

Deep Expectation-Consistent Approximation for Phase Retrieval

Saurav K. Shastri

Dept. ECE, The Ohio State University
Columbus, OH, 43210
shastri.19@osu.edu

Rizwan Ahmad

Dept. BME, The Ohio State University
Columbus, OH, 43210
ahmad.46@osu.edu

Philip Schniter

Dept. ECE, The Ohio State University
Columbus, OH, 43210
schniter.1@osu.edu

Abstract—The expectation consistent (EC) approximation framework is a state-of-the-art approach for solving (generalized) linear inverse problems with high-dimensional random forward operators and i.i.d. signal priors. In image inverse problems, however, both the forward operator and image pixels are structured, which plagues traditional EC implementations. In this work, we propose a novel incarnation of EC that exploits deep neural networks to handle structured operators and signals. For phase-retrieval, we propose a simplified variant called “deepECpr” that reduces to iterative denoising. In experiments recovering natural images from phaseless, shot-noise corrupted, coded-diffraction-pattern measurements, we observe accuracy surpassing the state-of-the-art prDeep (Metzler et al., 2018) and Diffusion Posterior Sampling (Chung et al., 2023) approaches with two-orders-of-magnitude complexity reduction.

Index Terms—Expectation-Consistent Approximation, Generalized Linear Model, Phase Retrieval, Message-Passing Algorithms, Plug-and-Play Algorithms

I. INTRODUCTION

In nonlinear inverse problems, one observes corrupted measurements $\mathbf{y} \in \mathcal{Y}^m$ of a signal/image $\mathbf{x} \in \mathbb{R}^d$ or $\mathbb{C}^d$ and they would like to recover $\mathbf{x}$ from $\mathbf{y}$. We consider problems for which the relationship between $\mathbf{y}$ and $\mathbf{x}$ can be described using a likelihood model of the form

$$p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^m p_{y_i|z_i}(y_i|z_i) \text{ for } \mathbf{z} \triangleq \mathbf{A}\mathbf{x}, \quad (1)$$

where the forward operator $\mathbf{A} \in \mathbb{C}^{m \times d}$ and scalar measurement channel $p_{y|z}$ are both known. In the statistics literature, (1) is known as the generalized linear model (GLM). Versions of $p_{y|z}$ exist for, e.g., additive noise of an arbitrary distribution, logistic regression [1], Poisson regression [2], noisy quantization [3], and phase retrieval [4], [5]. In this work, we focus on phase retrieval, although many of the ideas that we describe can be applied more generally.

For phase retrieval, although many forward operators $\mathbf{A}$ can be considered, the most common are the (possibly oversampled) 2D Fourier transform and the coded diffraction pattern (CDP) [6] shown in (2), where $\mathbf{F}$ is the unitary 2D Fourier transform and $\{\mathbf{D}_k\}_{k=1}^K$ are diagonal matrices with entries drawn independently and uniformly from the unit circle in the complex plane. Although several choices of $p_{y|z}$ have been

applied to phase retrieval, we will focus on the choice in (2) since it has been shown to work well in a variety of real-world settings [7].

$$\mathbf{A} = \frac{1}{\sqrt{K}} \begin{bmatrix} \mathbf{F}\mathbf{D}_1 \\ \vdots \\ \mathbf{F}\mathbf{D}_K \end{bmatrix}, \quad p_{y|z}(y|z) = \exp\left(-\frac{1}{2v}(y - |z|)^2\right). \quad (2)$$

Various computational approaches have been proposed for phase retrieval in imaging applications. Classical methods, like the Gerchberg-Saxton [8] and Hybrid Input-Output [9] algorithms are based on iterative projection. A more modern approach is to minimize the negative log-likelihood, i.e., $\arg \min_{\mathbf{x}} \{-\ln p_{\mathbf{y}|\mathbf{z}}(\mathbf{y}|\mathbf{A}\mathbf{x})\}$, using gradient-based iterative methods with a spectral initialization [10]–[12]. Although convex-relaxation-based methods like PhaseLift [13] have also been proposed (see the overview in [14]), they tend to be computationally impractical at typical image sizes. Approximate message-passing (AMP) algorithms have been proposed for phase retrieval [15], [16] that are near-optimal for high-dimensional i.i.d. or rotationally invariant random $\mathbf{A}$ [17], [18], but they tend to diverge with Fourier or CDP $\mathbf{A}$.

The aforementioned phase-retrieval methods do not exploit prior knowledge about $\mathbf{x}$. To exploit the knowledge that $\mathbf{x}$ is a natural image, several approaches have been proposed that involve deep neural networks. For example, by using the plug-and-play (PnP) [19] or RED [20] frameworks, one can iteratively alternate between negative-log-likelihood reduction and neural-network based image denoising [21], [22]. Or, inspired by CSGM [23], when given an image-generator $g_{\theta}(\mathbf{z})$, one could search for the code vector $\mathbf{z}$ such that the elementwise magnitude of $\mathbf{A}g_{\theta}(\mathbf{z})$ matches the phaseless measurements $\mathbf{y}$ [24]. A variation on this idea, inspired by DIP [25], is to optimize the generator parameters θ for a fixed code vector $\mathbf{z}$ [26], [27]. A more recent trend is to use diffusion methods, like DPS [28], for phase retrieval.

In this work, we propose a novel approach to phase retrieval that builds on the expectation consistent (EC) approximation algorithm from [29]. Although there are connections to AMP algorithms like [16], our approach does not require random $\mathbf{A}$. And although our approach performs iterative denoising (like with plug-and-play, RED, and diffusion methods) it converges two orders-of-magnitude faster.

The authors are funded in part by the National Science Foundation under grant CCF-1955587.

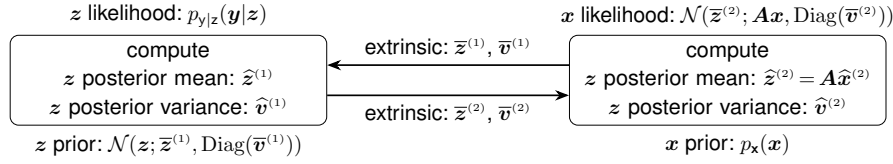


Fig. 1. Summary of EC applied to the GLM in (1). The algorithm iteratively estimates the transform outputs $z = \mathbf{A}x$ by alternating between two estimation modules that exchange extrinsic means and variances.

II. PROPOSED APPROACH

A. Review of EC

The expectation-consistent (EC) approximation algorithm [29] is a message-passing algorithm for iterative inference of vector-valued variables. Unlike the sum-product algorithm, it passes mean and variance messages rather than full probability distributions. More precisely, EC is a parallel version of the expectation propagation (EP) algorithm [30] that (locally) minimizes a known cost function. Over the last decade, EC has become famous as a method that admits rigorous analysis, and in some cases optimal recovery, in high-dimensional random settings. Perhaps the best known setting is the standard linear model (i.e., the GLM (1) with additive Gaussian $p_{y|z}$) with high-dimensional rotationally invariant $\mathbf{A}$ and i.i.d. x . This application of EC is sometimes referred to as Vector AMP (VAMP), which obeys a rigorous state-evolution whose fixed points are minimum mean-squared error (MMSE)-optimal whenever they are unique [31], [32]. EC has also been proposed for the GLM [16], [33], [34] and rigorously analyzed in [35] under the same random- $\mathbf{A}$ and i.i.d.- x assumptions. Although the i.i.d.- x assumption was circumvented in [36], random- $\mathbf{A}$ remains an important ingredient in existing applications/analyses of EC.

In image inverse problems, applying EC is challenging because $\mathbf{A}$ is not random. Consequently, EC behaves unpredictably and may not even converge [37]. Although workarounds have been proposed for the special case of magnetic resonance imaging [38], [39], successful applications of EC to general imaging inverse problems remain elusive.

Figure 1 illustrates how EC iteratively estimates the transform outputs $z = \mathbf{A}x$ by alternating between two estimation modules. The left module computes the posterior mean and (pixel-wise) variance of $\{z_i\}_{i=1}^m$ using the likelihoods $p_{y|z}(y_i|\cdot)$ and a Gaussian prior informed by the extrinsic message $\{(\bar{z}_i^{(1)}, \bar{v}_i^{(1)})\}_{i=1}^m$ received from the right module, i.e.,

$$\forall i: \begin{cases} \hat{z}_i^{(1)} = \mathbb{E}\{z_i|y_i; \bar{z}_i^{(1)}, \bar{v}_i^{(1)}\} \\ \hat{v}_i^{(1)} = \text{var}\{z_i|y_i; \bar{z}_i^{(1)}, \bar{v}_i^{(1)}\} \end{cases} \quad (3)$$

$$\text{via } p(z_i|y_i; \bar{z}_i^{(1)}, \bar{v}_i^{(1)}) = \frac{p_{y|z}(y_i|z_i)\mathcal{N}(z_i; \bar{z}_i^{(1)}, \bar{v}_i^{(1)})}{\int p_{y|z}(y_i|z_i)\mathcal{N}(z_i; \bar{z}_i^{(1)}, \bar{v}_i^{(1)})dz_i}.$$

Meanwhile, the right module computes the posterior mean and covariance of x using the prior p_x and a linear-Gaussian likelihood informed by the extrinsic message from the left

module, i.e.,

$$\begin{aligned} \hat{x}^{(2)} &= \mathbb{E}\{x|y; \bar{z}^{(2)}, \bar{v}^{(2)}\} \\ \hat{C}^{(2)} &= \text{Cov}\{x|y; \bar{z}^{(2)}, \bar{v}^{(2)}\} \end{aligned} \quad (4)$$

via $p(x|y; \bar{z}^{(2)}, \bar{v}^{(2)}) = \frac{\mathcal{N}(\bar{z}^{(2)}; \mathbf{A}x, \text{Diag}(\bar{v}^{(2)}))p_x(x)}{\int \mathcal{N}(\bar{z}^{(2)}; \mathbf{A}x, \text{Diag}(\bar{v}^{(2)}))p_x(x)dx}$, and then uses those quantities to compute the posterior mean and (pixel-wise) variance of z , i.e.,

$$\begin{aligned} \hat{z}^{(2)} &= \mathbf{A}\hat{x}^{(2)} \\ \hat{v}^{(2)} &= \text{diag}(\mathbf{A}\hat{C}^{(2)}\mathbf{A}^H). \end{aligned} \quad (5)$$

Finally, the messages passed between the two modules take the form

$$\forall i: \begin{cases} \bar{v}_i^{(2)} = (1/\hat{v}_i^{(1)} - \gamma^{(1)}/\bar{v}_i^{(1)})^{-1} \\ \bar{z}_i^{(2)} = (\hat{z}_i^{(1)}/\hat{v}_i^{(1)} - \gamma^{(1)}\bar{z}_i^{(1)}/\bar{v}_i^{(1)})\bar{v}_i^{(2)} \\ \bar{v}_i^{(1)} = (1/\hat{v}_i^{(2)} - \gamma^{(2)}/\bar{v}_i^{(2)})^{-1} \\ \bar{z}_i^{(1)} = (\hat{z}_i^{(2)}/\hat{v}_i^{(2)} - \gamma^{(2)}\bar{z}_i^{(2)}/\bar{v}_i^{(2)})\bar{v}_i^{(1)} \end{cases}, \quad (6)$$

which are “extrinsic” when $\gamma^{(1)} = 1 = \gamma^{(2)}$. When x is i.i.d. and $\mathbf{A}$ is a large, rotationally invariant random matrix, it is possible to avoid the high-dimensional integral and posterior covariance matrix $\hat{C}^{(2)}$ in (4), as detailed in [16], [33]–[35]. But what can be done when x and $\mathbf{A}$ are structured and non-random, as in most imaging applications?

B. The Proposed deepEC and deepECpr

For GLM image recovery, we propose to use EC as above, but with (4)–(5) approximated as follows:

- 1) Compute the posterior mean $\hat{x}^{(2)}$ using a neural network $f_\theta(\bar{z}^{(2)}; \bar{v}^{(2)})$ trained to minimize $J_f(\theta) = \sum_{t=1}^T \mathbb{E} \|x_t - \hat{x}_t^{(2)}\|^2$, with training $\{x_t\}_{t=1}^T$, output $\hat{x}_t^{(2)} = f_\theta(\mathbf{A}x_t + e_t; \bar{v}^{(2)})$, noise $e_t \sim \mathcal{N}(\mathbf{0}, \text{Diag}(\bar{v}^{(2)}))$, and variances $\bar{v}_i^{(2)} \sim \text{i.i.d. Unif}[0, v_{\max}]$ for some $v_{\max}$.
- 2) Compute the posterior variances $\hat{v}^{(2)}$ using a neural network $h_\phi(\bar{z}^{(2)}; \bar{v}^{(2)})$ trained to minimize $J_h(\phi) = \sum_{t=1}^T \mathbb{E} \|\bar{z}_t - \hat{z}_t^{(2)}\|^2 - \hat{v}_t^{(2)}\|^2$, with true $z_t = \mathbf{A}x_t$, estimated $\hat{z}_t^{(2)} = \mathbf{A}\hat{x}_t^{(2)}$, elementwise square $(\cdot)^{\odot 2}$, and $\hat{v}_t^{(2)} = h_\phi(\bar{z}_t^{(2)}; \bar{v}^{(2)})$. This avoids computing $\hat{C}^{(2)}$.

For phase retrieval, we make additional simplifications, some of which exploit $\mathbf{A}^H\mathbf{A} = \mathbf{I}$, which holds for both the (possibly oversampled) Fourier and CDP (2) incarnations of $\mathbf{A}$.

- 1) Use the Laplace approximation [40] of $(\hat{z}_i^{(1)}, \hat{v}_i^{(1)})$ in (3), computable in closed-form.
- 2) Approximate the posterior variance vectors $\hat{v}^{(j)}$ by 1 times their average value $\hat{v}^{(j)} \triangleq \frac{1}{m} \mathbf{1}^T \hat{v}^{(j)}$, for $j = 1, 2$. Consequently, $\bar{v}^{(j)}$ has the form $\mathbf{1}\bar{v}^{(j)}$ for scalar $\bar{v}^{(j)}$.

TABLE I
AVERAGE PSNR AND SSIM AT VARIOUS SHOT-NOISE LEVELS α FOR SET12 TEST IMAGES

method	$\alpha = 9$		$\alpha = 18$		$\alpha = 27$		denoiser calls
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
HIO	30.92	0.9557	25.79	0.8568	22.44	0.7413	-
prDeep	38.90	0.9900	34.60	0.9778	32.04	0.9633	800
deepECpr	39.27	0.9913	34.69	0.9785	32.02	0.9626	30

TABLE III
AVERAGE PSNR AND SSIM AT VARIOUS SHOT-NOISE LEVELS α FOR FFHQ TEST IMAGES

method	$\alpha = 9$		$\alpha = 18$		$\alpha = 27$		denoiser calls
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
HIO	20.84	0.7821	20.02	0.5519	18.89	0.3959	-
DPS	41.98	0.9803	38.4	0.9619	36.14	0.9432	1000
deepECpr	43.68	0.9865	39.92	0.9724	37.70	0.9595	30

TABLE II
IMPROVEMENT OF DEEPECPR OVER PRDEEP
VERSUS α

α	$\Delta\text{PSNR} \pm \text{SE}$	$\Delta\text{SSIM} \pm \text{SE}$
9	0.37 ± 0.06	0.0013 ± 0.0003
18	0.09 ± 0.05	0.0007 ± 0.0002
27	-0.02 ± 0.06	-0.0007 ± 0.0007

TABLE IV
IMPROVEMENT OF DEEPECPR OVER DPS
VERSUS α

α	$\Delta\text{PSNR} \pm \text{SE}$	$\Delta\text{SSIM} \pm \text{SE}$
9	1.70 ± 0.03	0.0063 ± 0.0007
18	1.52 ± 0.03	0.0105 ± 0.0009
27	1.57 ± 0.03	0.0163 ± 0.0012

- 3) Replace the linearized measurement $\bar{z}_t^{(2)} = \mathbf{A}\mathbf{x}_t + \mathbf{e}_t$, where $\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \bar{v}^{(2)}\mathbf{I})$, with the sufficient statistic $\bar{\mathbf{r}}_t^{(2)} \triangleq \mathbf{A}^H \bar{z}_t^{(2)} = \mathbf{x}_t + \boldsymbol{\epsilon}_t$, where again $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \bar{v}^{(2)}\mathbf{I})$. Consequently, estimating $\mathbf{x}_t$ from $\bar{\mathbf{r}}_t^{(2)}$ becomes a *denoising* task, for which we use $\hat{\mathbf{x}}^{(2)} = f_{\theta}(\bar{\mathbf{r}}_t^{(2)}, \bar{v}^{(2)}\mathbf{1})$.
- 4) Approximate $\hat{\mathbf{v}}^{(2)} = h_{\phi}(\bar{z}_t^{(2)}; \bar{v}^{(2)}\mathbf{1})$ by $\hat{\mathbf{v}}^{(2)} = \beta \bar{\mathbf{v}}^{(2)}$ for some scalar $\beta \in (0, 1]$. This is admittedly heuristic, but it works well empirically and avoids the need to use (and hence train) a variance-estimation network h_{ϕ} .

The resulting phase-retrieval algorithm, “deepECpr,” performs iterative denoising, similar to PnP or RED. However, it requires many fewer iterations, as we show next.

III. NUMERICAL EXPERIMENTS

To compare to prDeep [21], we repeat one of their experiments, where 128×128 grayscale images from Set12 [41] were recovered from phaseless CDP measurements $\mathbf{y}$ under the noise model

$$y_i^2 = |z_i|^2 + w_i \text{ with } w_i \sim \mathcal{N}(0, \alpha^2 |z_i|^2). \quad (7)$$

This approximates shot-noise corruption at noise level α since y_i^2/α^2 is approximately Poisson($|z_i|^2/\alpha^2$) [21].

For deepECpr, we use the $p_{y|z}$ from (2) with bank of five DnCNN [41] denoisers trained on BSD400 using the bias-free approach from [42]. These denoisers were trained to remove σ -std AWGN with σ drawn uniformly over $[0, 5]$, $[5, 10]$, $[10, 20]$, $[20, 40]$, and $[40, 60]$, respectively, where 255 is the maximum pixel intensity. At each iteration, deepECpr uses the denoiser whose σ interval includes the current value of $\sqrt{\bar{v}^{(2)}}$. For message passing, we used $\gamma^{(1)} = 1$ and $\gamma^{(2)} = 0$ in (6). We initialized with $\bar{z}^{(1)} = \mathbf{A}\mathbf{1}\bar{x}$, where $\bar{x}$ is the average pixel value over the training data, and with $\bar{v}^{(1)}$ set to the variance of $z_i - \bar{z}_i^{(1)}$ averaged over the training data and measurement indices i . We set $\beta = 0.5$.

For prDeep, we used the Python implementation [43] under the settings recommended in [21], which includes using $p_{y|z}$ from (2) and a bank of four DnCNN denoisers trained on the BSD400 over the σ intervals $[0, 10]$, $[10, 20]$, $[20, 40]$, and $[40, 60]$, respectively. First prDeep is run for 200 iterations using the denoiser trained for $\sigma \in [40, 60]$, then the result is improved

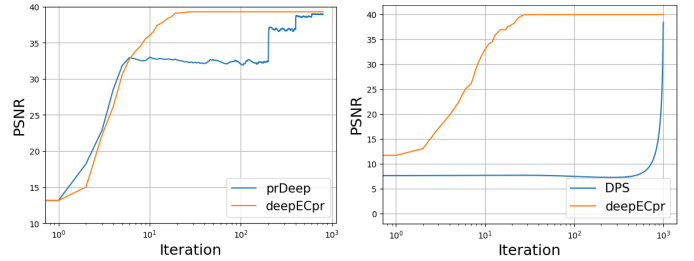


Fig. 2. Average PSNR versus iteration for Set12 test data with shot-noise level $\alpha = 9$ (left) and FFHQ test data with shot-noise level $\alpha = 18$ (right).

by running another 200 iterations using the denoiser trained for $\sigma \in [20, 40]$, and so on, for a total of 800 iterations. We did not further split the $[0, 10]$ interval, as we did with deepECpr, because we found that this degraded the performance of prDeep. Like with deepECpr, we trained DnCNN using the bias-free approach from [42] and initialized prDeep with $\hat{\mathbf{x}} = \mathbf{1}\bar{x}$, where $\bar{x}$ is the average pixel value over the training data. For both prDeep and deepECpr, v in (2) is set at the variance of $y_i - |z_i|$ averaged over the training data and measurement indices i .

We also compare to the classical HIO algorithm [9], as implemented in [44], using 1000 iterations.

Table I shows PSNR and SSIM averaged over the Set12 test images, where the proposed deepECpr outperformed prDeep and HIO at noise levels $\alpha = 9$ and 18. At $\alpha = 27$, deepECpr and prDeep exhibit comparable performance, with both vastly outperforming HIO. Table II¹ shows that the PSNR and SSIM gains at $\alpha = 9$ and 18 are statistically significant. Furthermore, Fig. 2 shows that deepECpr converges two orders-of-magnitude faster than prDeep. The example reconstructions in Fig. 3 show that deepECpr obtained superior visual quality by removing more noise than prDeep and HIO while still preserving fine details.

In a second experiment, we compare to the recent “DPS” conditional diffusion method from [28], where first 30 images from the FFHQ dataset [45] were recovered from phaseless CDP measurements at $K = 4$ with shot noise from (7).

¹In Table II, IV, and VI, the abbreviation ‘SE’ denotes ‘standard error.’

TABLE V
AVERAGE PSNR AND SSIM AT GAUSSIAN NOISE LEVEL $\sqrt{v}$ FOR FFHQ TEST IMAGES

method	$\sqrt{v} = 0.04$		$\sqrt{v} = 0.06$		$\sqrt{v} = 0.08$		denoiser calls
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
HIO	18.98	0.5154	18.05	0.3692	16.87	0.2697	-
DPS	37.95	0.9584	35.91	0.9392	34.47	0.9206	1000
deepECpr	39.28	0.9679	37.34	0.9537	35.82	0.9359	30

TABLE VI
IMPROVEMENT OF DEEPECPR OVER DPS
VERSUS $\sqrt{v}$

$\sqrt{v}$	Δ PSNR $\pm$ SE	Δ SSIM $\pm$ SE
0.04	1.33 ± 0.04	0.0095 ± 0.0009
0.06	1.43 ± 0.04	0.0145 ± 0.0010
0.08	1.35 ± 0.06	0.0154 ± 0.0014

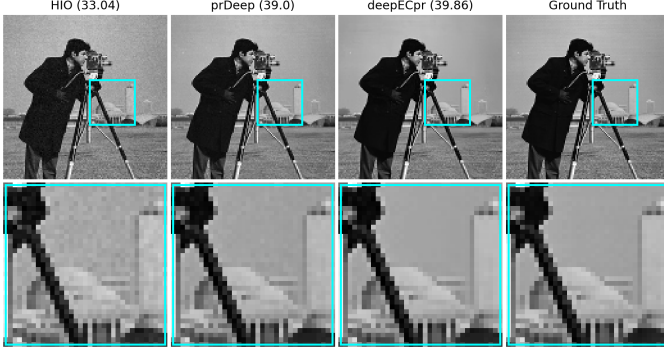


Fig. 3. Top: Reconstructions of a 128×128 Set12 image from phaseless CDP measurements at $K = 4$ in shot noise with $\alpha = 9$. PSNR is shown in parentheses. Bottom: Zoomed versions of the cyan squares in the top row.

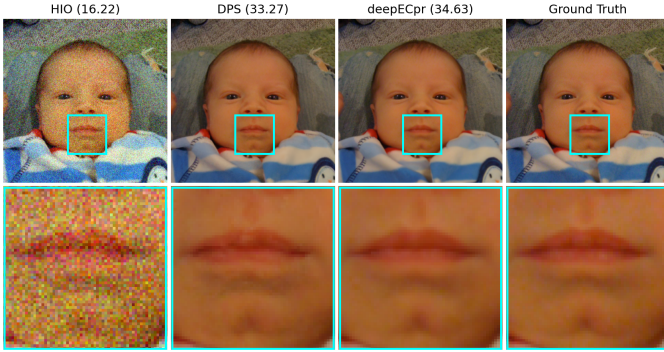


Fig. 4. Top: Reconstructions of a 256×256 FFHQ image from phaseless CDP measurements at $K = 4$ in shot noise with $\alpha = 27$. PSNR is shown in parentheses. Bottom: Zoomed versions of the cyan squares in the top row. Note how DPS generated artifacts in and around the lips.

DPS was run using the authors' code from [46], modified to accommodate the $p_{y|z}$ from (2). The codebase includes a pre-trained unconditional FFHQ diffusion model. In DPS, tuning the hyperparameter v in $p_{y|z}$ is equivalent to tuning the step-size on the gradient of the log likelihood, which we did by maximizing the PSNR on validation data that comprised FFHQ images 31-60. For deepECpr, the same pre-trained unconditional diffusion model is rescaled to act as a denoiser, as in [28, eq. (9)]. Since this denoiser accepts a noise variance, we feed it $\bar{v}^{(2)}$. As before, the average pixel value from the training data is used to initialize deepECpr, v in (2) is set at the variance of $y_i - |z_i|$ averaged over the training data and measurement indices i , deepECpr used $\beta = 0.5$, $\gamma^{(1)} = 1$, and $\gamma^{(2)} = 0$, and 1000-iteration HIO is used as a baseline.

Tables III and IV show PSNR and SSIM averaged over the

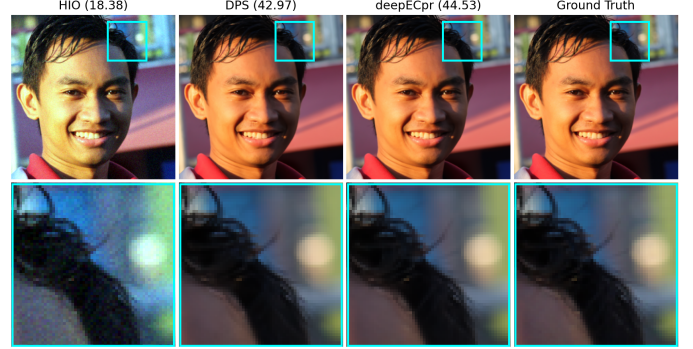


Fig. 5. Top: Reconstructions of a 256×256 FFHQ image from phaseless CDP measurements at $K = 4$ in shot noise with $\alpha = 9$. PSNR is shown in parentheses. Bottom: Zoomed versions of the cyan squares in the top row. Note how DPS failed to reconstruct the fine hair stands near the center of the zoomed plot.

FFHQ test images, where the proposed deepECpr outperformed DPS and HIO at all three noise levels α . Furthermore, Fig. 2 shows that deepECpr converged two orders-of-magnitude faster than DPS. Example reconstruction plots in Fig. 4 and Fig. 5 show that deepECpr obtained superior visual quality over HIO and DPS.

In a third experiment, we compare deepECpr to DPS and HIO using phaseless CDP measurements at $K = 4$ corrupted by additive white Gaussian noise:

$$y_i = |z_i| + w_i \text{ with } w_i \sim \mathcal{N}(0, v), \quad (8)$$

where v controls the variance of the additive white Gaussian noise. Table V demonstrates that deepECpr performed better than both HIO and DPS in this scenario, and Table VI suggests that the performance gap between deepECpr and DPS is statistically significant.

IV. CONCLUSION

For generalized linear models, we proposed a novel variant of expectation consistent (EC) approximation [29] that exploits deep neural networks. Unlike the traditional EC implementations, the proposed "deepEC" framework does not require random forward operators nor an i.i.d. signal prior. For phase retrieval, we proposed a simplified variant called "deepECpr." In experiments recovering natural images from phaseless, shot-noise corrupted, coded-diffraction-pattern outputs, we observed deepECpr outperforming the state-of-the-art prDeep [21] and the DPS [28] methods in reconstruction accuracy, while reducing complexity by two orders of magnitude.

REFERENCES

- [1] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd ed., 2009.
- [2] R. M. Willett, R. F. Marcia, and J. M. Nichols, "Compressed sensing for practical optical imaging systems: A tutorial," *Optical Eng.*, vol. 50, July 2011.
- [3] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Prentice-Hall, 1984.
- [4] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 87–109, 2015.
- [5] J. Dong, L. Valzania, A. Maillard, T.-A. Pham, S. Gigan, and M. Unser, "Phase retrieval: From computational imaging to machine learning: A tutorial," *IEEE Signal Process. Mag.*, vol. 40, no. 1, pp. 45–57, 2023.
- [6] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Appl. Comput. Harmonic Anal.*, vol. 39, no. 2, pp. 277–299, 2015.
- [7] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, "Experimental robustness of Fourier ptychography phase retrieval algorithms," *Optical Eng.*, vol. 23, no. 26, pp. 33214–33240, 2015.
- [8] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optika*, vol. 35, no. 2, pp. 237–246, 1972.
- [9] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Optics*, vol. 21, pp. 2758–2769, Aug. 1982.
- [10] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.
- [11] M. Mondelli and A. Montanari, "Fundamental limits of weak recovery with applications to phase retrieval," in *Proc. Conf. Learn. Theory*, pp. 1445–1450, 2018.
- [12] W. Luo, W. Alghamdi, and Y. M. Lu, "Optimal spectral initialization for signal recovery with applications to phase retrieval," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2347–2356, 2019.
- [13] E. J. Candès, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure & Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.
- [14] A. Fannjiang and T. Strohmer, "The numerics of phase retrieval," *Acta Numerica*, vol. 29, pp. 125–228, 2020.
- [15] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.*, vol. 63, pp. 1043–1055, Feb. 2015.
- [16] P. Schniter, S. Rangan, and A. K. Fletcher, "Vector approximate message passing for the generalized linear model," in *Proc. Asilomar Conf. Signals Syst. Comput.*, pp. 1525–1529, 2016.
- [17] A. Maillard, B. Loureiro, F. Krzakala, and L. Zdeborová, "Phase retrieval in high dimensions: Statistical and computational phase transitions," in *Proc. Neural Inf. Process. Syst. Conf.*, vol. 33, pp. 11071–11082, 2020.
- [18] M. Mondelli and R. Venkataramanan, "Approximate message passing with spectral initialization for generalized linear models," in *Proc. Int. Conf. Artificial Intell. Statist.*, pp. 397–405, 2021.
- [19] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Global Conf. Signal Inf. Process.*, pp. 945–948, 2013.
- [20] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (RED)," *SIAM J. Imag. Sci.*, vol. 10, no. 4, pp. 1804–1844, 2017.
- [21] C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, "prDeep: Robust phase retrieval with flexible deep neural networks," in *Proc. Int. Conf. Mach. Learn.*, pp. 3501–3510, 2018.
- [22] X. Chang, L. Bian, and J. Zhang, "Large-scale phase retrieval," *eLight*, vol. 1, no. 1, pp. 1–12, 2021.
- [23] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models," in *Proc. Int. Conf. Mach. Learn.*, pp. 537–546, 2017.
- [24] P. Hand, O. Leong, and V. Voroninski, "Phase retrieval under a generative prior," in *Proc. Neural Inf. Process. Syst. Conf.*, vol. 31, 2018.
- [25] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proc. IEEE Conf. Comp. Vision Pattern Recog.*, 2018.
- [26] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light: Science & Applications*, vol. 9, no. 1, p. 77, 2020.
- [27] E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, "Deep phase decoder: Self-calibrating phase microscopy with an untrained deep neural network," *Optica*, vol. 7, no. 6, pp. 559–562, 2020.
- [28] H. Chung, J. Kim, M. T. McCann, M. L. Klasky, and J. C. Ye, "Diffusion posterior sampling for general noisy inverse problems," in *Proc. Int. Conf. on Learn. Rep.*, 2023.
- [29] M. Opper and O. Winther, "Expectation consistent approximate inference," *J. Mach. Learn. Res.*, vol. 1, pp. 2177–2204, 2005.
- [30] T. Minka, *A Family of Approximate Algorithms for Bayesian Inference*. PhD thesis, Dept. Comp. Sci. Eng., MIT, Cambridge, MA, USA, 2001.
- [31] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Inf. Theory*, vol. 65, pp. 6664–6684, Oct. 2019.
- [32] K. Takeuchi, "Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements," *IEEE Trans. Inf. Theory*, vol. 66, pp. 368–386, Jan. 2020.
- [33] H. He, C.-K. Wen, and S. Jin, "Generalized expectation consistent signal recovery for nonlinear measurements," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 2333–2337, 2017.
- [34] X. Meng, S. Wu, and J. Zhu, "A unified Bayesian inference framework for generalized linear models," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 398–402, 2018.
- [35] P. Pandit, M. Sahraee-Ardakan, S. Rangan, P. Schniter, and A. K. Fletcher, "Inference with deep generative priors in high dimensions," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 1, pp. 336–347, 2020.
- [36] A. K. Fletcher, P. Pandit, S. Rangan, S. Sarkar, and P. Schniter, "Plug-in estimation in high-dimensional linear inverse problems: A rigorous analysis," in *Proc. Neural Inf. Process. Syst. Conf.*, pp. 7440–7449, 2018.
- [37] A. K. Fletcher, M. Sahraee-Ardakan, S. Rangan, and P. Schniter, "Expectation consistent approximate inference: Generalizations and convergence," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 190–194, 2016.
- [38] C. Millard, A. T. Hess, B. Mailhé, and J. Tanner, "Approximate message passing with a colored aliasing model for variable density Fourier sampled images," *IEEE Open J. Signal Process.*, vol. 1, pp. 146–158, 2020.
- [39] S. K. Shastri, R. Ahmad, C. A. Metzler, and P. Schniter, "Denoising generalized expectation-consistent approximation for mr image recovery," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 3, pp. 528–542, 2022.
- [40] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2007.
- [41] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, pp. 3142–3155, July 2017.
- [42] S. Mohan, Z. Kadkhodaie, E. P. Simoncelli, and C. Fernandez-Granda, "Robust and interpretable blind image denoising via bias-free convolutional neural networks," in *Proc. Int. Conf. on Learn. Rep.*, 2020.
- [43] D. Hekstra, I. Hunt-Isaak, J. Greisman, and J. Russell, "phase-retrieval." <https://github.com/Hekstra-Lab/phase-retrieval>, 2018.
- [44] C. A. Metzler, "prdeep." <https://github.com/ricedsp/prDeep/tree/master>, 2018.
- [45] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE Conf. Comp. Vision Pattern Recog.*, pp. 4396–4405, 2019.
- [46] H. Chung, J. Kim, M. T. McCann, M. L. Klasky, and J. C. Ye, "diffusion-posterior-sampling." <https://github.com/ricedsp/prDeep/tree/master>, Mar. 2023.