A08 Sparse exit wave reconstruction via deep unfolding

Benjamin Berkels

Review Meeting SFB 1481
Sparsity and Singular Structures
February 3, 2022

Variational exit wave reconstruction

Given TCCs $T_j : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ and corresponding TEM images $g_j : \mathbb{R}^d \to \mathbb{R}$ for $j = 1, \ldots, K$, exit wave reconstruction aims to find $\Psi : \mathbb{R}^d \to \mathbb{C}$ that minimizes

$$\mathcal{E}[\Psi] = \frac{1}{2K} \sum_{j=1}^{K} \left\| \Psi \star_{T_j} \Psi - \mathcal{F}(g_j) \right\|_{L^2}^2 + \mathcal{R}(\Psi).$$

Here, for $f, g : \mathbb{R}^d \to \mathbb{C}$ and $w : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$,

$$\big(f \star_w g\big)(x) := \int_{\mathbb{R}^d} \overline{f(y)} g(x+y) w(x+y, y) \,\mathrm{d}y$$

Simplified exit wave reconstruction - Phase retrieval

Simplifying assumption: Separability, i.e.

$$T_j(x,y) = \overline{w_j(x)}w_j(y) \text{ for } w_j : \mathbb{R}^d \to \mathbb{C}.$$

With this, we get $\Psi \star_{T_j} \Psi = (\Psi w_j) \star (\Psi w_j)$.

$$\Rightarrow \mathcal{F}^{-1}(\Psi \star_{T_j} \Psi) = (2\pi)^{\frac{d}{2}} \left|\mathcal{F}^{-1}(\Psi w_j)\right|^2.$$

Simplified objective functional

$$\mathcal{E}[\Psi] = \frac{1}{2K} \sum_{j=1}^{K} \left\| (2\pi)^{\frac{d}{2}} \left|\mathcal{F}^{-1}(\Psi w_j)\right|^2 - g_j \right\|_{L^2}^2 + \mathcal{R}(\Psi).$$

Discretization and minimization

A discretization with a Cartesian grid leads to

$$E(\psi) = \underbrace{\frac{1}{2KN}\sum_{j=1}^{K} \|\varphi(A_j\psi) - G_j\|_2^2}_{=:D(\psi)} + R(\psi),$$

with $\varphi : \mathbb{C} \to \mathbb{R}, z \mapsto |z|^2$ applied element-wise.

Minimization by forward-backward splitting

$$\psi^{k+1} = \operatorname{prox}_{\tau_k R}(\psi^k - \tau_k \nabla D(\psi^k)).$$

Gradient structure

$$\nabla D(\psi) = \frac{1}{KN}\sum_{j=1}^{K} \overline{A_j^T}\,(\varphi(A_j\psi) - G_j) \odot \varphi'(A_j\psi)$$

Many approaches covered by the splitting

Unregularized case ($R = 0$)

WF: $\varphi(z) = |z|^2$

RWF: $\varphi(z) = |z|$

TAF: $\varphi(z) = |z|$ + truncated ∇D (encoded in φ')

Regularized case ($R \neq 0$)

SPARTA: TAF + $R = I_{C_s}$, where C_s is the set of s-sparse vectors

IHTA $R = \lambda \|\cdot\|_0$ $\varphi(z) = z$, $K = 1$

ISTA $R = \lambda \|\cdot\|_1$ $\varphi(z) = z$, $K = 1$
+ $A_1 = A\Phi$ with a measurement matrix A and a dictionary matrix Φ $\longrightarrow$ [Behboodi, Rauhut, Schnoor '20]

Potentially learnable parameters

- (scalar) step sizes τ_k
- shared positive semi-definite matrix S
  (change step sizes to $\tau_k S$)
- measurement weights w_j
- input weights c_j
  (changing the data term to $\sum_{j=1}^{K} \|c_j \odot (\varphi(A_j \psi) - G_j)\|_2^2$)
- operator conjugate $\overline{A_j^T}$
  (combined with scaling s.t. $\|(\varphi(A_j\psi) - G_j) \odot \varphi'(A_j\psi)\|_2 \leq 1$)
- regularizes weight λ
- dictionary matrix Φ (replacing A_j with $A_j\Phi$, **WP 2**)
- focus values z_j (only **WP 3**)
- ...

Extension of results to non-separable TCCs

Many TCCs can be represented as limits of the form [Doberstein '20]

$$T_{j,L}(x,y) = \sum_{l=1}^{L} \overline{w_j^l(x)} w_j^l(y) \text{ for } L \to \infty.$$

1. Relax separability assumption $T_j(x,y) = \overline{w_j(x)} w_j(y)$ to

$$T_j(x,y) = \sum_{l=1}^{L} \overline{w_j^l(x)} w_j^l(y)$$

$\to$ Instead of $\varphi\left(\mathcal{F}^{-1}(\Psi w_j)\right)$, phase-less measurements are

$$\tilde{\varphi}\left(\sum_{l=1}^{L} \left|\mathcal{F}^{-1}(\Psi w_j^l)\right|^2\right).$$

$\to$ Nonlinearity structure changes from $\mathbb{C} \to \mathbb{C}$ to $\mathbb{C}^L \to \mathbb{C}$.

2. Investigate convergence properties for the case $L \to \infty$.

Nonlinearites in the unfolded network

Forward part of the splitting $(\psi^k - \tau_k \nabla D(\psi^k)) \to \varphi\varphi'$ and φ'

	$\varphi(z)$	$\varphi'(z)$	$\varphi(z)\varphi'(z)$
WF	$\lvert z\rvert^2$	$2z$ (linear)	$2\lvert z\rvert^2 z$
RWF	$\lvert z\rvert$	$\frac{z}{\lvert z\rvert}$	z (linear)
	Huber	nonlinear	nonlinear

Backward part of the splitting $\to \mathrm{prox}_{\tau_k R}$

R	$\mathrm{prox}_{\tau_k R}(y)$
$R = 0$	identity (linear)
$R = \lambda \lVert\cdot\rVert_1$	soft thresholding
$R = \lambda \lVert\cdot\rVert_0$	hard thresholding
$R = I_{C_s}$	top-s operator

Generalization erros bounds

- Hypothesis $h : \mathcal{X} \to \mathcal{Y}$, $h \in \mathcal{H}$
- Training samples $\mathcal{S} = (z_1, \ldots, z_m) \subset \mathcal{Z} := \mathcal{X} \times \mathcal{Y}$
- Loss function $l : \mathcal{H} \times \mathcal{Z} \to \mathbb{R}$
- Empirical loss $\hat{\mathcal{L}}(h) = \frac{1}{m} \sum_{j=1}^{m} l(h, z_j)$
- True loss $\mathcal{L}(h) = \mathbb{E}_{z \sim \mathcal{D}}(l(h, z))$
- Generalization error $|\hat{\mathcal{L}}(h_{\mathcal{S}}) - \mathcal{L}(h_{\mathcal{S}})|$

Generalization error for unfolded ISTA with dictionary (with high probability) [Behboodi, Rauhut, Schnoor '20]

$$|\hat{\mathcal{L}}(h) - \mathcal{L}(h)| \lesssim \sqrt{\frac{Nn \log(L) + N^2 \log(L)}{m}}$$

L number of layers, n linear measurements, N signal length
Trainable parameters: Dictionary Φ

References I

[1] S. Arridge, P. Maaß, O. Ötem, and C.-B. Schölieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, May 2019.

[2] A. Beck. *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, Oct. 2017.

[3] A. Behboodi, H. Rauhut, and E. Schnoor. Compressive sensing and neural networks from a statistical learning perspective, Oct. 2020.

[4] B. Berkels. An unconstrained multiphase thresholding approach for image segmentation. In *Proceedings of the Second International Conference on Scale Space Methods and Variational Methods in Computer Vision (SSVM 2009)*, volume 5567 of *Lecture Notes in Computer Science*, pages 26–37. Springer, 2009.

[5] B. Berkels, P. Binev, D. A. Blom, W. Dahmen, R. C. Sharpley, and T. Vogt. Optimized imaging using non-rigid registration. *Ultramicroscopy*, 138:46–56, Mar. 2014.

[6] B. Berkels, A. Effland, and M. Rumpf. Time discrete geodesic paths in the space of images. *SIAM Journal on Imaging Sciences*, 8(3):1457–1488, 2015.

[7] B. Berkels, A. Effland, and M. Rumpf. A posteriori error control for the binary Mumford-Shah model. *Mathematics of Computation*, 86(306):1769–1791, 2017.

[8] B. Berkels, M. Kotowski, M. Rumpf, and C. Schaller. Sulci detection in photos of the human cortex based on learned discriminative dictionaries. In *Third International Conference on Scale Space Methods and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, pages 326–337. Springer, 2011.

[9] B. Berkels and B. Wirth. Joint denoising and distortion correction of atomic scale scanning transmission electron microscopy images. *Inverse Problems*, 33(9):095002, 2017.

[10] T. Blumensath and M. E. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14(5-6):629–654, Sept. 2008.

References II

[11] E. J. Candes, X. Li, and M. Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, Apr. 2015.

[12] A. Chambolle and T. Pock. An introduction to continuous optimization for imaging. *Acta Numer.*, 25:161–319, 2016.

[13] Y. Chen and E. J. Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *Communications on Pure and Applied Mathematics*, 70(5):822–883, Apr. 2016.

[14] W. Coene, A. Thust, M. Op de Beeck, and D. Van Dyck. Maximum-likelihood method for focus-variation image reconstruction in high resolution transmission electron microscopy. *Ultramicroscopy*, 64:109–135, 1996.

[15] I. Daubechies, M. Defrise, and C. de Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, Aug. 2004.

[16] C. Doberstein. *Joint exit wave reconstruction and multimodal registration of transmission electron microscopy image series*. PhD thesis, RWTH Aachen University, 2020.

[17] C. Doberstein and B. Berkels. A least-squares functional for joint exit wave reconstruction and image registration. *Inverse Problems*, 35(5), 2019.

[18] T. Duden, A. Thust, C. Kumpf, and F. S. Tautz. Focal-series reconstruction in low-energy electron microscopy. *Microscopy and Microanalysis*, 20(3):968–973, Apr. 2014.

[19] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse Problems*, 23(3):947–968, Apr. 2007.

[20] K. Gregor and Y. LeCun. Learning fast approximations of sparse coding. In J. Fürnkranz and T. Joachims, editors, *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 399–406, Haifa, Israel, June 2010. Omnipress.

References III

[21] A. Griewank and A. Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Society for Industrial and Applied Mathematics, 2008.

[22] P. Grohs, S. Koppensteiner, and M. Rathmair. Phase retrieval: Uniqueness and stability. *SIAM Review*, 62(2):301–350, Jan. 2020.

[23] J. R. Hershey, J. L. Roux, and F. Weninger. Deep unfolding: Model-based inspiration of novel deep architectures, Sept. 2014.

[24] K. Jaganathan, Y. C. Eldar, and B. Hassibi. Phase retrieval: An overview of recent developments. In *Optical Compressive Imaging*, pages 263–296. Taylor & Francis, Oct. 2015.

[25] E. J. Kirkland. Improved high resolution image processing of bright field electron micrographs: I. Theory. *Ultramicroscopy*, 15(3):151–172, Jan. 1984.

[26] E. Kobler, A. Effland, K. Kunisch, and T. Pock. Total deep variation for linear inverse problems. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2020.

[27] A. Y. Kruger. On Fréchet subdifferentials. *Journal of Mathematical Sciences*, 116(3):3325–3358, 2003.

[28] J.-W. Liu, Z.-J. Cao, J. Liu, X.-L. Luo, W.-M. Li, N. Ito, and L.-C. Guo. Phase retrieval via Wirtinger flow algorithm and its variants. In *2019 International Conference on Machine Learning and Cybernetics (ICMLC)*. IEEE, July 2019.

[29] N. Mevenkamp and B. Berkels. Variational multi-phase segmentation using high-dimensional local features. In *IEEE Winter Conference on Applications of Computer Vision (WACV 2016)*. IEEE, 2016.

[30] V. Monga, Y. Li, and Y. C. Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2):18–44, Mar. 2021.

[31] N. Naimipour, S. Khobahi, and M. Soltanalian. Unfolded algorithms for deep phase retrieval, Dec. 2020.

References IV

[32] N. Naimipour, S. Khobahi, and M. Soltanalian. UPR: A model-driven architecture for deep phase retrieval. In *2020 54th Asilomar Conference on Signals, Systems, and Computers*. IEEE, Nov. 2020.

[33] U. Naumann. *The Art of Differentiating Computer Programs*. Society for Industrial and Applied Mathematics, 2011.

[34] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[35] T. Pinetz, E. Kobler, C. Doberstein, B. Berkels, and A. Effland. Total deep variation for noisy exit wave reconstruction in transmission electron microscopy. In *Eighth International Conference on Scale Space and Variational Methods in Computer Vision*, volume 12679 of *Lecture Notes in Computer Science*, pages 491–502. Springer, 2021.

[36] R. Remmert. *Theory of Complex Functions*. Springer New York, 1998.

[37] Y. Song, Z. Cao, K. Wu, Z. Yan, and C. Zhang. Learning fast approximations of sparse nonlinear regression, Oct. 2020.

[38] R. Tatano, B. Berkels, and T. M. Deserno. Mesh-to-raster based non-rigid registration of multi-modal images. *Journal of Medical Imaging*, 4(4):044002–1–044002–14, Oct. 2017.

[39] A. Thust, W. Coene, M. Op de Beeck, and D. Van Dyck. Focal-series reconstruction in HRTEM: simulation studies on non-periodic objects. *Ultramicroscopy*, 64(1-4):211–230, Aug. 1996.

[40] C.-J. Wang, C.-K. Wen, Shang-Ho, Tsai, S. Jin, and G. Y. Li. Phase retrieval using expectation consistent signal recovery algorithm based on hypernetwork, Jan. 2021.

References V

[41] C.-J. Wang, C.-K. Wen, S.-H. Tsai, and S. Jin. Phase retrieval with learning unfolded expectation consistent signal recovery algorithm. *IEEE Signal Processing Letters*, 27:780–784, 2020.

[42] G. Wang, G. B. Giannakis, J. Chen, and M. Akcakaya. SPARTA: Sparse phase retrieval via truncated amplitude flow. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Mar. 2017.

[43] G. Wang, G. B. Giannakis, and Y. C. Eldar. Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Transactions on Information Theory*, 64(2):773–794, Feb. 2018.

[44] Z. Yuan and H. Wang. Phase retrieval via reweighted Wirtinger flow. *Applied Optics*, 56(9):2418, Mar. 2017.

[45] H. Zhang, Y. Liang, and Y. Chi. A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms. *Journal of Machine Learning Research*, 18(141):1–35, 2017.