

# Super-Resolution of Point Sources Down to the Rayleigh Limit from Multiple Observations

March 29, 2019

## 1 Formal Setup

### 1.1 Single Observation Problem Statement

Suppose we observe a true signal whose representation in the physical domain is

$$x^*(t) = \sum_{j=1}^s a_j \delta_{t_j}(t), \quad (1)$$

for  $t \in \mathbb{T}$  where  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ , which we will usually think of as the unit interval  $[0, 1]$  with its endpoints identified, and  $T = \{t_j\}_{j=1}^s \subset \mathbb{T}$  a discrete support set. In the Fourier domain, this signal takes the form

$$\widehat{x^*}(k) = \sum_{j=1}^s a_j \exp(-2\pi i k t_j) \quad (2)$$

for  $k \in \mathbb{Z}$ .

Now, we are interested in recovering this true signal  $x$  from an observation that suffers from low resolution, which we represent as convolution of  $x(t)$  with a point-spread function (PSF) denoted by  $\phi(t)$ . We then observe the signal

$$(\phi * x^*)(t) = \sum_{j=1}^s a_j \phi(t - t_j). \quad (3)$$

In the simplest case,  $\phi$  is a Dirichlet kernel with cutoff frequency  $f_c$ , in which case the Fourier transform of the above is simply the truncation of the Fourier transform of  $x^*$ . Denoting by  $y$  the signal we observe in the Fourier domain, we have

$$y(k) = \left( \sum_{j=1}^s a_j \exp(-2\pi i k t_j) \right) \mathbb{1}_{\{|k| \leq f_c\}}. \quad (4)$$

We may then think of the data we observe as simply the  $2f_c + 1$  values  $y(-f_c), \dots, y(f_c)$ . We write  $\mathcal{F}_n$  for the sensing operator mapping  $x^*$  to these  $n$  Fourier coefficients. A popular technique for solving this problem is *total variation minimization*, which attempts to recover  $x^*$  by solving the

convex problem

$$\begin{aligned} & \text{minimize} \quad \|x\|_{\text{TV}} \\ & \text{subject to} \quad \mathcal{F}_n x = y. \end{aligned} \tag{5}$$

## 1.2 Extension to Multiple Observations

We now consider a generalization of the problem presented in the previous section, where we make *several* observations signals sharing the support  $T = \{t_1, \dots, t_s\}$  of  $x^*$ , but having varying amplitudes  $a_j$ . To formalize this, our true signal is now

$$x_\ell^*(t) = \sum_{j=1}^s a_{\ell,j} \delta_{t_j} \text{ for } \ell \in \{1, \dots, m\}, \tag{6}$$

and our observations are  $y_{\ell,k} = (\mathcal{F}_n x_\ell^*)_k$  for  $k \in \{-f_c, \dots, f_c\}$  and  $\ell \in \{1, \dots, m\}$ . We think of the  $a_{\ell,j}$  as organized into vectors  $\mathbf{a}_j \in \mathbb{R}^m$  for  $j \in \{1, \dots, s\}$ , and the  $y_{\ell,k}$  as organized into a matrix  $\mathbf{Y} \in \mathbb{C}^{m \times n}$ . The *group total variation minimization* is the natural extension of the previous convex problem to this setting, where we solve

$$\begin{aligned} & \text{minimize} \quad \|x\|_{\text{gTV}} \\ & \text{subject to} \quad \begin{bmatrix} \mathcal{F}_n x_1 & \mathcal{F}_n x_2 & \dots & \mathcal{F}_n x_m \end{bmatrix}^\top = \mathbf{Y}. \end{aligned} \tag{7}$$

## 1.3 Dual Certificates

In this section, we briefly describe the results of applying Lagrangian duality to TV and gTV norm minimization, the main tool for theoretical analysis of the performance of these algorithms.

**Definition 1** Let  $\mu^0 \subset \mathbb{C}$  denote the complex unit circle. A sign pattern on a set is an assignment of points of  $\mu^0$  to each point of the set.

**Definition 2** For a sign pattern  $v \in (\mu^0)^T$ , a low-pass trigonometric polynomial  $q : \mathbb{T} \rightarrow \mathbb{C}$ ,

$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{2\pi i k t}, \tag{8}$$

is a single-observation dual certificate for  $v$  if  $q(t_j) = v_j$  for  $t_j \in T$  and  $|q(t)| < 1$  for  $t \notin T$ .

**Proposition 1 (TV Norm Minimization Duality)** If a dual certificate for the sign pattern  $v_j = a_j/|a_j|$  exists, then  $x^*$  is the unique solution of TV minimization for the super-resolution problem.

Therefore, to prove the effectiveness of TV minimization it suffices to show that a dual certificate exists under some conditions on  $T$ . The main result of [1] establishes that this is true under a minimum separation condition.

**Definition 3** The minimum separation of a set  $T \subset \mathbb{T}$  is

$$\Delta(T) = \min_{\substack{s, t \in T \\ s \neq t}} |s - t|, \tag{9}$$

where we measure the wrap-around distance as on a circle.

**Theorem 1 (Proposition 2.3 of [1])** *If  $\Delta(T) \geq 1.26\lambda_c$  and  $f_c \geq 10^3$ , then a dual certificate exists for any sign pattern on  $T$ .*

As also observed in [1], the same ideas extend to the multiple observation case in a straightforward fashion.

**Definition 4** *Let  $\mu^{m-1} \subset \mathbb{C}^m$  denote the  $m$ -dimensional complex unit sphere. A  $m$ -dimensional sign pattern on a set is an assignment of points of  $\mu^{m-1}$  to each point of the set.*

**Definition 5** *For a sign pattern  $\mathbf{v} \in (\mu^{m-1})^T$ , a low-pass trigonometric polynomial  $q : \mathbb{T} \rightarrow \mathbb{C}^m$  given by  $q(t) = (q_1(t), \dots, q_m(t))$  and*

$$q_\ell(t) = \sum_{k=-f_c}^{f_c} c_k \exp(2\pi i k t) \text{ for } \ell \in \{1, \dots, m\} \quad (10)$$

*is an  $m$ -observation dual certificate for  $\mathbf{v}$  if  $q(t_j) = \mathbf{v}_j$  for  $t_j \in T$ , and  $\|q(t)\|_{\ell^2} < 1$  for  $t \notin T$ .*

**Proposition 2 (gTV Norm Minimization Duality)** *If a dual certificate for the sign pattern  $\mathbf{v}_j = \mathbf{a}_j / \|\mathbf{a}_j\|_{\ell^2}$  exists, then  $x^*$  is the unique solution of gTV minimization for the super-resolution problem.*

However, while numerical experiments in [1] suggest that when the amplitude vectors  $\mathbf{a}_j$  are taken randomly then recovery is possible down to a lower critical minimum separation which accumulates at  $\frac{1}{2}\lambda_c$  as  $m \rightarrow \infty$ , the same uniform argument cannot apply—indeed, it is always possible that the amplitudes  $\mathbf{a}_j$  are all identical, in which case clearly it cannot be possible to recover  $x^*$  any more effectively than in the one observation case. Thus, to show that gTV norm minimization substantially improves on TV norm minimization, it is necessary to make some incoherence assumptions about the data generating process of the  $\mathbf{a}_j$ . One simple such assumption that captures the idea of these points being in “general position” is to assume that the signs  $\mathbf{v}_j = \mathbf{a}_j / \|\mathbf{a}_j\|_{\ell^2}$  are i.i.d. isotropically random unit vectors.

## 2 Dual Certificate Construction

Recall that we are interested in building a low-pass trigonometric polynomial  $q : \mathbb{T} \rightarrow \mathbb{C}^m$  interpolating the points  $(t_j, \mathbf{v}_j)$  while remaining strictly inside the unit sphere in  $\mathbb{C}^m$  elsewhere. The general idea of such constructions is usually to choose some *interpolation basis* of polynomials, and build  $q$  as a linear combination of them under some constraints. We first describe a simple approach to choosing interpolation bases based on constraints and a variational heuristic, that justifies several existing techniques and motivates our own.

### 2.1 Equivalence of $L^2$ Norm Minimization and Interpolation

Minimizing  $\|q\|_{L^2([0,1])}$  under some interpolation constraints is equivalent to interpolating with a restricted basis, the basis arising from the nature of the constraints. This general principle can be

implemented in specific cases with simple Lagrange multiplier calculations. Consider, for instance, the problem

$$\begin{aligned} & \text{minimize} \quad \|q\|_{L^2([0,1])} \\ & \text{subject to} \quad q(t_j) = \mathbf{v}_j \quad \forall j \in [s] \end{aligned} \quad (11)$$

Let us write

$$q_{\mathbf{c}}(t) = \sum_{f=-f_c}^{f_c} c_f \exp(2\pi i f t), \quad (12)$$

and define the matrix  $\mathbf{F}_T \in \mathbb{C}^{s \times (2f_c+1)}$  by  $(\mathbf{F}_T)_{jf} = \exp(2\pi i f t_j)$ , with  $-f_c \leq f \leq f_c$ . Then, the calculation of the values of  $q_{\mathbf{c}}$  on  $T$  is a matrix-vector multiplication,

$$\begin{bmatrix} q_{\mathbf{c}}(t_1) \\ \vdots \\ q_{\mathbf{c}}(t_s) \end{bmatrix} = \mathbf{F}_T \mathbf{c}, \quad (13)$$

and  $\|q_{\mathbf{c}}\|_{L^2([0,1])} = \|\mathbf{c}\|_2$ . Thus, the above problem is equivalently

$$\begin{aligned} & \text{minimize} \quad \|\mathbf{c}\|_2 \\ & \text{subject to} \quad \mathbf{F}_T \mathbf{c} = \mathbf{v} \end{aligned} \quad (14)$$

We write a Lagrangian  $L(\mathbf{c}, \boldsymbol{\alpha}) = \frac{1}{2} \|\mathbf{c}\|_2^2 - \langle \boldsymbol{\alpha}, \mathbf{F}_T \mathbf{c} - \mathbf{v} \rangle$ , and upon differentiating and solving the first-order optimality condition obtain  $\mathbf{c} = \mathbf{F}_T^* \boldsymbol{\alpha}$ . Generally, we have

$$q_{\mathbf{F}_T^* \boldsymbol{\alpha}}(t) = \sum_{f=-f_c}^{f_c} \left( \sum_{j=1}^s \exp(-2\pi i f t_j) \alpha_j \right) \exp(2\pi i f t) = \sum_{j=1}^s \alpha_j K(t - t_j) \quad (15)$$

where  $K$  is the Dirichlet kernel. We therefore have found the formal equivalence:

$$\begin{aligned} & \text{minimize} \quad \|q\|_{L^2([0,1])} \\ & \text{subject to} \quad q(t_j) = \mathbf{v}_j \quad \forall j \in [s] \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} & \text{find} \quad q(t) = \sum_{j=1}^s \alpha_j K(t - t_j) \\ & \text{subject to} \quad q(t_j) = \mathbf{v}_j \quad \forall j \in [s] \end{aligned} \quad (16)$$

In the same way, one obtains a justification for the popular trick for allowing one to force the derivative of  $q$  to be zero at each  $t_j$  by adding terms  $\beta_j K'(t - t_j)$  to the sum above: an analogous calculation shows

$$\begin{aligned} & \text{minimize} \quad \|q\|_{L^2([0,1])} \\ & \text{subject to} \quad q(t_j) = v_j \quad \forall j \in [s] \\ & \quad \quad \quad q'(t_j) = 0 \quad \forall j \in [s] \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} & \text{find} \quad q(t) = \sum_{j=1}^s \alpha_j K(t - t_j) + \sum_{j=1}^s \beta_j K'(t - t_j) \\ & \text{subject to} \quad q(t_j) = v_j \quad \forall j \in [s] \\ & \quad \quad \quad q'(t_j) = 0 \quad \forall j \in [s] \end{aligned} \quad (17)$$

## 2.2 Relaxing the Derivative Condition

As mentioned before, in the multidimensional setting, we do not expect setting the derivative to equal zero at each interpolation point will allow us to lower the minimum separation threshold substantially beyond where the analogous one-dimensional dual certificate is feasible. Instead, we propose a new construction where this condition is relaxed. Rather than setting the derivative at each interpolation point to equal zero, we only constrain it to the tangent plane to the complex unit sphere at that interpolation point. [TODO: something about not dividing real from complex for clarity though this helps in numerics]. Following the ideas in the previous section, we propose as a dual certificate the polynomial solving the following optimization problem:

$$\begin{aligned} & \text{minimize} \quad \|q\|_{L^2([0,1])} \\ & \text{subject to} \quad q(t_j) = \mathbf{v}_j \quad \forall j \in [s] \\ & \quad \quad \langle q'(t_j), \mathbf{v}_j \rangle = 0 \quad \forall j \in [s] \end{aligned} \tag{L2-MIN-FULL}$$

Performing the translation to an interpolation problem as before gives the following:

$$\begin{aligned} & \text{find} \quad q_k(t) = \sum_{j=1}^s \alpha_{j,k} K(t - t_j) + \sum_{j=1}^s \beta_j V_{j,k} K'(t - t_j) \\ & \text{subject to} \quad q_k(t_j) = V_{j,k} \quad \forall j \in [s], k \in [m] \\ & \quad \quad \langle q'(t_j), \mathbf{v}_j \rangle = 0 \quad \forall j \in [s] \end{aligned} \tag{INT-FULL}$$

where we organize the sign pattern vectors  $\mathbf{v}_j$  as the rows of the matrix  $\mathbf{V} \in \mathbb{C}^{s \times m}$ . Note that if we instead took the constraint  $q'(t_j) = \mathbf{0}$ , then the second term would involve terms  $\beta_{j,k} K'(t - t_j)$  in coordinate  $k$ , but here we are only allowed *one* set of coefficients  $\beta_j$  to use across all coordinates. This, along with the randomness of the  $V_{j,k}$ , essentially ensures that the second term cannot often be large: a single vector  $\boldsymbol{\beta}$  cannot align with many weakly dependent random vectors  $\mathbf{V}_{\bullet,k}$  at once.

## 2.3 Explicit Solution

Next, we derive concise formulae solving for the coefficients  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$  (when this is possible, which will later need some justification). Let

$$\mathbf{K}_r := \left( \frac{d^r}{dt^r} K(t_i - t_j) \right)_{i,j=1}^s. \tag{18}$$

Then, we may write

$$(q_k(t_j))_{j \in [s], k \in [m]} = \mathbf{K}_0 \boldsymbol{\alpha} + \mathbf{K}_1 \text{diag}(\boldsymbol{\beta}) \mathbf{V} \tag{19}$$

$$(q'_k(t_j))_{j \in [s], k \in [m]} = \mathbf{K}_1 \boldsymbol{\alpha} + \mathbf{K}_2 \text{diag}(\boldsymbol{\beta}) \mathbf{V} \tag{20}$$

Thus, the interpolation equations may be written

$$\mathbf{K}_0 \boldsymbol{\alpha} + \mathbf{K}_1 \text{diag}(\boldsymbol{\beta}) \mathbf{V} = \mathbf{V} \tag{21}$$

$$\text{diag}((\mathbf{K}_1 \boldsymbol{\alpha} + \mathbf{K}_2 \text{diag}(\boldsymbol{\beta}) \mathbf{V}) \mathbf{V}^*) = \mathbf{0} \tag{22}$$

From the first equation we obtain

$$\boldsymbol{\alpha} = \mathbf{K}_0^{-1}(\mathbf{I}_s - \mathbf{K}_1 \text{diag}(\boldsymbol{\beta}))\mathbf{V}, \quad (23)$$

and substituting this into the second obtain

$$\boldsymbol{\beta} = \frac{\text{diag}(\mathbf{V}\mathbf{V}^*\mathbf{K}_0^{-1}\mathbf{K}_1)}{\text{diag}(\mathbf{V}\mathbf{V}^*\mathbf{K}_2) - \text{diag}(\mathbf{V}\mathbf{V}^*\mathbf{K}_1\mathbf{K}_0^{-1}\mathbf{K}_1)}, \quad (24)$$

where division in the latter is coordinatewise, and these formulae hold provided  $\mathbf{K}_0$  is invertible and  $\text{diag}(\mathbf{V}\mathbf{V}^*(\mathbf{K}_2 - \mathbf{K}_1\mathbf{K}_0^{-1}\mathbf{K}_1))$  is nowhere zero.

## 2.4 Pointwise Evaluation

Let

$$\mathbf{k}_r(t) := \left( \frac{d^r}{dt^r} K(t - t_j) \right)_{j=1}^s. \quad (25)$$

$$\mathbf{q}(t)^\top = \mathbf{k}_0(t)^\top \boldsymbol{\alpha} + \mathbf{k}_1(t)^\top \text{diag}(\boldsymbol{\beta})\mathbf{V} \quad (26)$$

$$= \mathbf{k}_0(t)^\top \mathbf{K}_0^{-1}(\mathbf{I}_s - \mathbf{K}_1 \text{diag}(\boldsymbol{\beta}))\mathbf{V} + \mathbf{k}_1(t)^\top \text{diag}(\boldsymbol{\beta})\mathbf{V} \quad (27)$$

$$= \mathbf{k}_0(t)^\top \mathbf{K}_0^{-1}\mathbf{V} + \left( \mathbf{k}_1(t)^\top - \mathbf{k}_0(t)^\top \mathbf{K}_0^{-1}\mathbf{K}_1 \right) \text{diag}(\boldsymbol{\beta})\mathbf{V} \quad (28)$$

[NEXT] Is triangle equality ok in the above?

## 3 Special Case: Orthogonal Signs

Consider the special case where the  $\mathbf{v}_j$  are orthogonal, and for the sake of simplicity  $s = m$ , so that they form a complete basis. Then, without loss of generality we further consider the case where the  $\mathbf{v}_j$  are rotated to equal the standard basis, so that  $\mathbf{V} = \mathbf{I}_s$ . One may then simplify the equations from before, but it is even easier to note directly that the equations governing each polynomial coordinate  $q_k(t)$  become decoupled: we are now interested in the interpolation problem

$$\begin{aligned} \text{find } q_k(t) &= \sum_{j=1}^s \alpha_{j,k} K(t - t_j) + \beta_k K'(t - t_k) \\ \text{subject to } q_k(t_j) &= \delta_{j,k} \quad \forall j \in [s], k \in [m] \\ q'_k(t_k) &= 0 \quad \forall k \in [m] \end{aligned} \quad (\text{INT-ORTH})$$

Note that this decomposes into  $s$  independent interpolation problems, each of the form

$$\begin{aligned} \text{find } q(t) &= \sum_{j=1}^s \alpha_j K(t - t_j) + \beta K'(t - t_1) \\ \text{subject to } q(t_j) &= \delta_{1,j} \quad \forall j \in [s] \\ q'(t_1) &= 0 \end{aligned} \quad (\text{INT-1D})$$

### 3.1 Explicit Solution

From the second constraint we may directly solve

$$\beta_k = -\frac{1}{K''(0)} \sum_{j=1}^s \alpha_{j,k} K'(t_k - t_j),$$

and therefore

$$q_k(t) = \sum_{j=1}^s \alpha_{j,k} \left( K(t - t_j) - \frac{K'(t_k - t_j)K'(t - t_k)}{K''(0)} \right).$$

We define the matrix corresponding to this linear system,  $\mathbf{M}_k$ , by

$$(\mathbf{M}_k)_{ij} = K(t_i - t_j) + \frac{K'(t_k - t_j)K'(t_k - t_i)}{K''(0)},$$

so that, writing  $\alpha_k = \alpha_{\bullet,k}$ , we find  $\alpha_k = \mathbf{M}_k^{-1} \mathbf{e}_k$ . We may further decompose

$$\mathbf{M}_k = \mathbf{K}_0 + \Delta_k,$$

where

$$(\Delta_k)_{ij} = \frac{K'(t_k - t_j)K'(t_k - t_i)}{K''(0)}.$$

This matrix is rank one:

$$\Delta_k = \frac{1}{K''(0)} \kappa_{1,k} \kappa_{1,k}^\top \tag{29}$$

$$(\kappa_{1,k})_i = K'(t_k - t_i) \tag{30}$$

Therefore,  $\mathbf{M}_k^{-1}$  may be computed by the Sherman-Morrison formula:

$$\mathbf{M}_k^{-1} = \mathbf{K}_0^{-1} - \frac{1}{K''(0) + \kappa_{1,k}^\top \mathbf{K}_0^{-1} \kappa_{1,k}} \mathbf{K}_0^{-1} \kappa_{1,k} \kappa_{1,k}^\top \mathbf{K}_0^{-1},$$

and we obtain the formula

$$\alpha_k = \mathbf{M}_k^{-1} \mathbf{e}_k = \mathbf{K}_0^{-1} \mathbf{e}_k - \frac{\kappa_{1,k}^\top \mathbf{K}_0^{-1} \mathbf{e}_k}{K''(0) + \kappa_{1,k}^\top \mathbf{K}_0^{-1} \kappa_{1,k}} \mathbf{K}_0^{-1} \kappa_{1,k},$$

expressing  $\alpha_k$  as a perturbation of  $\mathbf{K}_0^{-1} \mathbf{e}_k$ , which is what we would obtain if we were simply doing Dirichlet kernel interpolation on the sign pattern  $\mathbf{e}_k$ . Indeed, if only the first term were present, we would simply have  $\alpha \approx \mathbf{K}_0^{-1}$ .

### 3.2 Approximation of $\mathbf{K}_0^{-1}$

We build an approximation of  $\mathbf{K}_0^{-1}$  column by column. Without loss of generality, suppose we are interested in approximating the second column of  $\mathbf{K}_0^{-1}$ . Our approximation will consist in estimating the first, second, and third entries, and claiming that the remaining entries must be

small. Let us write

$$\mathbf{K}_0 = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}, \quad (31)$$

with  $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ ,  $\mathbf{B} \in \mathbb{R}^{3 \times (s-3)}$ , and  $\mathbf{C} \in \mathbb{R}^{(s-3) \times (s-3)}$ . Inverting the block matrix, we find that the first three columns of the inverse are given by

$$\mathbf{K}_0^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{B} (\mathbf{K}_0 / \mathbf{A})^{-1} \mathbf{B}^\top \mathbf{A}^{-1} & \cdots \\ -(\mathbf{K}_0 / \mathbf{A})^{-1} \mathbf{B}^\top \mathbf{A}^{-1} & \cdots \end{bmatrix}, \quad (32)$$

where  $\mathbf{K}_0 / \mathbf{A} = \mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}$  is the Schur complement. We will argue that the terms besides  $\mathbf{A}^{-1}$  are small by showing that  $\mathbf{B}$  is small and the other terms are not too large. We have

$$\|\mathbf{B}\|_\infty = \max_{i \in \{1,2,3\}} \sum_{j=1}^{n-3} |B_{ij}| \quad (33)$$

### 3.3 Three-Point Approximation

We now introduce a means of approximating the solution to (INT-1D). The idea is to solve a truncated version of the same problem in closed form, and hope that the resulting solution is close to the true solution of (INT-1D). The truncated problem is as follows, obtained by simply ignoring all constraints on points other than  $t_s, t_1, t_2$  (the three-point neighborhood surrounding  $t_1$ , where the interpolant must equal 1). Here, it is convenient to assume (without loss of generality) that  $t_1 = 0$ , and to rename  $t_s = -\Delta_-$  and  $t_2 = \Delta_+$  where  $\Delta_\pm > 0$ , for the sake of more symmetrical notation.

$$\begin{aligned} \text{find } q(t) &= \alpha_0 K(t) + \alpha_+ K(t - \Delta_+) + \alpha_- K(t + \Delta_-) + \beta K'(t) \\ \text{subject to } q(0) &= 1 \\ q(\Delta_+) &= q(-\Delta_-) = 0 \\ q'(0) &= 0 \end{aligned} \quad (\text{INT-1D-TRUNC})$$

The coefficients of the solution solve the following linear system of equations in four variables:

$$\begin{bmatrix} K(0) & K(\Delta_+) & K(\Delta_-) & K'(0) \\ K(\Delta_+) & K(0) & K(\Delta_+ + \Delta_-) & K'(\Delta_+) \\ K(\Delta_-) & K(\Delta_+ + \Delta_-) & K(0) & -K'(\Delta_-) \\ K'(0) & K'(\Delta_+) & -K'(\Delta_-) & K''(0) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_+ \\ \alpha_- \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (34)$$

We may solve this in closed form, obtaining an explicit (if complicated) function that gives the solution:

$$q^{3\text{pt}}(\Delta_+, \Delta_-; t) : [0, 1] \times [0, 1] \times \mathbb{T} \rightarrow \mathbb{R} \quad (35)$$

Let us similarly set some notation for the solution of (INT-1D):

$$q^{1\text{D}}(T; t) : [0, 1]^s \times \mathbb{T} \rightarrow \mathbb{R}, \quad (36)$$



then our main hypothesis is that the following approximation is sound: suppose  $T = \{0 = t_1, t_2, \dots, t_s\}$  with these points ordered along  $\mathbb{T}$ , then

$$q^{1D}(\{0 = t_1, t_2, \dots, t_s\}; \bullet) \approx q^{3pt}(|t_2|, |t_s|; \bullet). \quad (37)$$

## 4 Further Special Case: Equally Spaced Points

Consider further a special case where the  $t_j$  are all equally spaced in  $\mathbb{T}$ , so that  $\Delta = 1/s$ , and every pair of  $t_j, t_{j+1}$  achieves this spacing exactly in the wraparound metric. The point of this simplification is to induce some extra structure in the matrix  $\mathbf{K}_0$  to further ease the formal analysis performed in the previous section. If we do this, then if we have  $1/s = \Delta = \gamma\lambda_c = \gamma/f_c$ , we must have  $f_c = \gamma s$  is an integer.

Indeed, in this case we have  $t_i - t_j = (i - j)\Delta$  (with equality modulo 1), and therefore

$$(\mathbf{K}_0)_{ij} = K((i - j)\Delta),$$

so  $\mathbf{K}_0$  is circulant. In case  $K$  is the Dirichlet kernel, we have  $\mathbf{K}_0 = \mathbf{F}\mathbf{F}^*$  where  $\mathbf{F} \in \mathbb{C}^{s \times (2f_c + 1)}$  having, assuming  $t_1 = \Delta$  as we may without loss of generality,

$$F_{jf} = \exp(2\pi i t_j f) = \exp(2\pi i j f \Delta) = \exp\left(2\pi i \frac{jf}{s}\right).$$

That is,  $\mathbf{F}$  is nothing but a discrete Fourier transform (DFT) matrix with some redundant frequencies included. In this formulation, we see immediately a simple significance of the threshold  $\gamma = \frac{1}{2}$ : it is the threshold below which  $\mathbf{F}$  is no longer a complete DFT matrix in dimension  $s$ .

In this case, the inverse of  $\mathbf{K}_0$  is related to  $\mathbf{K}_0$  itself, which may be seen by computing  $\mathbf{F}^*\mathbf{F}$ , where we see that the usual orthogonality of Fourier harmonics of order  $s$  appears:

$$(\mathbf{F}^*\mathbf{F})_{fg} = \sum_{j=0}^{s-1} \exp\left(2\pi i \frac{j(g-f)}{s}\right) = s[f \equiv g \pmod{s}].$$

We have  $(\mathbf{K}_0)_{jj} = 2f_c + 1$ , and

$$\begin{aligned} (\mathbf{K}_0)_{jk} &= \sum_{f=-f_c}^{f_c} \exp\left(2\pi i \frac{(k-j)f}{s}\right) \\ &= \frac{\sin((f_c + \frac{1}{2})(k-j))}{\sin(\frac{1}{2}(k-j))} \end{aligned}$$

## 5 Local Analysis

**Proposition 3** *Let  $q : [0, 1] \rightarrow \mathbb{C}^m$  be a smooth function. Write  $N(t) = \|q(t)\|^2 = \sum_{k=1}^m |q_k(t)|^2$ . Then,*

$$N'(t) = 2\operatorname{Re}(\langle q(t), q'(t) \rangle) \quad (38)$$

$$N''(t) = 2\operatorname{Re}(\langle q(t), q''(t) \rangle) + 2\|q'(t)\|^2. \quad (39)$$

Thus, for the local analysis of  $N(t)$  near one of the  $t_j$  in the special case of orthogonal signs, we have

$$N(t_k) = 1 \quad (40)$$

$$N'(t_k) = 0 \quad (41)$$

$$N''(t_k) = 2\text{Re}(q_k''(t_k)) + 2 \sum_{j \neq k} |q_j'(t_k)|^2 \quad (42)$$

Thus, the local concavity of  $N(t)$  near  $t_k$  is a battle between, on the one hand, the concavity of  $q_k$  near  $t_k$ , and on the other hand the total magnitude of the derivatives in all other dimensions. Since  $K'''(0) = 0$ , we may express

$$q_k''(t_k) = \sum_{j=1}^s \alpha_{jk} K''(t_k - t_j) = \langle \alpha_k, \kappa_{2,k} \rangle,$$

where we substitute in the leading behavior derived in the previous section.

For the derivative terms, we have

$$q_j'(t_k) = \sum_{i=1}^s \alpha_{ij} \left( K'(t_k - t_i) - \frac{K'(t_j - t_i) K''(t_k - t_j)}{K''(0)} \right) \quad (43)$$

$$= \langle \alpha_j, \kappa_{1,k} \rangle - \frac{K''(t_k - t_j)}{K''(0)} \langle \alpha_j, \kappa_{1,j} \rangle \quad (44)$$

The last inner product here may be written fairly simply,

$$\langle \alpha_j, \kappa_{1,j} \rangle = \kappa_{1,j}^\top \mathbf{K}_0^{-1} \mathbf{e}_j \left( 1 - \frac{1}{\frac{\kappa_{1,k}^\top \mathbf{K}_0^{-1} \kappa_{1,k}}{K''(0)} + 1} \right).$$

**Heuristics** If we assume the second term above is negligible, then the “adversary term” for local concavity would be of the form

$$\sum_{j \neq k} |\langle \alpha_j, \kappa_{1,k} \rangle|^2.$$

We may think heuristically of  $\kappa_{1,k}$  as having an equal positive and negative weight on entries  $k \pm 1$  respectively, whereby this computes the total “gradient” of each  $\alpha_j$  between entries  $k \pm 1$ . Thus, to the extent that  $\mathbf{K}_0^{-1}$ , the solution operator of Dirichlet interpolation, operates in a “local” way, the only significant contributions should come from the first few neighbors on either side of dimension  $k$ . (Indeed, away from these neighbors we expect the second correction term to be small as well, since the leading factor will be small according to the decay of  $K''$  away from zero.)

## 6 Tools

### 6.1 Bernstein Polynomial Inequality

The idea of using this comes from [2]. The inequality is as follows.

**Lemma 1** *Let  $p(z)$  be a complex polynomial of degree  $N$ . Then,*

$$\sup_{|z| \leq 1} |p'(z)| \leq N \sup_{|z| \leq 1} |p(z)|.$$

The way this is used in that paper is basically to argue that a well-bounded polynomial must also have a good Lipschitz property over the unit circle:

$$\begin{aligned} |p(e^{ix}) - p(e^{iy})| &\leq |e^{-ix} - e^{iy}| \sup_{|z|=1} |p'(z)| \\ &\leq \left( CN \sup_{|z|=1} |p(z)| \right) |x - y| \end{aligned}$$

There, this is used to show that random polynomials are close to their mean, by showing this on a grid and then applying this Lipschitz property to pass from the grid to the entire interval. It's not clear how this can help us, since there is no natural choice of “comparison” polynomial.

### 6.2 Nazarov Inequality

**Lemma 2** *Let  $p(t)$  be a complex trigonometric polynomial of degree  $n$ , and  $E \subset [0, 1]$ . Then,*

$$\max_{t \in [0, 1]} |p(t)| \leq \left( \frac{C}{m(E)} \right)^{n-1} \max_{t \in E} |p(t)|.$$

## References

- [1] Carlos Fernandez-Granda. Super-resolution of point sources via convex programming. *Information and Inference: A Journal of the IMA*, 5(3):251–303, 2016.
- [2] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *IEEE transactions on information theory*, 59(11):7465–7490, 2013.