

Primal-Dual Approach for Uniform Noise Removal

Zhan Zhang, Wen-You Wei

Department of Mathematics, Faculty of Science, Kunming University of Science and Technology, Kunming, 650500, China

Abstract—In this paper, we consider the problem of uniform noise removal, which can be formulated as a minimization problem with L^∞ norm based constraints. A numerical difficulty arises due to the property of the non-differentiability of the L^∞ norm. In this paper, we apply first-order primal dual approach to solve the problem of uniform noise removal. Numerical results are given to demonstrate the performance of the proposed method.

Keywords- L^∞ fitting; uniform noise; inverse problem

I. INTRODUCTION

In many applications such as signal processing, the observed data is generally corrupted by the uniform noise, which has an approximately uniform distribution. The noise may be caused by quantizing the signal to a number of discrete levels or by the round off error. The model of noise removal can be formulated as

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \quad (1.1)$$

where \mathbf{f} is the original true data or noise-free data, \mathbf{H} is a linear operator, \mathbf{n} is the noise with uniform distribution in the between the interval $[\delta, \delta]$, \mathbf{g} is the observed data. The aim of uniform noise removal is to estimate the original data \mathbf{f} from the observed data \mathbf{g} .

The simplest approach of performing the inverse transformation is not feasible. The linear operator may be singular and therefore it is not reversible even though the absence of noise. Moreover, due to the presence of the noise and ill-condition of linear operator \mathbf{H} , a small perturbation in the estimated data. Regularization theory is often used to handle such ill-conditioned problems. When the noise in (1.1) is Gaussian white noise, the data \mathbf{f} can be estimated by solving the following least square regularization minimization problem [1]

$$\min_{\mathbf{f}} \|\mathbf{H}\mathbf{f} - \mathbf{b}\|_2^2 + \alpha \|\mathbf{R}\mathbf{f}\|^2. \quad (1.2)$$

Here $\alpha > 0$ is called the regularization parameter and \mathbf{R} is the regularization matrix. The first term in (1.2) is the data fitting term and the second term is the regularization term. Usually, \mathbf{R} is the identity matrix, in which the original data is assumed to be not large, or \mathbf{R} is the finite difference matrix, in which case the original data should not rapidly.

The minimization problem in (1.2) can be also derived by a Maximum a Posteriori (MAP) approach [11]. The data fitting term in (1.2) is derived by the assumption of Gaussian noise in the corrupted noise \mathbf{n} . When the noise in the observed data follows the uniform distribution, this assumption is not hold. We can not use the L^2 norm since it

will give a sub-optimal solution. In order to satisfy the assumption of uniform noise, we need to modify the L^2 norm in the data fitting term of (1.2) into the L^∞ norm:

$$\min_{\mathbf{f}} \|\mathbf{H}\mathbf{f} - \mathbf{b}\|_{L^\infty} + \alpha \|\mathbf{R}\mathbf{f}\|^2. \quad (1.3)$$

Difficulties on computational stability and efficiency for solving the minimization problem (1.3) arise since the L^∞ norm is non-quadratic and is non-differentiable. Therefore, we cannot obtain the Euler-Lagrange equation associated with the minimization problem. Numerical methods for minimizing L^∞ functional have been investigated in [4,5] for optimal control of partial differential equations. There has also been some recent interest in L^∞ functional in the context of geometric vision [6]. Recently, Clason [1] applied semi-smooth Newton method to solve the L^∞ norm problem.

According to the convex optimization theorem, the L^∞ norm based minimization problem in (1.3) can be rewritten as

$$\min_{\mathbf{f}} \|\mathbf{R}\mathbf{f}\|^2 \quad \text{subject to} \quad \|\mathbf{H}\mathbf{f} - \mathbf{b}\|_{L^\infty} \leq \delta. \quad (1.4)$$

Here δ is the positive parameter. It is well known that there exists a particular choice for the parameter δ such that the solution of (1.4) is also the solution of (1.3). Compare with the parameter α in the constrained problem (1.3), the parameter δ in the constrained problem of (1.4) is easy to be estimated. Therefore, we consider the solution of (1.4) in this paper.

Recently, alternating direction method of multipliers (ADMM) [12] and the primal-dual approach [10,13] had been successively applied in the problem of Gaussian noise removal. However, to our best knowledge, there are few literatures to extend the ADMM or the primal-dual approach to deal with the problem of uniform noise removal. The main contribution in this paper is to apply the first order primal-dual approach to solve the problem of uniform noise removal in (1.4). We should highlight the major difference between (1.3) and (1.4) though they are equivalent in mathematics. In general, the regularization α in (1.3) is chosen by trial and error, it is very difficult to tune. However, the parameter δ in (1.4) is general given or can be estimated since the noise in the observed δ follows the uniform distribution. Hence we don't need to tune the regularization parameter if we consider the minimization problem (1.4).

The outline of this paper is as follows. In Section 2, we will first introduce the first-order primal-dual model and then formulate the problem of uniform noise into the framework of

the primal-dual approach. In Section 3, numerical results are given to demonstrate the performance of the proposed method. Final, we give a short conclusion in Section 4.

II. PRIMAL-DUAL MODEL

A. First-Order Primal-Dual Model

In [10], Chambolle and Pock proposed a first order primal dual method to solve the following minimax problem:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in A} \phi(\mathbf{x}) + \langle \mathbf{K}\mathbf{x}, \mathbf{y} \rangle - \psi(\mathbf{y}). \quad (2.1)$$

Here ϕ, ψ are convex and lower semi-continuous functions, and \mathbf{K} is a linear operator, $\langle \cdot, \cdot \rangle$ denotes the inner product. The iterative scheme associated with the minimax problem is given by

$$\begin{cases} \mathbf{x}^{k+1} = \arg \min_{\mathbf{x} \in X} \phi(\mathbf{x}) + \langle \mathbf{K}\mathbf{x}, \mathbf{u} \rangle + \frac{1}{2s} \|\mathbf{x} - \mathbf{x}^k\|_2^2, \\ \mathbf{x}^{-k+1} = \mathbf{x}^{k+1} + \theta(\mathbf{x}^{k+1} - \mathbf{x}^k), \\ \mathbf{u}^{k+1} = \arg \max_{\mathbf{u}} \langle \mathbf{K}\mathbf{x}^{-k+1}, \mathbf{u} \rangle - \psi(\mathbf{u}) - \frac{1}{2t} \|\mathbf{u} - \mathbf{u}^k\|_2^2. \end{cases} \quad (2.2)$$

The parameter $s, t > 0$ are step sizes of the primal and dual steps respectively, and θ is the combination parameter. The convergence can be guaranteed when appropriated step sizes were chosen. When the parameter $\theta = 0$, (2.2) can also be equivalently interpreted as a first-order primal-dual relaxed Arrow-Hurwitz algorithm. When the parameter $\theta = 1$ and the step satisfy $st < 1/\|\mathbf{K}^T \mathbf{K}\|$, iterative algorithm convergence rate is $O(1/k)$. In fact, the algorithm (2.2) is to seek the saddle point of the minimax problem.

Definition 2.1 Let the function $Q(\mathbf{x}, \mathbf{y})$ is a convex function on the variable \mathbf{x} and is a concave function on \mathbf{y} . For any \mathbf{x}, \mathbf{y} , such that

$$Q(\mathbf{x}^*, \mathbf{y}) \leq Q(\mathbf{x}^*, \mathbf{y}^*) \leq Q(\mathbf{x}, \mathbf{y}^*).$$

Then $(\mathbf{x}^*, \mathbf{y}^*)$ is a saddle point about $(\mathbf{x}^*, \mathbf{y}^*)$.

The saddle point of is defined as

$$\begin{cases} \mathbf{x}^{k+1} = \arg \min_{\mathbf{x} \in X} Q(\mathbf{x}, \mathbf{y}^{k+1}) + \frac{1}{2s} \|\mathbf{x} - \mathbf{x}^k\|_2^2, \\ \mathbf{x}^{-k+1} = \mathbf{x}^{k+1} + \theta(\mathbf{x}^{k+1} - \mathbf{x}^k), \\ \mathbf{y}^{k+1} = \arg \min_{\mathbf{y} \in Y} Q(\mathbf{x}^{-k+1}, \mathbf{y}) - \frac{1}{2t} \|\mathbf{y} - \mathbf{y}^k\|_2^2. \end{cases} \quad (2.3)$$

B. Primal-Dual Approach for Uniform Noise Removal

Now we consider how to find the solution of the constrained optimization problem (1.4). We introduce an auxiliary variable \mathbf{u} such that $\mathbf{u} = \mathbf{H}\mathbf{f} - \mathbf{b}$. Hence the Lagrange function associated with the equation constraints is give by

$$Q(\mathbf{u}, \mathbf{f}, \mathbf{y}) \equiv \frac{1}{2} \|\mathbf{R}\mathbf{f}\|_2^2 + \mathbf{y}^T (\mathbf{u} - \mathbf{H}\mathbf{f} + \mathbf{b}). \quad (2.4)$$

Here \mathbf{y} is the Lagrangian multiplier. Define the set $A = \{\mathbf{u} : \|\mathbf{u}\|_\infty \leq \delta\}$, the image restoration problem (1.4) can be written as

$$\min_{\mathbf{f}, \|\mathbf{f}\|_\infty \leq \delta} \max_{\mathbf{y}} Q(\mathbf{u}, \mathbf{f}, \mathbf{y}). \quad (2.5)$$

We note that $Q(\mathbf{u}, \mathbf{f}, \mathbf{y})$ is convex in \mathbf{f} and concave in \mathbf{y} . Hence, there exists a saddle point $(\mathbf{u}^*, \mathbf{f}^*, \mathbf{y}^*)$ of $Q(\mathbf{u}, \mathbf{f}, \mathbf{y})$ [7]. Using the existence of the saddle point of $Q(\mathbf{u}, \mathbf{f}, \mathbf{y})$, convex analysis can be applied to show that the minimization and the maximum in (2.5) can be swapped, i.e.,

$$\min_{\mathbf{f}, \|\mathbf{f}\|_\infty \leq \delta} \max_{\mathbf{y}} Q(\mathbf{u}, \mathbf{f}, \mathbf{y}) = Q(\mathbf{u}^*, \mathbf{f}^*, \mathbf{y}^*) = \max_{\mathbf{y}} \min_{\mathbf{f}, \|\mathbf{f}\|_\infty \leq \delta} Q(\mathbf{u}, \mathbf{f}, \mathbf{y}).$$

In order to apply Chambolle-Pock's algorithm (2.2) to the uniform noise removal problem in (2.5), we need to formulate (2.5) in the form of (2.1)

Denote

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{f} \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} \mathbf{I} & -\mathbf{H} \end{pmatrix}$$

$\phi(\mathbf{f}) = \frac{1}{2} \|\mathbf{R}\mathbf{f}\|_2^2$, and $\psi(\mathbf{y}) = -\langle \mathbf{y}, \mathbf{b} \rangle$. We see that (2.5) completely fits into the framework of (2.1). Hence (2.2) can readily be used to solve (2.5). By using [7, Proposition 2.6.1], a pair $(\mathbf{u}^*, \mathbf{f}^*, \mathbf{y}^*)$ is a saddle point of (2.4) if and only if $\mathbf{f}^*, \mathbf{y}^*$ and \mathbf{u}^* are the optimal solution of the problems, i.e.,

$$\mathbf{u}^* = \arg \min_{\mathbf{u} \in A} Q(\mathbf{u}, \mathbf{f}^*, \mathbf{y}^*) \quad (2.6)$$

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} Q(\mathbf{u}^*, \mathbf{f}, \mathbf{y}^*) \quad (2.7)$$

$$\mathbf{y}^* = \arg \max_{\mathbf{y}} Q(\mathbf{u}^*, \mathbf{f}^*, \mathbf{y}) \quad (2.8)$$

respectively.

Seeking the saddle point of (2.5) is equivalent to solving systems (2.6), (2.7) and (2.8). More specifically, starting from some initial values for $\mathbf{f}^0, \mathbf{u}^0, \mathbf{y}^0$, let $s > 0, t > 0$, and $\theta = 1$ be fixed value, the iterative scheme to calculate the solution of (2.4), produces the following sequence of update

$$\mathbf{y}^{k+1} = \arg \max_{\mathbf{y}} \psi(\mathbf{u}^k, \mathbf{f}^k, \mathbf{y}) \quad (2.9)$$

$$\mathbf{u}^{k+1} = \arg \min_{\mathbf{u} \in A} \psi(\mathbf{u}, \mathbf{f}^k, \mathbf{y}^{k+1}) \quad (2.10)$$

$$\mathbf{f}^{k+1} = \arg \min_{\mathbf{f}} \psi(\mathbf{u}^{k+1}, \mathbf{f}, \mathbf{y}^{k+1}) \quad (2.11)$$

where

$$\psi(\mathbf{u}, \mathbf{f}, \mathbf{y}) \equiv Q(\mathbf{u}, \mathbf{f}, \mathbf{y}) - \frac{1}{2s} \|\mathbf{y} - \mathbf{y}^k\|_2^2 \quad (2.12)$$

$$\psi(\mathbf{u}, \mathbf{f}, \mathbf{y}) \equiv Q(\mathbf{u}, \mathbf{f}, \mathbf{y}) + \frac{1}{2t} \|\mathbf{u} - \mathbf{u}^k\|_2^2 \quad (2.13)$$

$$\psi(\mathbf{u}, \mathbf{f}, \mathbf{y}) \equiv Q(\mathbf{u}, \mathbf{f}, \mathbf{y}) + \frac{1}{2t} \|\mathbf{f} - \mathbf{f}^k\|_2^2 \quad (2.14)$$

With constants $s, t > 0$. We remark that our method is different from the primal-dual method proposed in [8] and [9], where a gradient-descent method is employed to the primal and dual variables alternatively. Our method resembles in some way the dual method in [10]. Next, we discuss how to solve subproblems (2.9)-(2.11).

We consider the solution of (2.9) now. The problem (2.9) can be written as

$$\begin{aligned} \mathbf{y}^{k+1} &= \arg \max_{\mathbf{y}} \left\{ \frac{1}{2} \|\mathbf{R}\mathbf{f}^k\|_2^2 + \mathbf{y}^T (\mathbf{u}^k - \mathbf{H}\mathbf{f}^k + \mathbf{b}) \right. \\ &\quad \left. - \frac{1}{2s} \|\mathbf{y} - \mathbf{y}^k\|_2^2 \right\} \\ &= \mathbf{y}^k + s(\mathbf{H}\mathbf{f}^k - \mathbf{u}^k - \mathbf{b}). \end{aligned} \quad (2.15)$$

To make \mathbf{y} convergence, we should

$$\mathbf{y}^{k+1} = \mathbf{y}^{k+1} + \theta(\mathbf{y}^{k+1} - \mathbf{y}^k). \quad (2.16)$$

where $\theta = 1$ and iterative algorithm converges at a rate of $O(1/k)$. Similarly, we can compute (2.10)

$$\begin{aligned} \mathbf{u}^{k+1} &= \arg \min_{\mathbf{u} \in A} \left\{ \frac{1}{2} \|\mathbf{R}\mathbf{f}^k\|_2^2 + (\mathbf{y}^{k+1})^T (\mathbf{u} - \mathbf{H}\mathbf{f}^k - \mathbf{b}) \right. \\ &\quad \left. + \frac{1}{2t} \|\mathbf{u} - \mathbf{u}^k\|_2^2 \right\} \\ &= \arg \min_{\mathbf{u} \in A} \frac{1}{2t} \{ 2 \langle \mathbf{u} - \mathbf{u}^k, -t\mathbf{y}^{k+1} \rangle + \|\mathbf{u} - \mathbf{u}^k\|_2^2 \} \\ &= \arg \min_{\mathbf{u} \in A} \left\| \mathbf{u} - (\mathbf{u}^k + t\mathbf{y}^{k+1}) \right\|_2^2. \end{aligned} \quad (2.17)$$

The minimization problem is equivalent to computing the projection of $(\mathbf{u}^k + t\mathbf{y}^{k+1})$ onto set A . Therefore

$$\mathbf{u}^{k+1} = P_A(\mathbf{u}^k + t\mathbf{y}^{k+1}). \quad (2.18)$$

Subproblem for Primal Variable \mathbf{f} : From (2.11), we see that the objection in (2.7) is quadratic with respect to \mathbf{f} .

Hence, \mathbf{f}^{k+1} can be easily computed by

$$\mathbf{f}^{k+1} = (\mathbf{I} + t\mathbf{R}^T \mathbf{R})^{-1} (\mathbf{f} - t\mathbf{H}^T \mathbf{y}^{k+1}), \quad (2.19)$$

where \mathbf{R} is the finite difference matrix.

Therefore, the resulting algorithm using the primal-dual model is summarized in Algorithm 1.

Algorithm 1: Primal-Dual Model for Image Restoration Algorithm(PDM)

Function: $\mathbf{f} = \text{PDM}(\mathbf{b}, \mathbf{R})$.

Input: \mathbf{b}, \mathbf{R} .

1: Initialize $\mathbf{f}^0, \mathbf{u}^0$ and \mathbf{y}^0 . Set step sizes s and t, N .

2: **For** iteration numbers is less than N

3: $\mathbf{y}^{k+1} = \mathbf{y}^k + s(\mathbf{H}\mathbf{f}^k - \mathbf{u}^k - \mathbf{b})$;

4: $\mathbf{y}^{k+1} = \mathbf{y}^{k+1} + \theta(\mathbf{y}^{k+1} - \mathbf{y}^k)$;

5: $\mathbf{u}^{k+1} = P_A(\mathbf{u}^k + t\mathbf{y}^{k+1})$;

6: $\mathbf{f}^{k+1} = (\mathbf{I} + t\mathbf{R}^T \mathbf{R})^{-1} (\mathbf{f} - t\mathbf{H}^T \mathbf{y}^{k+1})$;

7: **end**

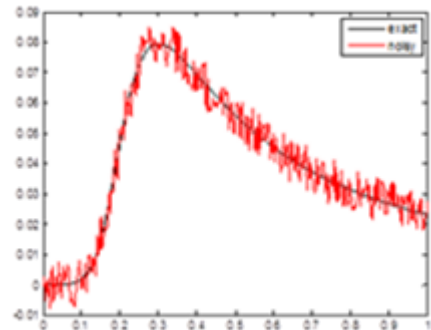
8: **return** $\mathbf{f} = \mathbf{f}^{k+1}$.

III. NUMERICAL RESULTS

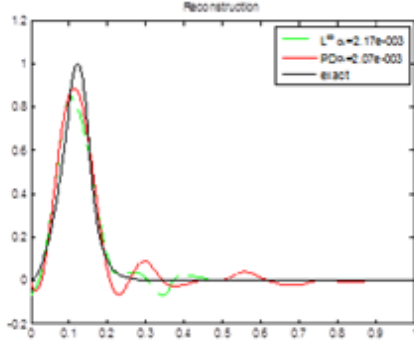
In this section, we illustrate the performance of the proposed algorithms(primal-dual method) and compare it with the Semi-smooth Newton method in [1]. The experiments were performed under Windows 7 and MATLAB R2007a. The stopping criterion of all the methods is that the relative difference between the successive iterate of the number of iterations is larger than 100. We will give a one-dimensional example, the example of the noise image is added to uniform noise in the original image. A typical realization of noisy data is displayed in Figure 1a for $d = 0.1$, Figure 1c for $d = 0.5$ and Figure 1e for $d = 1.0$. The respective reconstructions \mathbf{f} are shown in Figure 1b, 1d, 1f. For comparison, we also show the solution to the data L^∞ fitting problem.

Figure 1 Results for inverse heat conduction problem

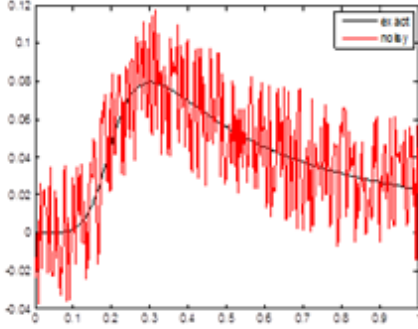
(a) exact and noisy data($d=0.1$)



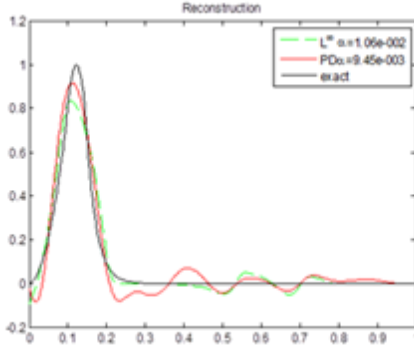
(b) comparison of two methods($d=0.1$)



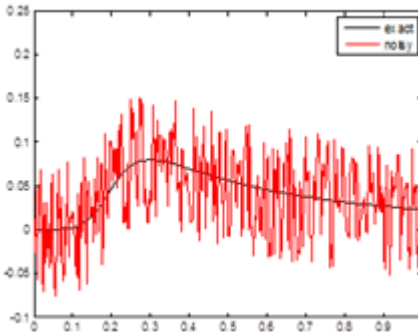
(c) exact and noisy data($d = 0.5$)



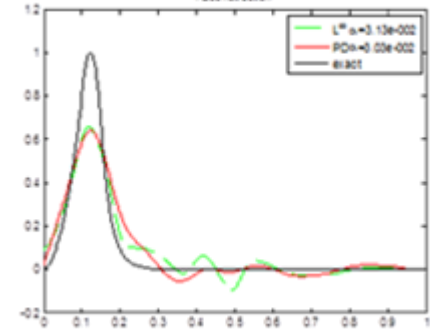
(d) comparison of two methods($d=0.5$)



(e) exact and noisy data($d = 1.0$)



(f) comparison of two methods($d=1.0$)



IV. CONCLUSION

In this paper we applied primal-dual method to remove the uniform noise in the corrupted data. We have compared the proposed method with the semi-smooth Newton method and L2 norm based method. Numerical results were shown the efficient of the proposed method.

ACKNOWLEDGMENT

The research supported in part by NSFC Grant 11361030 and a SRF for ROCS of SEM.

References

- [1] C. Clason, "L[∞] fitting for inverse problems with uniform noise," *Inverse Problems* 28(2012), 1044007.
- [2] M. Burger and S. Osher(2004), "Convergence rates of convex variational regularization," *Inverse Problems* 20.5, pp. 1411-1421.
- [3] J. Williams, and Z. Kalogiratos (1993a). "Least squares and Chebyshev fitting for parameter estimation in ODEs," *Adv. Comput. Math.* 1.3-4, pp. 357-366.
- [4] J. Williams, and Z. Kalogiratos (1993b). "Nonlinear Chebyshev fitting from the solution of ordinary differential equations," pp. 325-337.
- [5] Sim. K. and Hartley. R(2006), "Removing outliers using the L[∞] norm," *IEEE Computer Society*, Washington, DC, USA, pp.485-494.
- [6] D. Bertsekas, A.Nedic, and E. Ozdaglar, "Convex analysis and optimization." Belmont, MA: Athena Scientific, 2003.
- [7] M. Zhu, "Fast numerical algorithms for total variation based image restoration," Ph.D. dissertation, Univ. California, Los Angeles, 2008.
- [8] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," Univ. California, Los Angeles, UCLA CAM Rep. 08-34, 2007.
- [9] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17-40, 1976.
- [10] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, 40(1):120-145, 2011.
- [11] A. Ben Hamza, Hamid Krim, "A variational approach to maximum a posteriori estimation for image denoising".
- [12] M. Yang, "Total variation regularization and fast algorithms based on alternating direction method".
- [13] Benning, Martin, Calatroni, Luca, Düring, Bertram, "A primal-dual approach for a total variation Wasserstein flow", 2013.