Fast and memory-efficient reconstruction of sparse Poisson data in listmode with non-smooth priors with application to time-of-flight PET

1Georg Schramm, 2Martin Holler

1Department of Imaging and Pathology, Division of Nuclear Medicine, KU/UZ Leuven, Belgium
2Institute of Mathematics and Scientific Computing, University of Graz, Austria
The authors have no financial interests to disclose.

The main results of this talk are published in

**Physics in Medicine & Biology**

**PAPER**

**Fast and memory-efficient reconstruction of sparse Poisson data in listmode with non-smooth priors with application to time-of-flight PET**

Georg Schramm\(^1\) and Martin Holler\(^{3,2}\)

Published 27 July 2022 • © 2022 Institute of Physics and Engineering in Medicine

*Physics in Medicine & Biology, Volume 67, Number 15*

**Citation** Georg Schramm and Martin Holler 2022 *Phys. Med. Biol.* 67 155020

**DOI** 10.1088/1361-6560/ac71f1
Iterative TOF PET reconstruction with non-smooth priors

\[ \arg\min_{x \geq 0} \sum_i \bar{y}_i(x) - d_i \log \bar{y}_i(x) + \beta R(\nabla x) \]

forward model
\[ \tilde{y} = Px + s \]

negative Poisson logL
non-smooth prior
e.g. TV, GTV, DTV

Challenges

1. **standard gradient-based methods cannot be applied** if \( R \) is non-smooth

2. **computation** of complete **TOF forward model** is **slow** (20s ... several minutes)

3. TOF data **sinograms are huge**
Stochastic Primal-Dual Hybrid Gradient (SPDHG)

1: Initialize $x(=0), y(=0), (S_i)_i, T, (p_i)_i,$
2: $\bar{z} = z = P^T y$
3: repeat
4: $x = \text{proj}_{\geq 0}(x - T\bar{z})$
5: Select $i \in \{1, \ldots, n+1\}$ randomly according to $(p_i)_i$
6: if $i \leq n$ then
7: $y_i^+ \leftarrow \text{prox}_{D_i}^S(y_i + S_i(p_i x + s_i))$
8: $\delta z \leftarrow P_i^T(y_i^+ - y_i)$
9: else
10: $y_i^+ \leftarrow \text{prox}_{R_i}^S(y_i + S_i \nabla x)$
11: $\delta z \leftarrow \nabla^T(y_i^+ - y_i)$
12: end if
13: $y_i \leftarrow y_i^+$
14: $z \leftarrow z + \delta z$
15: $\bar{z} \leftarrow z + (\delta z / p_i)$
16: until stopping criterion fulfilled
17: return $x$

M. J. Ehrhardt, P. Markiewicz, and C. B. Schönlieb. "Faster PET reconstruction with non-smooth priors by randomization and preconditioning", PMB 2019
**SPDHG - a closer look**

1. Initialize \( x(=0), y(=0), (S_i)_i, T, (p_i)_i, \)
2. \( \bar{z} = z = P^T y \)
3. repeat
4. \( x = \text{proj}_{\geq 0}(x - T\bar{z}) \)
5. Select \( i \in \{1, \ldots, n+1\} \) randomly according to \( (p_i)_i \)
6. if \( i \leq n \) then
7. \( y_i^+ \leftarrow \text{prox}_{D_i}^{S_i}(y_i + S_i(P_i x + s_i)) \)
8. \( \delta z \leftarrow P_i^T(y_i^+ - y_i) \)
9. else
10. \( y_i^+ \leftarrow \text{prox}_{R_i}^{S_i}(y_i + S_i \nabla x) \)
11. \( \delta z \leftarrow \nabla^T(y_i^+ - y_i) \)
12. end if
13. \( y_i \leftarrow y_i^+ \)
14. \( z \leftarrow z + \delta z \)
15. \( \bar{z} \leftarrow z + (\delta z/p_i) \)
16. until stopping criterion fulfilled
17. return \( x \)

**PROS**
- guaranteed convergence (“almost surely”)
- huge number of subsets possible
  - “reasonable” convergence after e.g. 10 it. / 252 ss.
- applicable to many convex priors (TV, DTV, GTV …) – also non-smooth

**CONS**
- only works in sinogram space (binned data)
- need to store 2nd complete (TOF) sinogram during iterations (y)
  - not efficient for sparse and huge TOF data

M. J. Ehrhardt, P. Markiewicz, and C. B. Schönlieb. “Faster PET reconstruction with non-smooth priors by randomization and preconditioning”, PMB 2019
Sparsity of TOF PET sinograms

- **modern TOF emission sinograms** are huge, but very sparse

- **sparsity ~ 1/(n. TOF bins) ~ 1/(TOF resolution)**
  → further increase of sparsity in future
  with better TOF resolution

- **reconstruction** in “sinogram/histogram mode” very inefficient
  → sparse sinogram or listmode processing

emission sinogram, 80s liver bed position
323 MBq [$^{18}$F]FDG, 70min p.i., $5 \times 10^7$ prompt counts

4 ring GE Discovery DMI (400ps TOF FWHM, 169ps TOF bin width)
sinogram dim. (425, 272, 1261, 29) → $10^9$ bins
Reducing the memory requirements of SPDHG
A better initialization → no need for empty data bins during iterations

1: Initialize $x(= 0), y(d \neq 0), (S_i)_i, T, (p_i)_i,$
2: $\bar{z} = z = P^T y$
3: repeat
4: $x = \text{proj}_{\geq 0}(x - T\bar{z})$
5: Select $i \in \{1, \ldots, n + 1\}$ randomly according to $(p_i)_i$
6: if $i \leq n$ then
7: $y_i^+ \leftarrow \text{prox}_{D_i^*}^{S_i}(y_i + S_i(P_i x + s_i))$
8: $\delta z \leftarrow P_i^T(y_i^+ - y_i)$
9: else
10: $y_i^+ \leftarrow \text{prox}_{R_i^*}^{S_i}(y_i + S_i \nabla x)$
11: $\delta z \leftarrow \nabla^T(y_i^+ - y_i)$
12: end if
13: $y_i \leftarrow y_i^+$
14: $z \leftarrow z + \delta z$
15: $\bar{z} \leftarrow z + (\delta z/p_i)$
16: until stopping criterion fulfilled
17: return $x$

Schramm & Holler “Fast and memory-efficient reconstruction of sparse TOF PET data with non-smooth priors”, Proceedings of the 16th Virtual International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine, https://arxiv.org/abs/2110.04143

$$(\text{prox}_{D_j^*}^{S_i}(y))_j = \frac{1}{2} \left( y_j + 1 - \sqrt{(y_j - 1)^2 + 4(S_i)_j d_j} \right)$$
“Listmode” SPDHG
accelerate TOF fwd/back projections
Schramm and Holler: “Fast and memory-efficient reconstruction of sparse Poisson data in listmode with non-smooth priors with application to time-of-flight PET”
Phys Med Biol 2022

```
1: Input event list N
2: Calculate event counts $\mu_e$ for each e in N
3: Split event list N into m sublists $N_i$
4: Initialize m sub lists $l_{N_i}$ with 0s
5: Initialize $x$, $(S_i)_i$, $T$, $(p_i)_i$, g
6: Preprocessing $\tilde{z} = z = P^T(d \neq 0)$
7: repeat
8:   $x = \text{proj}_{\geq 0}(x - T\tilde{z})$
9:   Select $i \in \{1, \ldots, m + 1\}$ randomly accord. to $(p_i)_i$
10:  if $i \leq m$ then
11:      $l_{N_i}^+ \leftarrow \text{prox}_{\delta D} \left( l_{N_i} + S_i \left( P_{N_i}^L x + s_{N_i} \right) \right)$
12:      $\delta z \leftarrow P_{N_i}^L \left( \frac{l_{N_i}^+ - l_{N_i}}{\mu_{N_i}} \right)$
13:  else
14:      $g^+ \leftarrow \text{prox}_{1/2\delta} (g + S_i \nabla x)$
15:      $\delta z \leftarrow \nabla^T (g^+ - g)$
16:      $g \leftarrow g^+$
17:  end if
18: end repeat
19: $z \leftarrow z + \delta z$
20: $\tilde{z} \leftarrow z + (\delta z/p_i)$
21: until stopping criterion fulfilled
22: return x
```
Schramm and Holler: “Fast and memory-efficient reconstruction of sparse Poisson data in listmode with non-smooth priors with application to time-of-flight PET” Phys Med Biol 2022

listmode fwd / back projections instead of sinogram projections

\[
(prox_{D_j}^{S_i} (y))_j = \frac{1}{2} \left( y_j + 1 - \sqrt{(y_j - 1)^2 + 4(S_i)_j \mu_j} \right)
\]
LM-SPDHG

| event num | det 1 | det 2 | TOF bin |
|-----------|-------|-------|---------|
| 1         | 1     | 3     | 1       |
| 2         | 1     | 4     | 1       |
| 3         | 2     | 4     | 2       |
| 4         | 1     | 3     | 2       |
| 5         | 1     | 3     | 1       |

**Time to calc $\mu_e$**

0.23s (1e7 counts) (single V100 GPU)

2.76s (1e8 counts)

Schramm and Holler: “Fast and memory-efficient reconstruction of sparse Poisson data in listmode with non-smooth priors with application to time-of-flight PET”

Phys Med Biol 2022

1. **Input** event list $N$

2. **Calculate** event counts $\mu_e$ for each $e$ in $N$

3. **Split** event list $N$ into $m$ sublists $N_i$

4. **Initialize** $m$ sub lists $l_{N_i}$ with $0$s

5. Initialize $x, (S_i)_i, T, (p_i)_i, g$

6. Preprocessing $\bar{z} = z = P^T (d \neq 0)$

7. **repeat**

8. $x = \text{proj}_{\geq 0}(x - T\bar{z})$

9. Select $i \in \{1, \ldots, m + 1\}$ randomly accord. to $(p_i)_i$

10. **if** $i \leq m$ **then**

11. $l^+_{N_i} \leftarrow \text{prox}_{D_{p_i}} \left( l_{N_i} + S_i \left( p_{LM_{N_i}} x + s_{N_i} \right) \right)$

12. $\delta z \leftarrow P_{LM}^T \left( \frac{l^+_{N_i} - l_{N_i}}{\mu_{N_i}} \right)$

13. $l_{N_i} \leftarrow l^+_{N_i}$

14. **else**

15. $g^+ \leftarrow \text{prox}_{||.||_1} (g + S_i \nabla x)$

16. $\delta z \leftarrow \nabla^T (g^+ - g)$

17. $g \leftarrow g^+$

18. **end if**

19. $z \leftarrow z + \delta z$

20. $z \leftarrow z + (\delta z / p_i)$

21. **until** stopping criterion fulfilled

22. **return** $x$
Listmode projectors are (almost always) faster

Timing (s) for TOF sinogram fwd+back projection
(1 out of 28 subsets, GE 4 ring DMI, 400ps TOF)

\[ 28 \times 0.2s = \textbf{5.6s} \text{ needed for complete sino fwd + back projection} \]

Timings (s) for TOF listmode fwd+back projection

\[ 4\times10^7 \text{ events (80s liver scan)} \text{ can be fwd + back projected in } \textbf{0.6s} \]
## Memory requirements of SPDHG vs LM-SPDHG

| Algorithm      | 5e8 Prompts | 7e7 Prompts | 1e7 Prompts |
|----------------|-------------|-------------|-------------|
| SPDHG          | 60.0 GB     | 60.0 GB     | 60.0 GB     |
| LM-SPDHG       | 12.5 GB     | 2.1 GB      | 0.8 GB      |

GE DMI-4 (20cm axial FOV) geometry – using “span 1” TOF sinograms
400ps TOF resolution, 29 TOF bins
Methods
Methods

- **reconstruction** of simulated 2D TOF PET data from brain phantom using
  - PDHG (10000 iterations) → reference solution \((x^*)\)
  - SPDHG (100 iterations / diff. num. subsets)
  - LM-SPDHG (100 iterations / diff. num. subsets)
  - LM-EMTV (100 iterations / diff. num. subsets)

- different count levels, prior strength and two priors: TV and DTV (directional TV)

- 2D/3D data simulation including attenuation, smooth contamination, finite resolution

- reconstruction of real 3D TOF data from GE DMI (NEMA IQ phantom)

Convergence monitored via

\[
\text{relative cost } c_{\text{rel}}(x) = \frac{c(x) - c(x^*)}{c(x^0) - c(x^*)} \\
\text{PSNR} = 20 \log \frac{|x^*|_\infty}{\text{MSE}(x, x^*)}
\]
Results
LM SPDHG converges as fast as sinogram SPDHG

(a) 3e5 true (5e5 prompt) counts, TV prior, $\beta = 0.03$
LM SPDHG vs EM-TV in 2D simulations

Sawatzky et al. “Accurate EM-TV algorithm in PET with low SNR”, IEEE NSS/MIC 2008

(a) $3e5$ true ($5e5$ prompt) counts, TV prior, $\beta = 0.03$
Speed of Convergence vs Number of Subsets in 2D simulations

(a) $3e5$ true ($5e5$ prompt) counts, TV prior, $\beta = 0.03$
Reconstructions of NEMA IQ phantom scan
Discussion and Conclusion
Discussion

• convergence speed **LM-SPDHG** very similar to (sinogram) **SPDHG**

• for “normal count” acquisitions @ 400ps systems:
  → **LM-SPDHG** much faster and **memory efficient** than **SPDHG**

• all **PDHG** versions are **non-monotonic**
  → stopping (very) early not recommended

• behaviour of all PDHG-variants in **early iterations very sensitive** to:
  - **initialization** of primal and dual variable
  - **step size ratio** (“S vs T”)
Impact of the step size ratio on (LM-S)PDHG

\[ \gamma = 0.03 / \text{max(img)} \]  
\[ \gamma = 3 / \text{max(img)} \]  
\[ \gamma = 300 / \text{max(img)} \]

20 it / 56 ss
Impact of the step size ratio on (LM-S)PDHG

\[ \gamma = 0.03 / \text{max(img)} \]

\[ \gamma = 3 / \text{max(img)} \]

\[ \gamma = 300 / \text{max(img)} \]

Adaptive Primal-Dual Splitting Methods for Statistical Learning and Image Processing

Thomas Goldstein*  
Department of Computer Science  
University of Maryland  
College Park, MD

Min Li†  
School of Economics and Management  
Southeast University  
Nanjing, China

Xiaoming Yuan‡  
Department of Mathematics  
Hong Kong Baptist University  
Kowloon Tong, Hong Kong

Journal of Mathematical Imaging and Vision (2024) 66:294–313  
https://doi.org/10.1007/s10851-024-01174-1

Stochastic Primal–Dual Hybrid Gradient Algorithm with Adaptive Step Sizes

Antonin Chambolle¹,²  
Claire Delplancke³  
Matthias J. Ehrhardt⁴  
Carola-Bibiane Schönlieb⁵  
Junqi Tang⁶
PDHG and LM-SPHG to optimize the Poisson logL and total variation

This example demonstrates the use of the primal dual hybrid gradient (PDHG) algorithm, the listmode stochastic PDHG (LM-SPDHG) to minimize the negative Poisson log-likelihood function combined with a total variation regularizer:

$$f(x) = \sum_{i=1}^{m} \tilde{d}_i(x) - d_i \log(\tilde{d}_i(x)) + \beta \|\nabla x\|_{1,2}$$

subject to

$$x \geq 0$$

using the linear forward model

$$\tilde{d}(x) = Ax + s$$
Y = 0.03 / max(img)

10000 it / 56 ss