Joint Inverse Problems for Signal Reconstruction via Dictionary Splitting

Elham Sakhaee  
University of Florida, Gainesville, FL. USA  
Email: e.sakhaee@ufl.edu

Alireza Entezari  
University of Florida, Gainesville, FL. USA  
Email: entezari@ufl.edu

Abstract—Sparse signal recovery from limited and/or degraded samples is fundamental to many applications such as medical imaging, remote sensing, astronomical and seismic imaging. Discrete Wavelet Transform (DWT) has been commonly used for sparse representation of signals; nevertheless, due to its shift-variant nature, pseudo-Gibbs artifacts are present in the recovered signals. Using the redundant Shift-invariant Wavelet Transform (SWT) is the ideal solution to obtain shift-invariance; however, high redundancy factor of SWT limits its application in practical settings.

We propose a dictionary splitting approach for sparse recovery from incomplete data, that leverages the ideas of cycle spinning in combination with Bregman splitting. The proposed method significantly improves conventional signal reconstruction with DWT, offers the advantages of SWT, and overcomes high redundancy factor of SWT. We solve parallel sparse recovery problems with orthogonal dictionaries (DWT and its permuted versions), while we impose consistency between the results by updating the recovered image at each iteration. Our experiments demonstrate that few shifts are sufficient to achieve reconstruction accuracy as high as recovery with SWT, and significantly reduces its computational cost and redundancy factor.

I. INTRODUCTION

Sparse recovery of signals from limited and/or degraded samples is key to several image processing applications, such as denoising, reconstruction, and deblurring. Discrete Wavelet Transform (DWT) has been widely used in image processing tasks for sparse representation of the underlying signal, owing to its effectiveness in sparsifying natural/biomedical images. DWT in combination with Total Variation (TV) has been shown effective for medical imaging applications and reconstruction from under-sampled data [1]. While DWT is efficient in sparse representation of signals, it introduces pseudo-Gibbs artifacts to the recovered image, mostly attributed to its shift-variant nature [2].

In order to overcome these artifacts, ideally, one replaces DWT with Shift-invariant Wavelet Transform (SWT) to achieve translation-invariance. While SWT has been successfully applied to denoising and reconstruction of images [3], [4], its high redundancy factor limits its application in practical settings.

Cycle spinning was introduced in [2] as an alternative to achieve translation-invariant denoising. In cycle spinning, the noisy signal is translated to multiple shifts. Each translated signal is then denoised separately. The resulting signals are translated back, and linearly averaged to obtain the final (denoised) signal. This approach, in fact, leverages the shift-variance of wavelet transform. In other words, wavelet coefficients of the original signal are, in general, different from wavelet coefficients of translated versions, which makes denoising problems independent from each other.

While cycle spinning is efficient in denoising (and recent work has shown its effectiveness in reconstruction [3]), since each shifted signal is tackled separately, there is no guarantee that resulting signals are consistent (agree with each other and represent the same signal). Kamilov et al., [5] discussed the inconsistency issue, proposed a denoising approach for consistent cycle spinning with 1-level Haar transform, and established its equivalence to TV minimization.

Another shortcoming of cycle spinning is linear averaging for obtaining the final result from multiple translated versions. Fletcher et al., [6] discussed that linear averaging is not optimal for finding the final image, and proposed recursive cycle spinning, in which they generate a sequence of estimates of the signal by iteratively cycling through the shifts. At each iteration a shrinkage operator is applied on the input signal translated to a different shift, the result is translated back, and used as the input for the next iteration. They discuss that cycle spinning, in essence, projects the noisy signal onto several lower dimensional subspaces and averages the projections, while the method in [6] projects the signal onto the intersection of these subspaces. We note that, SWT spans the intersection of all possible subspaces in cycle spinning, hence, spans the optimum lower dimensional subspace in this context.

In contrast to existing work, we propose a parallel approach for image reconstruction/denoising that allows for obtaining results as accurate as full SWT with only few number of shifts, and significantly reduces the computational cost compared to SWT and conventional cycle spinning. We formulate the problem of signal recovery with sparsity in SWT, as multiple subproblems with sparsity in DWT and its permuted versions, hence, benefit from projection onto the intersection of lower dimensional subspaces. Our formulation inherently enforces consistency between the recovered signals by updating the final image at each algorithm iteration. In our method, the linear averaging stage of cycle spinning is replaced by an (efficient) quadratic problem. Unlike [6], our method allows for parallel computation of sparse representations at different shifts; therefore, adding number of shifts does not adversely impact the computational cost.
II. CONSISTENT DICTIONARY SPLITTING

The problem of recovering a signal in $\mathbb{R}^N$ from limited/degraded measurements, $f$, can be written as an $\ell_1$ minimization problem:

$$\arg\min_{u \in \mathbb{R}^N} \|\Psi u\|_1 + \frac{\mu}{2} \|Au - f\|_2^2,$$  \hspace{1cm} (1)

where $A$ is the sensing matrix (e.g., partial Fourier matrix restricted to a subset of frequencies, Radon transform, identity, or blur kernel), $\Psi$ represents a sparsifying transform, and $\mu$ is the penalty parameter that controls the level of sparsity in the solution. Recent work has provided efficient methods for solving (1) (e.g., [7], [8], [9]).

Since sparsity of the transform coefficients, $\Psi u$, is crucial to success of the recovery problem (1), ample research has been carried out on the choice of $\Psi$. For example, finite differences have been shown to be effective for compressed sensing magnetic resonance imaging (CS-MRI) [10], [11]. The composite of DWT and finite differences is most commonly used for reconstruction from incomplete data [1], [12]. Geometric transforms, such as Contourlets and Curvelets have also been used for MRI reconstruction [13], [14]; however, long computational time due to their high redundancy factor has motivated research on improved algorithms [15].

The optimal alternative to DWT, to achieve shift-invariance and to avoid pseudo-Gibbs artifacts in reconstruction, is to use SWT [4]; however, SWT with $p$ levels of decomposition has redundancy factor $3p + 1$, which for a $256 \times 256$ image with full decomposition amounts to a redundancy factor of 25.

In our method, we split the signal recovery problem with SWT into several subproblems with orthogonal DWT and its permuted versions. In this regard, we consider the cycle spinning approach [2]. In cycle spin denoising, $M$ problems in the following form are solved separately:

$$\arg\min_{\theta_i \in \mathbb{R}^N} \|\theta_i\|_1 + \frac{\mu_i}{2} \|A\Psi^{-1}\theta_i - f_i\|_2^2 \hspace{1cm} i = 1, \ldots, M$$ \hspace{1cm} (2)

where $\theta_i$ and $f_i$ denote the DWT coefficients and measurements corresponding to the $i$th shift, and $A$ is the identity matrix. This is equivalent to translating the noisy signal to $M$ shift locations (via operator $S_i$, for $i = 1, \ldots, M$, and applying wavelet thresholding at each shift separately. The final signal is then obtained by averaging: $\frac{1}{M} \sum_{i=1}^M S_i^{-1} \Psi^{-1}\theta_i$, where $S_i$ is the shift operator corresponding to the $i$th shift.

We note that for denoising and deconvolution (deblurring) tasks, obtaining $f_i$ is straightforward. For reconstruction tasks such as reconstruction from sub-sampled Fourier data, we benefit from equivalence of translation in spatial domain and phase shift in frequency domain. In other words, $f_i$ can be obtained by multiplying $f$ with corresponding Fourier multiplier for a specific phase shift [11].

Since the $M$ problems in (2) are solved separately for $\theta_i$, there is no guarantee that $S_i^{-1} \Psi^{-1}\theta_i$ represent the same signal for all $i$. We address this inconsistency issue by writing the problems in the form of (1):

$$\arg\min_{u \in \mathbb{R}^N} \|\Psi S_i u\|_1 + \frac{\mu}{2} \|Au - f\|_2^2 \hspace{1cm} i = 1, \ldots, M$$ \hspace{1cm} (3)

We then solve the $M$ problems in (3) simultaneously:

$$\arg\min_{u \in \mathbb{R}^N} \frac{\mu M}{2} \|Au - f\|_2^2 + \|\Psi S_1 u\|_1 + \cdots + \|\Psi S_M u\|_1$$ \hspace{1cm} (4)

An efficient approach to solve (4) is via Bregman iterations [16], in which we iterate between updating the image and sparse representation variables $\theta_i$:

$$\arg\min_{\theta_1, \ldots, \theta_M, u \in \mathbb{R}^N} \frac{\mu M}{2} \|Au - f\|_2^2 + \sum_{i=1}^M \|\theta_i\|_1 + \lambda \sum_{i=1}^M \|\theta_i - \Psi S_i u - b_{\theta_i}\|_2^2$$ \hspace{1cm} (5)

where $b_{\theta_i}$ ($i = 1, \ldots, M$) denotes the vector of Lagrange multipliers (also updated at each iteration) and $\lambda$ is called the augmented Lagrangian penalty parameter. We note that (5) is an optimization problem in $M + 1$ variables. It can be solved via Alternating Directions Method of Multipliers (ADMM) as described in Algorithm 1, which facilitates enforcing consistency between the shifted versions of the signal at each iteration by updating $u$. 

---

**Fig. 1.** Reconstruction of Shepp-Logan phantom from 3.06% of Fourier data. While 4-shift conventional cycle spinning (cycle spinning) and RecPF (TV-DWT) clearly improve over DWT, they do not achieve exact reconstruction as in SWT and the proposed approach (proposed method). The proposed method achieves reconstruction accuracy as high as SWT, while reducing its computational cost by about 20 times.
**Algorithm 1** Dictionary Splitting Cycle Spin Reconstruction for recovery from limited data

<table>
<thead>
<tr>
<th>Initialize:</th>
<th>$f^0 = f$, $\theta^0_1, \ldots, \theta^0_M = 0$, $b^0_1, \ldots, b^0_M = 0$, $k = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>while</td>
<td>$</td>
</tr>
<tr>
<td>for $l = 1$ to $L$</td>
<td></td>
</tr>
<tr>
<td>$u^{k+1} = \arg\min_{u \in \mathbb{R}^N} \frac{M}{2}</td>
<td></td>
</tr>
<tr>
<td>$\theta^{k+1} = \arg\min_{\theta_i \in \mathbb{R}^N}</td>
<td></td>
</tr>
<tr>
<td>$b^{k+1}_i = b^{k}_i + (\Psi S_i u^{k+1} - \theta^{k+1}_i)$</td>
<td></td>
</tr>
<tr>
<td>$k = k + 1$</td>
<td></td>
</tr>
<tr>
<td>return $u^k$.</td>
<td></td>
</tr>
</tbody>
</table>

We note that the update step for $u$ is in quadratic form and has a closed-form solution:

$$M(\mu A^* A + \lambda I)u = \mu M A^* f + \lambda \sum_{i=1}^M S^T_i \Psi^T (\theta^k_i - b^k_i),$$

where $A^*$ is the conjugate transpose of $A$.

When $A$ is partial Fourier (as in MRI reconstruction), the above equation can be efficiently solved by Hadamard product and fast Fourier transform. Denote partial Fourier matrix by $RF$, where $R$ is the row-selector operator that selects random rows of Fourier matrix $F$. Then:

$$F^T (\mu M R^T R + \lambda M I)Fu =$$

$$\mu M R^T f + \lambda \sum_{i=1}^M S^T_i \Psi^T (\theta^k_i - b^k_i).$$

Let $\Lambda = \mu M R^T R + \lambda M I$, which is a diagonal matrix, hence, its inverse is trivial.

$$u = F^\Lambda^{-1} F (\mu M R^T f + \lambda \sum_{i=1}^M S^T_i \Psi^T (\theta^k_i - b^k_i))$$

When $A$ is identity (as in denoising), the update stage for $u$ reduces to:

$$u = \frac{1}{M(\mu + \lambda)} (\mu M f + \lambda \sum_{i=1}^M S^T_i \Psi^T (\theta^k_i - b^k_i))$$

Updates for $\theta_i$ are performed via shrinkage operator (wavelet thresholding). Since updates for $\theta_i$ and $b^i$ can be done in parallel for all $M$ problems, one can choose to add many additional shifts. However, as we show in the experiments section, $M$ can be as small as 4, yet the algorithm achieves the accuracy of SWT overcomplete dictionary which has redundancy factor $3p + 1$.

### III. Results and Discussion

We compare the proposed approach against reconstruction with sparsity in DWT, conventional cycle spinning, reconstruction with full SWT (the optimal case), and reconstruction with sparsity in both TV and DWT (TV--DWT) which is most commonly used for reconstruction from limited data [12].

We fix the system matrix to partial Fourier matrix, therefore, the problem is to recover an image from sub-sampled Fourier data as in CS-MRI [1]. Sampling in Fourier space is performed with radial sampling. The amount of under-sampling in Fourier space is reported in terms of sampling rate which is percentage of retained Fourier samples. In all experiments, Haar is used as DWT and SWT basis.

Cycle spinning is carried out by solving multiple optimization problems as in (2) for different shifts of the signal separately, and then averaged. The implementation used for TV--DWT is the publicly available code for RecPF.

The parameters for each method are obtained empirically, and tuned for best results for fair comparison. The accuracy of reconstruction is evaluated in terms of Signal to Noise Ratio (SNR) and reported in logarithmic scale.

The first experiment is carried out with MATLAB Shepp-Logan phantom of size $256 \times 256$ at $3.06\%$ sampling rate. We compare the proposed method with conventional cycle spinning with 4 shifts, and the RecPF method [12]. The parameters are tuned empirically and set to $\lambda = 10^2$, $\mu = 2 \times 10^5$ with 100 iterations for each method. For RecPF more iterations were required, and number of iterations is set to 1000. Fig. 1 shows that while cycle spinning (cycle spinning) and RecPF clearly improve the reconstruction accuracy compared to DWT, SWT and proposed method offer exact reconstruction at this sampling rate. We note that proposed method with four 2D shifts (shift vectors=$\{[0,0]^T, [1,1]^T, [2,2]^T, [3,3]^T\}$) offers an approximately 22 times reduction in computational cost compared to SWT, which is crucial for large-scale recovery problems.

Accuracy of reconstruction for DWT, cycle spinning, SWT and the proposed method is compared in Fig. 2(a) for Shepp-Logan phantom of size $128 \times 128$. proposed method and cycle spinning are carried out with 8 shifts (shift set $\{0,1,2,3,4,5,6,7,8\}$). For all methods the number iterations is set to $10^3$. $\lambda = 10^3$ and $\mu = 2 \times 10^5$. As the plot suggests, the expected phase transition behavior of CS reconstruction occurs at much lower sampling rates for proposed method and SWT, compared to DWT and cycle spinning. However, proposed method achieves similar accuracy as SWT.

---

1. [http://www.caam.rice.edu/~optimization/L1/RecPF](http://www.caam.rice.edu/~optimization/L1/RecPF)

2. For simplicity of presentation, we denote such shifts set as $\{0,1,2,3,4\}$. 
Ground Truth

DWT
SNR: 18.30 dB
Time: 4.64s

cycle spinning
SNR: 21.25 dB
Time: 15.36s

SWT
SNR > 80 dB
Time: 211.77s

proposed method
SNR > 80 dB
Time: 10.83s

Fig. 3. Reconstruction of the brain phantom from 11.6% of Fourier data. SWT and proposed method both achieve exact reconstruction while SWT requires approximately 20 times more computational time.

Ground Truth

DWT
SNR: 21.38 dB
Time: 23.14s

TV-DWT
SNR: 22.69 dB
Time: 18.98s

proposed method
SNR: 26.03 dB
Time: 56.08s

Fig. 4. Reconstruction of a real MRI image from 7.66% of Fourier data. TV-DWT promotes piece-wise constancy in the reconstructed image; therefore, fine features are recovered as flat regions. For example, the fork-like feature is smeared in the TV-DWT image.

Fig. 5. The impact of different shift sets on reconstruction accuracy. The fact that certain shift sets result in identical reconstruction accuracy implies the redundancy of shifts in SWT.

with far lower computational cost (reported as elapsed time in Fig. 2(b)). It is worth noting that ADMM machinery allows for distributed optimization for large-scale problems. In this paper, our implementation is not distributed or parallelized, yet significantly reduces computational cost.

The next experiment is carried out with a realistic brain phantom [17] of size $256 \times 256$, as shown in Fig. 3. While proposed method offers similar accuracy as SWT (both achieve exact reconstruction), it requires approximately 22 times less computational cost.

Fig. 4 shows the reconstruction of a real (noisy) MRI image from 7.66% of Fourier data. proposed method achieves reconstruction accuracy as high as SWT with 4 shifts, while artifacts are evident in RecPF reconstruction.

The plot in Fig. 5 presents the impact of different shift sets on reconstruction accuracy for Shepp-Logan phantom of size $128 \times 128$. The observation that certain shift sets result in identical reconstruction accuracy implies the redundancy of shifts in SWT. This verifies the benefits of the proposed approach in reducing the redundancy factor of SWT.

IV. SUMMARY AND FUTURE WORK

We propose an approach for reconstruction from limited/degraded data using an ADMM formulation for cycle spinning. The benefits of our approach is two-fold. First, compared to cycle spinning, our method ensures consistency between reconstructions from shifted versions of the signal. Second, the quadratic update formula for obtaining the signal at each iteration is superior to linear averaging of cycle spinning. Moreover, the computational cost of our method is significantly lower than reconstruction using SWT.

In this paper, we presented results for reconstruction from sub-sampled Fourier data, however, the proposed method is readily applicable to denoising, deconvolution, and tomographic reconstruction, where measurements corresponding to shifted signals can be obtained in closed-form [18].

Since the optimal subset of shifts is signal-dependent, previous work on cycle spinning [19] have suggested use of a set of random shifts. In our future work, we plan to investigate the impact of choosing different random shift sets at each iteration. Another promising direction of research is to apply our dictionary splitting method for other overcomplete dictionaries, such as Curvelets or Contourlets.

ACKNOWLEDGMENT

This research has been partially funded by the NSF grant IIS-1617101, ONR grant N00014-14-1-0762, and UFII graduate fellowship.
REFERENCES


