ABSTRACT
Plug-and-play (PnP) methods have recently emerged as a powerful framework for image reconstruction that can flexibly combine different physics-based observation models with data-driven image priors in the form of denoisers, and achieve state-of-the-art image reconstruction quality in many applications. In this paper, we aim to further improve the computational efficacy of PnP methods by designing a new algorithm that makes use of stochastic variance-reduced gradients (SVRG), a nascent idea to accelerate runtime in stochastic optimization. Compared with existing PnP methods using batch gradients or stochastic gradients, the new algorithm, called PnP-SVRG, achieves comparable or better accuracy of image reconstruction at a much faster computational speed. Extensive numerical experiments are provided to demonstrate the benefits of the proposed algorithm through the application of compressive imaging using partial Fourier measurements in conjunction with a wide variety of popular image denoisers.

Index Terms— plug-and-play, stochastic optimization, variance reduction, image reconstruction

1. INTRODUCTION

We consider the reconstruction of an image \( x^* \in \mathbb{R}^n \) from a set of noisy measurements \( y \in \mathbb{C}^m \), formulated as

\[
y = A(x^*) + \epsilon,
\]

where \( A : \mathbb{R}^n \rightarrow \mathbb{C}^m \) characterizes the forward model of the imaging system known a priori, \( y = \{y_i\}_{1 \leq i \leq m} \) is the set of measurements, and \( \epsilon = \{\epsilon_i\}_{1 \leq i \leq m} \) is additive (random) noise present in the system. In many applications, the forward model is ill-posed with more unknowns than the number of measurements, i.e. \( m \ll n \), due to limited sampling budget or acquisition time. Without prior knowledge about the structure of the ground truth image, the problem is generally impossible to solve. Therefore, one would like to derive image reconstruction algorithms that exploit prior knowledge in the image in order to generate a proper estimate, \( \hat{x} \), of the image.

Image denoising and reconstruction algorithms have been thoroughly studied. While classically image priors are designed using hand-crafted features such as wavelets, many data-driven image priors based on deep learning and neural networks have been proposed in recent years [1, 2, 3, 4]. A number of techniques have been proposed, such as algorithm unrolling, using neural networks as generative models of images, end-to-end training, and so on; see [5, 6, 7] for recent overviews.

This paper focuses on the plug-and-play (PnP) methods for image reconstruction, first proposed in [8]. PnP methods have recently emerged as a powerful framework for image reconstruction that can flexibly combine different physics-based observation models with data-driven image priors in the form of denoisers, and achieve state-of-the-art image reconstruction quality in many applications [9, 10]. To motivate the methodology of PnP, note that many classical approaches can be viewed as solving an optimization problem that is composed of a data-fidelity term that incorporates the observation model and a regularization term that incorporates image priors. Proximal algorithms such as proximal gradient methods and alternating direction method of multipliers are popular solvers, where the proximal map avoids the need to find derivatives of complicated image regularizers which may be non-differentiable. In [8], the authors made the keen observation that the proximal map is equivalent to finding the maximum a posteriori (MAP) image denoising estimate upon a suitable probability prior. Bearing this perspective, PnP methods replace this proximal update by “plugging in” an image denoiser of the user’s choice, such as non-local means (NLM) [11], block matching 3D filtering (BM3D) [12], or of particular interest, a state-of-the-art convolutional neural network (CNN) based image denoiser [13].

Generally, the PnP framework allows the injection of sophisticated denoisers into iterative reconstruction algorithms, where first-order algorithms are particularly appealing due to their scalability to large-scale problem instances and amenability to computing advances. When combined with gradient descent (GD), this leads to the PnP-GD algorithm which alternates between taking a GD step with respect to the data-fidelity term, followed by a denoising step using the plugged-in denoiser of choice [14]. Moreover, the data-fidelity term is often given as a finite sum of sample losses over each observation, e.g. when the observations are collected in an independently and identically distributed manner. Consequently, one can leverage the finite sum structure to design PnP methods that take advantage of stochastic algorithms such as Stochastic Gradient Descent (SGD), which was done in [15] to derive the PnP-SGD algorithm for image reconstruction. In [15], the authors demonstrated that PnP-SGD achieves a considerable speed up in computation time, while achieving the same reconstruction quality as PnP-GD.

To further accelerate the reconstruction, the main contribution of this paper is to introduce the use of stochastic variance-reduced gradient (SVRG) methods [16] into the PnP framework. It is well-
known from the optimization literature that SGD tends to converge with a much larger number of iterations compared with GD, due to the higher variability in its search direction, while the per-iteration cost of GD is expensive when the data size is large. The crux of SVRG [16] is a carefully-designed stochastic gradient with reduced variance by using a reference batch gradient that is periodically updated to save computation. Therefore, it is of great interest to investigate whether SVRG brings additional benefits to the PnP framework. We propose a new algorithm called PnP-SVRG, which can be regarded as replacing the proximal map in the proximal SVRG algorithm [17] by a wide variety of popular image denoisers including NLM, CNN and BM3D. We demonstrate through extensive numerical experiments that compared with PnP-SGD and PnP-GD, PnP-SVRG achieves comparable or better accuracy of image reconstruction at a faster computational speed.

The rest of this paper is organized as follows. First, we provide backgrounds on GD-based algorithms for PnP in Section 2. We develop the stochastic variance-reduction based PnP algorithm, called PnP-SVRG, in Section 3. We then evaluate the performance of the proposed PnP-SVRG algorithm in the context of compressive image reconstruction using partial 2D Fourier measurements in Section 4, with comparisons to PnP-GD and PnP-SGD. In Section 5, we conclude our findings and discuss future directions. Throughout the paper, the symbols $\mathbb{R}$ and $\mathbb{C}$ denote the set of real and complex numbers, respectively. We use boldfaced symbols to represent vectors and matrices. For any matrix $M$, we let $\|M\|_F$ denote the Frobenius norm, $M^T$ the transpose, and $\odot$ the Hadamard product.

2. BACKGROUND

In this section, we first introduce the PnP-GD and PnP-SGD algorithms, and then discuss the use of stochastic variance-reduced gradient methods in finite-sum stochastic optimization.

2.1. Plug-and-Play with GD and SGD

To solve for the image from the observations (1), we start with a data-fidelity loss function that takes the following finite-sum form:

$$
\ell(x) = \frac{1}{m} \sum_{i=1}^{m} \ell_i(x),
$$

where $\ell_i(x) = \ell_i(x; y_i)$ depends on the $i$th measurement. A popular choice is minimizing the least-squares difference between the observed measurements and those produced by the current image estimate. A common approach to incorporate image priors is to solve the regularized optimization problem:

$$
\min x \ell(x) + \lambda g(x),
$$

where $g(\cdot)$ is the regularizer and $\lambda > 0$ is some regularization parameter. The above optimization problem can be solved by proximal gradient methods, which update the image estimate as

$$
x_t = \text{prox}_{\lambda g}(x_{t-1} - \eta \nabla \ell(x_{t-1})),
$$

where $\text{prox}_\lambda$ is the proximal map of $g(\cdot)$, $\nabla \ell(\cdot)$ is the gradient of $\ell(x)$ and $\eta > 0$ is a proper learning rate. The key step of PnP methods is to replace the proximal map with respect to some choice of regularization function $g(\cdot)$ by an image denoising routine of choice. More specifically, we define an image denoiser

$$
denoise_x : \mathbb{R}^n \mapsto \mathbb{R}^n,
$$

which is parameterized by some estimate of the noise level present in the image $\sigma > 0$. Plugging the above denoiser into (4) leads directly to the PnP-GD algorithm in Alg. 1. Similarly, leveraging the finite-sum form of the data-fidelity term leads to the PnP-SGD algorithm [15] in Alg. 2, where a mini-batch of $B$ measurements can be used to construct the search direction to balance the convergence speed and per-iteration computational cost.

While the PnP method yields a simple form, its theoretical performance is still largely unknown and remains an active field of study. The authors of [10] proposed a set of conditions such that a denoising routine can be handled similarly to a Moreau proximal mapping. In [13], the authors borrowed from operator theory to prove convergence for strongly convex loss functions with properly trained CNN-based image denoisers. In a related approach, regularization by denoising [18] can be used to prove convergence of a slightly broader class of algorithms including PnP methods.

2.2. Stochastic variance-reduced gradient methods

SVRG [16], together with many variants and offsprings, has become a popular algorithmic approach in finite-sum optimization to achieve a better runtime complexity than both GD and SGD. In a nutshell, to train a popular algorithmic approach in finite-sum optimization to achieve a better runtime complexity than both GD and SGD. In a nutshell, to train a
and a proximal map at each iteration, and achieves a similar runtime speed-up.

3. PROPOSED ALGORITHM

Motivated by the success of SVRG in accelerating finite-sum optimization, we propose a novel algorithm that combines SVRG with PnP methods to solve image reconstruction. The proposed algorithm, dubbed PnP-SVRG, is presented in Alg. 3. Specifically, the algorithm contains two loops. The outer loop, sometimes referred to as an epoch, is used to update the reference point \( \hat{x} \) as well as its full gradient \( w = \ell(\hat{x}) \) in a periodic manner. The inner loop, on the other end, performs stochastic gradient updates using the variance-reduced stochastic gradient \( \tilde{v}_t \), which can be computed in mini-batches to further reduce variance and improve performance.

The proposed PnP-SVRG algorithm can be regarded as replacing the proximal map in the proximal SVRG algorithm [17] by a wide variety of popular image denoisers such as CNN, BM3D, NLM, and so on, which will be examined in detail in Sec. 4.

**Algorithm 3 PnP-SVRG**

Input: \( x_0, \eta, T_1, T_2, B, \hat{s} \).

1: Initialize: \( x_0 \).
2: for \( s = 1, 2, \ldots, T_1 \) do
3: \( \tilde{x} = x_{s-1} \);
4: \( w = \nabla \ell(\tilde{x}) \);
5: \( z_0 = \tilde{x} \).
6: for \( t = 1, 2, \ldots, T_2 \) do
7: \( \text{pick a set } I_t \subset \{1, \ldots, m\} \text{ of cardinality } B \)
8: \( \text{uniformly at random.} \)
9: \( v_t = \frac{1}{T} \sum_{i \in I_t} (\nabla \ell_i(z_{t-1}) - \nabla \ell_i(\tilde{x})) + \w \);
10: \( z_t = \text{denoise}_{\hat{s}}(z_{t-1} - \eta v_t) \).
11: \( x_s = z_{T_2} \).
12: end for
13: Output: \( \hat{x} = x_{T_1} \).

It is worth emphasizing that Alg. 3 is composed of \( T_1 \) epochs (outer loops) which each contains \( T_2 \) stochastic gradient and denoising updates. Within each epoch, we compute the full gradient of a reference point once which is used within the inner loop for stochastic gradient updates. Overall, we perform \( T_1 \) full gradient computations and \( 2T_1T_2 \) stochastic gradient computations and \( T_1T_2 \) denoising operations. It is critical to leverage a parameter configuration such that the benefit of the SVRG methods outweighs the computation cost of calculating a full gradient in each epoch, particularly when the denoising operation is expensive.

4. NUMERICAL EXPERIMENTS

In this section, we examine the performance of the proposed PnP-SVRG algorithm for compressive image reconstruction using partial 2D DFT measurements, a common setting used in compressive magnetic resonance imaging (MRI).

4.1. Measurement model and experimental settings

Consider the problem where our goal is to recover an image from bilinear measurements of an orthonormal matrix and its matrix transpose. This captures examples using the discrete Fourier matrix, discrete cosine matrix, discrete wavelet matrix, and many more. We will assume that images are square, but the process can be easily generalized. Within the context of this section, the image \( X^* \in \mathbb{R}^{n \times n} \) will be treated as a matrix, and the measurements \( Y \in \mathbb{R}^{p \times n} \) will be zero-padded to correspond with the image size. The matrix \( F \in \mathbb{C}^{n \times n} \) is the discrete Fourier transform (DFT) matrix, \( M \in \{0, 1\}^{n \times n} \) is a binary mask which determines the measurements that are sampled, and \( \varepsilon \) represents the noise in the system. We define the sampling rate \( p \) to be equal to the ratio of the number of observed measurements to the size of the image. We arrive at the following measurement model:

\[
Y = M \odot (FX^*F^T + \varepsilon). \tag{6}
\]

We formulate a natural squared loss function based on the Frobenius norm difference between the observed measurement matrix and the 2-D Fourier transform of the image, as

\[
\ell(X) = \frac{1}{2m} \| Y - M \odot FXF^T \|_F^2. \tag{7}
\]
Clearly, this loss function can be written as a finite sum over the observed measurements, and thus fits the algorithmic framework of PnP-SVRG.

We evaluate the reconstruction quality using the standard metric Peak Signal-to-Noise Ratio (PSNR) between the original image $X^*$ and the recovered image $\hat{X}$, which is calculated as

$$\text{PSNR} = 10 \log_{10} \left( \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (X_{i,j}^* - \hat{X}_{i,j})^2 \right),$$

assuming that each image is normalized such that values are within the range of $[0, 1]$.

Throughout the experiments, we take an image of size $256 \times 256$, and the noise in (6) is generated as additive white Gaussian noise with a standard deviation of $\sigma = 5$. For PnP-SVRG and PnP-SGD, we utilize a mini-batch size of $B = 2000$. We consider a few choices of image denoisers, including NLM, BM3D, TV-norm based wavelet denoising, and our own trained CNN. For NLM, we pick a patch size of 5 with a patch distance of 6 and a filter size of 0.015. For BM3D, we pick a noise estimate of $\hat{\sigma} = 0.015$. We implement a decay on the learning rate and the noise level estimate to improve stability. For CNN, we use a denoising CNN (DnCNN) [19], which follows a U-Net structure composed of a series of successive convolutions that first increase then decrease the input image's channels in order to separate the additive noise from the noisy image. The DnCNN denoiser we train is based off of the “realSN-DnCNN” introduced in [13], which uses a variation of layer-wise spectral norm to train a DnCNN. We train our denoiser on a truncated version of the Flickr30k dataset [20], using only 1000 images to train and 600 images to validate our results, where all RGB images are first converted to grayscale before training. We train the DnCNN on a variety of noise levels: $\sigma \in [0.1, 0.25, 0.5, 1, 2, 3, 4, 5, 10]$ and find that $\sigma = 0.25$ produces the best results when denoising. Each network is trained on a Titan RTX for 20 epochs, taking no more than 5 minutes to train.

4.2. Performance comparisons

In Fig. 1, we evaluate the performance of the proposed PnP-SVRG algorithm in comparison with PnP-GD and PnP-SGD for image reconstruction at a sampling rate $p = 0.5$, using the NLM denoiser and the CNN denoiser. It can be seen that PnP-SVRG achieves slightly higher PSNR than PnP-GD and PnP-SGD. Fig. 2 further examines the performance of image reconstruction via the proposed PnP-SVRG algorithm with the NLM denoiser in comparison to PnP-GD and PnP-SGD in terms of the runtime in (a), and sampling rate in (b). It can be seen that the proposed PnP-SVRG achieves a significantly improved runtime compared to its competitors, as well as slightly better reconstruction quality when the sampling rate is larger than 0.5. Similar trends are also observed for other denoisers. Lastly, Fig. 3 demonstrates the reconstruction quality of PnP-SVRG with various image denoisers for an RGB image when $p = 0.5$. Each RGB channel is processed separately and then combined for rendering. Here, the CNN, NLM, and BM3D based denoisers all achieve high visual quality. In conclusion, PnP-SVRG can be successfully implemented with a wide range of denoising routines.

5. CONCLUSIONS

In this paper, we proposed a new algorithm, dubbed PnP-SVRG, for image reconstruction from compressive measurements by carefully employing stochastic variance reduction techniques within the PnP framework. We demonstrated that substantially improved computational efficiency can be obtained without sacrificing image reconstruction quality. In the future, we plan to further investigate the performance of PnP-SVRG for other imaging tasks such as deblurring, superresolution, and inpainting. In addition, other stochastic variance reduction algorithms such as SARAH [21] and SAGA [22] will also be examined within the PnP framework. Last but not least, we are interested in evaluating the performance of the proposed PnP-SVRG algorithm for real data applications such as electron microscopy imaging [10].
6. REFERENCES


