Stochastic Primal-Dual Hybrid Gradient Algorithm

UNIVERSITY OF CAMBRIDGE

M. J. Ehrhardt[1], P. Markiewicz[2], A. Chambolle[3], P. Richtárik[4], J. Schott[2] and C.-B. Schönlieb[1]

1 University of Cambridge, UK 2 University College London, UK 3 Ecole Polytechnique, France
4 KAUST, Saudi Arabia and University of Edinburgh, UK

contact: m.j.ehrhardt@damtp.cam.ac.uk

Motivation

- **Convex optimization** problems can often be efficiently solved as **saddle point problems** [1].
- In many applications (e.g. CT, parallel MRI, PET) saddle point problems are **separable** in the dual variable $y = (y_1, \dots, y_n)$:

$$(x^\sharp, y^\sharp) \in \arg\min_x \max_y \left\{ \sum_{i=1}^n \langle \mathbf{A}_i x, y_i \rangle - f_i^*(y_i) + g(x) \right\}$$

- Exploit structure by extending **primal-dual hybrid gradient** [1] to randomization over dual variables.

Stochastic PDHG Algorithm (SPDHG) [2]

- probability of the ith variable being selected: $p_i = \mathbb{P}(i \in \mathbb{S}^k) > 0$
- evaluation of $\mathbf{A}_i$ and $\mathbf{A}_i^*$ only for $j \in \mathbb{S}^k$

  - $x^{k+1} = \operatorname{prox}_{\tau g}(x^k + \tau \sum_{i=1}^n \mathbf{A}_i^* \overline{y}_i^k)$
  - Select $\mathbb{S}^k \subset \{1, \dots, n\}$
  - $y_i^{k+1} = \begin{cases} \operatorname{prox}_{\sigma_i f_i^*}(y_i^k - \sigma_i \mathbf{A}_i x^{k+1}) & \text{if } i \in \mathbb{S}^k \\ y_i^k & \text{else} \end{cases}$
  - $\overline{y}_i^{k+1} = y_i^{k+1} + \frac{\theta}{p_i}(y_i^{k+1} - y_i^k) \quad i = 1, \dots, n$

Efficient Implementation of SPDHG [3]

- example for **separable** samplings: $\mathbb{S}^k = \{i\}$
- auxiliary variable $z^k = \sum_{i=1}^n \mathbf{A}_i^* y_i^k$

  Initialize: $z^0 = \sum_{i=1}^n \mathbf{A}_i^* y_i^0, \quad \overline{z}^0 = z^0$
  - $x^{k+1} = \operatorname{prox}_{\tau g}(x^k + \tau \overline{z}^k)$
  - Select $j \in \{1, \dots, n\}$
  - $y_i^{k+1} = \begin{cases} \operatorname{prox}_{\sigma_i f_i^*}(y_i^k - \sigma_i \mathbf{A}_i x^{k+1}) & \text{if } i = j \\ y_i^k & \text{else} \end{cases}$
  - $\Delta z = \mathbf{A}_j^*(y_j^{k+1} - y_j^k)$
  - $z^{k+1} = z^k + \Delta z, \quad \overline{z}^{k+1} = z^{k+1} + \frac{\theta}{p_j} \Delta z$

Convergence of SPDHG [2]

Theorem. Let $(x^\sharp, y^\sharp)$ be a saddle point and the random sampling $\mathbb{S}^k$ be iid and separable. Choose $\theta = 1$ and σ_i, τ such that

$$\sigma_i \tau \|\mathbf{A}_i\|^2 < p_i \, .$$

Then the iterates of SPDHG **converge** to a saddle point in a Bregman sense almost surely:

$$D_g^{r^\sharp}(x^k, x^\sharp), \; D_{f^*}^{q^\sharp}(y^k, y^\sharp) \to 0 \quad \text{a.s.}$$

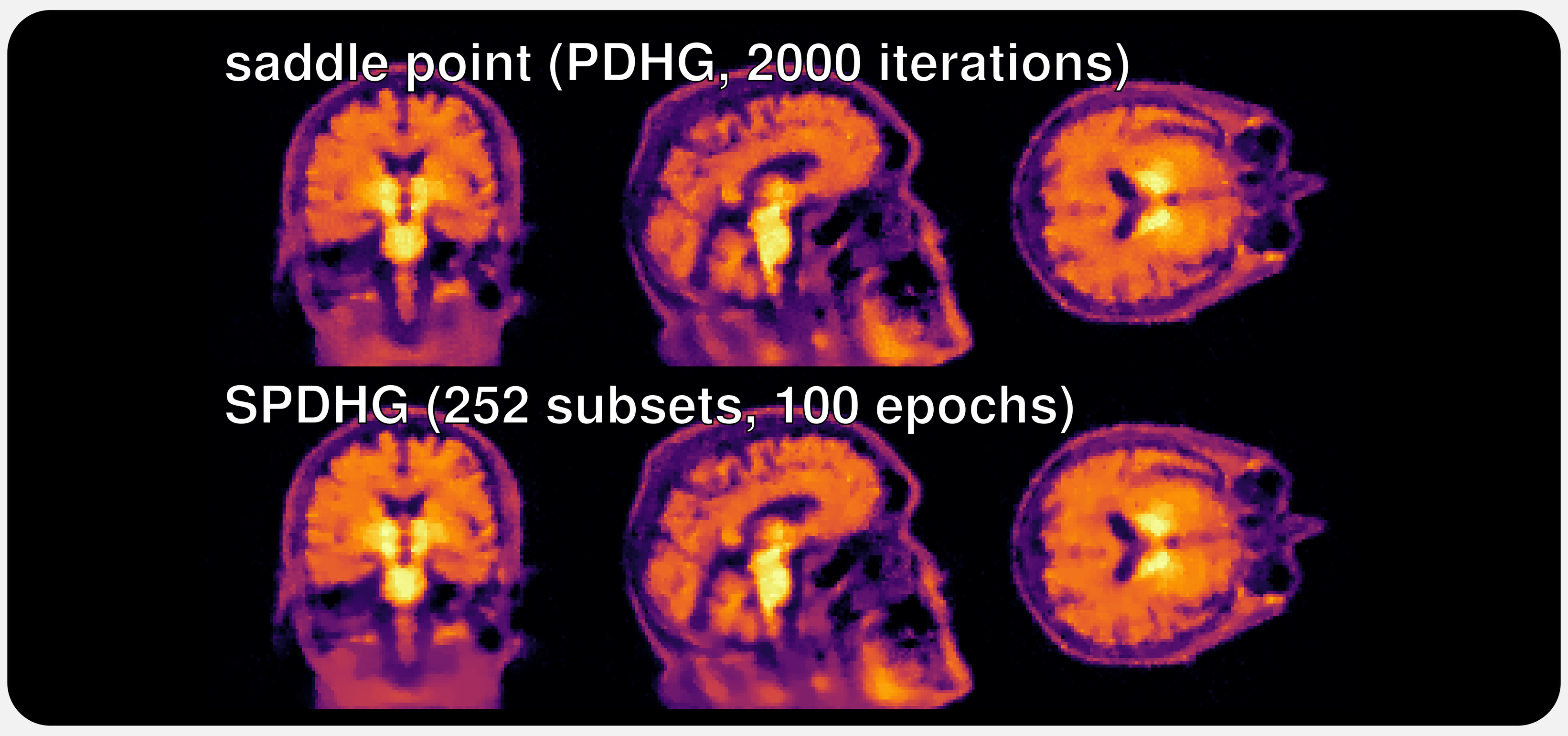
Moreover, the ergodic sequence $(x_K, y_K) = \frac{1}{K} \sum_{k=1}^K (x^k, y^k)$ converges to a saddle point with **rate** $\mathcal{O}(1/K)$:

$$\mathbb{E}\left\{ D_g^{r^\sharp}(x_K, x^\sharp) + D_{f^*}^{q^\sharp}(y_K, y^\sharp) \right\} \leq \frac{C}{K}$$

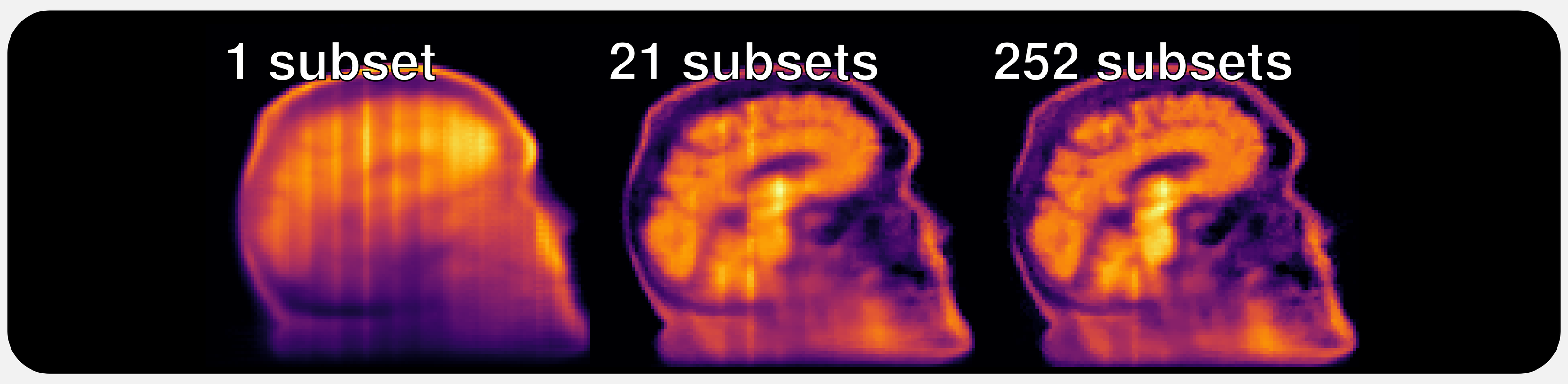
- see [2] for other samplings, acceleration and linear convergence

Faster PET Reconstruction [3]

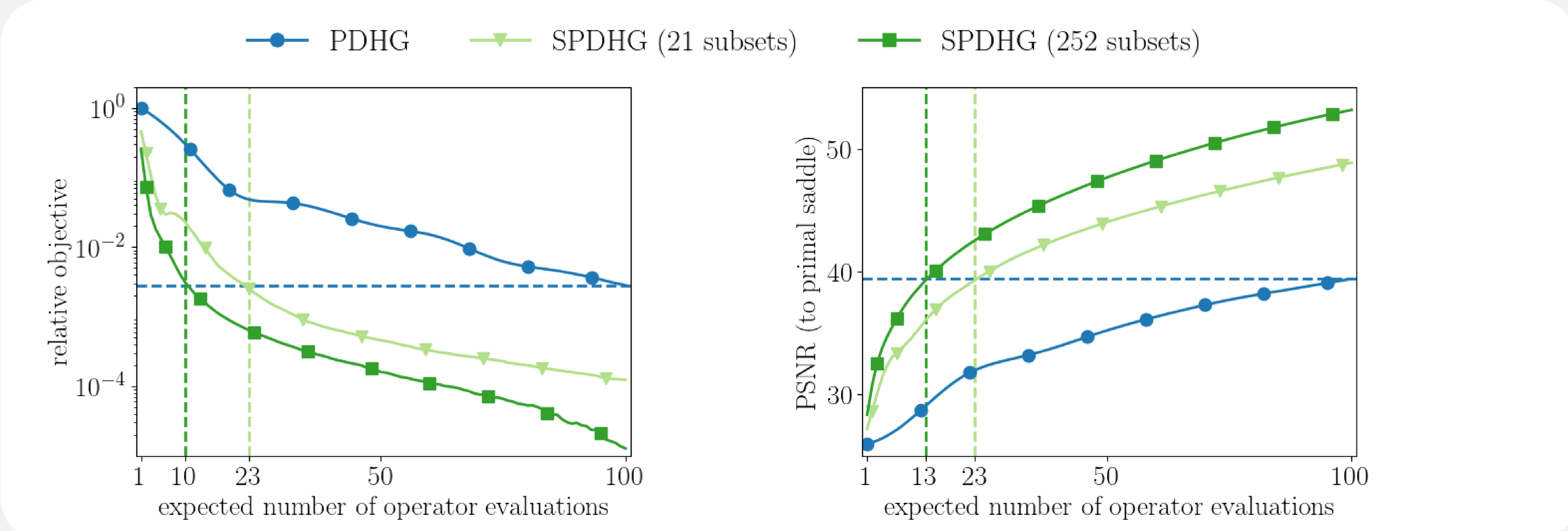
- SPDHG **converges** to solution of **deterministic** problem.



- SPDHG is **faster** with more subsets (example for 20 epochs).

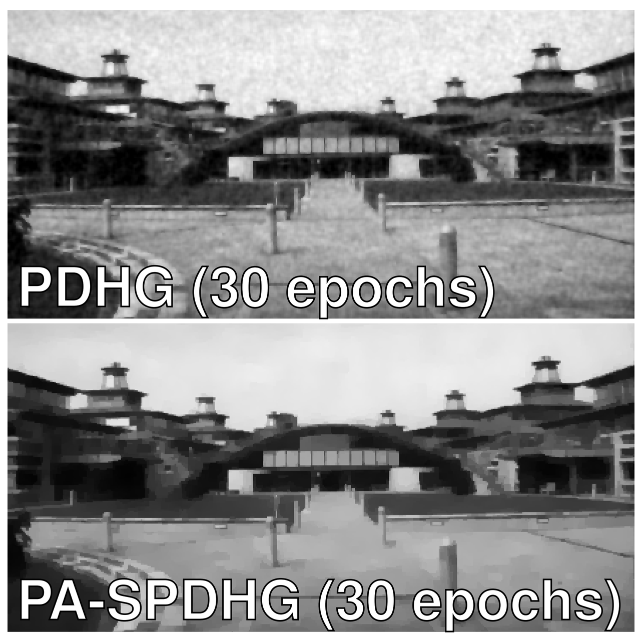


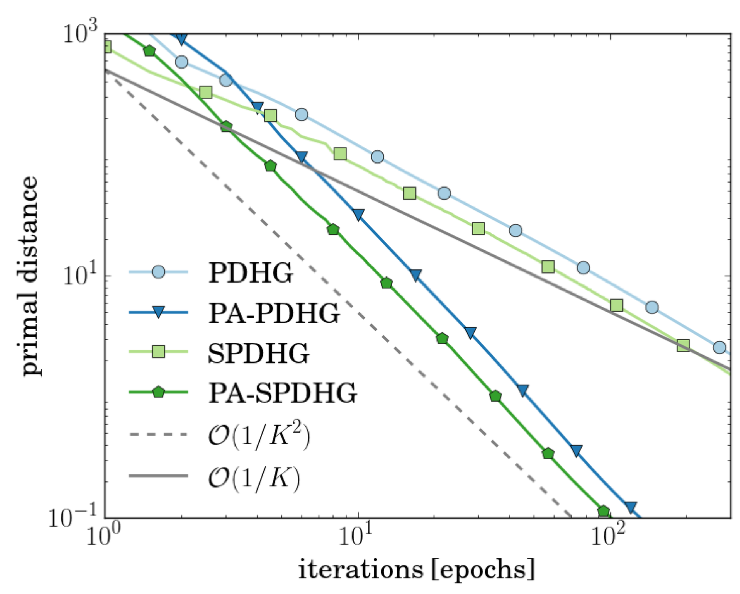
- SPDHG is **faster** in terms of objective decrease and PSNR.



Faster by Acceleration [2]

- **Accelerated** SPDHG converges as $\mathcal{O}(1/k^2)$ for ROF denoising.





Conclusions

- **Stochastic optimization** for dual separable cost functionals.
- **Generalization** of PDHG as it is recovered for $n = 1$, $p_1 = 1$.
- **Convergence rates** for SPDHG in expectation.
- **Faster PET reconstruction**: 10x speed-up on clinical data.
- SPDHG can be **accelerated** to achieve $\mathcal{O}(1/k^2)$ convergence.

[1] A. Chambolle and T. Pock. "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging". *J Math Imaging Vision* 40.1 (2011), pp. 120–145.
[2] A. Chambolle, M. J. Ehrhardt, P. Richtárik and C.-B. Schönlieb. "Stochastic Primal-Dual Hybrid Gradient Algorithm with Arbitrary Sampling and Imaging Applications". *arXiv* (2017).
[3] M. J. Ehrhardt, P. Markiewicz, A. Chambolle, P. Richtárik, J. Schott and C.-B. Schönlieb. "Faster PET Reconstruction with a Stochastic Primal-Dual Hybrid Gradient Method". *SPIE Proc* (2017).

CCIMI CANTAB CAPITAL INSTITUTE FOR THE MATHEMATICS OF INFORMATION
cmih centre for mathematical imaging in healthcare
EPSRC Engineering and Physical Sciences Research Council
The Leverhulme Trust