An Interior Point Method for Nonnegative Sparse Signal Reconstruction

Xiang Huang[1,*], Kuan He[2], Seunghwan Yoo[2], Oliver Cossairt[2], Katsaggelos Aggelos[2], Nicola Ferrier[1], and Mark Hereld[1]

[1]Mathematics and Computer Science Division, Argonne National Lab., USA [2]Department of EECS, Northwestern University, USA

*Email: xianghuang@gmail.com *Web: http://xianghuang.net

This project was supported by U.S. Dept. of Energy, Contract# DE-AC02-06CH11357

Key Contributions

A fast primal-dual Interior Point Method (IPM) with

- *A novel pre-conditioner to compute Newton direction*
- *A careful starting point suitable for our IPM*
- *A simple yet effective prediction-correction scheme*

Problem Formulation

Goal: reconstruct a sparse signal $\mathbf{x} \in \mathbb{R}^N$ from its linear observation $\mathbf{b} = \mathbf{Ax} \in \mathbb{R}^M$, by solving the L_1-norm regularized optimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^N} \quad \frac{1}{2}\|\mathbf{Ax}-\mathbf{b}\|^2 + \tau\|\mathbf{x}\|_1, \qquad \text{(BPDN)}$$

We are interested in the case of $x \geq 0$,

$$\min_{\mathbf{x}\in\mathbb{R}^N} \quad \frac{1}{2}\|\mathbf{Ax}-\mathbf{b}\|^2 + \tau\mathbf{e}^T\mathbf{x} \qquad \text{(BPDN+)}$$
$$\text{s.t.} \quad \mathbf{x} \geq \mathbf{0}.$$

where $\mathbf{e}$ is an all one vector. This is because:

- pixel intensities from photon counts are ≥ 0
- algorithms for (BPDN+) can be used for (BPDN) by splitting $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$, where $\mathbf{x}^+, \mathbf{x}^- \geq \mathbf{0}$.

Solvers Overview

- 1st order gradient methods: Iterative Shrinkage / Thresholding (TwIST, FISTA), Gradient Projection (GPSR), Augmented Lagrange Multiplier (ADMM).
- 2nd order Hessian methods: IPM, Primal-dual IPM.

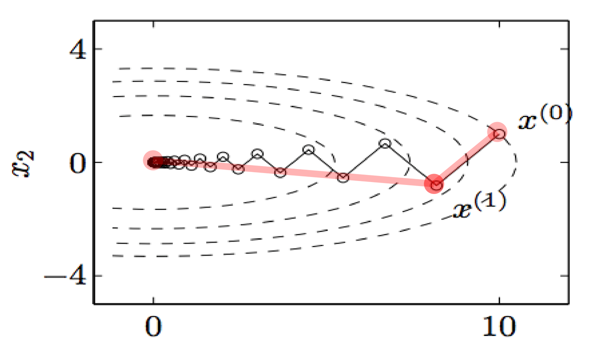


Fig. 1. Gradient based (cheap step but more iters) vs Hessian based methods (expensive step but less iters). Courtesy of S. Boyd etc.

Our Primal-dual IPM

We solve (BPDN+) from its modified KKT system:

$$\mathbf{A}^T\mathbf{Ax} - \mathbf{s} - \mathbf{A}^T\mathbf{b} + \tau\mathbf{e} = \mathbf{0}, \qquad (2a)$$
$$\mathbf{XSe} = \sigma\mu\mathbf{e}, \qquad (2b)$$
$$(\mathbf{x},\mathbf{s}) \geq \mathbf{0}, \qquad (2c)$$

where $\mathbf{s} \in \mathbb{R}^N$ is the dual variable, $\mathbf{X} = \text{diag}(\mathbf{x}), \mathbf{S} = diag(\mathbf{s})$, $\sigma \in [0,1]$ is a centering parameter, and $\mu = \frac{\mathbf{x}^T\mathbf{s}}{N} \to 0$ when converges.

We find the Newton direction $(\Delta\mathbf{x}, \Delta\mathbf{s})$ at $(\mathbf{x},\mathbf{s})$ from:

$$\mathbf{A}^T\mathbf{A}\Delta\mathbf{x} - \Delta\mathbf{s} = \mathbf{r}_d, \qquad (3a)$$
$$\mathbf{S}\Delta\mathbf{x} + \mathbf{X}\Delta\mathbf{s} = \mathbf{r}_c, \qquad (3b)$$

where $\mathbf{r}_d$ and $\mathbf{r}_c$ are gradient and slackness residuals:

$$\mathbf{r}_d := \mathbf{s} - \nabla h(\mathbf{x}), \qquad (4a)$$
$$\mathbf{r}_c := \sigma\mu\mathbf{e} - \mathbf{XSe}. \qquad (4b)$$

$h(\mathbf{x})$ is the objective function of (BPDN+), its gradient

$$\nabla h(\mathbf{x}) = \mathbf{A}^T\mathbf{Ax} - \mathbf{A}^T\mathbf{b} + \tau\mathbf{e}$$

We iteratively update $(\mathbf{x} \leftarrow \mathbf{x} + \alpha_p\Delta\mathbf{x}, \mathbf{s} \leftarrow \mathbf{s} + \alpha_d\Delta\mathbf{s})$ use Algorithm 1.

Algorithm 1 Primal Dual Preconditioned IPM Framework

Inputs: choose $(\mathbf{x}^0, \mathbf{s}^0) > \mathbf{0}$ from Section 2.1, stop accuracy ϵ (e.g. $1e-6$), and maximum iteration number $k_{\max}$.
for $k = 1, 2, \ldots, k_{\max}$ **do**
  Perform Prediction Step: set $\sigma \leftarrow 0.01$.
  $(\mathbf{x}^k, \mathbf{s}^k, \alpha_p, \alpha_d)$ = UPDATE$(\mathbf{x}^{k-1}, \mathbf{s}^{k-1}, \sigma)$
  if $\min(\alpha_p, \alpha_d) \leq 0.1$ **then**
    Perform Correction Step: set $\sigma \leftarrow 0.99$.
    $(\mathbf{x}^k, \mathbf{s}^k, \alpha_p, \alpha_d)$ = UPDATE$(\mathbf{x}^{k-1}, \mathbf{s}^{k-1}, \sigma)$
  if $\mu_k \leq \epsilon h(\mathbf{x}^k)$ and $\|\mathbf{r}_d^k\| \leq \epsilon$ **then**
    Break
Output: $\mathbf{x}^k$
function UPDATE$(\mathbf{x}^{k-1}, \mathbf{s}^{k-1}, \sigma)$
  Compute $\Delta\mathbf{x}, \Delta\mathbf{s}$ with $\sigma, \mathbf{x}^{k-1}, \mathbf{s}^{k-1}$ use Section 2.2
  Compute α_p, α_d with $\mathbf{x}^{k-1}, \mathbf{s}^{k-1}, \Delta\mathbf{x}, \Delta\mathbf{s}$ use Section 2.3
  Update $(\mathbf{x}^k, \mathbf{s}^k) \leftarrow (\mathbf{x}^{k-1} + \alpha_p\Delta\mathbf{x}, \mathbf{s}^{k-1} + \alpha_d\Delta\mathbf{s})$.
  return $(\mathbf{x}^k, \mathbf{s}^k, \alpha_p, \alpha_d)$

Our Preconditioner

Most computation is spent on solving Newton Eq.(3). We first eliminate $\Delta\mathbf{s}$ and solve $\Delta\mathbf{x}$:

$$(\mathbf{D}^{-1} + \mathbf{A}^T\mathbf{A})\Delta\mathbf{x} = \sigma\mu\mathbf{X}^{-1}\mathbf{e} - \nabla h(\mathbf{x})$$

where $\mathbf{D} = \mathbf{XS}^{-1}$. We solve it by conjugate gradient method with a diagonal plus rank 1 preconditioner $\mathbf{M}$:

$$\mathbf{M} = \mathbf{D}^{-1} + \mathbf{vv}^\mathbf{T},$$

where $\mathbf{v}$ is the scaled eigenvector of $\mathbf{A}^\mathbf{T}\mathbf{A}$. This is good for deconvolution case where $\mathbf{A}$ is a circulant matrix. Note $\mathbf{M}^{-1}$ is easy by Sherman-Morrison formula:

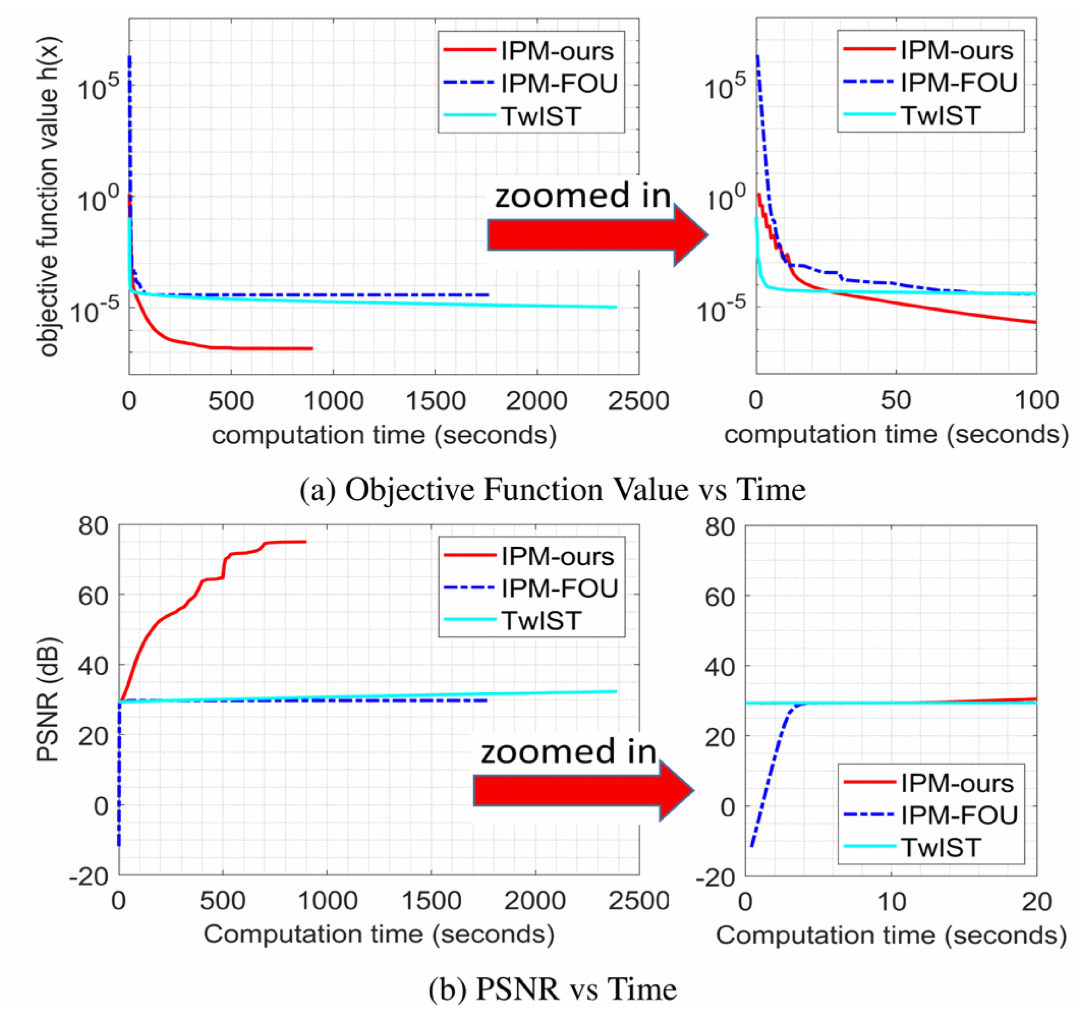
$$\mathbf{M}^{-1} = \mathbf{D} - \frac{\mathbf{Dvv}^\mathbf{T}\mathbf{D}}{1 + \mathbf{v}^\mathbf{T}\mathbf{Dv}}.$$

Experiments

Reconstruct 3D volumetric image of $256 \times 256 \times 16$ from its 2D observation of 256×256: $(N = 106, M = 6.5 \times 104)$

$$b(u,v) = \sum_{z=1}^{N_z} psf(u,v;z) * x(u,v;z)$$

Quantitative Comparison

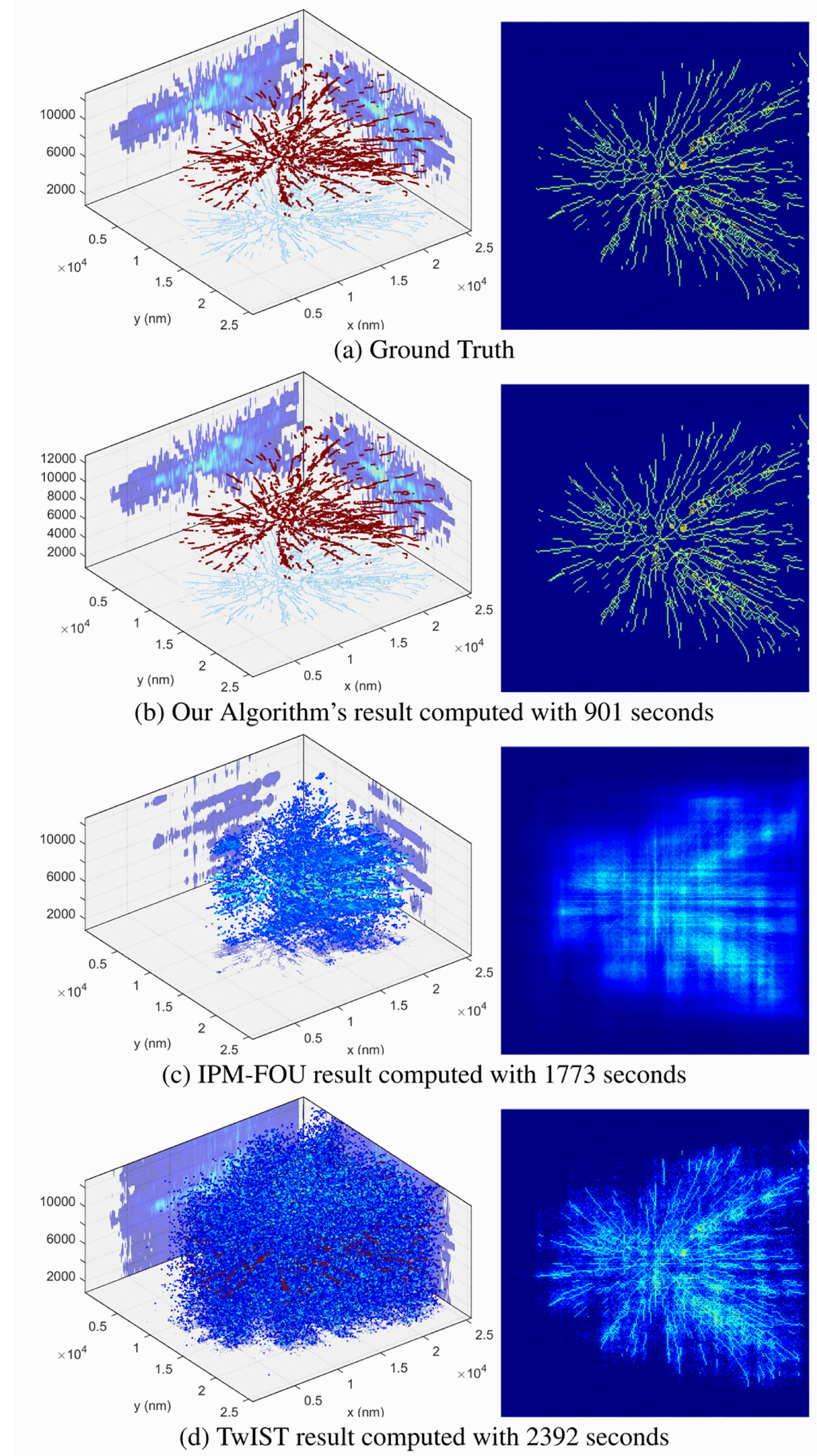


(a) Objective Function Value vs Time

(b) PSNR vs Time

Fig. 2: (a) Comparison of the objective function values in Eq. (BPDN+) over time. (b) Comparison of the PSNRs over time.

Qualitative Comparison



(a) Ground Truth

(b) Our Algorithm's result computed with 901 seconds

(c) IPM-FOU result computed with 1773 seconds

(d) TwIST result computed with 2392 seconds

Fig. 3: Comparison of the results for the three algorithms to the ground truth of 3D microtubule. Left column: 3D volumetric image results, showing as both point clouds and 2D projections to xy,yz,xz planes. Right column: the projection onto the xy plane. Intensities from low to high are mapped from blue to red.

[IPM-Fou] K. Fountoulakis, J. Gondzio, and P. Zhlobich, “Matrixfree interior point method for compressed sensing problems,” Mathematical Programming Computation, vol. 6, no. 1, pp. 1–31, 2014

[[TwIST]] J. Bioucas-Dias and M. Figueiredo, “A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration,” IEEE Trans. on Image Processing, vol. 16, no. 12, pp. 2992–3004, Dec 2007.