}} f(x) w(x) \,\mathrm{d}x \quad \text{for } f \in H_K$$

by a quadrature rule

$$Q(f, \boldsymbol{p}) := \lambda \sum_{i=1}^{M} f(p_i), \qquad \lambda := \frac{1}{M} \int_{\mathcal{X}} w(x) \,\mathrm{d}x \tag{2.5}$$

for appropriately chosen points $p_j \in \mathcal{X}$. This quadrature rule appears to play a key role in our paper. In the literature mainly the case $w \equiv 1$ was considered, see [36] and the references therein. The *worst case quadrature error* is given by

$$\mathrm{err}_K(\boldsymbol{p}) := \sup_{\substack{f \in H_K \\ \|f\|_{H_K} \leq 1}} |I_w(f) - Q(f,\boldsymbol{p})| = \|I_w - Q(\cdot, \boldsymbol{p})\|, \tag{2.6}$$

where the later norm is the operator norm of the linear functionals on H_K. In particular, we see that $I_w(f) = Q(f,\boldsymbol{p})$ for some $\boldsymbol{p} \in \mathcal{X}^M$ and all $f \in H_K$ if and only if $\mathrm{err}_K(\boldsymbol{p}) = 0$.
The following proposition shows a relation between this error functional and the halftoning functional (1.3).

THEOREM 2.1. *Let K be a positive semi-definite function and H_K the associated RKHS. Then the relation*

$$\mathrm{err}_K(\boldsymbol{p})^2 = 2\lambda E_K(\boldsymbol{p}) + \|h_w\|_{H_K}^2$$

holds true, where

$$E_K(\boldsymbol{p}) := \frac{\lambda}{2} \sum_{i,j=1}^{M} K(p_i, p_j) - \sum_{i=1}^{M} \int_{\mathcal{X}} w(x) K(p_i, x) \,\mathrm{d}x. \tag{2.7}$$

In particular, the minimizers of err_K *and* E_K *coincide.*
Proof: We rewrite I_w as

$$\begin{aligned} I_w(f) &= \int_{\mathcal{X}} \langle f, K(\cdot,x)\rangle_{H_K} w(x)\,\mathrm{d}x \\ &= \langle f, \int_{\mathcal{X}} w(x) K(\cdot,x)\,\mathrm{d}x\rangle_{H_K} \;=\; \langle f, h_w\rangle_{H_K}, \end{aligned}$$

so that by (2.2) and (2.5)

$$I_w(f) - Q(f,\boldsymbol{p}) = \langle f, h_w - \lambda \sum_{i=1}^{M} K(\cdot,p_i)\rangle_{H_K}$$

and consequently

$$\mathrm{err}_K(\boldsymbol{p}) = \|h_w - \lambda \sum_{i=1}^{M} K(\cdot,p_i)\|_{H_K}.$$

Now the squared worst case error reads

$$\begin{aligned} \mathrm{err}_K(\boldsymbol{p})^2 &= \|h_w\|_{H_K}^2 - 2\lambda \langle h_w, \sum_{i=1}^{M} K(\cdot,p_i)\rangle_{H_K} + \lambda^2 \sum_{i,j=1}^{M} K(p_i,p_j) \\ &= \|h_w\|_{H_K}^2 - 2\lambda \sum_{i=1}^{M} h_w(p_i) + \lambda^2 \sum_{i,j=1}^{M} K(p_i,p_j) \\ &= \|h_w\|_{H_K}^2 - 2\lambda \sum_{i=1}^{M} \int_{\mathcal{X}} w(x) K(p_i,x)\,\mathrm{d}x + \lambda^2 \sum_{i,j=1}^{M} K(p_i,p_j) \end{aligned} \tag{2.8}$$

and the minimizers of this functional coincide with those of E_K. ■

By the following proposition, slight modifications of the kernel do not change the minimizers of the functional.

PROPOSITION 2.2. *Let* $K : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ *be a symmetric function and* $\tilde{K}(x,y) := aK(x,y) + b(K(x,0)+K(0,y)) + c$ *with* $a > 0$ *and* $b, c \in \mathbb{R}$. *Then the minimizers of* E_K *and* $E_{\tilde{K}}$ *coincide.*
Proof: By (2.7) and the definition of λ, we obtain with constants C_i, $i = 1, 2$ independent of $\boldsymbol{p}$ that

$$\begin{aligned} E_{\tilde{K}}(\boldsymbol{p}) &= \frac{\lambda}{2} \sum_{i,j=1}^{M} \big(aK(p_i,p_j) + b\,(K(p_i,0)+K(0,p_j))\,\big) \\ &\quad - \sum_{i=1}^{M} \int_X w(x)\big(aK(p_i,x) + b\,(K(p_i,0)+K(0,x))\,\big)\,\mathrm{d}x \;+\; C_1 \\ &= \frac{a\lambda}{2} \sum_{i,j=1}^{M} K(p_i,p_j) - a\sum_{i=1}^{M} \int_X w(x) K(p_i,x)\,\mathrm{d}x \\ &\quad + \frac{b\lambda}{2} 2M \sum_{i=1}^{M} K(p_i,0) - b\sum_{i=1}^{M} K(p_i,0) \int_X w(x)\,\mathrm{d}x \;+\; C_2, \\ &= \frac{a\lambda}{2} \sum_{i,j=1}^{M} K(p_i,p_j) - a\sum_{i=1}^{M} \int_X w(x) K(p_i,x)\,\mathrm{d}x + C_2 \;=\; aE_K(\boldsymbol{p}) + C_2. \end{aligned}$$

Hence the functionals $E_{\tilde{K}}$ and E_K have the same minimizers. ∎

The functional E_K looks as those in (1.3) if K is a radial kernel $K(x,y) = \varphi(\|x-y\|_2)$. Therefore we ask for radial kernels which are positive semi-definite in the next subsection.

2.2. Relation to Halftoning Functionals on $\mathbb{R}^d$. In this subsection, we consider the relation between E_K and (1.3) in detail. To this end, let in this subsection $\mathcal{X} := \mathbb{R}^d$ and $w : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ be a nontrivial, continuous function with compact support in $[0,1]^d$.
A kernel $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is called a *radial kernel* if $K(x,y) = \varphi(\|x-y\|_2)$ for some function $\varphi : [0,\infty) \to \mathbb{R}$. We are looking for positive semi-definite, radial kernels K. Note that since $\mathbb{R}^l$ is a subspace of $\mathbb{R}^d$ for $l \leq d$ positive semi-definiteness of a kernel on $\mathbb{R}^d$ implies its positive semi-definiteness on $\mathbb{R}^l$. There are positive definite, radial kernels on $\mathbb{R}^d$ as for example the inverse multiquadric $K(x,y) := (\varepsilon^2 + \|x-y\|)_2^2)^{-\tau}$, $\varepsilon > 0, \tau > d/2$ related to $\varphi(r) = r^{-\tau}$ in (1.3) or the 'hat function' in $\mathbb{R}^1$. For other examples see [54, 17].
To get the kernel in (1.2) we have to consider *conditionally positive definite, radial functions* $\Phi(x) := \varphi(\|x\|_2)$ *of order 1.* These functions are determined to be continuous with the property that for any $M \in \mathbb{N}$ points $x_1, \ldots, x_M \in \mathbb{R}^d$ the relation

$$a^{\mathrm{T}} \left(\Phi(x_i - x_j)\right)_{i,j=1}^M a > 0 \quad \forall a = (a_1, \ldots, a_M)^{\mathrm{T}} \not\equiv 0 \quad \text{with} \quad \sum_{i=1}^M a_i = 0$$

holds true. For conditionally positive definite, radial functions of higher order which are not relevant in this paper, we refer to [54]. Examples of conditionally positive definite, radial functions of order 1 in $\mathbb{R}^d$ are

$$\begin{aligned} &\Phi(x) := -\|x\|_2^\tau, \quad 0 < \tau < 2, \\ &\Phi(x) := -(\varepsilon^2 + \|x\|_2^2)^\tau, \quad 0 < \tau < 1, \quad \text{(multiquadrics)}. \end{aligned}$$

Of course our dithering functional (1.2) is exactly $E_{\Phi(x-y)}$ for the first function with $\tau = 1$, while the multiquadrics is related to (1.3) with $\varphi(r) = -r^\tau$. Unfortunately the above kernels $\Phi(x-y)$ are not positive semi-definite. However, the slight modification of conditionally positive definite radial kernels Φ of order 1 given by

$$K_\Phi(x,y) := \Phi(x-y) - \Phi(y) - \Phi(x) + \Phi(0) + 1$$

defines again a positive semi-definite kernel which gives rise to a RKHS H_{K_Φ}. These spaces can be characterized as in [54, Theorem 10.18 - 10.21]. However, by Proposition 2.2, we see that $E_{\Phi(x-y)}$ and E_{K_Φ} have the same minimizers, so that we can work with the original kernel Φ.
REMARK 2.3. **(Halftoning in $\mathbb{R}^1$)**
For $\mathcal{X} = \mathbb{R}$, the minimizers of (1.3) with $\varphi(r) = -r$ can be described analytically. In one dimension we can suppose that the point positions are ordered by $p_1 \leq \ldots \leq p_M$ such that our functional simplifies to the strictly convex functional

$$\begin{aligned} E_\varphi(\boldsymbol{p}) &:= \sum_{i=1}^M \int_0^1 w(x)|p_i - x| \,\mathrm{d}x - \frac{\lambda}{2} \sum_{i,j=1}^M |p_i - p_j| \\ &= \sum_{i=1}^M \int_0^1 w(x)|p_i - x| \,\mathrm{d}x + \lambda \sum_{i=1}^M (M - (2i-1))p_i. \end{aligned}$$

The minimizer $\hat{\boldsymbol{p}}$ of this functional are computed componentwise by

$$\begin{aligned} 0 \;=\; \frac{\partial}{\partial p_i} E_\varphi(\hat{\boldsymbol{p}}) &= \lambda(M - (2i-1)) + \int_0^{\hat{p}_i} w(x) \,\mathrm{d}x - \int_{\hat{p}_i}^1 w(x) \,\mathrm{d}x \\ &= \lambda(M - (2i-1)) + \int_0^{\hat{p}_i} w(x) \,\mathrm{d}x - (\lambda M - \int_0^{\hat{p}_i} w(x) \,\mathrm{d}x) \end{aligned}$$

which leads to

$$\int_0^{\hat{p}_i} w(x)\,\mathrm{d}x = \lambda(i-\frac{1}{2}), \quad i=1,\dots,M.$$

In other words, $\hat{p}_1$ is determined by $\int_0^{\hat{p}_1} w(x)\,\mathrm{d}x = \lambda/2$ and the other points by $\int_{\hat{p}_i}^{\hat{p}_{i+1}} w(x)\,\mathrm{d}x = \lambda$, $i=1,\dots,M-1$.
There is an interesting connection to the Sobolev spaces

$$H_{K^\beta} := \{f:[0,1]\to\mathbb{R} : f(\beta)=0, f \text{ abs. continuous}, f'\in L_2([0,1])\}$$

anchored at $\beta\in[0,1]$ which were considered in [36]. These RKHSs have the reproducing kernels

$$K^\beta(x,y) := \frac{1}{2}\big(|x-\beta|+|y-\beta|-|x-y|\big).$$

Using similar arguments as in the proof of Proposition 2.2 one can check that our functional E_φ and the functionals E_{K^β}, $\beta\in[0,1]$ have the same minimizers. □

3. Discrepancies. The quadrature errors considered in the previous section are closely related to discrepancies which adds another interesting point of view. We consider in the following $\mathcal{X}\in\{[0,1]^d,\mathbb{S}^1,\mathbb{T}^2,\mathbb{S}^2\}$ with the Lebesgue measure and the spherical measure $\mu_\mathcal{X}$ respectively. Let $D:=\mathcal{X}\times[0,R]$ and let $\mathcal{B}(c,r):=\{x\in\mathcal{X} : d_\mathcal{X}(c,x)\le r\}$ be the ball centered at $c\in\mathcal{X}$ with radius $0\le r\le R$. By $1_{\mathcal{B}(c,r)}$ we denote the characteristic function of $\mathcal{B}(c,r)$. Then we define

$$\begin{aligned} K_\mathcal{B}(x,y) &:= \int_0^R\int_\mathcal{X} 1_{\mathcal{B}(c,r)}(x)1_{\mathcal{B}(c,r)}(y)\,\mathrm{d}c\,\mathrm{d}r \\ &= \int_0^R \mu_\mathcal{X}(\mathcal{B}(x,r)\cap\mathcal{B}(y,r))\,\mathrm{d}r, \end{aligned} \tag{3.1}$$

where we used the for $\mathrm{d}\mu_\mathcal{X}(c)$ the abbreviation $\mathrm{d}c$. Since

$$a^\mathrm{T}\left(K_\mathcal{B}(x_i,x_j)\right)_{i,j=1}^M a = \int_0^R\int_\mathcal{X}\left(\sum_{j=1}^M a_j 1_{\mathcal{B}(c,r)}(x_j)\right)^2 \mathrm{d}c\,\mathrm{d}r \ge 0$$

we see that $K_\mathcal{B}$ is a positive semi-definite function. Integration on the RKHSs $H_{K_\mathcal{B}}$ is related to the notation of discrepancy, see [36] and the references therein. Set $t:=(c,r)\in D$ and $\mathrm{d}t:=\mathrm{d}c\,\mathrm{d}r$. We define the *$L_2$-discrepancy* as

$$\mathrm{disc}_2^\mathcal{B}(\boldsymbol{p}) := \left(\int_D\left(\int_\mathcal{X} w(x)1_{\mathcal{B}(t)}(x)\,dx - \lambda\sum_{i=1}^M 1_{\mathcal{B}(t)}(p_i)\right)^2\mathrm{d}t\right)^{\frac{1}{2}}. \tag{3.2}$$

The expression in the inner brackets relates the integral of w on $\mathcal{B}(c,r)$ with the number of points contained in $\mathcal{B}(c,r)$ for fixed $(c,r)\in D$. The discrepancy is then the squared error of their differences taken over all $t\in D$. This point of view is closely related to *capacity-constrained methods* used in [3, 4].
The relation between the discrepancy $\mathrm{disc}_2^\mathcal{B}$ and the quadrature error $\mathrm{err}_{K_\mathcal{B}}$ is given by the following theorem.

THEOREM 3.1. *Let $K_\mathcal{B}$ be defined by* (3.1) *and let $H_{K_\mathcal{B}}$ be the associated RKHS of functions on $\mathcal{X}$. Then* $\mathrm{err}_{K_\mathcal{B}}$ *given by* (2.6) *and* $\mathrm{disc}_2^\mathcal{B}$ *determined by* (3.2) *coincide*

$$\mathrm{err}_{K_\mathcal{B}}(\boldsymbol{p}) = \mathrm{disc}_2^\mathcal{B}(\boldsymbol{p}).$$

Proof: Using the definition (3.1) of $K_{\mathcal{B}}$ and (3.2) we obtain

$$\begin{aligned}
(\mathrm{disc}_2^{\mathcal{B}}(\boldsymbol{p}))^2 &= \int_D \left(\int_{\mathcal{X}} w(x) 1_{\mathcal{B}(t)}(x) \mathrm{d}x\right)^2 \mathrm{d}t - 2\lambda \int_D \int_{\mathcal{X}} w(x) 1_{\mathcal{B}(t)}(x) \mathrm{d}x \big(\sum_{i=1}^M 1_{\mathcal{B}(t)}(p_i)\big) \mathrm{d}t \\
&\quad + \lambda^2 \sum_{i,j=1}^M \int_D 1_{\mathcal{B}(t)}(p_i) 1_{\mathcal{B}(t)}(p_j) \mathrm{d}t \\
&= \int_D \left(\int_{\mathcal{X}} w(x) 1_{\mathcal{B}(t)}(x)\, dx\right)^2 \mathrm{d}t - 2\lambda \sum_{i=1}^M \int_{\mathcal{X}} w(x) \int_D 1_{\mathcal{B}(t)}(x) 1_{\mathcal{B}(t)}(p_i) \mathrm{d}t\, \mathrm{d}x \\
&\quad + \lambda^2 \sum_{i,j=1}^M K_{\mathcal{B}}(p_i, p_j) \\
&= \int_D \left(\int_{\mathcal{X}} w(x) 1_{\mathcal{B}(t)}(x) \mathrm{d}x\right)^2 \mathrm{d}t - 2\lambda \sum_{i=1}^M \int_{\mathcal{X}} w(x) K_{\mathcal{B}}(p_i, x) \mathrm{d}x + \lambda^2 \sum_{i,j=1}^M K_{\mathcal{B}}(p_i, p_j).
\end{aligned}$$

Finally we see by (2.4) that

$$\begin{aligned}
\int_D \left(\int_{\mathcal{X}} w(x) 1_{\mathcal{B}(t)}(x)\, dx\right)^2 \mathrm{d}t &= \int_{\mathcal{X}} \int_{\mathcal{X}} w(x) w(y) \int_D 1_{\mathcal{B}(t)}(x) 1_{\mathcal{B}(t)}(y) \mathrm{d}t\, \mathrm{d}x\, \mathrm{d}y \\
&= \int_{\mathcal{X}} \int_{\mathcal{X}} w(x) w(y) K_{\mathcal{B}}(x, y) \mathrm{d}t\, \mathrm{d}x\, \mathrm{d}y \;=\; \|h_w\|^2_{H_{K_{\mathcal{B}}}}.
\end{aligned}$$

and we are done by (2.8). ∎

Next we want to examine the relation between the distance kernel $K(x,y) = \Phi(x-y) = -\|x-y\|_2$ considered in Subsection 2.2 and the 'discrepancy kernel' $K_{\mathcal{B}}$ defined in (3.1) for $\mathcal{X} \in \{\mathbb{S}^1, \mathbb{T}^2, \mathbb{S}^2\}$.

Kernels on $\mathbb{S}^1$. The circle $\mathbb{S}^1$ is naturally embedded in $\mathbb{R}^2$ by

$$\mathbb{S}^1 := \{x := (2\pi)^{-1}(\cos 2\pi\alpha, \sin 2\pi\alpha)^{\mathrm{T}} \in \mathbb{R}^2 : \alpha \in [0,1)\},$$

where the correspondence between $x \in \mathbb{S}^1$ and $\alpha \in [0,1)$ is one-to-one. The geodesic distance is given for $x := (\cos 2\pi\alpha, \sin 2\pi\alpha)^{\mathrm{T}}$ and $y := (\cos 2\pi\beta, \sin 2\pi\beta)^{\mathrm{T}}$ by

$$\mathrm{d}_{\mathbb{S}^1}(x,y) = (2\pi)^{-1} \arccos\left(\cos 2\pi(\alpha-\beta)\right) = \min\{|\alpha-\beta|, 1-|\alpha-\beta|\} \le \frac{1}{2}.$$

The restriction of the negative Euclidean distance on $\mathbb{S}^1$ is

$$\begin{aligned}
\Phi(x-y) = -\|x-y\|_2 &= -(2\pi)^{-1}\sqrt{(\cos 2\pi\alpha - \cos 2\pi\beta)^2 + (\sin 2\pi\alpha - \sin 2\pi\beta)^2} \\
&= -\frac{1}{\pi}|\sin \pi(\alpha-\beta)| = -\frac{1}{\pi}|\sin \pi \mathrm{d}_{\mathbb{S}^1}(x,y)|\,.
\end{aligned}$$

For the discrepancy kernel we use the balls $\mathcal{B}(c,r)$, $c \in \mathbb{S}^1$, $0 \le r \le 1/2$ and obtain with $d := \mathrm{d}_{\mathbb{S}^1}(x,y)$ that

$$\begin{aligned}
K_{\mathcal{B}}(x,y) &= \int_0^{\frac{1}{2}} \mu_{\mathbb{S}^1}(\mathcal{B}(x,r) \cap \mathcal{B}(y,r))\, \mathrm{d}r \\
&= \int_{\frac{d}{2}}^{\frac{1}{2}} 2\left(r - \frac{d}{2}\right) \mathrm{d}r + \int_{\frac{1}{2}-\frac{d}{2}}^{\frac{1}{2}} 2\left(r - \frac{1}{2} + \frac{d}{2}\right) \mathrm{d}r \\
&= \frac{1}{4} + \frac{1}{2} d(d-1).
\end{aligned}$$

Note that Wahba's smoothing spline kernel R of order $m = 1$ on $\mathbb{S}^1$, cf., [52, pp. 21], is given by

$$\begin{aligned} R(x,y) &= 1 + 2\sum_{k=1}^{\infty} \frac{1}{(2\pi k)^2} \cos(2\pi k d) \\ &= \frac{13}{12} + \frac{1}{2} d(d-1) \\ &= \frac{5}{6} + K_{\mathcal{B}}(x,y) \end{aligned}$$

so that $E_{K_{\mathcal{B}}}$ and E_R have the same minimizers. The kernel R is the reproducing kernel of the Hilbert space consisting of the functions

$$f(x) = \tilde{f}(\alpha) := f_0 + 2\sum_{k=1}^{\infty} f_k \cos(2\pi k\alpha) \quad \text{with} \quad \sum_{k=1}^{\infty} k^2 f_k^2 < \infty,$$

$x := (\cos 2\pi\alpha, \sin 2\pi\alpha)^{\mathrm{T}}$, with inner product

$$\langle f, g \rangle_{H_R} = \int_0^1 \tilde{f}(\alpha)\,\mathrm{d}\alpha \int_0^1 \tilde{g}(\alpha)\,\mathrm{d}\alpha + \int_0^1 \tilde{f}'(\alpha)\tilde{g}'(\alpha)\,\mathrm{d}\alpha.$$

The kernels $\Phi(x - (-(2\pi)^{-1}, 0)^{\mathrm{T}})$ and $K_{\mathcal{B}}(x, (-(2\pi)^{-1}, 0)^{\mathrm{T}})$ as functions of α are plotted in Figure 3.1. Since adding a constant and multiplying the kernel by a positive constant does not change the local minimizers of our functional (2.7), we compare the different kernels after an appropriate affine scaling. That is the maximum and the minimum of the scaled kernels are set without loss of generality to 1 and 0, respectively. The figure shows that both kernels are quite similar.

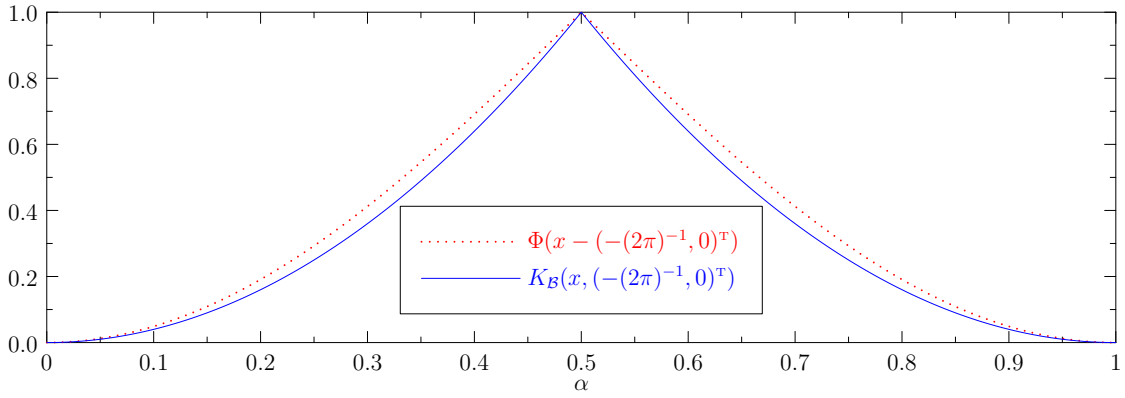

FIG. 3.1. *Scaled kernels* $\Phi(x - (-(2\pi)^{-1}, 0)^T)$ *and* $K_{\mathcal{B}}(x, (-(2\pi)^{-1}, 0)^T)$ *on* $\mathbb{S}^1$ *as functions of* α.

By the following remark, the minimizers of E_K for the discrepancy kernel K can be characterized analytically. The arguments are similar as those in Remark 2.3.

REMARK 3.2. **(Halftoning in $\mathbb{S}^1$)**
The discrepancy kernel reads up to a constant as $K(x,y) = \frac{1}{2}([x-y]_1^2 - [x-y]_1)$, where

$$[x-y]_1 := \begin{cases} |x-y| & \text{if } |x-y| \le 1/2, \\ 1 - |x-y| & \text{otherwise,} \end{cases} \qquad x, y \in [0,1].$$

We are looking for minimizers of

$$\sum_{i=1}^{M} \int_0^1 w(x)[p_i - x]_1\,\mathrm{d}x - \frac{\lambda}{2} \sum_{i,j=1}^{M} [p_i - p_j]_1 - \sum_{i=1}^{M} \int_0^1 w(x)[p_i - x]_1^2\,\mathrm{d}x + \frac{\lambda}{2} \sum_{i,j=1}^{M} [p_i - p_j]_1^2.$$

If $|p_i - p_j| \geq 1/2$ we obtain that

$$-[p_i - p_j]_1 + [p_i - p_j]_1^2 = -1 + |p_i - p_j| + 1 - 2|p_i - p_j| + |p_i - p_j|^2 = -|p_i - p_j| + |p_i - p_j|^2.$$

Thus we can replace $[\cdot]_1$ by $|\cdot|$ in the functional. Now an ordering of the point positions $0 \leq p_1 \leq \ldots \leq p_M \leq 1$ results as in Remark 2.3 in

$$\lambda \sum_{i=1}^{M} (M - (2i-1)) p_i + \sum_{i=1}^{M} \int_0^1 w(x)|p_i - x| \,\mathrm{d}x + \frac{\lambda}{2} \sum_{i,j=1}^{M} (p_i - p_j)^2 - \sum_{i=1}^{M} \int_0^1 w(x)(p_i - x)^2 \,\mathrm{d}x.$$

Setting the gradient to zero we obtain

$$\int_0^{\hat{p}_i} w(x) \,\mathrm{d}x = \lambda(i - \frac{1}{2}) + \lambda\big(\sum_{i=1}^{M} \hat{p}_i - M \frac{\int_0^1 x w(x) \,\mathrm{d}x}{\int_0^1 w(x) \,\mathrm{d}x}\big), \quad i = 1, \ldots, M.$$

Subtracting the i-th equation from the $(i+1)$-st one we see that the points have to fulfill $\int_{\hat{p}_i}^{\hat{p}_{i+1}} w(x) \,\mathrm{d}x = \lambda$, $i = 1, \ldots, M-1$. For a constant weight $w \equiv c$, it follows that $\hat{p}_1$ can be chosen arbitrarily in $[0, \lambda]$, where $\lambda = c/M$. □

Kernels on $\mathbb{T}^2$. The torus $\mathbb{T}^2 := \mathbb{S}^1 \times \mathbb{S}^1 \subset \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$ is naturally embedded in $\mathbb{R}^4$ by

$$\mathbb{T}^2 := \{x := (2\pi)^{-1}(\cos 2\pi\alpha_1, \sin 2\pi\alpha_1, \cos 2\pi\alpha_2, \sin 2\pi\alpha_2)^{\mathrm{T}} \in \mathbb{R}^4 : \alpha_1, \alpha_2 \in [0,1)\}$$

with geodesic distance

$$\mathrm{d}_{\mathbb{T}^2}(x, y) = \sqrt{\mathrm{d}_{\mathbb{S}^1}(\alpha_1, \beta_1)^2 + \mathrm{d}_{\mathbb{S}^1}(\alpha_2, \beta_2)^2}$$

for

$$\begin{aligned} x &:= (2\pi)^{-1}(\cos 2\pi\alpha_1, \sin 2\pi\alpha_1, \cos 2\pi\alpha_2, \sin 2\pi\alpha_2)^{\mathrm{T}}, \\ y &:= (2\pi)^{-1}(\cos 2\pi\beta_1, \sin 2\pi\beta_1, \cos 2\pi\beta_2, \sin 2\pi\beta_2)^{\mathrm{T}}. \end{aligned}$$

The restriction of the negative Euclidean distance is

$$\begin{aligned} \Phi(x - y) &:= -\|x - y\|_2 = -\frac{1}{\pi}\sqrt{\sin^2 \pi(\alpha_1 - \beta_1) + \sin^2 \pi(\alpha_2 - \beta_2)} \\ &= -\frac{1}{\pi}\sqrt{1 - \cos \pi(\alpha_1 + \alpha_2 - \beta_1 - \beta_2) \cos \pi(\alpha_1 - \alpha_2 - \beta_1 + \beta_2)}. \end{aligned}$$

Since the torus is flat, the balls $\mathcal{B}(x, r)$ on $\mathbb{T}^2$ with radius $r \leq R \leq 1/2$ can be considered as two-dimensional Euclidean balls $\tilde{\mathcal{B}}((\alpha_1, \alpha_2), r)$. In the Euclidean plane $\mathbb{R}^2$ the area of intersection of two balls of radius r with distance d between their centers is

$$a(r, d) := \begin{cases} 2r^2 \arccos(d/(2r)) - d\sqrt{r^2 - d^2/4}, & r \geq d/2, \\ 0, & \text{else.} \end{cases}$$

With the integral

$$\begin{aligned} A_R(d) &:= \int_0^R a(r, d) \,\mathrm{d}r \\ &= \begin{cases} \frac{R}{3}(2R^2 \arccos(d/(2R)) - \sqrt{4R^2 - d^2}) - \frac{d^3}{12} \log\left(d/(2R + \sqrt{4R^2 - d^2})\right), & d \leq 2R, \\ 0, & \text{else,} \end{cases} \end{aligned} \tag{3.3}$$

we obtain the kernel

$$K_{\mathcal{B}}(x, y) = \int_0^R \mu_{\mathbb{T}^2}(\mathcal{B}_r(x) \cap \mathcal{B}_r(y)) \mathrm{d}r = \sum_{i=1}^{4} \int_{d_i/2}^{R} a(r, d_i) \,\mathrm{d}r = \sum_{i=1}^{4} A_R(d_i),$$

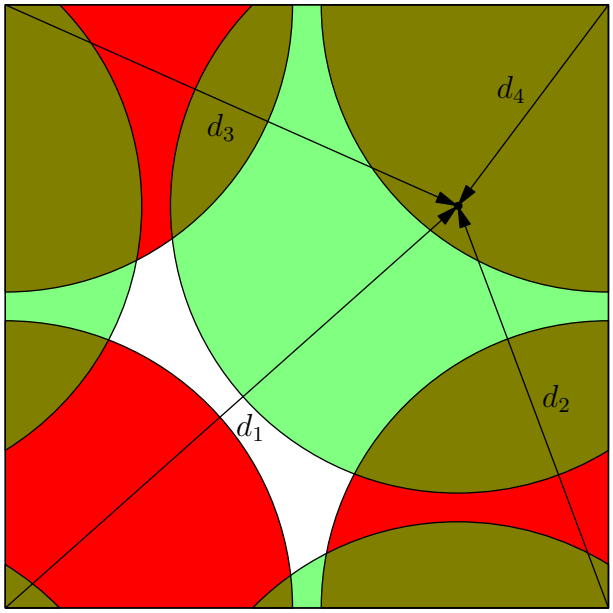

FIG. 3.2. *Visualization of the intersection of two balls on the torus* $\mathbb{T}^2$, *cf.* (3.3) - (3.4).

where

$$\begin{aligned} d_1 &:= \sqrt{|\alpha_1-\beta_1|^2+|\alpha_2-\beta_2|^2},\ d_2 := \sqrt{(|\alpha_1-\beta_1|-1)^2+|\alpha_2-\beta_2|^2}, \\ d_3 &:= \sqrt{|\alpha_1-\beta_1|^2+(|\alpha_2-\beta_2|-1)^2},\ d_4 := \sqrt{(|\alpha_1-\beta_1|-1)^2+(|\alpha_2-\beta_2|-1)^2}. \end{aligned} \tag{3.4}$$

For an illustration of the above relations see Figure 3.2, where one easily observes that this kernel is the periodization of the radial kernel $A_{1/2}(d)$. We remark that this kernel can not be written in the form $K_{\mathcal{B}}(x,y) = \tilde{K}(d_{\mathbb{T}^2}(x,y))$, hence it is not rotationally invariant. Figure 3.3, left shows the appropriately scaled kernels $\Phi(x)$ and $K_{\mathcal{B}}(x,0)$ as functions of (α_1,α_2) which have nearly the same shape.
For our implementations it will be necessary to approximate the kernels by a trigonometric polynomial. Therefore we expand the kernel $K_{\mathcal{B}}$ in a Fourier series

$$K_{\mathcal{B}}(x,y) = \sum_{(n_1,n_2)^{\mathrm{T}}\in\mathbb{Z}^2} \hat{K}_{\mathcal{B}}(n_1,n_2)\mathrm{e}^{-2\pi\mathrm{i}(n_1,n_2)^{\mathrm{T}}\cdot(\alpha_1-\beta_1,\alpha_2-\beta_2)^{\mathrm{T}}}.$$

Since the kernel $K_{\mathcal{B}}$ is the periodization of the radial kernel $A_{1/2}(d)$ the Fourier coefficients of $K_{\mathcal{B}}$ are given by

$$\hat{K}_{\mathcal{B}}(n_1,n_2) = \hat{k}_{\mathcal{B}}(\sqrt{n_1^2+n_2^2}),$$

where

$$\begin{aligned} \hat{k}_{\mathcal{B}}(0) &= 2\pi\int_0^1 xA_R(x)\mathrm{d}x, \\ \hat{k}_{\mathcal{B}}(r) &= 2\pi\int_0^1 xA_R(x)J_0(2\pi xr)\mathrm{d}x, \qquad r>0, \end{aligned}$$

and J_0 is the *Bessel function of first kind* and order 0. We further approximate $\hat{k}_{\mathcal{B}}(r)$ by a function

$$\hat{\tilde{k}}_{\mathcal{B}}(r) := a/r^3 + (b+c\sin(d+2\pi r))/r^4,$$

where the parameters a, b, c, d are determined such that the least squares fit $\|(\hat{\tilde{k}}_{\mathcal{B}}(r_j))_{i=0}^{1200} - (\hat{k}_{\mathcal{B}}(r_i))_{i=0}^{1200}\|_2^2$ with $r_i := 1 + i/100$ becomes minimal. This results in

$$\hat{\tilde{k}}_{\mathcal{B}}(r) = 0.00599 \left(1/r^3 + (0.21 + 0.35 \sin(1.87 + 2\pi r))/r^4\right). \tag{3.5}$$

The right-hand side of Figure 3.3 depicts the scaled versions of the kernel $K_{\mathcal{B}}(x, 0)$ and the truncation of $\tilde{K}_{\mathcal{B}}(x, 0)$ with bandwidth $N = 40$.

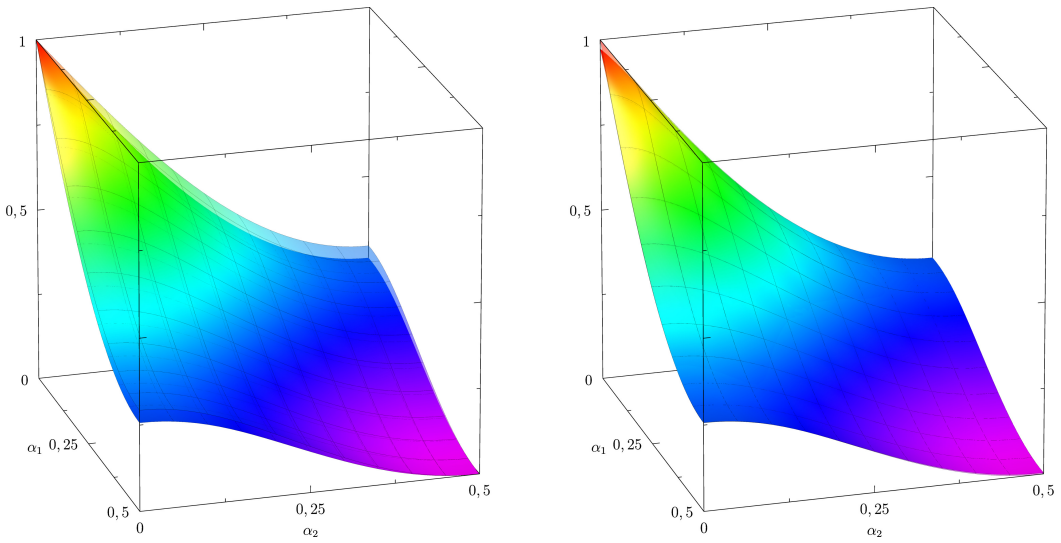

FIG. 3.3. *Left: Scaled kernels $\Phi(x)$ and $K_{\mathcal{B}}(x, 0)$ on $\mathbb{T}^2$. Right: Scaled kernels $K_{\mathcal{B}}(x, 0)$ and $\tilde{K}_{\mathcal{B}}(x, 0)$ as functions of (α_1, α_2).*

Kernels on $\mathbb{S}^2$. The sphere $\mathbb{S}^2$ is embedded in $\mathbb{R}^3$ by $\mathbb{S}^2 := \{x \in \mathbb{R}^3 : \|x\|_2 = 1\}$. For $x \in \mathbb{S}^2$ we make use of the parameterization in spherical coordinates

$$x = x(\theta, \varphi) := (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^{\mathrm{T}}, \qquad (\varphi, \theta) \in [0, 2\pi) \times [0, \pi].$$

The geodesic distance is given by

$$\mathrm{d}_{\mathbb{S}^2}(x, y) = \arccos(x \cdot y), \qquad x, y \in \mathbb{S}^2.$$

The restricted distance kernel has the form

$$\Phi(x - y) = -\|x - y\|_2 = -2\sin(\mathrm{d}_{\mathbb{S}^2}(x, y)/2).$$

On the sphere there is no special direction. Hence the discrepancy kernel $K_{\mathcal{B}}$ obtained from the spherical caps $\mathcal{B}(c, r)$, $c \in \mathbb{S}^2$ with radius $r \le \pi$ is rotationally invariant, i.e., $K_{\mathcal{B}}(x, y) = \tilde{K}_{\mathcal{B}}(d_{\mathbb{S}^2}(x, y))$. For computing the function $\tilde{K}_{\mathcal{B}}$ we need the area of intersection of two spherical caps with center distance d and radius r which is given by

$$a(r, d) = \begin{cases} 0, & 0 \le r \le d/2, \\ 4\left[\arccos\left(\sin(d/2)/\sin r\right) - \cos r \arccos\left(\tan(d/2)\cot r\right)\right], & d/2 < r < \pi/2, \\ 4r - 2d, & r = \pi/2, \\ 4\left[\arccos\left(\sin(d/2)/\sin r\right) - \cos r \arccos\left(\tan(d/2)\cot r\right)\right], & \pi/2 < r < \pi - d/2, \\ -4\pi\cos r, & \pi - d/2 \le r < \pi. \end{cases}$$

Then the discrepancy kernel is given by

$$\tilde{K}_{\mathcal{B}}(d) = \int_0^{\pi} a(r, d)\,\mathrm{d}r.$$

Figure 3.4 shows a plot of the scaled distance and discrepancy kernels as functions of $\mathrm{d}_{\mathbb{S}^2}(x, (0, 0, 1)^{\mathrm{T}})$.

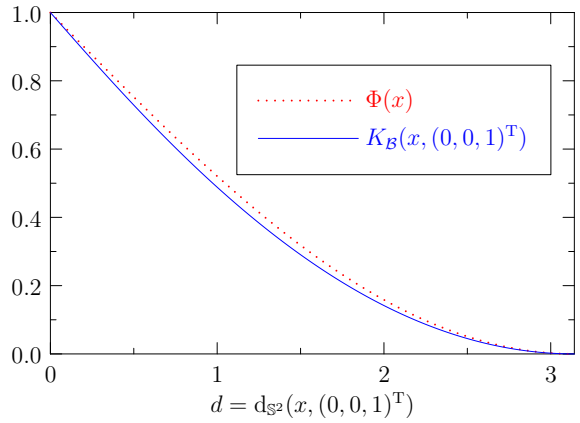

FIG. 3.4. *Scaled kernels* $\Phi(x)$ *and* $K_{\mathcal{B}}(x,(0,0,1)^T)$ *on* $\mathbb{S}^2$ *as functions of* $\mathrm{d}_{\mathbb{S}^2}(x,(0,0,1)^T)$.

4. Computation of Minimizers on $\mathbb{S}^1$, $\mathbb{T}^2$ and $\mathbb{S}^2$. In this section, we develop algorithms for the efficient computation of minimizers $\hat{\boldsymbol{p}}$ of functionals E_K for given functions w on $\mathcal{X} \in \{\mathbb{S}^1, \mathbb{T}^2, \mathbb{S}^2\}$. The case $\mathcal{X} = \mathbb{S}^1$ is only included for convenience. In the next section, we will use the resulting coordinates $\hat{p}_i$, $i = 1, \dots, M$ as point positions for halftoning w on the torus and the sphere. Our algorithms rest upon bandlimited kernels $K = K_N$ which approximate the distance/discrepancy kernels from the previous section. First, we reformulate E_{K_N} as a nonlinear least squares functional $\mathcal{E}_N$. We will see that the evaluation of this functional, its gradient and the vector multiplication with its Hessian at any point $\boldsymbol{p} \in \mathcal{X}^M$ can be realized in an efficient way. This will be used within the nonlinear CG method to compute a minimizer of $\mathcal{E}_N$.

4.1. A Least Squares Setting. Let $\mathcal{X} \in \{\mathbb{S}^1, \mathbb{T}^2, \mathbb{S}^2\}$ and let $\{\psi_l : l \in \mathbb{N}\}$ be an orthonormal basis of $L_2(\mathcal{X})$. Then any real-valued function $w \in L_2(\mathcal{X})$ can be written in the form

$$w(x) = \sum_{l=1}^{\infty} \hat{w}_l \psi_l(x), \quad \hat{w}_l = \langle w, \psi_l \rangle_{L_2} = \int_{\mathcal{X}} w(x) \overline{\psi_l(x)} \, \mathrm{d}x. \tag{4.1}$$

We will work in spaces of bandlimited functions

$$\Pi_N(\mathcal{X}) := \operatorname{span}\{\psi_l : l = 1, \dots, d_N\}$$

of dimension $d_N := \dim \Pi_N(\mathcal{X})$. More precisely, we will use the following settings:

$$\begin{aligned}
&\Pi_N(\mathbb{S}^1) := \operatorname{span}\{\, \mathrm{e}^{-2\pi \mathrm{i} n(\cdot)} : n = -N/2, \dots, N/2\}, && d_N = N+1,\\
&\Pi_N(\mathbb{T}^2) := \operatorname{span}\{\, \mathrm{e}^{-2\pi \mathrm{i} n(\cdot)} : n = (n_1, n_2), n_j = -N/2, \dots, N/2,\ j = 1,2\}, && d_N = (N+1)^2,\\
&\Pi_N(\mathbb{S}^2) := \operatorname{span}\{Y_n^k : n = 0, \dots, N; k = -n, \dots, n\}, && d_N = (N+1)^2,
\end{aligned}$$

where N is supposed to be even in the first two cases. Here Y_n^k denote the *spherical harmonics* of degree n and order k, cf. [35],

$$Y_n^k(x) = Y_n^k(\theta, \varphi) := \sqrt{\frac{2n+1}{4\pi}} P_n^{|k|}(\cos\theta) \, \mathrm{e}^{\mathrm{i}k\varphi},$$

where the *associated Legendre functions* $P_n^k : [-1,1] \to \mathbb{R}$ and the *Legendre polynomials* $P_n : [-1,1] \to \mathbb{R}$ are given by

$$\begin{aligned}
P_n^k(x) &:= \left(\frac{(n-k)!}{(n+k)!}\right)^{1/2} \left(1-x^2\right)^{k/2} \frac{\mathrm{d}^k}{\mathrm{d}x^k} P_n^k(x), \qquad n \in \mathbb{N}_0,\ k = 0, \dots, n,\\
P_n(x) &:= \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(x^2 - 1\right)^n, \qquad n \in \mathbb{N}_0.
\end{aligned}$$

We will apply bandlimited kernels of the form

$$K_N(x,y) := \sum_{l=1}^{d_N} \lambda_l \psi_l(x) \overline{\psi_l(y)} \tag{4.2}$$

with $\lambda_l > 0$. Note that these kernels are reproducing kernels for the RKHSs $H_{K_N} := \Pi_N(\mathcal{X})$ with the inner product

$$\langle f, g \rangle_{H_{K_N}} = \sum_{l=1}^{d_N} \frac{\hat{f}_l \overline{\hat{g}_l}}{\lambda_l}.$$

We are interested in minimizers of E_{K_N} for given $w \in L_2(\mathcal{X})$. For the efficient computation of these minimizers it is useful to rewrite the functional as a weighted least squares problem.

THEOREM 4.1. *Let the kernel K_N be given by (4.2) and let $H_{K_N} := \Pi_N(\mathcal{X})$ be the associated RKHS. Then the relation $\mathrm{err}_{K_N}(\boldsymbol{p})^2 = \mathcal{E}_N(\boldsymbol{p})$ holds true, where*

$$\mathcal{E}_N(\boldsymbol{p}) := \sum_{l=1}^{d_N} \lambda_l \left| \lambda \sum_{i=1}^{M} \overline{\psi_l(p_i)} - \hat{w}_l \right|^2 = \| \boldsymbol{\Lambda}^{\frac{1}{2}} F(\boldsymbol{p}) \|_2^2 \tag{4.3}$$

with $\boldsymbol{\Lambda} := \mathrm{diag}(\lambda_l)_{l=1}^{d_N}$ and $F(\boldsymbol{p}) = (F_l(\boldsymbol{p}))_{l=1}^{d_N}$, $F_l(\boldsymbol{p}) := \lambda \sum_{i=1}^{M} \overline{\psi_l(p_i)} - \hat{w}_l$. In particular, the functionals E_{K_N} and $\mathcal{E}_N$ have the same minimizers.

Proof: We rewrite the function in (4.3) as

$$\begin{aligned}
\mathcal{E}_N(\boldsymbol{p}) &= \lambda^2 \sum_{l=1}^{d_N} \lambda_l \left| \sum_{i=1}^{M} \left(\overline{\psi_l(p_i)} - \frac{\hat{w}_l}{M\lambda} \right) \right|^2 \\
&= \lambda^2 \sum_{l=1}^{d_N} \lambda_l \sum_{i=1}^{M} \left(\overline{\psi_l(p_i)} - \frac{\hat{w}_l}{M\lambda} \right) \sum_{j=1}^{M} \left(\psi_l(p_j) - \frac{\overline{\hat{w}_l}}{M\lambda} \right) \\
&= \lambda^2 \sum_{l=1}^{d_N} \lambda_l \sum_{i=1}^{M} \sum_{j=1}^{M} \left(\frac{|\hat{w}_l|^2}{M^2\lambda^2} - \frac{\hat{w}_l}{M\lambda} \psi_l(p_i) - \frac{\overline{\hat{w}_l}}{M\lambda} \overline{\psi_l(p_j)} + \overline{\psi_l(p_i)} \psi_l(p_j) \right) \\
&= \sum_{l=1}^{d_N} \lambda_l |\hat{w}_l|^2 - 2\lambda \mathrm{Re} \left(\sum_{l=1}^{d_N} \lambda_l \hat{w}_l \sum_{i=1}^{M} \psi_l(p_i) \right) + \lambda^2 \sum_{i=1}^{M} \sum_{j=1}^{M} K_N(p_i, p_j).
\end{aligned}$$

Using the relation $h_w(x) = \sum_{l=1}^{d_N} \lambda_l \hat{w}_l \psi_l(x)$, cf. (2.3), we further conclude that

$$\mathcal{E}_N(\boldsymbol{p}) = \|h_w\|_{H_{K_N}}^2 - 2\lambda \sum_{i=1}^{M} \int_{\mathcal{X}} w(x) K_N(p_i, x) \mathrm{d}x + \lambda^2 \sum_{i=1}^{M} \sum_{j=1}^{M} K_N(p_i, p_j)$$

which yields the assertion. ∎

REMARK 4.2. **(Relation to spherical designs)**
By Theorem 4.1 we have that $I_w(f) = Q(f, \boldsymbol{p})$ for some $\boldsymbol{p} \in \mathcal{X}^M$ and all $f \in H_{K_N}$ if and only if $\mathrm{err}_{K_N}(\boldsymbol{p})^2 = \mathcal{E}_N(\boldsymbol{p}) = 0$.
Consider the case $\mathcal{X} = \mathbb{S}^2$, $w \equiv 1 = \sqrt{4\pi} Y_0^0$ and

$$K_N(x,y) := \sum_{n=0}^{N} \sum_{k=-n}^{n} \lambda_n Y_n^k(x) \overline{Y_n^k(y)} = \sum_{n=0}^{N} \lambda_n \frac{2n+1}{4\pi} P_n(x \cdot y). \tag{4.4}$$

A set $\{p_i : i = 1, \dots, M\}$ satisfying

$$\int_{\mathbb{S}^2} f(x)\,\mathrm{d}x = \frac{4\pi}{M}\sum_{i=1}^{M} f(p_i), \qquad \text{for all } f \in \Pi_N(\mathbb{S}^2)$$

is called *spherical N-design*. The concept of spherical N-designs was introduced by Delsarte, Goethals and Seidel [13]. Up to now there is no theoretical result which proves the existence of an N-design with $M = (N+1)^2$ nodes for arbitrary $N \in \mathbb{N}$. But recently, in [11] it was verified that for $N = 1, \dots, 100$, spherical N-designs with $(N+1)^2$ nodes exist using the characterization of fundamental spherical N-designs and interval arithmetic. For further recent developments on spherical N-designs and related topics we refer to the survey article [5]. Finally, we remark that the equivalence between spherical N-designs and the relation

$$\mathcal{E}_N(\boldsymbol{p}) = \lambda^2 \sum_{n=1}^{N}\sum_{k=-n}^{n} \lambda_n \left|\sum_{i=1}^{M} Y_n^k(p_i)\right|^2 = 0$$

was applied by Sloan and Womersley in [42]. □

The rest of this section describes how to compute for given $\hat{w}_l$, $l = 1, \dots, d_N$ local minimizers of $\mathcal{E}_N(\boldsymbol{p})$ in an efficient way also for large numbers M of point positions. As our algorithm of choice we present the nonlinear CG method in the next subsection. However, the efficient computation of each CG step rests upon

- algorithms for the fast evaluation of bandlimited functions on $\mathbb{S}^1, \mathbb{T}^2$ and $\mathbb{S}^2$,
- simple representations of the gradient and the Hessian of $\mathcal{E}_N$.

In the following, we describe these two items in more detail. The evaluation of bandlimited functions

$$f(p_i) = \sum_{l=1}^{d_N} \hat{f}_l \psi_l(p_i), \quad i = 1, \dots, M$$

can be written in matrix-vector form as

$$\boldsymbol{f} = \boldsymbol{A}_N \hat{\boldsymbol{f}},$$

where $\boldsymbol{f} := (f(p_i))_{i=1}^{M}$, $\hat{\boldsymbol{f}} := \left(\hat{f}_l\right)_{l=1}^{d_N}$ appropriately ordered and

$$\boldsymbol{A}_N := \begin{cases} \boldsymbol{F}_N = \left(\mathrm{e}^{-2\pi \mathrm{i} n p_i}\right)_{i=1,\dots,M;n=-N/2,\dots,N/2} \in \mathbb{C}^{M,N+1} & \text{for } \mathbb{S}^1, \\ \boldsymbol{F}_{2,N} = \left(\mathrm{e}^{-2\pi \mathrm{i}(n_1,n_2)^{\mathrm{T}}\cdot p_i}\right)_{i=1,\dots,M;n_i=-N/2,\dots,N/2,i=1,2} \in \mathbb{C}^{M,(N+1)^2} & \text{for } \mathbb{T}^2, \\ \boldsymbol{Y}_N = (Y_k^n(p_i))_{i=1,\dots,M;n=0,\dots,N,\,|k|\le n} \in \mathbb{C}^{M,(N+1)^2} & \text{for } \mathbb{S}^2. \end{cases} \tag{4.5}$$

Recently fast algorithms for the matrix-vector multiplication with $\boldsymbol{A}_N$ and $\overline{\boldsymbol{A}}_N^{\mathrm{T}}$ were proposed. More precisely, the algorithms for the first two cases $\mathbb{S}^1$ and $\mathbb{T}^2$, called *nonequispaced fast Fourier transform* (NFFT) or *unequally spaced fast Fourier transform* can be found, e.g., in [8, 15, 29, 38]. The algorithms on the sphere $\mathbb{S}^2$, called *nonequispaced fast spherical Fourier transform* (NFSFT) were developed in [30, 33], see also [14, 26]. In our numerical examples we have applied the software package [28]. These algorithms for the multiplication with the matrices $\boldsymbol{A}_N$ and $\overline{\boldsymbol{A}}_N^{\mathrm{T}}$ given in (4.5) achieve the following arithmetic complexity:

$$\begin{array}{ll} \mathcal{O}(N\log N + M\log(1/\epsilon)) & \text{for } \mathbb{S}^1, \\ \mathcal{O}(N^2\log N + M\log^2(1/\epsilon)) & \text{for } \mathbb{T}^2, \\ \mathcal{O}(N^2\log^2 N + M\log^2(1/\epsilon)) & \text{for } \mathbb{S}^2, \end{array} \tag{4.6}$$

where ϵ is the prescribed accuracy.

Using these algorithms the same complexity as in (4.6) is required for the evaluation of $\mathcal{E}_N(\boldsymbol{p})$ as the following corollary states.

COROLLARY 4.3. **(Efficient evaluation of $F(\boldsymbol{p})$ and $\mathcal{E}_N(\boldsymbol{p})$)**
For a given point $\boldsymbol{p} \in \mathcal{X}^M$ and $\hat{w}_l$, $l = 1, \dots, d_N$, the computation of $F(\boldsymbol{p})$ and $\mathcal{E}_N(\boldsymbol{p})$ can be realized with the arithmetic complexity (4.6).

Proof: We have that

$$\left(\sum_{i=1}^{M} \overline{\psi_l(p_i)} \right)_{l=1}^{d_N} = \overline{\boldsymbol{A}}_N^{\mathrm{T}} \boldsymbol{e}, \qquad \boldsymbol{e} := (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^M.$$

Hence the vector on the left-hand side can be computed with (4.6) arithmetic operations by applying NFFT or NFSFT. The remaining operations do not increase this complexity. ∎

Next we consider the gradient of $\mathcal{E}_N$. By $\nabla_{\mathcal{X}} \psi$ we denote the gradient of ψ on $\mathcal{X}$ and by $\nabla_{\mathcal{X}}^i \mathcal{E}_N$ the derivative of $\mathcal{E}_N$ with respect to the i-th component vector p_i. Then straightforward computation shows that the gradient $\nabla = \nabla_{\mathcal{X}^M}$ of $\mathcal{E}_N$ at $\boldsymbol{p} \in \mathcal{X}^M$ is given by $\nabla \mathcal{E}_N(\boldsymbol{p}) = \big(\nabla_{\mathcal{X}}^i \mathcal{E}_N(\boldsymbol{p})\big)_{i=1}^M$, where

$$\nabla_{\mathcal{X}}^i \mathcal{E}_N(\boldsymbol{p}) = 2\lambda \mathrm{Re}\big[\sum_{l=1}^{d_N} \lambda_l \underbrace{\left(\lambda \sum_{j=1}^{M} \overline{\psi_l(p_j)} - \hat{w}_l\right)}_{F_l(\boldsymbol{p})} \nabla_{\mathcal{X}} \psi_l(p_i). \big]$$

Hence the gradient can be written as

$$\nabla \mathcal{E}_N(\boldsymbol{p}) = 2\mathrm{Re}\big[\overline{J_F(\boldsymbol{p})}^{\mathrm{T}} \boldsymbol{\Lambda} \, F(\boldsymbol{p}) \big], \tag{4.7}$$

where

$$J_F(\boldsymbol{p}) := \Big(\big(\nabla_{\mathcal{X}}^i F_l(p_i) \big)^{\mathrm{T}} \Big)_{l=1,i=1}^{d_N, M} = \lambda \left(\nabla_{\mathcal{X}} \overline{\psi_l(p_i)} \right)_{l=1,i=1}^{d_N, M}$$

denotes the *Jacobian matrix* of F. For our three settings the gradients specify as follows:

Gradient on $\mathbb{S}^1$. For

$$\mathcal{E}_N(\boldsymbol{p}) = \sum_{n=-N/2}^{N/2} \lambda_n |\lambda \sum_{i=1}^{m} \mathrm{e}^{-2\pi \mathrm{i} n p_i} - \hat{w}_n|^2$$

we obtain with $\nabla_{\mathbb{S}^1} \psi_n(p_i) = -2\pi \mathrm{i} n \, \mathrm{e}^{-2\pi \mathrm{i} n p_i}$ that

$$\overline{J_F(\boldsymbol{p})}^{\mathrm{T}} = \lambda \left(-2\pi \mathrm{i} n \, \mathrm{e}^{-2\pi \mathrm{i} n p_i} \right)_{i=1, n=-N/2}^{M, N/2} = \lambda \boldsymbol{F}_N \boldsymbol{D}_N, \quad \boldsymbol{D}_N := \mathrm{diag}\left(-2\pi \mathrm{i} n \right)_{n=-N/2}^{N/2}. \tag{4.8}$$

Gradient on $\mathbb{T}^2$. For

$$\mathcal{E}_N(\boldsymbol{p}) = \sum_{n_1, n_2 = -N/2}^{N/2} \lambda_n |\lambda \sum_{i=1}^{M} \mathrm{e}^{-2\pi \mathrm{i} n \cdot p_i} - \hat{w}_n|^2, \qquad n := (n_1, n_2)^{\mathrm{T}}$$

we get with $\nabla_{\mathbb{T}^2} \psi_n(p_i) = -2\pi \mathrm{i} n \, \mathrm{e}^{-2\pi \mathrm{i} n \cdot p_i}$ and an appropriate ordering that

$$\overline{J_F(\boldsymbol{p})}^{\mathrm{T}} = \lambda (\boldsymbol{I}_2 \otimes \boldsymbol{F}_{2,N}) \begin{pmatrix} \boldsymbol{I}_N \otimes \boldsymbol{D}_N \\ \boldsymbol{D}_N \otimes \boldsymbol{I}_N \end{pmatrix}, \tag{4.9}$$

where $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix and $\otimes$ the Kronecker product.

Gradient on $\mathbb{S}^2$. On the sphere we will only work with kernels of the form (4.4) such that

$$\mathcal{E}_N(\boldsymbol{p}) = \sum_{n=0}^{N} \sum_{k=-n}^{n} \lambda_n |\lambda \sum_{i=1}^{m} \overline{Y_n^k(p_i)} - \hat{w}_n^k|^2.$$

Then we have to clarify the definition of the gradient of a function on $\mathbb{S}^2$. To this end, let $\mathrm{T}_x\mathbb{S}^2 := \{\boldsymbol{v} \in \mathbb{R}^3 : \langle \boldsymbol{v}, \boldsymbol{x} \rangle = 0\}$ be the *tangent space* at a point $x \in S^2$. For $x := x(\theta, \varphi) \in \mathbb{S}^2 \backslash \{(0,0,\pm 1)^{\mathrm{T}}\}$ this tangent space is spanned by the orthonormal vectors $e_\varphi := (-\sin\varphi, \cos\varphi, 0)^{\mathrm{T}}$ and $e_\theta := (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta)^{\mathrm{T}}$. Then the *spherical gradient operator* is defined as

$$\nabla_{\mathbb{S}^2} := e_\varphi \frac{1}{\sin\theta} \frac{\partial}{\partial\varphi} + e_\theta \frac{\partial}{\partial\theta},$$

cf. [20, 49]. Note that this is the orthogonal projection of $\nabla\tilde{\psi}(x) \in \mathbb{R}^3$ onto $\mathrm{T}_x\mathbb{S}^2$, where $\tilde{\psi}$ denotes an extension of ψ to $\mathbb{R}^3$. In particular, the derivatives of the spherical harmonics can be computed by

$$\begin{aligned} \frac{\partial}{\partial\varphi} Y_n^k(\theta,\varphi) &= \mathrm{i}k Y_n^k(\theta,\varphi), \\ \sin\theta \frac{\partial}{\partial\theta} Y_n^k(\theta,\varphi) &= \underbrace{n\sqrt{\frac{(n+1)^2-k^2}{(2n+1)(2n+3)}}}_{a_{n+1}^k} Y_{n+1}^k(\theta,\varphi) - \underbrace{(n+1)\sqrt{\frac{n^2-k^2}{(2n+1)(2n-1)}}}_{b_{n-1}^k} Y_{n-1}^k(\theta,\varphi), \end{aligned}$$

where $Y_{n-1}^k := 0$ for $|k| > n-1$, see [51, pp. 146]. Using this relation we obtain for $p_i = x(\theta_i, \varphi_i) \in \mathbb{S}^2 \backslash \{(0,0,\pm 1)^{\mathrm{T}}\}$ that

$$\nabla_{\mathbb{S}^2} Y_n^k(p_i) = \frac{1}{\sin\theta_i} \mathrm{i}k\, Y_n^k(\varphi_i,\theta_i)\, e_{\varphi_i} + \frac{1}{\sin\theta_i} \left(a_{n+1}^k Y_{n+1}^k(\varphi_i,\theta_i) - b_{n-1}^k Y_{n-1}^k(\varphi_i,\theta_i)\right) e_{\theta_i}$$

and consequently

$$\begin{aligned} \nabla_{\mathbb{S}^2}^i \mathcal{E}_N(\boldsymbol{p}) &= \frac{2\lambda}{\sin\theta_i} \mathrm{Re}\left[\sum_{n=0}^{N} \sum_{k=-n}^{n} Y_n^k(\varphi_i,\theta_i)\, \mathrm{i}k\, \lambda_n F_n^k(\boldsymbol{p})\right] e_{\varphi_i} \\ &\quad + \frac{2\lambda}{\sin\theta_i} \mathrm{Re}\left[\sum_{n=0}^{N+1} \sum_{k=-n}^{n} Y_n^k(\varphi_i,\theta_i) \left(a_n^k \lambda_{n-1} F_{n-1}^k(\boldsymbol{p}) - b_n^k \lambda_{n+1} F_{n+1}^k(\boldsymbol{p})\right)\right] e_{\theta_i} \\ &= x_{\varphi_i} e_{\varphi_i} + x_{\theta_i} e_{\theta_i}, \end{aligned} \tag{4.10}$$

where $F_{N+1}^k(\boldsymbol{p}) = F_{N+2}^k(\boldsymbol{p}) = 0$. Hence the coordinate vectors $\boldsymbol{x}_\varphi := (x_{\varphi_i})_{i=1}^M$ and $\boldsymbol{x}_\theta := (x_{\theta_i})_{i=1}^M$ can be computed by

$$\begin{pmatrix} \boldsymbol{x}_\varphi \\ \boldsymbol{x}_\theta \end{pmatrix} = 2\lambda \mathrm{Re}\left[(\boldsymbol{I}_2 \otimes \boldsymbol{S}^{-1}) \begin{pmatrix} \boldsymbol{Y}_N \boldsymbol{D}_{N,\varphi} \\ \boldsymbol{Y}_{N+1} \tilde{\boldsymbol{D}}_{N,\theta} \end{pmatrix} \boldsymbol{\Lambda}\, F(\boldsymbol{p})\right] \tag{4.11}$$

where $\boldsymbol{S} := \mathrm{diag}\,(\sin\theta_i)_{i=1}^M$, $\boldsymbol{D}_{N,\varphi}$ is the diagonal matrix determined by the first summand in (4.10) and $\tilde{\boldsymbol{D}}_{N,\theta}$ the matrix with at most two non-zero entries in each row corresponding to the second summand in (4.10).

In summary we obtain the following corollary.

COROLLARY 4.4. **(Efficient evaluation of $\nabla\mathcal{E}_N(\boldsymbol{p})$)**
For a given point $\boldsymbol{p} \in \mathcal{X}^M$ and given $\hat{w}_l$, $l = 1, \ldots, d_N$, the gradient $\nabla\mathcal{E}_N(\boldsymbol{p})$ can be computed with the arithmetic complexity given by (4.6).

Proof: The proof follows by Corollary 4.3 and the relation (4.7) together with (4.8), (4.9), (4.11). ∎

Finally, we are interested in the Hessian $\mathrm{H} = \mathrm{H}_{\mathcal{X}^M}$ of $\mathcal{E}_N$. By $\mathrm{H}_{\mathcal{X}}\psi$ we denote the Hessian of ψ on $\mathcal{X}$. By straightforward computation we obtain that

$$\mathrm{H}\mathcal{E}_N(\boldsymbol{p}) = \left(\mathrm{H}^{i,j}\mathcal{E}_N(\boldsymbol{p})\right)_{i,j=1}^{M},$$

where

$$\mathrm{H}^{i,j}\mathcal{E}_N(\boldsymbol{p}) = 2\lambda^2 \mathrm{Re}\left[\sum_{l=1}^{d_N} \lambda_l \, \nabla_{\mathcal{X}}\psi_l(p_i) \, \left(\nabla_{\mathcal{X}}\overline{\psi_l(p_j)}\right)^{\mathrm{T}}\right] \tag{4.12}$$

$$+ \, \delta_{i,j} 2\lambda \mathrm{Re}\left[\sum_{l=1}^{d_N} \lambda_l \underbrace{\left(\lambda \sum_{m=1}^{M} \overline{\psi_l(p_m)} - \hat{w}_l\right)}_{F_l(\boldsymbol{p})} \mathrm{H}_{\mathcal{X}}\psi_l(p_i)\right]. \tag{4.13}$$

Instead of the Hessian of $\mathcal{E}_N$ we will also apply its approximation $\tilde{\mathrm{H}}\mathcal{E}_N$ which involves only the first summand (4.12) in the above expression, i.e., the diagonal part is neglected so that

$$\begin{aligned}\tilde{\mathrm{H}}\mathcal{E}_N(\boldsymbol{p}) &:= 2\lambda^2 \mathrm{Re}\left(\sum_{l=1}^{d_N} \lambda_l \, \nabla_{\mathcal{X}}\psi_l(p_i) \, \left(\nabla_{\mathcal{X}}\overline{\psi_l(p_j)}\right)^{\mathrm{T}}\right)_{i,j=1}^{M} \\ &= 2\mathrm{Re}\left[\overline{J_F(\boldsymbol{p})}^{\mathrm{T}} \, \boldsymbol{\Lambda} \, J_F(\boldsymbol{p})\right].\end{aligned} \tag{4.14}$$

This matrix does not depend on the values $\hat{w}_l$. Note that the approximate Hessian is also used in the Gauss-Newton method for solving nonlinear least squares problems. see [25, p. 185]. Let us specify the Hessian for our three settings.

Hessian on $\mathbb{S}^1$. Since $\mathrm{H}_{\mathbb{S}^1}\left(\mathrm{e}^{-2\pi \mathrm{i} n p_i}\right) = (2\pi \mathrm{i} n)^2 \, \mathrm{e}^{-2\pi \mathrm{i} n p_i}$, we obtain together with (4.14) and (4.8) that

$$\mathrm{H}\mathcal{E}_N(\boldsymbol{p}) = 2\lambda^2 \mathrm{Re}\left[\boldsymbol{F}_N \boldsymbol{D}_N^2 \boldsymbol{\Lambda} \overline{\boldsymbol{F}_N}^{\mathrm{T}}\right] + 2\lambda \mathrm{diag}\left(\mathrm{Re}\left[\boldsymbol{F}_N \boldsymbol{D}_N^2 \boldsymbol{\Lambda} F(\boldsymbol{p})\right]\right).$$

Hessian on $\mathbb{T}^2$. We have that

$$\mathrm{H}_{\mathbb{T}^2}\left(\mathrm{e}^{-2\pi \mathrm{i} n \cdot p_i}\right) = -4\pi^2 \begin{pmatrix} n_1^2 & n_1 n_2 \\ n_1 n_2 & n_2^2 \end{pmatrix} \mathrm{e}^{-2\pi \mathrm{i} n \cdot p_i}.$$

Hence, the block-diagonal part (4.13) of the Hessian is given by

$$2\lambda \mathrm{Re}\left[(\boldsymbol{I}_2 \otimes \boldsymbol{F}_{2,N}) \begin{pmatrix} \boldsymbol{I}_N \otimes \boldsymbol{D}_N^2 & \boldsymbol{D}_N \otimes \boldsymbol{D}_N \\ \boldsymbol{D}_N \otimes \boldsymbol{D}_N & \boldsymbol{D}_N^2 \otimes \boldsymbol{I}_N \end{pmatrix} (\boldsymbol{I}_2 \otimes \boldsymbol{\Lambda} F(\boldsymbol{p}))\right]$$

and corresponding sorting. Thus, the multiplication of the first (4.12) and second part (4.13) of the Hessian with a vector can be realized in a fast way by applying (4.14) and (4.9).

Hessian on $\mathbb{S}^2$. The Hessian $\mathrm{H}_{\mathbb{S}^2}\psi(x)$ is linear operator on $\mathrm{T}_x\mathbb{S}^2$. For $x := x(\theta, \varphi)$ we consider the corresponding matrix with respect to the basis $\{e_\theta, e_\varphi\}$. Then, the Hessian reads as

$$\sin^2\theta \mathrm{H}_{\mathbb{S}^2} = \left(\sin\theta \frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right)^{\mathrm{T}} \left(\sin\theta \frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) - \cos\theta \begin{pmatrix} \sin\theta \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial\varphi} & -\sin\theta \frac{\partial}{\partial\theta} \end{pmatrix},$$

cf. [49]. Using the relations for the derivatives of the spherical harmonics again, we conclude that the block-diagonal part (4.13) of the Hessian can be expressed by

$$2\lambda \mathrm{Re}\left[(\boldsymbol{I}_2 \otimes \boldsymbol{S}^{-2}) \, \boldsymbol{M} \, (\boldsymbol{I}_2 \otimes \boldsymbol{\Lambda} F(\boldsymbol{p}))\right], \tag{4.15}$$

where

$$\boldsymbol{M} := \begin{pmatrix} \boldsymbol{Y}_{N+2}\tilde{\boldsymbol{D}}_{N+1,\theta}\tilde{\boldsymbol{D}}_{N,\theta} - \boldsymbol{C}\boldsymbol{Y}_{N+1}\tilde{\boldsymbol{D}}_{N,\theta} & \boldsymbol{Y}_{N+1}\tilde{\boldsymbol{D}}_{N,\theta}\boldsymbol{D}_{N,\varphi} - \boldsymbol{C}\boldsymbol{Y}_N\boldsymbol{D}_{N,\varphi} \\ \boldsymbol{Y}_{N+1}\tilde{\boldsymbol{D}}_{N,\theta}\boldsymbol{D}_{N,\varphi} - \boldsymbol{C}\boldsymbol{Y}_N\boldsymbol{D}_{N,\varphi} & \boldsymbol{Y}_N\boldsymbol{D}_{N,\varphi}\boldsymbol{D}_{N,\varphi} + \boldsymbol{C}\boldsymbol{Y}_{N+1}\tilde{\boldsymbol{D}}_{N,\theta} \end{pmatrix}.$$

with $\boldsymbol{C} := \operatorname{diag}(\cos\theta_i)_{i=1}^M$ and corresponding sorting. Hence, the multiplication of the first part (4.12) and second one (4.13) of the Hessian with a vector can be realized in a fast way by applying (4.14) and (4.9).

We summarize our findings (4.12) - (4.15) in the following corollary.

COROLLARY 4.5. **(Efficient vector multiplication with $\mathrm{H}\mathcal{E}_N(\boldsymbol{p})$)**
For a given point $\boldsymbol{p} \in \mathcal{X}^M$ and given $\hat{w}_l$, $l = 1, \ldots, d_N$, the multiplication of a vector with the Hessian $\mathrm{H}\mathcal{E}_N(\boldsymbol{p})$ can be computed with the arithmetic complexity (4.6).

4.2. Nonlinear CG Algorithm. Among the various minimization strategies for weighted least squares functionals as the Newton method, the Levenberg-Marquardt algorithm and the nonlinear CG algorithm we restrict ourselves to the later one. The reason for this is that very good results were achieved by this method for the computation of spherical designs in [22]. Since we can mainly follow the lines of [22] we only briefly sketch the approach for our halftoning setting. For $\mathcal{X} = \mathbb{S}^1$ and $\mathcal{X} = \mathbb{T}^2$ we apply the nonlinear CG algorithm in the Euclidean space. A good survey of this topic was given in [24].

Algorithm: (CG algorithm in the Euclidean space)
Initialization: $\boldsymbol{p}^{(0)}$, $\boldsymbol{h}^{(0)} := \nabla\mathcal{E}_N(\boldsymbol{p}^{(0)})$, $\boldsymbol{d}^{(0)} = -\boldsymbol{h}^{(0)}$
For $r = 0, 1, \ldots$ repeat until a convergence criterion is reached

1. Determine the step size α_r by the search of a local minimum along the line $\boldsymbol{p}^{(r)} + t\boldsymbol{d}^{(r)}$, $t > 0$, i.e., by $(\boldsymbol{d}^{(r)})^{\mathrm{T}}\nabla\mathcal{E}_N\left(\boldsymbol{p}^{(r)} + \alpha_r\boldsymbol{d}^{(r)}\right) = 0$.
2. $\boldsymbol{p}^{(r+1)} := \boldsymbol{p}^{(r)} + \alpha_r\boldsymbol{d}^{(r)}$
3. $\boldsymbol{h}^{(r+1)} := \nabla\mathcal{E}_N(\boldsymbol{p}^{(r+1)})$
4. Compute β_r by

$$\beta_r := \frac{\langle \boldsymbol{h}^{(r+1)}, \mathrm{H}\mathcal{E}_N(\boldsymbol{p}^{(r+1)})\boldsymbol{d}^{(r)}\rangle}{\langle \boldsymbol{d}^{(r)}, \mathrm{H}\mathcal{E}_N(\boldsymbol{p}^{(r+1)})\boldsymbol{d}^{(r)}\rangle}.$$

5. $\boldsymbol{d}^{(r+1)} := -\boldsymbol{h}^{(r+1)} + \beta_r\boldsymbol{d}^{(r)}$

There exist other CG algorithms which differ by the choice of β_r. The above method is the one for exact conjugacy proposed by Daniel in [12]. In the numerical part we will also apply a CG variant, where the Hessian $\mathrm{H}\mathcal{E}_N$ is replaced by the approximate Hessian $\tilde{\mathrm{H}}\mathcal{E}_N$. Furthermore, we replace the first step of the algorithm by a one-dimensional Newton step, where the step size is determined by

$$\alpha_r := -\frac{\langle \boldsymbol{d}^{(r)}, \boldsymbol{h}^{(r)}\rangle}{\langle \boldsymbol{d}^{(r)}, \mathrm{H}\mathcal{E}_N(\boldsymbol{p}^{(r)})\boldsymbol{d}^{(r)}\rangle}. \tag{4.16}$$

Using the results from the previous subsection, we conclude that every CG iteration can be realized with the arithmetic complexity given in (4.6).

In the case $\mathcal{X} = \mathbb{S}^2$ we use the nonlinear CG algorithm on Riemannian manifolds $(\mathcal{M}, g_{\mathcal{M}})$ with Riemannian metric $g_{\mathcal{M}}$, see [16, 43]. In Riemannian geometry the addition of a tangent vector from $\mathrm{T}_x\mathcal{M}$ to the base point $x \in \mathcal{M}$ as required in step 2 of the CG algorithm is replaced by the *exponential map* $\exp_x : \mathrm{T}_x\mathcal{M} \to \mathcal{M}$. Furthermore, the translation of tangent vectors which is needed in steps 4 and 5 of the CG algorithm is replaced by the concept of *parallel transport of a vector along geodesics* which is itself based on the Levi-Civita connection.

In our applications we deal with $\mathcal{M} := (\mathbb{S}^2)^M$. In the following, we introduce the above concepts on $\mathbb{S}^2$ which generalize in a straightforward way to $(\mathbb{S}^2)^M$. For $x \in \mathbb{S}^2$ and $v \in \mathrm{T}_x\mathbb{S}^2$, we consider the geodesic curve $g : [0, T] \to \mathbb{S}^2$, $T > 0$ given by

$$g(t) = \cos(t\|v\|_2)\, x + \sin(t\|v\|_2) \underbrace{\frac{v}{\|v\|_2}}_{\tilde{v}}, \quad t > 0,$$

i.e., $g(0) = x$ and $\dot{g}(0) = v$. Note that the vectors $\tilde{v} := \frac{v}{\|v\|_2}$ and $x \times \tilde{v}$ form an orthonormal system of $\mathrm{T}_x\mathbb{S}^2$ and that any $w \in \mathrm{T}_x\mathbb{S}^2$ can be written as $w = \langle w, \tilde{v}\rangle\tilde{v} + \langle w, x \times \tilde{v}\rangle(x \times \tilde{v})$. The *exponential map* $\exp_x : \mathrm{T}_x\mathbb{S}^2 \to \mathbb{S}^2$ is explicitely parameterized by the geodesic due to $\exp_x(v) = g(1)$, see [49, p. 19]. Now the *parallel transport of a vector* $w \in \mathrm{T}_x\mathbb{S}^2$ *along the geodesic* g is realized by

$$\begin{aligned} P_{g(t)}(w) &:= \langle w, \tilde{v}\rangle \frac{\dot{g}(t)}{\|v\|_2} + \langle w, x \times \tilde{v}\rangle (g(t) \times \frac{\dot{g}(t)}{\|v\|_2}), \\ &= \langle w, \tilde{v}\rangle \left(\cos(\|v\|_2 t)\, \tilde{v} - \sin(\|v\|_2 t)\, x\right) + w - \langle w, \tilde{v}\rangle \tilde{v}, \quad t \geq 0. \end{aligned}$$

An illustration of the parallel transport is given in Figure 4.1.
After these preliminaries the CG algorithm to minimize $\mathcal{E}_N$ on $\mathcal{M} := (\mathbb{S}^2)^M$ reads as follows:

Algorithm: (CG algorithm on Riemannian manifolds)
Initialization: $\boldsymbol{p}^{(0)}$, $\boldsymbol{h}^{(0)} := \nabla\mathcal{E}_N(\boldsymbol{p}^{(0)})$, $\boldsymbol{d}^{(0)} = -\boldsymbol{h}^{(0)}$
For $r = 0, 1, \ldots$ repeat until a convergence criterion is reached

1. Determine the step size α_r by $(\boldsymbol{P}_{\boldsymbol{g}(\alpha_r)}\boldsymbol{d}^{(r)})^{\mathrm{T}} \nabla\mathcal{E}_N(\boldsymbol{g}(\alpha_r)) = 0$.
2. $\boldsymbol{p}^{(r+1)} := \exp_{\boldsymbol{p}^{(r)}}\left(\alpha_r \boldsymbol{d}^{(r)}\right)$
3. $\boldsymbol{h}^{(r+1)} := \nabla\mathcal{E}_N(\boldsymbol{p}^{(r+1)})$
4. Compute β_r by

$$\beta_r := \frac{\langle \boldsymbol{h}^{(r+1)}, \mathrm{H}\mathcal{E}_N(\boldsymbol{p}^{(r+1)})\tilde{\boldsymbol{d}}^{(r)}\rangle}{\langle \tilde{\boldsymbol{d}}^{(r)}, \mathrm{H}\mathcal{E}_N(\boldsymbol{p}^{(r+1)})\tilde{\boldsymbol{d}}^{(r)}\rangle}, \quad \tilde{\boldsymbol{d}}^{(r)} := \boldsymbol{P}_{\boldsymbol{g}(\alpha_r)}(\boldsymbol{d}^{(r)}).$$

5. $\boldsymbol{d}^{(r+1)} := -\boldsymbol{h}^{(r+1)} + \beta_r \tilde{\boldsymbol{d}}^{(r)}$

For an illustration of one CG iteration see Figure 4.1. Again, in our numerical examples, step 1 is replaced by a one-dimensional Newton step, where the step size α_r is determined by (4.16). Moreover, on the sphere $\mathbb{S}^2$ we prefer the approximate Hessian $\tilde{H}\mathcal{E}_N$ over the Hessian since this led to similar results but requires less computational effort. For a comparison of different minimization methods to produce spherical designs we refer to [22].

5. Numerical results. In the following, we present some numerical result on the torus $\mathbb{T}^2$ and the sphere $\mathbb{S}^2$. We apply our iterative optimization algorithm on the functional $\mathcal{E}_N(\boldsymbol{p})$ for randomly distributed starting points $\boldsymbol{p}^{(0)} \in \mathcal{X}^M$. To this end, we have to determine the Fourier coefficients $\hat{w}_l$ and λ_l of the function $w : \mathcal{X} \to \mathbb{R}_{\geq 0}$, cf. (4.1), and the bandlimited kernel $K_N : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, cf. (4.2), respectively. If the Fourier coefficients $\hat{w}_l$ are not given explicitely, we compute them approximately by well-known quadrature rules on $\mathcal{X}$. More precisely, we sample the function w on sampling points $\boldsymbol{x} := (x_i)_{i=1}^L \in \mathcal{X}^L$ and obtain approximate Fourier coefficients

$$\hat{w}_l := \sum_{i=1}^{L} \omega_i w(x_i)\overline{\psi_l(x_i)}, \qquad l = 1, \ldots, d_N, \tag{5.1}$$

where the weights ω_i are given such that

$$\int_{\mathcal{X}} f(x)\mathrm{d}x = \sum_{i=1}^{L} \omega_i f(x_i), \qquad f \in \Pi_N(\mathcal{X}).$$

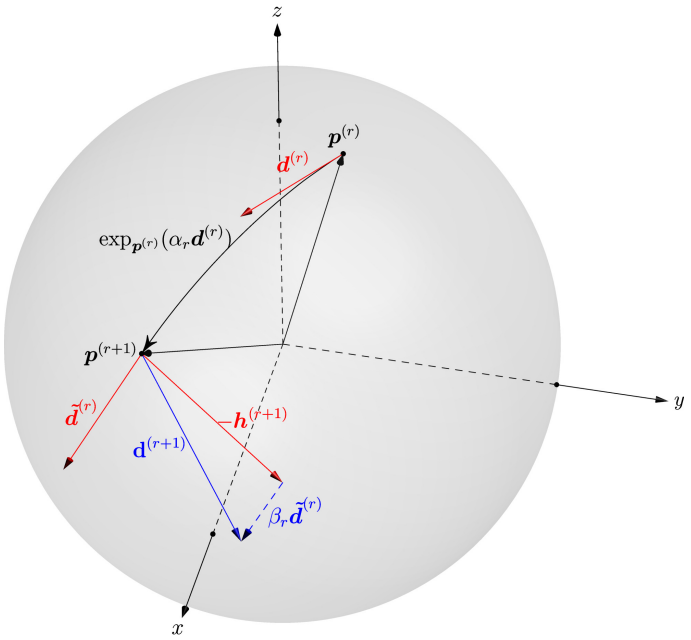

FIG. 4.1. *An iteration step of the nonlinear CG method on the sphere* $\mathbb{S}^2$.

We remind that the above sums can be evaluated in an efficient way by fast Fourier transforms.

In general we aim to use our method also for computing approximate minimizers of some arbitrarily given kernel K, which need not to be bandlimited. Then it is crucial to determine a good approximation, i.e., we need to know an appropriate bandwidth N of the truncation K_N. A quite simple heuristic rule is based on the following observation: For the constant function $w \equiv 1$, we are in the setting of standard quadrature rules. For example, on the torus $\mathbb{T}^2$ the standard Gauss quadrature with sampling nodes $\boldsymbol{x} := ((i/(N+1), j/(N+1))^{\mathrm{T}})_{i,j=0}^{N}$ and quadrature weights $\omega_{i,j} = 1/(N+1)^2$ is exact for trigonometric polynomials f up to degree N, i.e., $f \in \Pi_{2N}(\mathbb{T}^2)$. Obviously, these points $\boldsymbol{x}$ are minimizers of the functional $\mathcal{E}_{2N}$ for every kernel K_{2N}. Hence, for a sufficiently accurate kernel approximation a bandwidth $N \geq c\sqrt{M}$, $c \geq 2$ is required. Furthermore it is plausible that for regions of higher point densities a more accurate approximation is needed and we obtain for $w > 0$ the heuristic rule

$$N \geq c_{\mathcal{X}} \left(M \frac{\int_{\mathcal{X}} \mathrm{d}x}{\int_{\mathcal{X}} w(x)\mathrm{d}x} \max_{x \in \mathcal{X}} w(x) \right)^{1/d}, \qquad c_{\mathcal{X}} \geq 1,$$

where d is the dimension of the manifold $\mathcal{X}$. Of course, the 'optimal' constant $c_{\mathcal{X}}$ depends on the application and is given by a tradeoff between accuracy and computational costs. In the following, we have used bandwidths N which appear to be sufficiently large for good approximations.

The proposed algorithms are implemented in Matlab R2010a, where the mex-interface to the NFFT library [29] is used. The internal parameters in this library were set as follows: In both routines NFFT for $\mathbb{T}^2$ and NFSFT for $\mathbb{S}^2$ we set the cutoff parameter $m = 9$. In the NFSFT we set furthermore the threshold parameter $\kappa = 1000$ and used the flags PRE_PSI and PRE_PHI_HUT. The computations are performed on an Intel(R) Core(TM) i7 CPU 920 with 12GB RAM.

Examples on $\mathbb{T}^2$. For all test images we determine the Fourier coefficients $\hat{w}_l$ by the above mentioned Gauss quadrature rule. The underlying kernel is given by the bandlimited version of the approximated discrepancy kernel $\tilde{K}_{\mathcal{B}}$ with certain bandwidths, cf. (3.5). Moreover, in the CG algorithm we have used the Hessian of $\mathcal{E}_N$.

For a comparison of our results with other stippling and dithering methods we refer to the extensive experiments provided in [47]. Note that in contrast to [47] our assumption of periodic boundary conditions leads to some boundary artefacts.

Our **first example** deals with the image in Figure 1.1, left. The left image in Figure 5.1 shows the stippling result after $r = 500$ iterations with a kernel of bandwidth $N = 650$. At this stage the norm of the descent direction satisfies $\|\alpha_r \boldsymbol{d}^{(r)}\|_2 \approx$ 1e-3. This point distribution is far from being a local minimizer of $\mathcal{E}_{\tilde{K}_\mathcal{B}}$, but it looks quite nice, and the computation took less than 15min. The right image shows the result for the kernel with bandwidth $N = 1300$ after $r = 20000$ steps, where we have $\|\alpha_r \boldsymbol{d}^{(r)}\|_2 \approx$ 1e-11. One observes that the points are arranged in more regular patterns, but in this case the computation took about 1 day.

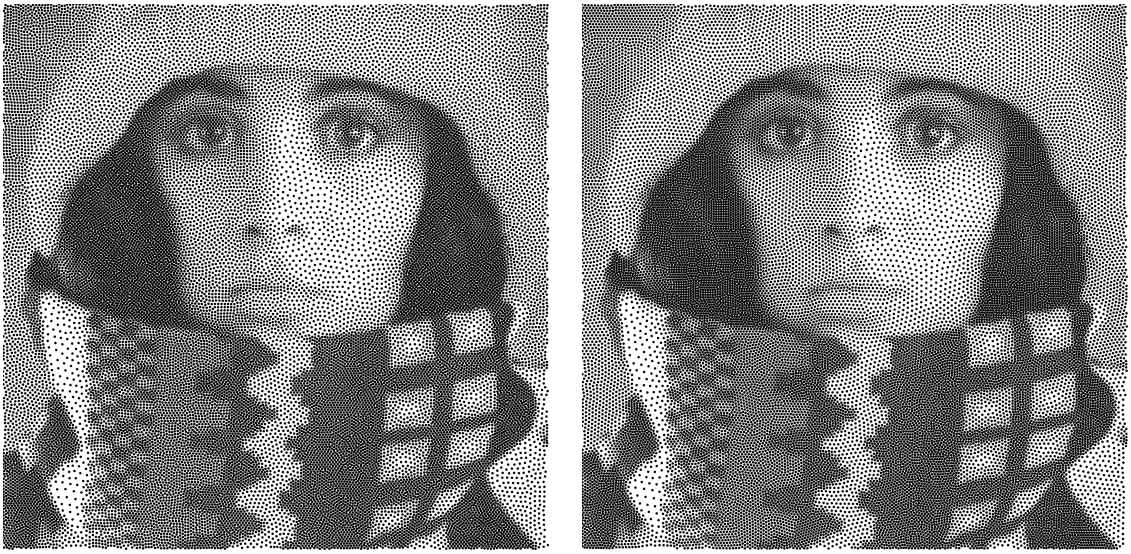

FIG. 5.1. *Halftoning results on the torus* $\mathbb{T}^2$ *for the image in Figure 1.1 with* $M = 30150$ *points. Left: kernel bandwidth* $N = 650$ *after* $r = 500$ *iterations. Right: kernel bandwidth* $N = 1300$ *after* $r = 20000$ *iterations.*

In the **second example** we consider a Gaussian weight w. Figure 5.2 depicts our stippling result for a kernel of bandwidth $N = 1300$ after $r = 20000$ iterations.

Examples on $\mathbb{S}^2$. In the following, we will use in our functional $\mathcal{E}_N$ the bandlimited version of the restricted kernel $\Phi(x - y) = -2\sin(\mathrm{d}_{\mathbb{S}^2}(x, y)/2)$, where the coefficients, cf. (4.4), are explicitly given by

$$\lambda_n = \frac{16\pi}{(2n+3)(2n+1)(2n-1)}, \quad n \in \mathbb{N}_0.$$

In the CG algorithm we apply the approximate Hessian of $\tilde{H}\mathcal{E}_N$, cf. (4.14).

The **first example** uses the topography map of the earth from Matlab. This map consists of the earth's elevation data. Since the values ranging from -7473 to 5731 we have scaled it to the range 0 to 1, in order to avoid negative values. The data is sampled on the grid $\boldsymbol{x} := (x\,(\pi i/180, \pi j/180))_{i=1,j=1}^{180,360}$. For this grid we have computed nonnegative quadrature weights $\omega_{i,j}$ for a polynomial degree $N = 179$ by the simple CG algorithm proposed in [21]. After applying the quadrature rule (5.1) we obtain a polynomial approximation $w = \sum_{n=0}^{179} \sum_{k=-n}^{n} \hat{w}_n^k Y_n^k$ of the earth's topography, see the left-hand side of Figure 5.3. We have applied our algorithm to $M = 200000$ random points $\boldsymbol{p} \in \mathcal{X}^M$ with a kernel of bandwidth $N = 1000$ and obtained after $r = 3600$ iterations the right image in Figure 5.3. Here an iteration takes about 1.5 min.

In our **second example**, we apply our halftoning procedure to three Gaussians on the sphere.

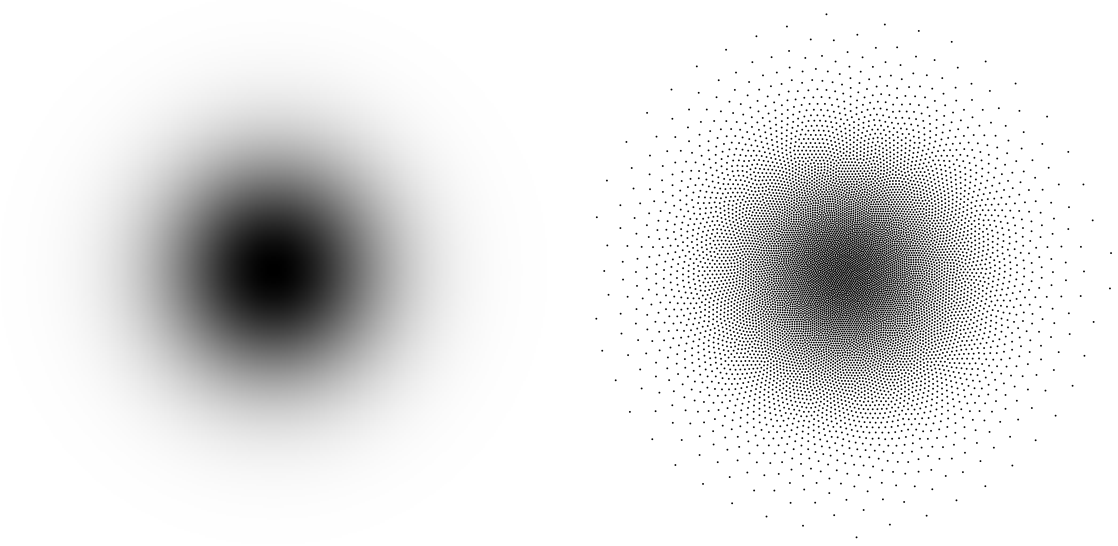

FIG. 5.2. *Left: Original Gaussian. Right: Halftoning result with* $M = 10023$ *points and kernel bandwidth* $N = 1300$ *after* $r = 20000$ *iterations.*

More precisely, the weight function is determined by

$$w(x) = \sum_{i=1}^{3} \exp(-5 \arccos(x \cdot q_i)^2),$$

where q_i are three orthonormal vectors which were produced by the Matlab command `[q,r]=qr(rand(3))`. Figure 5.4, right shows the result with $M = 5000$ points for a kernel of bandwidth $N = 400$ after $r = 1000$ iterations.

The **final example** is motivated by applications in geoscience. In [19] one is concerned with the problem of solving partial differential equations on the sphere by the method of radial basis functions. There the authors present an algorithm for placing sampling nodes adequately for some given partial differential equation to increase the accuracy and stability of the solvers. In a particular test case the nodes are distributed accordingly to the function

$$w(\theta, \varphi) = \begin{cases} \frac{\sqrt{3}}{2 \sin \theta} \operatorname{sech}^2(3 \sin \theta) \tanh(3 \sin \theta), & 0 < \theta < \pi \\ \frac{3\sqrt{3}}{2}, & \text{else,} \end{cases} \tag{5.2}$$

by a method based on electrostatic repulsion. In Figure 5.5 we see that our method produces similar point distributions as in [19]. Again we computed the Fourier coefficients $\hat{w}_n^k$, $n = 0, \ldots, 179$, $k = -n, \ldots, n$ as in the previous example. The result for $M = 1849$ points and kernel bandwidth $N = 400$ after $r = 1000$ iterations is presented in Figure 5.5.

Acknowledgment.. We like to thank E. Novak for pointing our attention to his interesting book [36]. MG and DP gratefully acknowledge the support by the German Research Foundation within the project PO 711/9-2. GS gratefully acknowledges partial support by the German Research Foundation, Grant STE 571/9-1.

REFERENCES

[1] M. Analoui and J. P. Allebach. Model-based halftoning using direct binary search. *Proc. SPIE*, 1666:96 – 108, 1992.
[2] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337 – 404, 1950.
[3] F. Aurehammer, F. Hoffmann, and B. Aronov. Minkowski-type theorems and least-squares clustering. *Algorithmica*, 20:61 – 76, 1998.

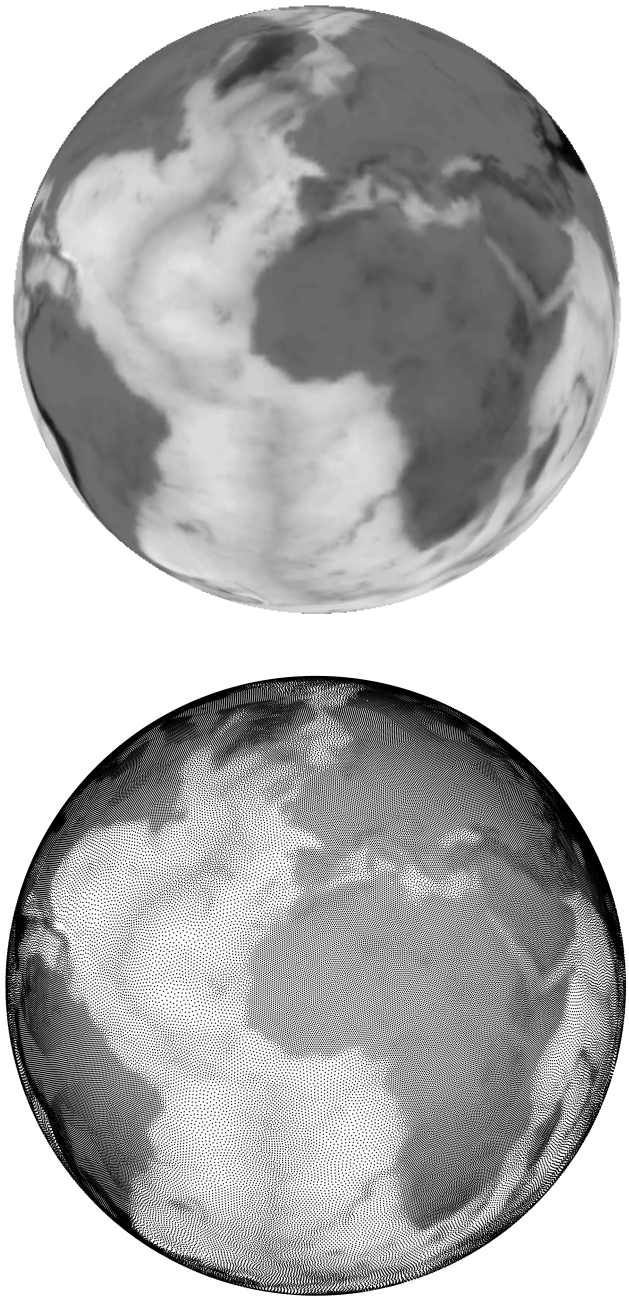

FIG. 5.3. *Top: Original image of the earth's topology adapted to our setting. Bottom: Halftoning result with with* $M = 200000$ *points for kernel bandwidth* $N = 1000$ *after* $r = 3600$ *iterations.*

[4] M. Balzer, T. Schlömer, and O. Deussen. Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Transactions on Graphics*, 28(3):Article 86, 2009.
[5] E. Bannai and E. Bannai. A survey on spherical designs and algebraic combinatorics on spheres. *European J. Combin.*, 30:1392 – 1425, 2009.
[6] F. A. Baqai and J. P. Allebach. Halftoning via direct binary search using analytical and stochastic printer models. *IEEE Trans. Image Process.*, 12(1):1 – 15, 2003.
[7] B. E. Bayer. An optimum method for two-level rendition of continuous-tone pictures. In *Conference Record, IEEE International Conference on Communications*, volume 1, pages (26–11)–(26–15), 1973.
[8] G. Beylkin. On the fast Fourier transform of functions with singularities. *Appl. Comput. Harmon. Anal.*, 2:363 – 381, 1995.

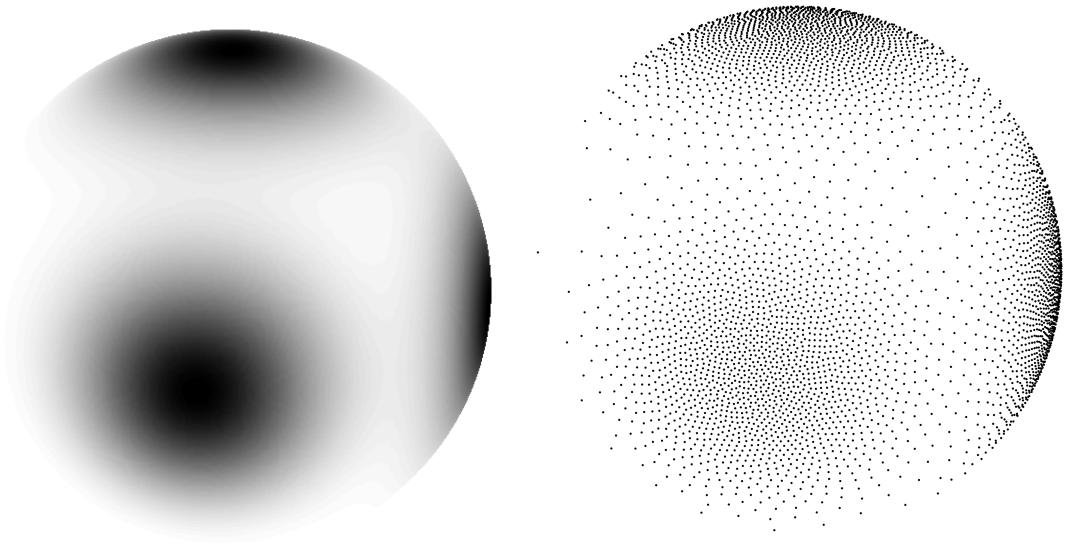

Fig. 5.4. *Left: Original Gaussians. Right: Halftoning result with* $M = 1000$ *points for kernel bandwidth* $N = 400$ *after* $r = 1000$ *iterations.*

[9] M. J. Bowick, D. R. Nelson, and A. Travesset. Interacting topological defects on frozen topographies. *Phys. Rev. B*, 62:8738–8751, 2000.

[10] J. Chang, B. Alain, and B. Ostromoukhov. Structure-aware error diffusion. *ACM Transactions on Graphics*, 28(5):Article 162, 2009.

[11] X. Chen, A. Frommer, and B. Lang. Computational existence proof for spherical t-designs. *Numer. Math.*, 117, 2010.

[12] J. W. Daniel. The conjugate gradient method for linear and nonlinear operator equations. *SIAM J. Numer. Anal.*, 4:10 – 26, 1967.

[13] P. Delsarte, J. M. Goethals, and J. J. Seidel. Spherical codes and designs. *Geom. Dedicata*, 6:363 – 388, 1977.

[14] J. R. Driscoll and D. Healy. Computing Fourier transforms and convolutions on the 2–sphere. *Adv. in Appl. Math.*, 15:202 – 250, 1994.

[15] A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data. *SIAM J. Sci. Stat. Comput.*, 14:1368 – 1393, 1993.

[16] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20:303 – 353, 1999.

[17] G. E. Fasshauer. *Meshfree approximation methods with MATLAB*. World Scientific Publishers, 2007.

[18] R. W. Floyd and L. Steinberg. An adaptive algorithm for spatial grey scale. *Proc. Society of Information Display*, 17:75 – 77, 1976.

[19] N. Flyer and E. Lehto. Rotational transport on a sphere: Local node refinement with radial basis functions. *J. Comput. Phys.*, 229:1954 – 1969, 2010.

[20] W. Freeden, T. Gervens, and M. Schreiner. *Constructive Approximation on the Sphere*. Oxford University Press, Oxford, 1998.

[21] M. Gräf, S. Kunis, and D. Potts. On the computation of nonnegative quadrature weights on the sphere. *Appl. Comput. Harmon. Anal.*, 27:124 – 132, 2009.

[22] M. Gräf and D. Potts. On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms. *TU Chemnitz, Fakultät für Mathematik, Preprint 10*, 2010.

[23] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Springer, LNM 1730, Berlin - Heidelberg - New York, 2000.

[24] W. W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.*, 2:35 – 58, 2006.

[25] M. Hanke-Bourgeois. *Grundlagen der numerischen Mathematik und des wissenschaftlichen Rechnens*. Vieweg + Teubner, Wiesbaden, third edition, 2009.

[26] D. M. Healy, P. J. Kostelec, and D. Rockmore. Towards Safe and Effective High-Order Legendre Transforms with Applications to FFTs for the 2-sphere. *Adv. Comput. Math.*, 21:59 – 105, 2004.

[27] J. F. Jarvis, C. N. Judice, and W. H. Ninke. A survey of techniques for the display of continuous tone pictures on bilevel displays. *Computer Graphics and Image Processing*, (5):13–40, 1976.

[28] J. Keiner, S. Kunis, and D. Potts. NFFT 3.0, C subroutine library. `http://www.tu-chemnitz.de/~potts/nfft`.

[29] J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.

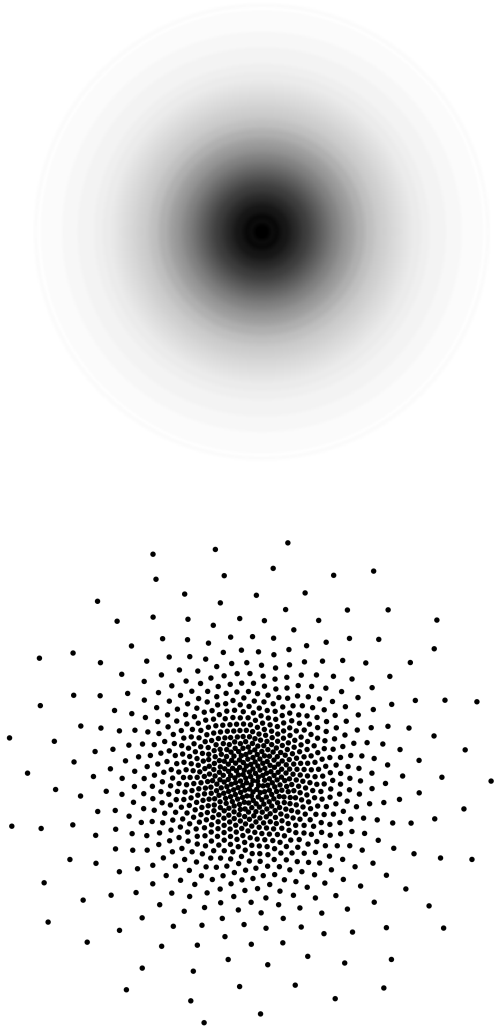

FIG. 5.5. *Left: Original weight w from* (5.2) *. Right: Distribution of $M = 1849$ points on the sphere $\mathbb{S}^2$ according to this weight function by our stippling algorithm with kernel bandwidth $N = 400$ after $r = 1000$ iterations.*

[30] J. Keiner and D. Potts. Fast evaluation of quadrature formulae on the sphere. *Math. Comput.*, 77:397 – 419, 2008.
[31] T. Kolling and A. Keller. Efficient illumination by high dynamic range images. In *Proceedings of the 14th Eurographics Workshop on Rendering*, volume 44 of *ACM International Conference Proceeding Series*, pages 45–50, 2003.
[32] A. B. J. Kuijlaars and E. B. Saff. Asymptotics for minimal discrete energy on the sphere. *Trans. Amer. Math. Soc.*, 350:523 – 538, 1998.
[33] S. Kunis and D. Potts. Fast spherical Fourier algorithms. *J. Comput. Appl. Math.*, 161:75 – 98, 2003.
[34] S. P. Lloyd. Least square quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129 – 137, 1982.
[35] C. Müller. *Spherical Harmonics*. Springer, Aachen, 1966.
[36] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems Volume II: Standard Information for Functionals*. Eur. Math. Society, EMS Tracts in Mathematics Vol 12, 2010.
[37] W.-M. Pang, Y. Qu, T.-T. Wong, D. Cohen-Or, and P.-A. Heng. Structure-aware halftoning. *ACM Transactions on Graphics*, 27:89:1 – 89:8, 2008.
[38] D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270, Boston, MA, USA, 2001. Birkhäuser.
[39] W. Purgathofer, R. F. Tobler, and M. Geiler. Forced random dithering: Improved threshold matrices for

ordered dithering. In *Proceedings of 1st IEEE International Conference on Image Processing*, volume 2, pages 1032–1035, Austin, Texas, 1994.

[40] C. Schmaltz, P. Gwosdek, A. Bruhn, and J. Weickert. Electrostatic halftoning. *Computer Graphics Forum*, 29:2313 – 2327, 2010.

[41] A. Secord. Weighted Voronoi stippling. In *Proceedings of the 2nd International Symposium on Non-Photorealistic Animation and Rendering*, pages 37–43. ACM Press, 2002.

[42] I. H. Sloan and R. S. Womersley. A variational characterisation of spherical designs. *J. Approx. Theory*, 159:308 – 318, 2009.

[43] S. T. Smith. Optimization techniques on Riemannian manifolds. In *Hamiltonian and gradient flows, algorithms and control*, volume 3 of *Fields Inst. Commun.*, pages 113 – 136. Amer. Math. Soc., Providence, RI, 1994.

[44] R. L. Stevenson and G. R. Arce. Binary display of hexagonally sampled continuous-tone images. *J. Opt. Soc. Amer. A*, 2(7):1009 – 1013, 1985.

[45] P. Stucki. MECCA - a multiple-error correcting computation algorithm for bilevel hardcopy reproduction. Tech. Report RZ1060, IBM Research Lab, Zürich, Switzerland, 1981.

[46] V. Surazhsky, P. Alliez, and C. Gotsman. Isotropic remeshing of surfaces: A local parameterization approach. In *Proceedings of the 12th International Meshing Roundtable*, pages 215–224, 2003.

[47] T. Teuber, G. Steidl, P. Gwosdek, C. Schmaltz, and J. Weickert. Dithering by differences of convex functions. *SIAM J. Imaging Sci.*, accepted.

[48] J. J. Thomson. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philos. Mag.*, 7:237 – 265, 1904.

[49] C. Udrişte. *Convex functions and optimization methods on Riemannian manifolds*, volume 297 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994.

[50] D. Vanderhaeghe and V. Ostromoukhov. Polyomino-based digital halftoning. In P. Isaías, editor, *IADIS International Conference on Computer Graphics and Visualization*, pages 11–18, 2008.

[51] D. Varshalovich, A. Moskalev, and V. Khersonskii. *Quantum Theory of Angular Momentum*. World Scientific Publishing, Singapore, 1988.

[52] G. Wahba. *Spline Models for Observational Data*. Springer, SIAM, 1990.

[53] D. J. Wales, H. McKay, and E. L. Altschuler. Defect motifs for spherical topologies. *Phys. Rev. B*, 79:224115, 2009.

[54] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, Cambridge, 2005.