Tikhonov Regularization of Circle-Valued Signals

Laurent Condat, Senior Member, IEEE

Abstract—It is common to have to process signals or images whose values are cyclic and can be represented as points on the complex circle, like wrapped phases, angles, orientations, or color hues. We consider a Tikhonov-type regularization model to smooth or interpolate circle-valued signals defined on arbitrary graphs. We propose a convex relaxation of this nonconvex problem as a semidefinite program, and an efficient algorithm to solve it.

Index Terms—circle-valued data, Tikhonov regularization, smoothing, convex relaxation, directional statistics

I. INTRODUCTION

In a wide range of applications, one has to deal with signals or images with cyclic, or circular, values, like phases, angles, orientations, or color hues, which are defined modulo $\pi$ or $2\pi$. Cyclic data appear, for instance, in interferometric synthetic aperture radar [1], color image restoration in HSV or LCh spaces, profilometry [2], Magnetic Resonance Imaging [3], biology, with data on the bacterial flagellar motor [4], in times series of wind directions [5], or in social sciences [6].

A cyclic value can be represented by a point on the complex circle; that is, a complex number of the form $e^{j\omega}$, for some phase $\omega \in \mathbb{R}$, where $j = \sqrt{-1}$. Equivalently, the value is represented by its wrapped phase $\omega \in (-\pi, \pi]$, and the signal presents artificial $2\pi$ jumps when the values cross the $\pi$ or $-\pi$ boundaries. Thus, to denoise or estimate circle-valued data, an option is to unwrap the phase map to remove these artificial discontinuities, by estimating the lost integer multiples of $2\pi$ in the phase values. Then the unwrapped signal or image can be processed using standard techniques for scalar data. Unfortunately, image unwrapping is a notoriously difficult problem [7]–[9], and the unwrapping process is prone to errors, so that it is preferable to process circle-valued data by keeping them on the circle.

We consider the general setting, where a signal is defined on a graph, with values located at the nodes. Two values are adjacent if there is an edge between their nodes. A 2-D image is a particular case with edges between every pair of neighboring pixels horizontally and vertically, forming a square grid. Then, to regularize signals on graphs, it is natural to promote the property that adjacent values are close to each other, in some sense. For scalar values, Tikhonov regularization consists in penalizing the squared differences of adjacent values and total variation (TV) regularization [10]–[12] consists, instead, in penalizing the absolute values of these differences. In this work, we focus on Tikhonov regularization for circle-valued signals. A few methods have been proposed for TV regularization of circle-valued signals [13]–[15]. There seems to be no available method for Tikhonov regularization of circle-valued data, with the exception of [14], where an iterative method based on the proximal point algorithm is proposed; it converges to the global solution for this type of problems on Hadamard manifolds, which the circle is not. Thus, we tackle the difficult nonconvex problem of Tikhonov regularization for circle-valued signals on graphs, by proposing a new convex relaxation.

The paper is organized as follows: in Section II, we discuss different formulations for the regularization of circle-valued signals. In Section III, we propose a convex relaxation of the considered nonconvex problem, and in Section IV, we propose an algorithm to solve it. In Section V, we illustrate the benefits of the proposed approach with several experiments.

II. TIKHONOV SMOOTHING FOR CIRCLE-VALUED SIGNALS

A. Circle-Valued Signals on Graphs

Let $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$ denote the complex unit circle.

We want to estimate a signal $x = (x_n)_{n \in V}$, with values $x_n \in \mathbb{S}$, defined on a connected undirected graph $(V, E)$, where $V$ is the set of nodes and $E$ is the set of edges, which are sets of two distinct nodes. Typically, we are given a noisy signal $y = (y_n)_{n \in V}$ defined on the same graph and the sought signal $x$ is a smoothed, or denoised, version of $y$, which achieves a tradeoff between closeness to $y$ and smoothness, in some sense. Another typical setting is interpolation, or inpainting: $y$ is defined on a subset $U \subset V$ of nodes and we want to estimate its missing samples; that is, $x$ is the smoothest signal defined on $V$ such that $x_n = y_n$, for every $n \in U$.

B. Classical Tikhonov Regularization

For real-valued signals, Tikhonov-regularized smoothing consists in solving the following convex optimization problem. Given $y = (y_n)_{n \in V}$ and nonnegative weights $(w_n)_{n \in V}$ and $(\lambda_{n,n'})_{(n,n') \in E}$, $x = (x_n)_{n \in V}$ is the solution to

$$\minimize_{x_n \in \mathbb{S}} \sum_{n \in V} \frac{w_n}{2} (x_n - y_n)^2 + \sum_{(n,n') \in E} \frac{\lambda_{n,n'}}{2} (x_n - x_{n'})^2.$$  

(1)

For the interpolating task with $y$ defined only on $U \subset V$, we want to solve, instead:

$$\minimize_{x_n \in \mathbb{S}} \sum_{(n,n') \in E} \frac{\lambda_{n,n'}}{2} (x_n - x_{n'})^2 \text{ s.t. } x_n = y_n, \forall n \in U.$$  

(2)

Formally, (2) can be viewed as a particular case of (1) with $w_n = \{+\infty \text{ if } n \in U, 0 \text{ otherwise}\}$, so that we can focus on the form (1), with the weights $w_n$ allowed to be $+\infty$. 

L. Condat is with the Visual Computing Center, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia. Contact: see https://lcondat.github.io/

Manuscript received XXX, 2021
We want to formulate an equivalent problem to (1) for signals \( x \) and \( y \) with values in \( \mathbb{S} \). Let us define the argument function \( \arg \), which maps \( z \in \mathbb{S} \) to \( \arg(z) \in (-\pi, \pi] \), such that \( z = e^{i\arg(z)} \). A natural adaptation to circle-valued signals is to replace the squared Euclidean distance \((t, t') \in \mathbb{R}^2 \mapsto (t - t')^2\) by the geodesic distance \((z, z') \in \mathbb{S}^2 \mapsto \min(|\arg(z) - \arg(z')|, 2\pi - |\arg(z) - \arg(z')|)\). This yields a nonconvex and nonsmooth, therefore very difficult, optimization problem to solve. In this work, we consider instead a statistical view of variables \( x \) for small deviations, the problems (3) and (1) behave similarly. The problem becomes:

\[
\min_{x_n \in \mathbb{S}: n \in \mathbb{V}} \sum_{n \in \mathbb{V}} w_n \left( 1 - \mathbb{R}(x_n, x^{\ast}_{n'}) \right) + \sum_{\{n, n'\} \in E} \lambda_{n, n'} \left( 1 - \mathbb{R}(x_n, x^{\ast}_{n'}) \right),
\]

where \( \mathbb{R} \) denotes the real part and \( * \) denotes the complex conjugation. Note that this problem is nonconvex for two reasons: the variables \( x_n \) are constrained to live in the nonconvex circle \( \mathbb{S} \) and the product \( \mathbb{R}(x_n, x^{\ast}_{n'}) \) is nonconvex. The second issue can be resolved by noticing that \( 1 - \mathbb{R}(x_n, x^{\ast}_{n'}) = \frac{1}{2} |x_n - x^{\ast}_{n'}|^2 \). Indeed \( |x_n - x^{\ast}_{n'}|^2 = |x_n x^{\ast}_{n'} - 1|^2 = (\mathbb{R}(x_n, x^{\ast}_{n'}) - 1)^2 + \Im(x_n x^{\ast}_{n'})^2 = 2 - 2\Re(x_n x^{\ast}_{n'}) \), where \( \Im \) denotes the imaginary part. Therefore, the problem (4) can be rewritten as:

\[
\min_{x_n \in \mathbb{S}: n \in \mathbb{V}} \sum_{n \in \mathbb{V}} w_n \left( |x_n - y_n|^2 \right) + \sum_{\{n, n'\} \in E} \lambda_{n, n'} \left( |x_n - x^{\ast}_{n'}|^2 \right),
\]

which is the natural extension of (1) in the complex plane, and where the objective function to minimize is convex; there remains the nonconvex circle constraint.

Another motivation for our model (4)–(5) is the following: suppose that \( y \) is a corrupted version of the unknown circle-valued signal \( x \) with complex Gaussian noise; that is, independent Gaussian noise with variance \( 1/w_n \) is added to the real and imaginary parts of each \( x_n \in \mathbb{S} \). Then, almost surely, the \( y_n \) are no longer in \( \mathbb{S} \). In that case, the MAP estimate of \( x \) is the natural extension of the solution to (5), where the \( y_n \) are now any complex numbers. Moreover, \( \frac{1}{2} |x_n - y_n|^2 = \frac{1}{2} (1 + |y_n|^2) - \Re(x_n y^{\ast}_{n'}) \), so that the problem (5) can be rewritten as:

\[
\min_{x_n \in \mathbb{S}: n \in \mathbb{V}} \sum_{n \in \mathbb{V}} w_n \left( \frac{1}{2} (1 + |y_n|^2) - \Re(x_n y^{\ast}_{n'}) \right)
+ \sum_{\{n, n'\} \in E} \lambda_{n, n'} \left( 1 - \Re(x_n x^{\ast}_{n'}) \right).
\]

The problem (6) generalizes (4) to any complex numbers \( y_n \), but since a constant value in the cost function to minimize does not change the solution, (4) and (6) are equivalent.

A natural idea to make the problem (5) convex is to remove the nonconvex circle constraint; we discuss this approach in Section II-F. Instead, we propose to stick with the formulation (6) and we propose a new convex relaxation of this problem, in Section III.

E. Limit Cases

Let us look at the two limit cases of (6), where the data-fit term overwhelms the regularization term, or the other way around. So, let us assume that all \( w_n \) are positive and that all \( \lambda_{n, n'} \) tend to zero. In that case, the solution is simply \( x_n = \{y_n | y_n \neq 0\} \) if \( y_n \neq 0 \), any point in \( \mathbb{S} \) otherwise, for every \( n \in \mathbb{V} \). The other limit case is more interesting: let us assume that all \( \lambda_{n, n'} \) tend to \( +\infty \). In that case, the regularization term is minimized, which means that the signal is constant: there exists \( x \in \mathbb{S} \) such that \( x_n = x \), for every \( n \in \mathbb{V} \). This point \( x \) minimizes the data-fit term \( \sum_{n \in \mathbb{V}} -w_n \Re(x y_n^{\ast}) = -\Re(x \sum_{n \in \mathbb{V}} w_n y^{\ast}_n) = -\Re(x x^{\ast}_{av}) \), where \( x_{av} = \sum_{n \in \mathbb{V}} w_n y_n \). The solution, called the (weighted) circular mean of the points, is \( x = \{x_{av} | x_{av} \neq 0\} \) if \( x_{av} \neq 0 \), any point in \( \mathbb{S} \) otherwise. That is, \( x \) is simply the weighted average of the \( y_n \), rescaled to be in \( \mathbb{S} \). Thus, when \( w_n \equiv 1 \), we have just recovered the well known property that the circular mean of a set of points on the circle is the maximum likelihood estimate of the mean for a von Mises distribution fitting the points.

F. Related Work

There is a large literature about optimization on manifolds. For instance, Bergmann and Tenbrinck [17] proposed
a generic approach for smoothing manifold-valued signals on graphs. Weimann et al. [14] proposed a method for a large class of functionals including the TV and Tikhonov costs, later extended to the Mumford–Shah functional for piecewise smooth reconstruction [18] and to the more general setting of inverse problems [19]. The Potts model can be used for the recovery of piecewise-constant signals [18]. In general, such approaches are heuristic and have no convergence guarantees; when an algorithm is proved to converge, this is typically to a local solution. Here we focus on estimating the exact global minimizer of the nonconvex problem (6). For this, we propose a convex relaxation of the problem, which is tight enough for its solution to coincide, not always but often in practice, with the one of the original problem; when this is the case, this can be certified.

Besides variational regularization, methods based on local averaging can be used. For instance, median filtering for circle-valued data has been proposed [20], which is robust to outliers. In the rest of this section, we focus on the structure tensor [21–25] a popular tool in image processing to analyze and process the local orientation of a vector field, typically the gradient field of an image. Identifying 2-D vectors with complex numbers, let us first recall that $z \in \mathbb{C}$ and $-z$ have same orientation. Defining the orientation as the angle $\omega = \arg(z)$, we can either restrict $\omega$ to the interval $[0, \pi)$, or define it on the real line modulo $\pi$; that is $\omega \in \mathbb{R}$ is the same orientation as $\omega + \pi$. Therefore, when processing orientations, to avoid cancellation effects that might happen when averaging numbers similar to $z$ with numbers similar to $-z$, it is better to multiply $\omega$ by two, so that it is in $(\pi/2, \pi)$, like a phase or angle; that is, one deals with $e^{j2\omega} = z^2$ instead of $z$. Thus, Tikhonov smoothing for orientations consists in solving (4) with the $y_n \in S$ replaced by $y_n^2$; then the smoothed values are the $\sqrt{M_n}$, to reverse the doubling operation.

Keeping this squaring effect for orientations in mind, the structure tensor method, to smooth a 2-D vector field identified with a complex-valued image $y$, works as follows: each value $y_n \in \mathbb{C}$ is mapped to the $2 \times 2$ real matrix of rank 1

$$ M_n = \begin{bmatrix} \Re(y_n)^2 & \Re(y_n)\Im(y_n) \\ \Re(y_n)\Im(y_n) & \Im(y_n)^2 \end{bmatrix}. $$

(7)

Note that $y_n$ and $-y_n$ are mapped to the same matrix, which is consistent with the discussion above: in this context, only the vector orientations matter, not their directions. Then the matrices $M_n$ are spatially averaged by applying a lowpass filter to them, elementwise. After filtering, the smoothed matrix $\overline{M}$ does not have rank 1, in general. Thus, the smoothed value $x_n$ is obtained by setting $[\Re(x_n) \Im(x_n)]^T$ as the principal eigenvector of $\overline{M}_n$. To understand this process, let us define $\alpha_n = \arg(y_n)$. We have, for every $n \in V$,

$$ M_n = \frac{|y_n|^2}{2} \begin{bmatrix} \cos(\alpha_n)^2 & \cos(\alpha_n)\sin(\alpha_n) \\ \cos(\alpha_n)\sin(\alpha_n) & \sin(\alpha_n)^2 \end{bmatrix} + \frac{|y_n|^2}{2} \begin{bmatrix} \cos(2\alpha_n) & \sin(2\alpha_n) \\ \sin(2\alpha_n) & -\cos(2\alpha_n) \end{bmatrix}. $$

(8)

After filtering, $\overline{M}_n$ is symmetric and can be written as

$$ \overline{M}_n = \begin{bmatrix} m_{n,1} & m_{n,3} \\ m_{n,3} & m_{n,2} \end{bmatrix} $$

$$ = d_n \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + a_n \begin{bmatrix} \cos(2\omega_n) & \sin(2\omega_n) \\ \sin(2\omega_n) & -\cos(2\omega_n) \end{bmatrix} $$

$$ = 2a_n \cos(\omega_n)^2 + d_n - a_n \begin{bmatrix} 2a_n \cos(\omega_n) \sin(\omega_n) \\ 2a_n \sin(\omega_n)^2 + d_n - a_n \end{bmatrix} $$

$$ = \begin{bmatrix} \cos(\omega_n) & \sin(\omega_n) \\ \sin(\omega_n) & -\cos(\omega_n) \end{bmatrix} \begin{bmatrix} d_n + a_n & 0 \\ 0 & d_n - a_n \end{bmatrix} \begin{bmatrix} \cos(\omega_n) & \sin(\omega_n) \\ \sin(\omega_n) & -\cos(\omega_n) \end{bmatrix}, $$

with $d_n = (m_{n,1} + m_{n,2})/2$, $a_n = ((m_{n,1} - m_{n,2})^2/4 + m_{n,3}^2)^{1/2} = |x_n|/2$, and $\omega_n = \arg(x_n)/2$, where we set

$$ x_n = (m_{n,1} - m_{n,2}) + 2jm_{n,3} = 2a_n e^{j2\omega_n}. $$

(10)

Thus, the two eigenvalues of $\overline{M}_n$, in decreasing order, are $d_n + a_n$ and $d_n - a_n$ and its principal eigenvector is $[\cos(\omega_n) \sin(\omega_n)]^T$. $a_n$ is a confidence indicator: if $a_n = 0$, there is no preferred direction locally, whereas if it is large, the direction $\omega_n$ is dominant. Thus, we do not need the matrix formalism: reasoning on the complex numbers $y_n$ and $x_n$ is equivalent and easier. Indeed, since $y_n^2 = |y_n|^2 e^{j2\omega_n} = (\Re(y_n)^2 - \Im(y_n)^2) + 2j\Re(y_n)\Im(y_n)$, $x_n$ is simply the result of spatial averaging applied to the $y_n^2$.

Tikhonov regularization amounts to lowpass filtering: with $w_n = 1$ and $\lambda_n, n' = \lambda$, the solution $x$ to (5) without the circle constraint is simply the result of a convolution applied to $y$, with inverse frequency response one plus $\lambda$ times the graph Laplacian. Therefore, the structure tensor method is essentially solving the Tikhonov problem (5), with the variables searched in $\mathbb{C}$ instead of $S$ and with the $y_n$ replaced by $y_n^2$; this second change being specific to the setting of orientations. Note that if $|y_n| = 1$ for all $n$, squaring the $y_n$ in (5), with the circle constraint, still corresponds to a MAP estimate with a scaled von Mises prior. However, if the amplitudes $|y_n|$ are arbitrary, squaring the $y_n$ also squares their amplitudes, so that the squared amplitudes are averaged by the regularization process; there seems to be no obvious Bayesian interpretation of (5), with or without the circle constraint, in that case. Thus, it is better to divide $M_n$ by $|y_n|$ in (7) before spatial averaging, which is the way the structure tensor is defined by Knutsson [24].

Finally, let us remark that if $|y_n| = 1$, when solving (5) without the circle constraint, the $x_n$ remain in the convex hull of the $y_n$, which is contained in the complex disk $D = \{z \in \mathbb{C} : |z| \leq 1\}$, the convex hull of $S$. Thus, there is no need to enforce the constraint that the $x_n$ belong to $D$, since it is automatically satisfied. In the sequel, we will refer to the following process, to find an approximate solution of (5), as the baseline method: (5) is solved without the circle constraint (which amounts to solving a linear system) and the $x_n$ are rescaled as $x_n/|x_n|$ afterwards, to make them lie in $S$. 


III. FOURIER LIFTING: CONVEX RELAXATION USING MOMENTS OF MEASURES

The method of moments -- There is a general recipe to reformulate, or lift, a nonconvex problem as a convex one: the minimization of a function $f$ over $\mathbb{S}$ is equivalent to minimizing $\int_{\mathbb{S}} f(z) d\mu(z)$ over $\mu$ in the set of probability measures (i.e., positive Borel measure with mass 1) on $\mathbb{S}$, assuming that $\text{arg}\min f$, the set of minimizers of $f$, is nonempty. This latter problem is convex, since it consists in minimizing a linear functional over a convex set. A minimizing measure $\mu^\ast$ will be concentrated over $\text{arg}\min f$; in particular, if the minimizer $z^\ast$ of $f$ is unique, $\mu^\ast = \delta_{z^\ast}$, the Dirac measure at $z^\ast$. This principled approach has a major downside, yet: the set of probability measures is infinite-dimensional, which prevents its numerical implementation in general. However, there is a case where the method can be implemented exactly: if the measure can be parameterized and recovered from a finite number $M$ of its moments $\mu_m = \int_{\mathbb{S}} f_m(z) d\mu(z)$, $m = 1, \ldots, M$, for some basis functions $\phi_m$, and if $f = \sum_{m=1}^{M} \alpha_m \phi_m$ is a linear combination of the $\phi_m$, then $\int_{\mathbb{S}} f(z) d\mu(z) = \sum_{m=1}^{M} \alpha_m \mu_m$, so that the problem becomes convex and finite-dimensional, in terms of the moments $\mu_m$: we want to minimize the linear term $\sum_{m=1}^{M} \alpha_m c_m$ with respect to the coefficients $c_m$, under the constraint that $c_m = \mu_m$ for every $m$, for some probability measure $\mu$. This approach is called the method of moments [26]. In this work, we use trigonometric moments, or Fourier coefficients. The characterization of a measure on the circle from a subset of its Fourier coefficients has a long history, rooted in Carathéodory’s work [27]; related theorems are often called Bochner’s theorems. In short, the constraint that the $c_m = \hat{\mu}_m$ for some positive measure $\mu$ on the circle is satisfied if the Toeplitz matrix formed by the $c_m$ is positive semidefinite [28], [29].

Convex relaxation of optimization over graphs using measures -- Our problem (6) features nonconvex pairwise costs $(x_n, x_{n'}) \mapsto \mathbb{R}(x_n, x_{n'})$. The minimization of a function $g(z, z')$ is equivalent to minimizing $\int_{\mathbb{S}} g(z, z') d\nu(z, z')$ over $\nu$ in the set of probability measures on $\mathbb{S}^2$. Hence, given an optimization problem over the graph $(V, E)$ with unary potential costs $f_n$ at the nodes and symmetric pairwise interaction costs $g_{n,n'}$ at the edges, all bounded from below and lower semicontinuous:

$$\min_{\nu \in \mathbb{S}^2} \sum_{n \in V} f_n(x_n) + \sum_{(n,n') \in E} g_{n,n'}(x_n, x_{n'})$$

we propose the following lifting technique: we introduce a probability measure $\nu_{n,n'}$ on $\mathbb{S}$ for each node $n \in V$, as well as a probability measure $\nu_{n,n'}$ on $\mathbb{S}^2$ for each edge $(n, n') \in E$, and we consider the lifted convex problem:

$$\min_{(\mu_n), (\nu_{n,n'})} \sum_{n \in V} \int_{\mathbb{S}} f_n(z) d\mu_n(z)$$

$$+ \sum_{(n,n') \in E} \int_{\mathbb{S}^2} g_{n,n'}(z, z') d\nu_{n,n'}(z, z')$$

s.t. the two marginals of $\nu_{n,n'}$ are $\mu_n$ and $\mu_{n'}$.

It is important to note that, in general, this convex relaxation is not tight: even if the solution $(x_n^\ast)_{n \in V}$ to (11) is unique, it is not guaranteed that the solution to (12) corresponds to $\mu_n^\ast = \delta_{x_n^\ast}$; it might be that measures which are not Diracs achieve a lower value of the objective function.

Interestingly, there is a strong connection with the theory of optimal transport [30]: the function

$$(\mu_n, \mu_{n'}) \mapsto \min_{\nu_{n,n'}} \int_{\mathbb{S}^2} g_{n,n'}(z, z') d\nu_{n,n'}(z, z')$$

s.t. the two marginals of $\nu_{n,n'}$ are $\mu_n$ and $\mu_{n'}$, is the Monge–Kantorovich optimal transport cost between the probability measures $\mu_n$ and $\mu_{n'}$, interpreting $g_{n,n'}(z, z')$ as the cost of moving one unit of mass from the point $z$ to the point $z'$; the minimizing measure $\nu_{n,n'}^\ast$ in (13), which exists, is called the optimal coupling measure. We refer to [31] for more details on optimal transport on the circle $\mathbb{S}$.

When the measures are restricted to live on a finite set of labels, instead of a continuous set like $\mathbb{S}$, the relaxation (12), which is a linear program, is well known in statistics, in the fields of graphical models, discrete inference and labeling, where it is called the local polytope relaxation [32]–[34]. In the present work, we do not want to discretize the circle $\mathbb{S}$. We will instead parameterize the measures by a finite number of their Fourier coefficients, like in the method of moments; hence, we name our convex relaxation approach as Fourier lifting.

Fourier lifting -- A probability measure $\nu$ defined on $\mathbb{S}^2$ has Fourier coefficients $\hat{\nu}_{m,m'} = \int_{\mathbb{S}^2} z^{-m} z'^{-m'} d\nu(z, z')$ for every $(m, m') \in \mathbb{Z}^2$. We have $\hat{\nu}_{m,-m'} = \hat{\nu}_{m,m'}$ and $\hat{\nu}_{0,0} = 1$. $\nu$ has two marginal measures $\mu = \int_{\mathbb{S}} d\nu(., z')$ and $\nu' = \int_{\mathbb{S}} d\nu(z, .)$, which are probability measures on $\mathbb{S}$. They have Fourier coefficients $\hat{\mu}_m = \int_{(-\pi, \pi)} e^{-j m \omega} \mu(e^{j \omega}) d\omega = \hat{\nu}_{m,0}$ and $\hat{\nu}_{m,m'} = \int_{(-\pi, \pi)} e^{-j (m_1 + n_2) \omega} \mu(e^{j \omega}) d\omega = \hat{\nu}_{m,m'}$ respectively, for every $(m, m') \in \mathbb{Z}^2$.

If $\nu$ is a 2-D Dirac in $(e^{j \omega_1}, e^{j \omega_2}) \in \mathbb{S}^2$, $c_{m,m'} = e^{-j (m_1 + n_2) \omega}$, $c_{m,n} = c_{n,m} \ast 1 = 1$ for every $(m, m')$. In this work, we will only parameterize $\nu$ using $c_{0,0} = 1$, $c_{1,0}$, $c_{0,1}$, $c_{-1,1}$. Indeed, if $\nu$ is a Dirac in $(e^{j \omega_1}, e^{j \omega_2})$, $c_{1,0} = e^{-j \omega_1}$, $c_{0,1} = e^{-j \omega_2}$, and

$$\cos(\omega_1 - \omega_2) = \int \cos(\omega - \omega') \nu(e^{j \omega}, e^{j \omega'}) d\omega, \omega'$$

$$= \frac{1}{2} (c_{1,-1} + c_{-1,1}) = \Re(c_{-1,1}).$$

Let us map these coefficients in the $3 \times 3$ Hermitian matrix

$$P = \begin{bmatrix}
1 & c_{1,0} & c_{0,1} \\
c_{1,0} & 1 & c_{-1,1} \\
c_{0,1} & c_{-1,1} & 1
\end{bmatrix}$$

$P$ is positive semidefinite, which we denote by $P \succ 0$, if and only if all its principal minors are nonnegative; that is,

$$\begin{bmatrix}
1 & c_{1,0} & c_{0,1} \\
c_{1,0} & 1 & c_{-1,1} \\
c_{0,1} & c_{-1,1} & 1
\end{bmatrix} \succeq 0,$$
\[ \begin{bmatrix} 1 & c_{0,1} \\ c_{0,1} & 1 \end{bmatrix} \geq 0, \quad \begin{bmatrix} 1 & c_{-1,1} \\ c_{-1,1} & 1 \end{bmatrix} \geq 0. \tag{18} \]

Equivalently, \(|c_{1,0}| \leq 1, |c_{0,1}| \leq 1, |c_{-1,1}| \leq 1, 1 + 2R(c_{1,0}c_{-1,1}c_{0,1}^2) - |c_{1,0}|^2 - |c_{0,1}|^2 - |c_{-1,1}|^2 \geq 0.

If \( \nu \) is a Dirac, \( P \gg 0 \). Thus, for every probability measure \( \nu \), by convexity of the positive semidefinite cone, \( P \gg 0 \). Moreover, \( P \gg 0 \) has rank 1 if and only if \( \nu \) is a Dirac; that is,

\[ P = \begin{bmatrix} 1 & c_{1,0} & c_{0,1} \\ c_{0,1} & 1 \end{bmatrix} \begin{bmatrix} 1 & c_{1,0} & c_{0,1} \end{bmatrix}, \tag{19} \]

with \(|c_{1,0}| = 1, |c_{0,1}| = 1|.

Hence, the proposed convex relaxation of (6) is:

\[
\begin{align*}
\text{minimize}_{x \in \mathbb{C}^N, n, n' \in \mathbb{C}} & \quad \Psi_{\text{conv}}(x, r) \\
\text{subject to} & \quad r_{n,n'} \in \mathbb{C}, (n, n') \in E \\
& \quad w_n \left( 1 + |y_n|^2 \right) - R(x_n y_n^*) + \lambda_{n,n'} \left( 1 - R(r_{n,n'}) \right) \geq 0, \forall (n, n') \in E. \tag{20}
\end{align*}
\]

A Hermitian positive semidefinite matrix with ones on its diagonal is a correlation matrix, the set of which is sometimes called an ellipse [35]. An ellipse is convex and compact. It is known that a linear function attains its minimum over a compact convex set at a point of its boundary. Since the rank-1 matrices associated to Diracs are extreme points of the ellipse, linear minimization over an ellipse is likely to yield a rank-1 matrix. The convex optimization problem (20) can be viewed as linear minimization over a product of ellipsoids with linear equality constraints; there is no guarantee that all matrices in (20) will be of rank 1 at a solution, but this is what we hope for.

Let us call \((x^*_n, r_{n,n'}) \in E \) and \((r^*_n, r_{n,n'}) \in E \) a solution obtained by solving (20) and \( \Psi_{\text{conv}} \), respectively, the corresponding minimal objective value. We also denote by \( \Psi_{\text{orig}} \) the minimal objective value of the original nonconvex problem (6). If all matrices appearing in (20) are of rank 1, or equivalently \(|x^*_n| = 1 \) for every \( n \in E \), and \( r_{n,n'} = r_{n,n'}^* \) for every \( (n, n') \in E \), \( \Psi_{\text{conv}} = \Psi_{\text{orig}} \) and we have obtained an exact solution of the original problem (6). Otherwise, we rescale the \( x_n^* \) as \( x_n^*/|x_n^*| \) and project them on \( \mathbb{S} \), and now we have an approximate solution \( x \) to (6). Let us denote by \( \Psi_{\text{approx}} = \Psi_{\text{orig}}(x) \) the objective value evaluated at this \( x \). We have \( \Psi_{\text{conv}} \leq \Psi_{\text{orig}} \leq \Psi_{\text{approx}} \), so that we can use \( \Psi_{\text{approx}} \) as a measure of relative suboptimality of the convex relaxation with respect to the original problem. We conjecture that the proposed relaxation is tight and yields the exact solution whenever the graph \((V, E)\) has no cycle, as is the case for a 1-D chain.

\[
V. \text{PROPOSED ALGORITHM}
\]

We endow \( \mathbb{C} \) with the inner product \( \langle z, z' \rangle = \Re(zz^*) \), to form a real Hilbert space. Then the problem (20) has the form

\[
\begin{align*}
\text{minimize}_{x \in \mathbb{C}^N} & \quad \Psi_{\text{conv}}(s, e) \equiv f(Ls), \tag{21}
\end{align*}
\]

where the variable \( s \) is the concatenation of all \( x_n \) and \( r_{n,n'} \), the dimension \( d \) is the total number of nodes and edges, \( e \in \mathbb{C}^d \) is the concatenation of all constants \(-w_n y_n\) and \(-\lambda_{n,n'}\), the linear operator \( L \) maps \( e \) to the concatenation of matrices

\[
(Le)_{n,n'} = \begin{bmatrix} 0 & x_n^* & x_{n,n'}^* \\ x_n & 0 & r_{n,n'} \\ x_{n,n'}^* & r_{n,n'} & 0 \end{bmatrix}
\]

for all \((n, n') \in E\). \( f(Q, Q') = \text{tr}(QQ') \), where \( \text{tr} \) denotes the trace.

To solve the problem (21), a well suited algorithm is the Proximal Method of Multipliers [36], [37], which, initialized with some variables \( U^{(0)} \in (\mathbb{C}^3 \times 3)^E \) and \( s^{(0)} \in \mathbb{C}^d \), consists in the iteration: for \( i = 0, 1, \ldots \)

\[
\begin{align*}
\alpha^{(i)} &= L \ast U^{(i)} + e \\
\gamma^{(i+1)} &= s^{(i)} - \tau \alpha^{(i)} \\
U^{(i+1)} &= \text{prox}_{d\tau}(U^{(i)}) + \sigma L_s^{(i+1)} - \tau \alpha^{(i)}
\end{align*}
\]

where \( L \ast \) denotes the adjoint operator of \( L \), \( f^* \) denotes the convex conjugate of \( f \) [38], \( \tau > 0 \) is a parameter, and we set \( \sigma = 1/(\|L\|^2 \tau) \), where the squared operator norm \( \|L\|^2 \) is twice the maximum number of edges per node. With this choice, the variable \( s^{(i)} \) in the algorithm converges to a solution \( s^* \) of (21) [37, Theorem 4.3]. In the algorithm, the proximity operator \( \text{prox}_{d\tau} \) maps each matrix \( Q_{n,n'} \) for \((n, n') \in E\), to the projection of \( Q_{n,n'} + \sigma I \) onto the cone of Hermitian negative semidefinite matrices, minus \( \sigma I \); this is achieved by computing the eigendecomposition and setting the positive eigenvalues to zero.

V. EXPERIMENTS

For the following experiments, MATLAB code implementing the algorithms and generating the images in the figures is available on the author’s webpage. The code was run in MATLAB R2022a on an Apple Macbook Pro 2019 laptop.

A. Denoising of a 1-D signal

In a first experiment, we denoise a 1-D signal of size \( N = 1000 \). That is, \( V = \{1, \ldots , N\} \) and \( E = \{(1, 2), \ldots , (N - 1, N)\} \). The ground-truth signal \( (e^{j\omega_n})_{n \in V} \) is generated using \( \omega_0 = 1 \) and i.i.d random increments \( \omega_{n+1} - \omega_n \) following the Gaussian law of standard deviation 0.1. Then the noisy signal \( y \) is formed by adding to the \( \omega_n \) white Gaussian noise of standard deviation \( \sqrt{1}/10 \), where \( \lambda = 50 \). These two signals are shown in Fig. 1 in green and black, respectively. We denoise \( y \) using the baseline method described at the end of Section II, with \( \omega_n \equiv 1 \) and \( \lambda_{n,n'} \equiv \lambda \); the denoised signal is shown in blue in Fig. 1. The corresponding cost value in (6),
or equivalently in (3), (4), or (5), is $\Psi_{\text{orig}}(x) \approx 227$. Then we apply the proposed approach by solving (20), with $\tau = 0.1$ in the algorithm, which converges to machine precision in about 300 iterations. The obtained denoised signal, shown in red in Fig. 1, satisfies $|x_n^\ast| = 1$ and $r_{n,n+1}^\ast = x_n^\ast x_{n+1}^\ast$ for every $n$, so that it is the exact solution to the original problem (6). The corresponding optimal cost value is $\Psi_{\text{conv}}^\ast = \Psi_{\text{orig}}^\ast \approx 226$. The quantitative and qualitative difference between the results of the baseline and proposed methods is small in this example, but it is satisfying to be able to solve the nonconvex problem of Tikhonov smoothing exactly.

B. Denoising of a 2-D image

In a second experiment, we denote a 2-D image: the phase of a smooth ground-truth image of size $97 \times 97$ is generated by cubic interpolation from a random $4 \times 4$ image and white Gaussian noise of standard deviation 0.5 is added to the phases to obtain a noisy version $y$; they are shown in Fig. 2 (a) and (b), respectively. The graph is the classical square grid: there is a node at each pixel and the edges connect all pairs of horizontally or vertically adjacent pixels. We denoise $y$ using the baseline method described at the end of Section II, with $\omega_n \equiv 1$ and $\lambda_{n,n'} \equiv \lambda = 5$ (with 400 iterations, computation time 0.09s); the denoised image is shown in Fig. 2 (c). The corresponding cost value in (6) is $\Psi_{\text{orig}}(x) \approx 2534$. Then we apply the proposed approach by solving (20), with $\tau = 0.1$ in the algorithm, which converges to machine precision in about 400 iterations (computation time 73s). The obtained denoised image, shown in Fig. 2 (d), satisfies $|x_n^\ast| = 1$ and $r_{n,n+1}^\ast = x_n^\ast x_{n+1}^\ast$ for every $(n,n')$, so that it is the exact solution to the original problem (6). The corresponding optimal cost value is $\Psi_{\text{conv}}^\ast = \Psi_{\text{orig}}^\ast \approx 2479$. By comparing the images in Fig. 2 (c) and (d), we can see that the image with the proposed method is a bit more regular, with less jagged level lines. We also show in Fig. 2 (e) the image obtained by replacing each pixel value by the weighted circular mean of its neighbors, with Gaussian weights, as described in Section II-E (computation time 0.003s); that is, we simply apply to $y$ a convolution with a Gaussian filter (of standard deviation 3 pixels) and we rescale each value $x_n$ to project it on the circle. The image is smooth and visually pleasant but its cost value is $\Psi_{\text{orig}}(x) \approx 2542$, similar to the one of the baseline method.

Our current implementation of the proposed algorithm is slow but it calls the eigendecomposition of every matrix to project it on the cone of positive semidefinite matrices. A careful implementation with a routine dedicated to this projection for $3 \times 3$ Hermitian matrices would reduce the computation time significantly.

It is not always the case that the proposed convex relaxation is tight and yields the solution to the original problem. For instance, if we keep the same experiment but with a noise standard deviation of 1 and $\lambda = 10$, the solution of (20) does not satisfy $|x_n^\ast| = 1$ for all $n$ any more. We have $\Psi_{\text{approx}} = 6797 > \Psi_{\text{conv}} \approx 5953$. The baseline method yields an image $x$ with a cost of $\Psi_{\text{orig}}(x) \approx 6749$. Both images are shown in Fig. 3. In particular, we can see that the proposed method has introduced an incorrect junction at the bottom right of the image in Fig. 3 (b). Elsewhere, it is more regular and probably closer to the exact solution of the problem (6) than with the baseline method. Thus, the proposed approach is best suited when the noise level is not too high.

C. Interpolation of a 1-D signal

We now consider interpolation at the intermediate indexes $n = 2, \ldots, 9$ of the 1-D signal $y$ defined at $n = 1$ and $n = 10$ by $y_1 = e^{-j}$ and $y_{10} = e^{2j}$. The problem we would like to solve is

$$\text{minimize}_{x_n \in S: n = 1, \ldots, 10} \sum_{n=1}^{9} \frac{1}{2} |x_{n+1} - x_n|^2 \quad \text{s.t.} \quad x_1 = e^{-j}, \ x_{10} = e^{2j}.$$  

(23)
TIKHONOV REGULARIZATION OF CIRCLE-VALUED SIGNALS

It has a closed form solution: the points are uniformly distributed on $S$, with $x_n = e^{\frac{j2\pi n}{10}}$; that is, the angles $\arg(x_n)$ linearly interpolate between $-1$ and $2$. The proposed convex relaxation is

$$\min_{x_n \in \mathbb{C}, n=1,\ldots,10} \sum_{n=1}^{9} \left| 1 - \Re(r_{n,n+1}) \right|$$

s.t. $x_1 = e^{-j}$, $x_{10} = e^{2j}$, and

$$\begin{bmatrix} 1 & x_n & x_{n+1} \\ x_n & 1 & r_{n,n+1} \\ x_{n+1} & r_{n,n+1} & 1 \end{bmatrix} \succeq 0, \forall n = 1, \ldots, 9.$$  \hspace{1cm} (24)

We solve the problem using the Chambolle–Pock algorithm [37], [40], which is similar to the algorithm shown in Section IV, with additional enforcement of $x_1 = e^{-j}$, $x_{10} = e^{2j}$ at every iteration. Again, it turns out that the relaxation (24) is tight and we obtain the exact solution to (23). The interpolated signal is shown in Fig. 4, in red, and is uniform on $S$, as predicted.

On the other hand, the baseline method consists in solving the convex problem

$$\min_{x_n \in \mathbb{C}, n=1,\ldots,10} \sum_{n=1}^{9} \left| x_{n+1} - x_n \right|^2$$

s.t. $x_1 = e^{-j}$, $x_{10} = e^{2j},$ \hspace{1cm} (25)

and then rescaling the obtained $x_n$ as $x_n/|x_n|$ to project them on $S$. We solved (25) using projected gradient descent, but the problem actually has a closed form solution, too: it is linear interpolation in $\mathbb{C}$, so that $x_n = \frac{n-1}{9}e^{2j} + \frac{10-n}{9}e^{-j}$. The interpolated signal is shown in Fig. 4, in blue. As we see, the angles of the $x_n$ are not uniform in $S$, so that the baseline method gives a bad approximate solution to the problem (23). This illustrates that the proposed convex relaxation is much tighter than the naive relaxation, which consists in reasoning in the disk $\mathbb{D}$ instead of the circle $S$.

VI. CONCLUSION

We proposed a new approach to smoothen or interpolate signals defined on the nonconvex complex circle, with a non-convex formulation translating Bayesian estimation with von Mises priors, and a convex relaxation based on semidefinite programming. We showed by experiments that the proposed relaxation is tight and yields the exact solution of the non-convex problem in several cases. This opens the door to solutions of better quality for many applications involving circular data.
