Multilevel approach for signal restoration problems with structured matrices

Malena Ines Español[1]

January 14, 2008

Abstract

We present a multilevel method for ill-posed problems. In this method we use the Haar wavelets as restriction and prolongation operators. The choice of the Haar wavelet operator has the advantage of preserving matrix structure such as Toeplitz or Toeplitz + Hankel, among grids, which can be exploited to obtain faster solvers on each level.

1 Introduction

Linear ill-posed problems in the form of Fredholm integral equations of the first kind occur frequently in the physical sciences. They model situations where hidden information is computed from external observations. The term *discrete ill-posed problem* refers to a discretization of such an integral equation, taking the form

$$Ax \approx b = b^{true} + e, \qquad b^{true} = Ax^{true}, \tag{1}$$

where A is a discretization of the integral operator, x^{true} is the true solution, and e denotes noise due to measurement and/or approximation error. In discrete ill-posed problems, the matrix A is ill-conditioned and its singular values decay to zero without significant spectral gap. In this paper, we restrict ourselves to the case when $A \in \mathbb{R}^{m\times m}$, with $m = 2^k$ for some integer k. We specifically consider the 1D signal restoration problem. In this case, A acts as a blurring operator on the 'signal' x^{true}, and b denotes the blurred, noisy signal. We are particularly interested in the case when the signal has 'edges' that we want to recover.

The noise e contained in the right-hand side b, in combination with the ill-conditioning of A, imply that the exact solution of (1) has little relationship to the noise-free solution x^{true} and is worthless. Thus, we need to apply a regularization method to determine a solution that approximates better the noise-free solution. Regularization methods [4] include Tikhonov regularization or early termination of certain iterative methods, such as conjugate gradient on the normal equations.

Multigrid methods [13] are well known as extremely efficient solvers for large-scale systems of equations resulting from the discretizations of certain classes of partial differential equations and integral equations of the second kind and have been extensively studied in recent years. However, for ill-posed problems, the classical multigrid approach is not immediately applicable. The first consideration is that vectors in the span of the singular vectors corresponding to the largest singular values are smooth, while the singular vectors corresponding to the small singular values represent the "high frequency" information. Consequently, since the integral operator is a smoothing operator, but the noise vector is assumed to be white, the expansion coefficients of b^{true} decay on average faster than the singular values while the expansion coefficients of the noise vector are roughly constant in absolute value. Thus, the noise perturbs the high frequency components of the right hand side and makes it impossible to recover, through the use of typical multigrid 'smoothers'

[1]Department of Mathematics, Tufts University, 503 Boston Ave., Medford, MA 02155
Email: malena.espanol@tufts.edu.
Research in collaboration with Misha E. Kilmer

such as Jacobi, Gauss-Seidel, or SOR because these techniques try to smooth out the higher frequency part of the error.

One recent approach to the deblurring problem for images (2D signals) was presented in [2]. In this work, the authors used a standard multigrid V-cycle and standard (with respect to the PDE literature) prolongation and coarse-grid correction operators. The smoothers advocated were those such as CGLS, which are known to capture the smooth component of the signal at each level without mixing in noise (high frequency information) as long as they are terminated early enough. Their approach did not include a residual correction (post-smoothing) step on the right-hand part of the V, and rather used only interpolation to move back up to the fine grid. They advocated for a very few iterations at each level (i.e. over-regularization), and multiple V-cycles. As a result, the edges in the restored images are not recovered. However, the work in [2] does illustrate the potential for a multi-level approach to regularized signal and image deblurring.

Therefore, in the present work, we turn to the difficult task of recovering edge information in 1D signals using a multilevel approach. It is well-known in the signal processing literature that one can restore edge information by using, for example, a Tikhonov based approach where the regularization functional is a Total-variation operator [7]. Another edge-preserving approach is the PPTSVD approach [11]. These approaches, although they can work very well, can be quite computationally expensive.

Therefore, we develop a multilevel edge preserving approach that has the capability to recover edge information at a smaller expense than other traditional methods. Our algorithm uses a different restriction operator than [2] and also uses a residual correction step which actually lives on a coarse grid but regularizes over the current grid. In particular, we choose to move between grids using a Haar wavelet operator. We show that the choice of the Haar wavelet operator has the advantage of preserving matrix structure among grids, which can be exploited to obtain faster solvers on each level, and allows for a straightforward analysis of errors and residuals as we move through the V-cycle.

The structure of the paper is as follows. In Section 2 we first give the outline of a generic multilevel algorithm for discrete ill-posed problems, and then describe how we will use the Haar wavelet operator to move between grids. In Section 3, using some matrix and residual analysis, we design our pre-smoothing, coarse-grid correction and so-called residual correction strategies. We address some of the computational issues in Section 4. Section 5 contains numerical results, and Section 6 summarizes conclusions and future work.

2 Multilevel Algorithm

We first describe the main characteristics of a multilevel method as it applies to a discrete ill-posed problem. The overall idea echos what is presented in [2]. Then, we give some details of how we will move between grids.

2.1 Basic framework of a multilevel method

The main idea of a multilevel method is to define a sequence of systems of equations decreasing in size,

$$A_i x_i = b_i, \qquad 0 \leq i \leq n,$$

where the subscript i denotes the i-th level we have processed. In particular, $i = 0$ corresponds to the finest level, and $i = n$ to the coarsest level. We call these systems *coarse grid equations*. The restriction operator R and the interpolation operator P define the transfers from finer to coarser grid and vice versa. The operator at each level is defined by

$$A_{i+1} = R_i A_i P_i.$$

Then, the solution process consists of computing a regularized solution to the fine scale system (Pre-smoothing), transferring of residual from fine to coarser grid, solving (in the regularized sense) the coarse grid correction equation (Coarse-grid Correction), interpolation of corrections from coarse grid, and computing a regularized solution to the updated residual equation at the finer grid (Residual Correction). We can summarize the method recursively as follows.

Algorithm 1. ***(Multilevel V-Cycle: MGM)***
Input: A_i, b_i
Output: x_{i+1}

1. If $i = n$
2.   *RegularizedSolve* $(A_i x_{i+1} = b_i)$ *(Coarse-grid Correction)*
3. *Else*
4.   If $i \neq 0$
5.   *RegularizedSolve* $(A_i x_i = b_i)$ *(Pre-smoothing)*
6.   End If
7. $r_i = b_i - A_i x_i$
8. $b_{i+1} = R_i r_i$
9. $A_{i+1} = R_i A_i P_i$
10. $y_{i+1} = MGM(A_{i+1}, b_{i+1})$
11. $x_i = x_i + P_i y_{i+1}$
12. $r_i = b_i - A_i x_i$
13.   *RegularizedSolve* $(A_i x_c = r_i)$ *(Residual Correction)*
14. $x_{i+1} = x_i + x_c$
15. End If

In order to give a complete definition of a multilevel method, we need to define the restriction and interpolation operators, and also what we mean by Pre-smoothing, Coarse-grid correction and Residual Correction, that is, we need to define how we solve the systems on lines 2, 5, and 13. We address these issues in the next section.

2.2 Haar Decomposition

Recall that in our problem, A is an $m \times m$ matrix with $m = 2^k$. Consider the Haar matrix Transform defined by

$$W^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & 1 & -1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} = \begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix}. \tag{2}$$

Our problem $Ax = b$ can be written in the wavelet domain as

$$\hat{A}\hat{x} = \hat{b}.$$

In other words, if W^T is the Haar matrix transform[2], then $\hat{A} = W^T A W$, $\hat{x} = W^T x$, and $\hat{b} = W^T b$.
We partition the transformed problem into blocks of size 2^{k-1} to obtain the following problem

$$\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix}. \tag{3}$$

We now discuss how to use this partitioning to our advantage.

It is clear from the definition that W_1^T applied to a vector does weighted averaging, while W_2^T does weighted difference. Indeed, in the wavelet literature [1] if v is a vector then $W_1^T v$ is called the vector *scaling coefficients* and $W_2^T v$ is called the vector of *wavelet coefficients*. In the interest of space, we will not give

[2]We use W^T rather than W for ease in reading the linear algebraic equation.

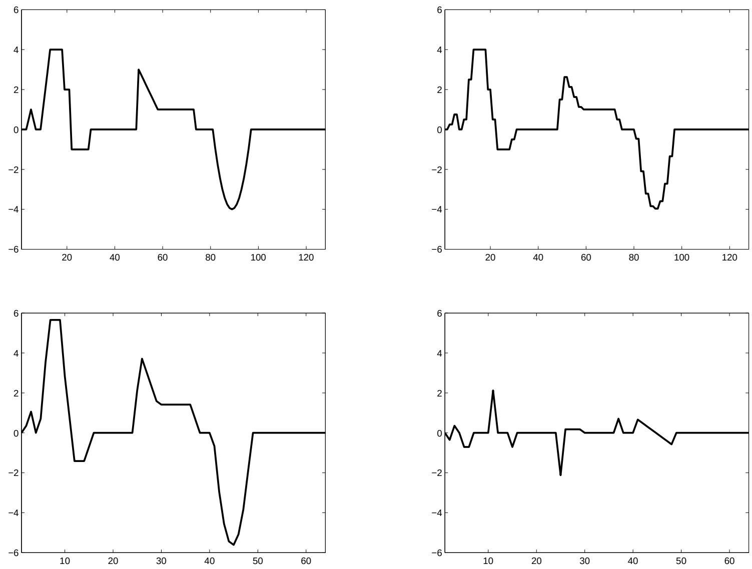

Figure 1: True solution x^{true} (top left), approximate solution $\tilde{x} = W[x_1; 0]$ (top right), scaled coefficients $\hat{x}_1$ (bottom left), and wavelet coefficients $\hat{x}_2$ (bottom right).

detailed background on wavelets here, but refer the interested reader to [1] as to why Haar wavelets are useful for decomposing signals with edges. We present only the facts most relevant to justifying the use of the Haar basis in our multilevel approach.

Notice that $W_1^T v$ is formed by weighted averaging of neighboring points, while $W_2^T v$ is formed by weighted difference. Not surprisingly, then, we observe from Figure 1 that the plot of the scaling coefficient vector $\hat{x}_1$ captures a similar amount of detail of the original image while the plot of the wavelet coefficient vector $\hat{x}_2$ is spiky. This leads us naturally to say that if we could recover the "downsampled" signal $\hat{x}_1$, we have a pretty decent approximation to the fine scale vector (in the sense that setting $\hat{x}_2$ to 0 and $\tilde{x} = W[x_1; 0]$ (see Figure 1) would already give an okay (but blocky) representation of the signal). So our first goal is to try to recover $\hat{x}_1$.

From the first block equation on (3), note that we have

$$\hat{A}_{11}\hat{x}_1 = \hat{b}_1 - \hat{A}_{12}\hat{x}_2. \tag{4}$$

We do not have $\hat{x}_2$, of course. In the next section, we discuss why it is appropriate to find a *regularized solution* to

$$\hat{A}_{11}\hat{x}_1 = b_1 \tag{5}$$

as our pre-smoothing and coarse grid correction problems, and we discuss what regularization methods are appropriate. Then, we will discuss how to recover $\hat{x}_2$ given our estimates of $\hat{x}_1$.

3 Analysis and Algorithm Details

Let us compare the submatrices $\hat{A}_{ij}$. Recall that A is a discretized version of a blurring operator. Let

$$A = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \Sigma \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

be the singular value decomposition (SVD) of A. The singular values of A, σ_i, are decreasing from 1 (assuming A has been normalized) to a very small number, with no gap. It is well known that the vectors in U and V increase in frequency from left to right, so we can think of the vectors in U_1 and V_1 as the lower-frequency vectors while U_2 and V_2 are the higher frequency. Then

$$\hat{A} = \begin{bmatrix} W_1^T U_1 & W_1^T U_2 \\ W_2^T U_1 & W_2^T U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T W_1 & V_1^T W_2 \\ V_2^T W_1 & V_2^T W_2. \end{bmatrix}, \tag{6}$$

If we look at the columns of $W_1^T U_1$, we find low-frequency vectors due to the fact that U_1 contains low-frequency column vectors and W_1^T smooths them even more. Since we know that the columns on U_1 are smoother that the columns on U_2, the $W_1^T U_2$ contains lower-frequency vectors than $W_1^T U_2$. As we mentioned above, W_2^T acts as a weighted difference. Then, the columns of $W_2^T U_1$ and $W_2^T U_2$ are higher-frequency vectors. This tells us that $W_1^T U_1$ retains the character of the lowest-frequency vectors of U. Similar analysis can be done for the the blocks on the right matrix of (6), and we therefore conclude that $V_1 W_1^T$ represents best, among all four subblocks, the characteristics of the lowest-frequency vectors of V.

In particular,

$$\hat{A}_{11} = W_1^T U_1 \Sigma_1 V_1^T W_1 + W_1^T U_2 \Sigma_2 V_2^T W_1 \tag{7}$$

can be written as $\hat{A}_{11} = W_1^T U_1 \Sigma_1 V_1^T W_1 + E$. We have the following lemma:

Lemma 1. *Given the partitioning above,*

$$\|\sigma_s(\hat{A}_{11}) - \sigma_s(W_1^T U_1 \Sigma_1 V_1^T W_1)\| \leq \|E\|_2 \leq \sigma_{2^{k-1}+1}, \qquad s = 1, \ldots, 2^{k-1},$$

where the notation $\sigma_s(B)$ means the s-largest singular value of the argument B.

Proof. Since W, U, V are matrices with orthonormal columns, it is easy to show $\|W_i^T U_i\| = 1$ and $\|V_i^T W_i\| = 1$ for $i = 1, 2$. Therefore, $\|E\|_2 \leq \sigma_{2^{k-1}+1}$ by Corollary 8.6.2 in [3]. □

Typically, in these problems, the singular value $\sigma_{2^{k-1}+1}$ of A is small, then we can say that the term $W_1^T U_1 \Sigma_1 V_1^T W_1$ is dominant in (7). This matrix is almost already in its SVD form as the matrices $W_1^T U_1$ and $V_1^T W_1$ are close to being orthogonal, and so if Σ_1 is somewhat ill-conditioned, it follows that $\hat{A}_{11}$ inherits these characteristics as well.

Specifically, let $\hat{A}_{11} = QDP^T$ be the SVD of $\hat{A}_{11}$. Let $T_1 = Q_{1:j}$ denote the first j columns of Q, and $T_2 = (W_1^T U_1)_{1:j}$ denote the first j columns of that matrix. Not surprisingly, since the first few (orthonormal) columns of U_1 were smooth and $W_1^T U_1$ is just the averaged version of U_1, the sine of the canonical angles between subspaces T_1, T_2 has to be small (where small is relative to the size of j). (A similar statement holds for the left singular vectors.) This is illustrated in Figure 2 for the blurring operator in our numerical results section. By Corollary I.5.4 in [12], define

$$cos\Theta(Range(T_1), Range(T_2)) = diag(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_j),$$

where $\tilde{\sigma}_l$ are the singular values of $T_1^T T_2$. Then,

$$sin\Theta = \sqrt{1 - \tilde{\sigma}_l^2} \qquad l = 1, \ldots, j.$$

In Figure 2, we plot the maximum sine of the canonical angles as a function of j.

This means that we expect $\hat{A}_{11}$ to still be somewhat ill-conditioned on the finest grids, but getting better conditioned the coarser we go, and that $\hat{A}_{11}$ always captures the lowest frequency characteristics of the original matrix. Now we are ready to discuss the regularization options for the pre-smoothing and coarse grid solve.

3.1 Pre-smoothing and Coarse Grid Correction

It is well known that Krylov subspace algorithms such as LSQR [10] work as a regularization methods if stopped early (i.e. before converging to the solution to the system). This is because the methods when applied to a discrete ill-posed problem try to reduce the residual in the first few iterations over the subspace

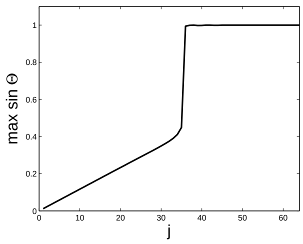

Figure 2: Maximum sine of the canonical angles between T_1 and T_2 for $j = 1, \ldots, 2^{k-1}$.

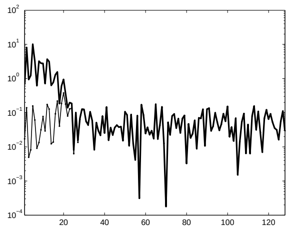

Figure 3: Spectral coefficients of b (thicker line) and spectral coefficients of $r = b - Ax$ (thinner line), where x is given by 3 iterations of LSQR.

corresponding to the largest singular values, and these are precisely the smooth modes [6]. In other words, if r denotes the residual vector after k steps of LSQR is applied to a discrete ill-posed problem with white noise, a plot of the components of $U^T r$ shows the first few have been decreased (compared to the initial residual) while the last few have remained untouched (meaning noise has not been mixed in to the solution), as shown in Figure 3.

The matrix analysis in the previous section has the following implication for our multigrid approach. Suppose we do 2 or 3 steps of LSQR on $Ax = b$, form $r \to b - Ax$, then form the residual equation in wavelet space:

$$\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \end{bmatrix}.$$

Then, consider the system

$$\hat{A}_{11}\hat{x}_1 = \hat{r_1}.$$

Looking at $Q^T \hat{r}_1$, this is equal to $Q^T W_1^T r = Q^T W_1^T U U^T r = [Q^T W_1^T U_1, Q^T W_1^T U_2](U^T r)$. So at least the first j components of this vector are the same in magnitude as they were before – that is, if they were made small on a previous grid they remain so.

We will consider the term $\hat{A}_{12}\hat{x}_2$ that appears in (4) as noise and use LSQR to find an (overregularized) solution to any intermediate system (we use b instead of r for consistency with the previous section)

$$\hat{A}_{11}\hat{x}_1 = \hat{b}_1. \tag{8}$$

We can consider the term $\hat{A}_{12}\hat{x}_2$ as noise for the following reason. Recall that $\hat{x}_2$ is the spiky wavelet coefficient vector. The operator $\hat{A}_{12}$ does a little smoothing over it (adding a slightly larger low frequency component to the it) but not much because $\hat{A}_{12}$ is not quite a "smoother" like $\hat{A}_{11}$ is. We also have that $\hat{b}_1 = W_1^T(b^{true} + e) = \hat{b}_1^{true} + \hat{e}_1$. Then, $\hat{A}_{12}\hat{x}_2$ is not white noise, but spectrally, it is below the additive white noise $\hat{e}_1$ in our applications, so we can ignore it.

At the coarsest level, though, while we may have adequately reduced the residual at the smoothest modes, we may not have obtained as much information over the higher frequency modes (some of the highest frequencies have disappeared, by this level, of course). We saw that if the original signal has edges, the coarse grid problem, since this is the system for the scaling coefficients, may still have edges. To recover that information here, we solve

$$\min_{\hat{x}_1} \|\hat{A}_{11}\hat{x}_1 - \hat{b}_1\|_2^2 + \lambda^q \|L\hat{x}_1\|_q^q, \tag{9}$$

where L is a first-derivative operator and $1 < q < 2$. This allows us to recover some edge information, but, since we are at a coarse level, this is a much more computationally tractable problem than if we had tried to employ it on the finest grid. On the other hand, if we are at a coarse enough grid, we can replace this simply by a direct solve, as there are no longer any small singular values to magnify noise in the right-hand side, and the noise is barely perceptible in the right-hand side at this level.

In summary, we have:

- Pre-smoothing: take a few (2 or 3) iterations of LSQR on $Ax = b$,
- Coarse-grid Correction: If $\hat{A}_{11}$ is somewhat ill-conditioned, solve (9), otherwise, solve $\hat{A}_{11}\hat{x}_1 = \hat{b}_1$ exactly by a direct or iterative method.

3.2 Residual Correction

We need to recover the wavelet coefficients $\hat{x}_2$ too in order to adequately recover edges but not make the solution look too artificially blocky. For this part, we define the residual correction step. Let us rewrite equation (3), assuming that we have an approximation $\hat{x}_1^*$ of $\hat{x}_1$,

$$\begin{bmatrix} \hat{A}_{12} \\ \hat{A}_{22} \end{bmatrix} \hat{x}_2 = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} - \begin{bmatrix} \hat{A}_{11} \\ \hat{A}_{21} \end{bmatrix} \hat{x}_1^*. \tag{10}$$

To solve this problem, we use again a Tikhonov formulation to keep certain smoothness in the solution.

$$\min_{\hat{x}_2} \left\{ \| \begin{bmatrix} \hat{A}_{12} \\ \hat{A}_{22} \end{bmatrix} \hat{x}_2 - (\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} - \begin{bmatrix} \hat{A}_{11} \\ \hat{A}_{21} \end{bmatrix} \hat{x}_1^*)\|_2^2 + \lambda^q \|L(W_1\hat{x}_1^* + W_2\hat{x}_2)\|_q^q \right\}, \tag{11}$$

where L is a first-derivative operator and $1 < q < 2$ as in Equation (9).

From a computational perspective, it is very important to note that the regularization is applied at a "finer" scale than the unknown vector $\hat{x}_2$ for which we are optimizing.

4 Computational issues

One nice feature of the Haar decomposition is that we can show [9] that it preserves the type of matrix structure we typically see in blurring operators (Toeplitz or Toeplitz-plus-Hankel, depending on which boundary conditions are used). This means we can have fast matrix-vector products and also preconditioners [**?**][**?**] for any intermediate interior linear least squares problems that must be solved iteratively (see next two subsections). Furthermore, we can show [9] that if the original matrix A is Toeplitz with a bandwidth less than $m/2$, the bandwidth of the coarser scale systems is reduced, and can be reduced as far as to a tridiagonal matrix depending on how many levels are used, again reducing the overall computational burden further down the V-cycle.

4.1 The Coarse-Scale Solve and Parameter Issue

Recall that we may need to solve Equation (9) on the coarsest grid if not enough levels are used. If $q = 2$, we could use a hybrid approach [8] to solve the Tikhonov problem and simultaneously select the regularization parameter in an efficient manner. However, we typically take $1 < q < 2$, and fairly close to 1 to preserve the edges at the coarsest grid. In this work we used the algorithm *Newton with Line Search* (newtonls) as implemented in [5], which actually uses direct factorization approaches in the inner iteration. We are only working with 1D signals and are assuming the coarse grid is coarse enough where a direct method is feasible; therefore, we do not consider this further. Choosing a good regularization parameter specifically for (9) is beyond the scope of this paper. In our numerical section, we present results for which λ was chosen to be optimal over a small, discrete set of choices. Choosing the parameter at this level is the subject of current research.

4.2 Wavelet-Correction Step

When $1 < q < 2$, we solve (11) using a damped Newton approach. Recall, as in Figure 1, the wavelet coefficients are 'spiky'. It is still worthwhile to use q as close to 1 as possible to recover wavelet coefficients accurately enough to represent the edges appropriately on the next finest grid level. At each step of the damped Newton approach, it is straightforward to show that we have to solve a linear least squares problem similar to the one in [5]. Recall that $\hat{x}_2$ lives on a coarser grid than where we are enforcing the regularization (at the next finest grid level). However, towards the end of the cycle if $\hat{x}_2$ is too large to use a direct method, we can use an iterative method to solve each of the linear least squares problems.

Clearly (11) is a harder problem to solve and therefore there is more computational work when q is closer to 1. We may therefore consider taking q equal to or at least closer to 2 at the last one or two correction steps, with some sacrifice of quality of the solution (See Figure 4 for an example with $q = 1.5$). As in the previous subsection, we consider the regularization parameter selection issue to be beyond the scope of the present work.

5 Numerical Results

We conclude with illustrations of the performance of our algorithm. All computations were performed in MATLAB. Our example is a signal restoration problem which consists of recovering the original signal from a blurred and noisy signal. The blur operator is defined by an $N \times N$ symmetric banded Toeplitz matrix A, where its first row is defined by the vector

$$z = \frac{1}{2\pi\sigma^2}[\exp(-([0 : band - 1].^2)/(2\sigma^2)); zeros(1, N - band)].$$

We use $N = 128$, and the parameters $\sigma = 3$ and $band = 7$. The condition number of A is $K(A) := \|A\|\|A^{-1}\| = 2.3 \times 10^6$. The exact solution, represented by x^{true}, is a vector of length 128 shown in Figure 1. The noise-free blurred signal , represented by b^{true}, is computed as $b^{true} = Ax^{true}$. The elements of the noise vector e are normally distributed with zero mean, and standard deviation is chosen such that $\frac{\|e\|_2}{\|b^{true}\|_2} = 5 \times 10^{-2}$. The noisy right-hand side of our system is defined by $b = b^{true} + e$; it is shown in Figure 4.

For the regularization term in Equations (9) and (11), we use the matrix

$$L = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{pmatrix},$$

which is a discrete approximation to the first derivative operator on a regular grid, with no assumptions on boundary conditions.

For the solutions shown in Figure 4, we solve the Coarse-Grid Correction by applying 3 iterations of LSQR. The operator at the coarsest level has size 16×16 and a direct solve is applied. We define 10 logarithmically spaced values of λ from 10^{-3} to 1. The value of λ is the same at all levels (which is not the optimal choice). The signals in Figure 4 show the optimal restorations chosen among the 10 candidates such that $\|x^{true} - x^{MGM}\|_2$ is minimum. We test the method with $q = 1.01$ and $q = 1.5$ in Equation 11. We see in Figure 4 that $q = 1.01$ recovers the edges better than method using $q = 1.5$ in the residual correction. As we see in Figure 4 without the residual correction the solution is blocky.

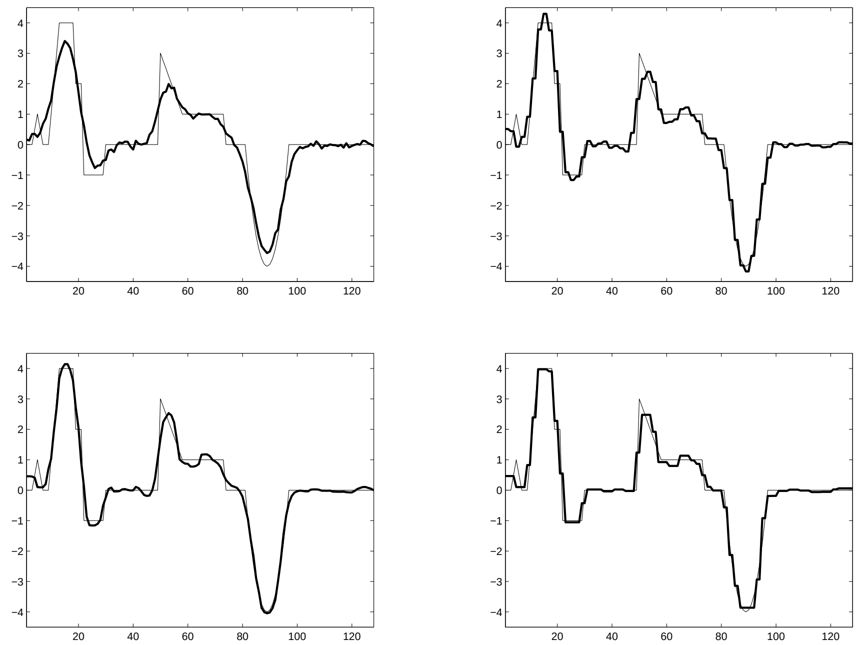

Figure 4: Blurred Noisy right-hand side, b (top left), and MGM solution without residual correction step (top right) MGM solution with $q = 1.5$ (bottom left), and MGM solution with $q = 1.01$ (bottom right). In all figures the exact solution x^{true} appears in thin line.

6 Conclusion

In this paper, a new multilevel algorithm for discrete ill-posed problems was presented. The transformation of the problem to the wavelet domain using the Haar Wavelets permits to define a multilevel algorithm that keeps matrix structures. Numerical experiments illustrate that this is a promising technique for ill-posed problems. Extension of these results to 2D and 3D image deblurring should be straightforward but non-trivial, and are the subject of our ongoing research. Future work includes efficient parameter selection techniques, number of smoothing steps at intermediate grids, and number of V-cycles.

References

[1] A. Boggess and F. J. Narcowich. *First Course in Wavelets with Fourier Analysis*. Prentice Hall, 2001.

[2] M. Donatelli. A Multigrid for image deblurring with Tikhonov regularization. *Numerical Linear Algebra with applications*, 12:715–729, 2005.

[3] G. H. Golub and C. F. Van Loan. *Matrix Computations (3rd. Ed.)*. The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[4] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, 1998.

[5] M. Jacobsen. DIP - Final Assignment Tikhonov regularization and non-Euclidean Norms.

[6] M. E. Kilmer and G. W. Stewart. Iterative regularization and minres. *SIAM Journal Matrix Analysis and Applications*, 21:613–628, 1992.

[7] S. Osher L. I. Rudin and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[8] P. C Hansen M. E. Kilmer and M. I. Español. A projection-based approach to general-form Tikhonov regularization. *SIAM Journal on Scientific Computing*, 29:315–330, 2007.

[9] M. E. Kilmer M. I. Español and D. P. O'leary. Multilevel approaches for the total least-squares method in deblurring problems. In progress. Preliminary work presented at iciam 07.

[10] C. C. Paige and M. A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transaction on Mathematical Software*, 8:43–71, 1982.

[11] J.M. Rasmussen P.C. Hansen, M. Jacobsen and H. Srensen. *Methods and Applications of Inversion*, chapter The PP-TSVD algorithm for image restoration problems, pages 171–186. Springer Berlin/Heidelberg, 2000.

[12] G. W. Stewart and J. G. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.

[13] V. E. Henson W. L. Briggs and S. F. McCormick. *A Multigrid Tutorial-2nd Ed.* SIAM, 2000.