COMPUTATION OF "BEST" INTERPOLANTS IN THE L_p SENSE

Pakshal Bohra and Michael Unser

Biomedical Imaging Group, École polytechnique fédérale de Lausanne, Switzerland

ABSTRACT

We study a variant of the interpolation problem where the continuously defined solution is regularized by minimizing the L_p-norm of its second-order derivative. For this continuous-domain problem, we propose an exact discretization scheme that restricts the search space to quadratic splines with knots on an uniform grid. This leads to a discrete finite-dimensional problem that is computationally tractable. Another benefit of our spline search space is that, when the grid is sufficiently fine, it contains functions that are arbitrarily close to the solutions of the underlying unrestricted problem. We implement an iteratively reweighted algorithm with a grid-refinement strategy that computes the solution within a prescribed accuracy. Finally, we present experimental results that illustrate characteristics, such as sparsity, of the L_p-regularized interpolants.

Index Terms— Interpolation, regularization, L_p-norm, splines

1. INTRODUCTION

Interpolation is the task of constructing a continuously-defined function that passes through a given set of data points. It is a fundamental operation in signal and image processing that finds use in a variety of other applications [1–3] where a continuous representation of the discrete data is required.

An elegant way of performing interpolation is to formulate it as the continuous-domain optimization problem

$$\min_{s\in\mathcal{X}} \|\mathrm{L}\{s\}\|_{\mathcal{Y}} \text{ s.t. } s(x_m) = y_m,\ m = 1, 2, \ldots, M, \quad (1)$$

where $\mathcal{X}$, $\mathcal{Y}$ are suitable function spaces, $\mathrm{L} : \mathcal{X} \to \mathcal{Y}$ is the regularization operator, $\|\cdot\|_{\mathcal{Y}}$ is the norm associated with the space $\mathcal{Y}$, and $(x_m, y_m)_{m=1}^{M}$ are the data points. For instance, the choice of the regularization $\|\mathrm{D}^2\{s\}\|_{L_2}^2$, where D^2 is the second-order derivative operator, leads to the widely used cubic-spline[1] interpolants [4]. On the other hand, the sparsity-promoting generalized total-variation (gTV) regularization $\|\mathrm{D}^2\{s\}\|_{\mathcal{M}}$ (where the $\mathcal{M}$-norm is an extension of the L_1-norm) results in sparse piecewise-linear solutions with few knots [5,6]. We illustrate these two interpolation schemes in Figure 1.

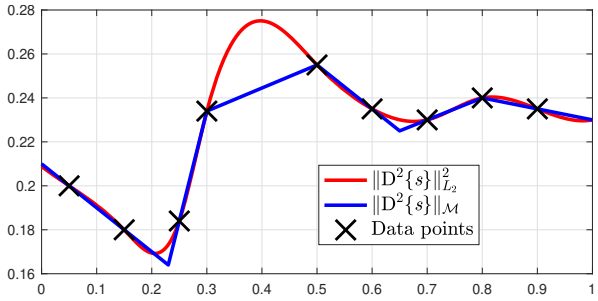

Fig. 1: Cubic-spline and linear-spline interpolation. Note that the knots of the linear spline solution are adaptive and fewer than the number of data points.

In this paper, we focus on L_p-regularized interpolation for a general $p \in (0, 2]$ and the second-order derivative operator *i.e.,* we consider the term $\|\mathrm{D}^2\{s\}\|_{L_p}^p$ in (1). For $p \geq 1$, this problem has been studied in [7, 8]. The work in [7] considers an infinite-dimensional search space of functions which have a second derivative with a finite L_p-norm and addresses theoretical questions such as existence and partial characterization of the solution. On the contrary, a finite-dimensional search space of cubic splines with knots at the data points is considered in [8]. We are interested in the former case which is more general, as well as the cases with $p < 1$ which have not been studied so far. Here, our aim is to develop an algorithm that numerically computes these interpolants and to investigate their characteristic features. To the best of our knowledge, there exists no numerical method for solving such problems, except for the special case of $p = 2$.

We first introduce our interpolation framework by defining the continuous-domain L_p-norm and the search space for the optimization task. Next, we present a spline-based exact discretization scheme that allows us to transform the continuous-domain problem into an equivalent finite-dimensional discrete one. We then describe an algorithm based on the iteratively reweighted least squares (IRLS) method [9–12], that computes the solution within a prescribed accuracy. Finally, we illustrate some characteristics of L_p-regularized interpolants in our experimental results.

This work was funded by the Swiss National Science Foundation under Grant 200020_184646 / 1.

[1] A polynomial spline of degree N_0 is a piecewise polynomial of degree N_0. It is defined such that its first $(N_0 - 1)$ derivatives are continuous. The points where the pieces are joined are called knots.

2. INTERPOLATION PROBLEM

We consider the problem of constructing a 1D signal $s : \mathbb{R} \to \mathbb{R}$ that honors the constraints $s(x_m) = y_m$. We assume that the sampling points $(x_m)_{m=1}^M$ lie in the interval $[0, T]$ and that they are distinct.

We now specify the continuous-domain L_p-norm which we use in the regularization term. For a measurable function $w : \mathbb{R} \to \mathbb{R}$, the L_p-norm ($0 < p < \infty$) is defined as[2]

$$\|w\|_{L_p} \triangleq \Big(\int_{\mathbb{R}} |w(x)|^p \, \mathrm{d}x \Big)^{1/p}. \tag{2}$$

The $\mathcal{M}$-norm used in gTV-regularization [6] is a generalization of the L_1-norm. In fact, for any function w with a finite L_1-norm, we have that $\|w\|_{\mathcal{M}} = \|w\|_{L_1}$. The major difference between the two is that the Dirac impulse, which is not included in the space of functions with a finite L_1-norm, has a unit $\mathcal{M}$-norm ($\|\delta\|_{\mathcal{M}} = 1$).

Next, we state the optimization tasks corresponding to our regularized interpolation problems. They are

$$\mathrm{S}_{\mathcal{M}} = \underset{s \in \mathcal{M}^{(2)}(\mathbb{R})}{\arg\min} \ \|\mathrm{D}^2\{s\}\|_{\mathcal{M}} \ \text{ s.t. } \ s(x_m) = y_m, \ \ \forall \, m \tag{3}$$

$$\mathrm{S}_p = \underset{s \in L_p^{(2)}(\mathbb{R})}{\arg\min} \ \|\mathrm{D}^2\{s\}\|_{L_p}^p \ \text{ s.t. } \ s(x_m) = y_m, \ \ \forall \, m, \tag{4}$$

where the spaces $\mathcal{M}^{(2)}(\mathbb{R}) = \{s : \|\mathrm{D}^2\{s\}\|_{\mathcal{M}} < +\infty\}$ and $L_p^{(2)}(\mathbb{R}) = \{s : \|\mathrm{D}^2\{s\}\|_{L_p} < +\infty\}$ are the largest ones for which the regularization term is finite.

It is well known that the L_2 problem has a unique cubic-spline solution with knots at the data points [4, 13]. Meanwhile, the gTV solution is not necessarily unique. The extreme points of $\mathrm{S}_{\mathcal{M}}$ are linear splines with a few adaptive knots [6]. The application of D^2 to the extreme points of $\mathrm{S}_{\mathcal{M}}$ uncovers Dirac impulses at the knot locations, thus demonstrating the sparsity-promoting effect of gTV-regularization. Henceforth, we refer to these extreme points as the sparse solutions of the gTV problem.

For $p > 1$, the L_p solution is unique. However, except for $p = 2$, it does not have a "simple" parametric form. As shown in [7], the second derivative of this solution is a nonlinear transformation of a linear spline. It is difficult to identify features of L_p-regularized interpolants from this partial characterization. Unlike the gTV and L_2 cases, the form of the solution here does not aid its numerical computation. Therefore, in this paper, we devise a spline-based algorithm to compute these solutions, as well as the more difficult scenario with $p < 1$.

[2]This definition corresponds to a quasinorm for $p < 1$.

3. DISCRETIZATION SCHEME

3.1. Search Space

We propose to discretize the continuous-domain problem (4) by restricting the search space to quadratic splines with knots on a uniform grid of size $h > 0$. Such splines can be uniquely expressed in the corresponding B-spline[3] basis [15]. This allows us to define the new search space

$$\mathcal{S}_{p,h}(\mathbb{R}) = \Big\{ \sum_{k \in \mathbb{Z}} c[k] \beta_h^2(\cdot - kh) : \ c * (1, -2, 1) \in \ell_p(\mathbb{Z}) \Big\}, \tag{5}$$

where $\beta_h^2(x)$ is the causal scaled quadratic B-spline given by

$$\beta_h^2(x) = \begin{cases} x^2/2h^2, & 0 \le x < h \\ (-2x^2 + 6xh - 3h^2)/2h^2, & h \le x < 2h \\ (3h - x)^2/2h^2, & 2h \le x < 3h \\ 0, & \text{otherwise.} \end{cases}$$

The second derivative of $\beta_h^2(x)$ is piecewise constant. It is given by

$$\mathrm{D}^2\{\beta_h^2\}(x) = \begin{cases} 1/h^2, & 0 \le x < h \\ -2/h^2, & h \le x < 2h \\ 1/h^2, & 2h \le x < 3h \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

A key property of $\mathcal{S}_{p,h}(\mathbb{R})$ is that it leads to an exact discretization. Further, the B-spline representation of $\mathcal{S}_{p,h}(\mathbb{R})$ ensures a well-conditioned discretization. This is because B-splines are compactly supported and form a Riesz basis [16]. Moreover, for $p > 1$, results in approximation theory [17] state that, when h is sufficiently small, $\mathcal{S}_{p,h}(\mathbb{R})$ is rich enough to contain functions that are arbitrarily close to the solution of the unrestricted continuous-domain problem (4).

3.2. Finite-Dimensional Problem

We now consider Problem (4) within the spline search space $\mathcal{S}_{p,h}(\mathbb{R})$. Since the sampling points satisfy $x_m \in [0, T]$, only a finite number of coefficients in $(c[k])_{k \in \mathbb{Z}}$ affect the constraints $s(x_m) = y_m$. Let $K = \{k_{\min}, \ldots, k_{\max}\}$ denote the set of their indices with $|K| = N$. We only need to optimize over these N coefficients. The remaining ones can be set such that they incur a vanishing regularization cost, leading to a solution that extends as a linear function outside of the interval $[0, T]$. Based on Property (6), we reformulate Problem (4) restricted to $\mathcal{S}_{p,h}(\mathbb{R})$ as the finite-dimensional problem

$$\min_{\mathbf{c} \in \mathbb{R}^N} \|\mathbf{L}\mathbf{c}\|_{\ell_p}^p \ \text{ s.t. } \ \mathbf{H}\mathbf{c} = \mathbf{y}, \tag{7}$$

[3]The quadratic causal B-spline with scaling factor h is the quadratic spline with knots in $h\mathbb{Z}$ that has the shortest support [14]. It is supported in [0, 3h].

where $\mathbf{y} = (y_1, \ldots, y_M)$, the matrix $\mathbf{H} : \mathbb{R}^N \to \mathbb{R}^M$ is

$$\mathbf{H} = \begin{bmatrix} \beta_h^2(x_1 - k_{\min}h) & \cdots & \beta_h^2(x_1 - k_{\max}h) \\ \vdots & & \vdots \\ \beta_h^2(x_M - k_{\min}h) & \cdots & \beta_h^2(x_M - k_{\max}h) \end{bmatrix}, \quad (8)$$

and the regularization matrix $\mathbf{L} : \mathbb{R}^N \to \mathbb{R}^{N-2}$ is

$$\mathbf{L} = \frac{1}{h^{2-\frac{1}{p}}} \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}. \quad (9)$$

The new formulation is exact. There are no discretization errors introduced in this step. Therefore, the solution to the finite problem (7) is the exact solution to the continuous-domain problem (4) when restricted to $\mathcal{S}_{p,h}(\mathbb{R})$. Further, if h is small enough, then the computed solution will be very close to the solution of the unrestricted problem (4).

The gTV problem (3) is discretized in a similar manner by using the search space $\mathcal{S}_{1,h}(\mathbb{R})$. This leads to the finite problem (7) with $p = 1$.

4. ALGORITHM

In this section, we discuss an algorithm that computes the solution to our interpolation problem. We begin with a grid size h_0. At each iteration t, we consider a finer grid with size $h_t = h_{t-1}/2$ and solve the resulting optimization task (7). The embedding property of the search spaces $\mathcal{S}_{p,h_{t-1}}(\mathbb{R}) \subset \mathcal{S}_{p,h_t}(\mathbb{R})$ guarantees that $\mathcal{J}^*_{h_t} \leq \mathcal{J}^*_{h_{t-1}}$, where $\mathcal{J}^*_h$ is the optimal cost corresponding to Problem (7). Thus, the solution can only improve in terms of the cost function. We stop this process when the relative decrease in cost is less than a prescribed accuracy. The embedded search spaces also allow us to use the solution from the previous grid as initialization for the current one, leading to faster convergence. This multiresolution strategy has been adapted from [18].

Since unconstrained optimization problems are usually easier to solve compared to their constrained counterparts, at each grid size h_t we find the solution to Problem (7) by solving[4]

$$\min_{\mathbf{c} \in \mathbb{R}^N} \|\mathbf{y} - \mathbf{H}\mathbf{c}\|_{\ell_2}^2 + \lambda \|\mathbf{L}\mathbf{c}\|_{\ell_p}^p \quad (10)$$

with a very small value of the parameter λ. We then solve Problem (10) using an iteratively reweighted method [9–12] which involves computing the sequence of iterates $\mathbf{c}^{(q)}$ (until convergence) given by

$$\mathbf{c}^{(q)} = \arg\min_{\mathbf{c} \in \mathbb{R}^N} \|\mathbf{y} - \mathbf{H}\mathbf{c}\|_{\ell_2}^2 + \lambda\, \mathbf{c}^T \mathbf{L}^T \mathbf{D}^{(q-1)} \mathbf{L}\mathbf{c}, \quad (11)$$

where $\mathbf{D}^{(q-1)}$ is an $(M \times M)$ diagonal matrix whose elements are $|(\mathbf{L}\mathbf{c}^{(q-1)})_m|^{p-2}$ for $m = 1, 2, \ldots, M$. Thus, the weights in Problem (11) depend on the previous iterate $\mathbf{c}^{(q-1)}$. The solution to (11) is obtained by solving the linear system of equations

$$(\mathbf{H}^T\mathbf{H} + \lambda \mathbf{L}^T \mathbf{D}^{(q-1)} \mathbf{L})\, \mathbf{c}^{(q)} = \mathbf{H}^T \mathbf{y}. \quad (12)$$

This can be done using various techniques such as direct inversion and QR or LU factorization.

For $p < 2$, the elements of the diagonal matrix $\mathbf{D}^{(q-1)}$ can grow unbounded. This is likely to happen especially for the small values of p which promote the sparsity of $\mathbf{L}\mathbf{c}$. To prevent this from happening, we follow the procedure in [11] and set the entries of $\mathbf{D}^{(q-1)}$ as $(((\mathbf{L}\mathbf{c}^{(q-1)})_m)^2 + \varepsilon)^{\frac{p}{2}-1}$, where $\varepsilon > 0$ is a hyperparameter. We also implement the damping technique of [11]. It improves the performance of our algorithm for the non-convex case $p < 1$. Our final procedure for solving Problem (7) is as follows: We begin with some initial estimate $\mathbf{c}^{(0)}$ and $\varepsilon = 1$. We then compute the sequence of iterates in (11) until convergence, as controlled by the distance between the solutions at successive iterations. At this point, the value of ε is decreased by a factor of 10 and the process is repeated with our new estimate as the starting point. We terminate the algorithm when $\varepsilon = 10^{-20}$.

This iteratively reweighted method (IRM) computes the unique solution of (7) when $p > 1$. For $p = 1$, it gives us one of the possibly many solutions and this solution is not always sparse. Alternately, one can recast the ℓ_1-problem as a linear program and use the simplex algorithm [19] to compute a sparse solution. For the non-convex case $p < 1$, IRM can reach a local minimum.

5. EXPERIMENTAL RESULTS

Our algorithm has been implemented in MATLAB. For all the experiments, the regularization parameter and grid tolerance level were set to 10^{-10} and 10^{-3}, respectively. While our method holds for all values of $p > 0$, we focus here on the range $p \in (0, 2]$.

We first consider the example in Figure 2. For the given data points, it is possible to prove that the gTV solution is unique. In fact, this unique, sparse gTV solution can be obtained by simply connecting the points linearly, *i.e.,* it is a linear spline with only two knots. In Figure 2, the computed gTV solutions[5] (for both IRM and simplex) look like a linear spline with two knots, in agreement with the theory. As we vary p from 2 to 1, we see that the solutions gradually converge towards the gTV solution. Note that the solution for $p = 1.001$ is very close to the sparse gTV solution. Further, we observe that choosing $p = 0.5$ results in a solution that also resembles the sparse linear spline solution.

In Figure 3, the data points are such that the gTV problem has infinitely many solutions. We observe that the simplex method computes a sparse solution that resembles a linear

[4] The solution to (7) can be obtained from (10) in the limit $\lambda \to 0$.

[5] As stated in Section 3.2, $p = 1$ corresponds to the gTV problem.

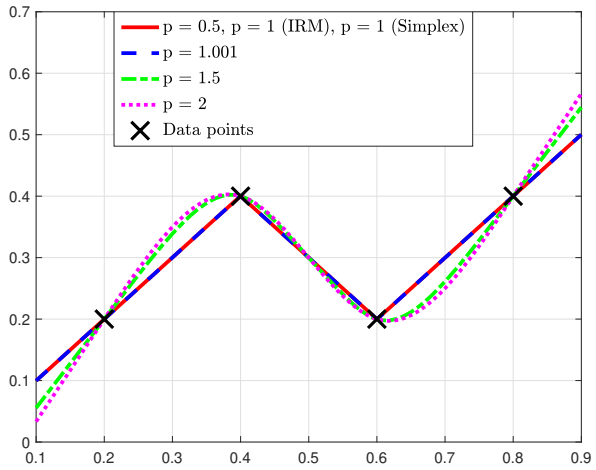

Fig. 2: Unique gTV solution

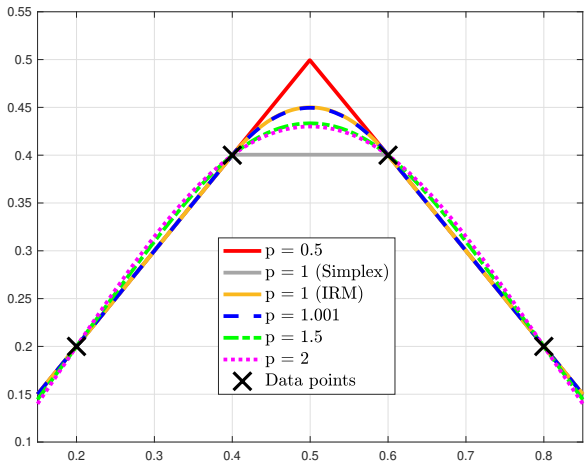

Fig. 3: Multiple gTV solutions

spline with two knots. At the same time, IRM produces a non-sparse gTV solution. In this case, as we vary p from 2 to 1, the solutions gradually converge towards the non-sparse gTV solution. Meanwhile, the solution corresponding to $p = 0.5$ looks like a linear spline but with a single knot. Therefore, by choosing $p = 0.5$ instead of $p = 1$, we are able to obtain a sparser solution in this case.

Based on the above examples and additional experiments of similar nature, we make a few claims.

- As p is varied from 2 to 1, there exists a continuum of solutions.
- When the gTV problem has a unique solution, the continuum converges to that unique, sparse, linear-spline solution as $p \to 1$.
- When the gTV problem has multiple solutions, the continuum converges to one of the non-sparse gTV solutions as $p \to 1$.
- For $p < 1$, some of the local minima solutions approach sparse linear splines with few knots. Moreover, in certain cases, our algorithm finds a solution that is sparser than the computed gTV one.

The existence of a continuum of solutions implies that one can use the value of p to obtain a tradeoff between the properties of the L_2 and gTV solutions. For example, one can control the smoothness of the interpolant by varying p.

Through this work, we are also able to draw a parallel between the continuous-domain L_p-norm regularization and the discrete ℓ_p-norm regularization. Based on our experiments, we claim that, in settings where the gTV solution is known to be unique, L_p-norm regularization with a small p acts as a sparsity-promoting prior. This mirrors the behavior of ℓ_p-norm regularization (with a small p) in compressed sensing [20, 21].

6. CONCLUSION

We have devised a method that solves the L_p-regularized interpolation problem with a second-order derivative regularization operator. Our method involves the use of quadratic splines, with uniformly spaced knots, for an exact discretization. We then solve the discrete problem using an iteratively reweighted method. We rely on a grid refinement strategy to select a small-enough grid size. Finally, through numerical experiments, we identify some interesting properties of such L_p-regularized interpolants.

7. REFERENCES

[1] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.

[2] P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Imaging, Processing and Analysis*, I.N. Bankman, Ed., chapter 25, pp. 393–420. Academic Press, San Diego CA, USA, 2000.

[3] E. Meijering, "A chronology of interpolation: From ancient astronomy to modern signal and image processing," *Proceedings of the IEEE*, vol. 90, no. 3, pp. 319–342, 2002.

[4] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, "The theory of splines and their applications," *Mathematics in Science and Engineering, New York: Academic Press, 1967*, 1967.

[5] E. Mammen and S. van de Geer, "Locally adaptive re-

gression splines," *The Annals of Statistics*, vol. 25, no. 1, pp. 387–413, 1997.

[6] M. Unser, J. Fageot, and J.P. Ward, "Splines are universal solutions of linear inverse problems with generalized TV regularization," *SIAM Review*, vol. 59, no. 4, pp. 769–793, 2017.

[7] C. de Boor, "On "best" interpolation," *Journal of Approximation Theory*, vol. 16, no. 1, pp. 28–42, 1976.

[8] J. E. Lavery, "Univariate cubic L_p splines and shape-preserving, multiscale interpolation by univariate cubic L_1 splines," *Computer Aided Geometric Design*, vol. 17, no. 4, pp. 319–336, 2000.

[9] M. R. Osborne, *Finite Algorithms in Optimization and Data Analysis*, John Wiley & Sons, Inc., New York, NY, USA, 1985.

[10] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 187–200, 1999.

[11] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 3869–3872.

[12] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 63, no. 1, pp. 1–38, 2010.

[13] H. Gupta, J. Fageot, and M. Unser, "Continuous-domain solutions of linear inverse problems with Tikhonov *versus* generalized TV regularization," *IEEE Transactions on Signal Processing*, vol. 66, no. 17, pp. 4670–4684, 2018.

[14] I. J. Schoenberg, *Cardinal Spline Interpolation*, vol. 12, SIAM, 1973.

[15] I. J. Schoenberg, "Contributions to the problem of approximation of equidistant data by analytic functions: Part A— On the problem of smoothing or graduation. A first class of analytic approximation formulae," *Quarterly of Applied Mathematics*, vol. 4, no. 1, pp. 45–99, 1946.

[16] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1425–1438, 2005.

[17] J. J. Lei, "L_p-approximation by certain projection operators," *Journal of Mathematical Analysis and Applications*, vol. 185, no. 1, pp. 1–14, 1994.

[18] T. Debarre, J. Fageot, H. Gupta, and M. Unser, "B-Spline-based exact discretization of continuous-domain inverse problems with generalized TV regularization," *IEEE Transactions on Information Theory*, pp. 1–1, 2019.

[19] G. B. Dantzig, A. Orden, and P. Wolfe, "The generalized simplex method for minimizing a linear form under linear inequality restraints," *Pacific Journal of Mathematics*, vol. 5, no. 2, pp. 183–195, 1955.

[20] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[21] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.