Inexact Alternating Direction Method of Multipliers for Separable Convex Optimization

Hongchao Zhang
hozhang@math.lsu.edu

Department of Mathematics
Center for Computation and Technology
Louisiana State University

U.S.-Mexico Workshop on Optimization and Applications
January 8th, 2016
Constrained Separable Convex Optimization

$$\min \sum_{i=1}^{m} f_i(x_i) + h_i(x_i) \quad \text{s. t.} \quad \sum_{i=1}^{m} A_i x_i = b$$

where $m \geq 2$ and

- $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ is convex, Lipschitz continuously differentiable.
- $h_i$ is a simple proper closed convex function on $\mathbb{R}^{n_i}$, but not necessarily smooth.
- To put constraint $x_i \in \mathcal{X}_i$, let $h_i$ be the indicator function for closed convex set $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$.
- Many applications in image processing, statistical learning and compressive sensing, etc.
Total variation image reconstruction

$$\min H(u) + \phi(Bu)$$

which is equivalent to

$$\min H(u) + \phi(w) \quad \text{s. t.} \quad Bu = w,$$

where

- \( H(u) = \frac{1}{2} \| Au - f \|^2 \) with \( A \) large, dense and ill-conditioned
- \( \phi(Bu) = \alpha \| u \|_{TV} = \alpha \sum_{i=1}^{N} \| (\nabla u)_i \| \) with \( \alpha > 0 \)
- Augmented Lagrangian

\[
\mathcal{L}^\rho(u, w, b) = H(u) + \phi(w) + \langle b, Bu - w \rangle + \frac{\rho}{2} \| Bu - w \|^2
\]
Motivation $m = 2$

- Augmented Lagrangian Method (Method of Multipliers)

$$ (u^{k+1}, w^{k+1}) = \arg\min_{u, w} \mathcal{L}^\rho(u, w, b^k), $$

$$ b^{k+1} = b^k + \rho(Bu^{k+1} - w^{k+1}). $$

- Alternating direction method of multipliers (ADMM)

$$ u^{k+1} = \arg\min_u \frac{1}{2} \|Au - f\|^2 + \frac{\rho}{2} \|Bu - w^k + \rho^{-1}b^k\|^2, $$

$$ w^{k+1} = \arg\min_w \phi(w) + \frac{\rho}{2} \|Bu^{k+1} - w + \rho^{-1}b^k\|^2, $$

$$ b^{k+1} = b^k + \rho(Bu^{k+1} - w^{k+1}). $$
Motivation $m = 2$

- The “$u$” subproblem:

$$u^{k+1} = \arg\min_{u} \frac{1}{2} \|Au - f\|^2 + \frac{\rho}{2} \|Bu - w^k + \rho^{-1}b^k\|^2$$

- Bregman operator splitting (BOS) scheme: $\delta_k > \|A^T A\|$

$$u^{k+1} = \arg\min_{u} \delta_k \|u - u^k + \delta_k^{-1}A^T(Au^k - f)\|^2 + \rho \|Bu - w^k + \rho^{-1}b^k\|^2$$

- Bregman operator splitting with variable stepsize (BOSVS): Apply nonmonotone line search with initial

$$\hat{\delta}_k = \max \left\{ \Delta_k, \frac{\|A(u^k - u^{k-1})\|^2}{\|u^k - u^{k-1}\|^2} \right\},$$

where $\Delta_k > 0$ is a lower bound and is adaptively increased.
BOSVS Algorithm

**Parameters:** $\tau, \eta > 1$, $\beta$, $C \geq 0$, $\rho$, $\delta_{\text{min}} > 0$, $\xi, \sigma \in (0, 1)$, $\delta_0 = 1$.

**Initialization:** $w^1$, $u^1$ and $b^1$. Set $Q_1 = 0$ and $\Delta_1 = \delta_{\text{min}}$.

**For** $k = 1, 2, \ldots$

**Step 1.** Set $\delta_k = \eta^j \hat{\delta}_k$ where $j \geq 0$ is the smallest integer such that $Q_{k+1} \geq -\frac{C}{k^2}$ where $Q_{k+1} := \xi_k Q_k + \Omega_k$ with

$$\Omega_k := \sigma(\delta_k \|u^{k+1} - u^k\|^2 + \rho\|Bu^{k+1} - w^k\|^2) - \|A(u^{k+1} - u^k)\|^2,$$

and

$$u^{k+1} = \arg\min_u \left\{ \delta_k \|u - u^k + \delta_k^{-1} A^T(Au^k - f)\|^2 + \rho\|w^k - Bu + b^k/\rho\|^2 \right\},$$

$0 \leq \xi_k \leq \min\{\xi, (1 - k^{-1})^2\}$.

**Step 2.** If $\delta_k > \max\{\delta_{k-1}, \Delta_k\}$ when $k > 1$, $\Delta_{k+1} := \tau \Delta_k$.

**Step 3.** $w^{k+1} = \arg\min_w \left\{ \phi(w) + \frac{\rho}{2}\|w - Bu^{k+1} + b^k/\rho\|^2 + \frac{\beta}{2}\|w - w^k\|^2 \right\}$.

**Step 4.** $b^{k+1} = b^k + \rho(Bu^{k+1} - w^{k+1})$.

**Step 5.** If a stopping criterion is satisfied, stop.

**End For**
Theorem. If the minimizer exists, the sequence $x^k = (u^k, w^k, b^k)$ generated by BOSVS converges to a solution $x^* = (u^*, w^*, b^*)$ satisfying the first-order optimality conditions

$$\nabla H(u^*) - B^T b^* = 0, \quad -b^* \in \partial \phi(w^*), \quad w^* = Bu^*.$$

Moreover, the ergodic mean $u_K$ given by $u_K := \frac{1}{K} \sum_{k=1}^{K} u^k$ satisfies

$$\phi(Bu_K) + H(u_K) - \min_u \{\phi(Bu) + H(u)\} = O\left(\frac{1}{K}\right).$$
Numerical Experiments (BOSVS)

Partially Parallel Imaging (PPI) with $L = 8$ coils

- Optimize
  
  $$\min_u \frac{1}{2} \sum_{l=1}^{L} \|F_p(s_l \odot u) - f_l\|^2 + \alpha \|u\|_{TV},$$

  where $F_p$ is the undersampled Fourier transform, $S_l$ is the sensitivity map of the $l$-th channel, and $\alpha = 10^{-5}$.

- Three data sets:
  - data 1 and data 2 uses a random Poisson mask (25% Fourier coefficients);
  - data 3 uses a radial mask (34% Fourier coefficients).

- Plot the error in the objective function versus the iteration number ($\rho = 10^{-4}$).
Figure: Plots of the objective error for data 1, data 2, and data 3.

Least squares fit of $y = cK^{-p}$ give $p$ as 1.6, 1.2, and 0.9 for data 1, data 2, and data 3 respectively.
Optimize

$$\min \sum_{i=1}^{m} f_i(x_i) + h_i(x_i) \quad \text{s. t.} \quad \sum_{i=1}^{m} A_i x_i = b$$

where

- $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ is convex, Lipschitz continuously differentiable.
- $h_i$ is a simple proper closed convex function on $\mathbb{R}^{n_i}$, but not necessarily smooth.

We assume

- $A_i^T A_i$ is nonsingular and the solution set of the problem is nonempty.
Direct extension of ADMM:

\[
\begin{align*}
&\text{For } i = 1, \ldots, m \\
&x_i^{k+1} = \text{arg min } \mathcal{L}^\rho(x_1^{k+1}, \ldots, x_{i-1}^{k+1}, x_i, x_{i+1}^{k}, \ldots, x_m^{k}; \lambda^k); \\
&\text{End} \\
&\lambda^{k+1} = \lambda^k + \rho(Ax^{k+1} - b),
\end{align*}
\]

where \( \mathcal{L}^\rho(x_1, \ldots, x_m; \lambda) = f(x) + \lambda^T(Ax - b) + \frac{\rho}{2} \|Ax - b\|^2 \),

\( f = \sum_{i=1}^m f_i + h_i \), \( Ax := \sum_{i=1}^m A_i x_i \) and \( \rho > 0 \) is a parameter.

Practical efficiency has been observed in many recent applications.

However, not necessarily converge! (Chen, He, Ye, Yuan 2013)

Moreover, each subproblem is solved exactly, which could be very expensive, not practical or even impossible when no closed formula exists.
Literature

- Han and Yuan (2012):
  * Assume all $f_i$ strongly convex and $\rho$ in a specific range
  * Each subproblem needs to be solved exactly.

- He, Tao, Xu and Yuan (2013)
  * Block Gaussian backward substitution is used to ensure convergence.
  * Each subproblem needs to be solved exactly.

- Hong and Luo (2013)
  $$\lambda^{k+1} = \lambda^k + \alpha \rho (Ax^{k+1} - b),$$
  * Stepsize $\alpha$ needs to be sufficiently small.
  * $f_i$ and $h_i$ needs to satisfy certain local error bound conditions.
  * Linear convergence rate is achieved.

- Much more recent works: randomization, $m - 2$ strongly convex assumption, ...
Motivation

- Extend BOSVS to the multi-block case to solve the subproblems inexacty.

- Analogous to the standard Augmented Lagrangian Method (ALM), i.e. $m = 1$, solve the subproblems to the adaptive accuracy relative to the current KKT error. (Do not require summable error! Ex. Eckstein-Bertsekas 1992: $\|x_i^k - x_i^*\| \leq \eta^k$ with $\sum_k^{\infty} \eta^k < \infty$)

- Apply accelerated optimal gradient methods to solve the subproblems.

- Global convergence is guaranteed.
Notation

- Let us define $H = \text{diag}(A_2^T A_2, A_3^T A_3, \ldots, A_m^T A_m)$ and

$$
M = \begin{pmatrix}
A_2^T A_2 & 0 & \cdots & 0 \\
A_3^T A_2 & A_3^T A_3 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
A_m^T A_2 & A_m^T A_3 & \cdots & A_m^T A_m
\end{pmatrix}.
$$

Note $H$ is positive definite and $M$ is nonsingular.

- Let us define

$$
\Phi_i^k(u, \bar{u}, \delta) = f_i(\bar{u}) + \nabla f_i(\bar{u})^T (u - \bar{u}) + \frac{\delta}{2} \|u - \bar{u}\|^2 + h_i(u)
+ \frac{\rho}{2} \|A_i u - b_i^k + \lambda^k / \rho\|^2,
$$

where $b_i^k = b - \sum_{j<i} A_j \bar{x}_j^k - \sum_{j>i} A_j \tilde{x}_j^k$. 

---

Separable Convex Optimization $m \geq 3$
Parameters: $\rho, \theta_1, \theta_2, \theta_3 \in \mathbb{R}^+; 0 < \delta_{\min} << \delta_{\max}, \alpha \in (0,1), 0 < \sigma < 1 < \eta, \{\epsilon^k\} \subset \mathbb{R}^+ \text{ with } \sum_{k=1}^{\infty} \epsilon^k < \infty.$

Step 0: Initialize $\lambda^1$. For $i = 1, \ldots, m$, set $\delta_{\min,i} = \delta_{\min}$, initialize $x^1_i$ and set $\tilde{x}^1_i = \bar{x}^1_i = x^1_i$. Let $k = 1$.

Step 1: For $i = 1, \ldots, m$

Choose $\delta^k_i = \eta^j \delta^k_{i,0}$, where $j \geq 0$ is the smallest integer such that

$$f_i(x^k_i) + \langle \nabla f_i(x^k_i), \bar{x}^k_i - x^k_i \rangle + \frac{(1-\sigma)\delta^k_i}{2} \|ar{x}^k_i - x^k_i\|^2 + \epsilon^k \geq f_i(\bar{x}^k_i),$$

where $\delta_{\min} \leq \delta^k_{i,0} \leq \delta_{\max}$ and $\bar{x}^k_i = \text{Arg min}_{u \in \mathbb{R}^{n_i}} \Phi^k_i(u, x^k_i, \delta^k_i)$.

Let $x^{k+1}_i = \bar{x}^k_i$ and $\bar{r}^k_i = (1/\delta^k_i)\|ar{x}^k_i - x^k_i\|^2$.

If $\delta^k_i > \max\{\delta^k_{i-1}, \delta_{\min,i}\}$ when $k > 1$, $\delta_{\min,i} := \eta \delta_{\min,i}$.

End

Step 2: Let $e^k = \theta_1 \|	ilde{v}^k - \tilde{v}^k\| + \theta_2 \| \sum_{i=1}^{m} A_i \bar{x}^k_i - b\| + \theta_3 \sqrt{\sum_{i=1}^{m} \bar{r}^k_i}$.

If $e^k$ is sufficiently small, stop the algorithm.

Step 3: Let $\bar{x}^{k+1}_1 = \bar{x}^k_1$ and $\lambda^{k+1} = \lambda^k + \alpha \rho (\sum_{i=1}^{m} A_i \bar{x}^k_i - b)$.

Compute $\tilde{v}^{k+1} = \tilde{v}^k + \alpha M^{-T} H(\bar{v}^k - \tilde{v}^k)$. (Backward Substitution)

Let $k := k + 1$ and go to Step 1.

Figure: L-ADMM-G Algorithm
Global Convergence of L-ADMM-G

Theorem. Let $\tilde{w}^k = (\tilde{x}_1^k, \tilde{x}_2^k, \ldots, \tilde{x}_n^k, \lambda^k)$, $\bar{w}^k = (\bar{x}_1^k, \bar{x}_2^k, \ldots, \bar{x}_n^k, \lambda^k)$ be the iterates generated by the L-ADM-G algorithm. Then,

$$\lim_{k \to \infty} \tilde{w}^k = \lim_{k \to \infty} \bar{w}^k = w^*,$$

where $w^* = (x_1^*, x_2^*, \ldots, x_n^*, \lambda^*)$ is an optimal primal-dual solution pair.

Proof. Denote $E_k = \rho \|\tilde{v}_e^k\|_G^2 + \frac{1}{\rho} \|\lambda_e^k\|^2 + \alpha \sum_{i=1}^m (\delta_i^k \|x_i^k\|^2)$. We can show

$$E_k \geq E_{k+1} + \tau_k,$$

for large $k$, where

$$\tau_k = \bar{c} \sum_{i=1}^m \|\tilde{x}_i^k - x_i^k\|^2 + \bar{c}(\rho \|\tilde{v}^k - \bar{v}^k\|_H^2 + \frac{1}{\rho} \|\lambda^k - \bar{\lambda}^k\|^2) - \bar{c} \epsilon^k.$$
Inexact ADMM with Gaussian backward substitution

Parameters: \( \rho, \theta_1, \theta_2, \theta_3 \in \mathbb{R}^{++}, \alpha \in (0,1), \{\epsilon^k\} \subset \mathbb{R}^+ \) with \( \sum_{k=1}^{\infty} \epsilon^k < \infty \).

Step 0: Initialize \( \lambda^1 \). For \( i = 1, \ldots, m \), initialize \( x^1_i \) and set \( \bar{x}^1_i = \bar{x}^1_i = x^1_i \). Let \( e^0 = \infty \), \( \Gamma^0_i = 0 \) and \( k = 1 \).

Step 1: For \( i = 1, \ldots, m \)

Use the Accelerated optimal Gradient (AOG) method to solve \( \min_{u \in \mathbb{R}^{n_i}} L^k_i(u) \) inexacty.

End

Step 2: Let \( e^k = \theta_1 \| \bar{v}^k - \tilde{v}^k \| + \theta_2 \| \sum_{i=1}^{m} A_i \bar{x}^k_i - b \| + \theta_3 \sqrt{\sum_{i=1}^{m} \bar{r}^k_i} \). If \( e^k \) is sufficiently small, stop the algorithm.

Step 3: Let \( \bar{x}^{k+1}_1 = \bar{x}^k_1 \) and \( \lambda^{k+1} = \lambda^k + \alpha \rho (\sum_{i=1}^{m} A_i \bar{x}^k_i - b) \).

Compute \( \bar{v}^{k+1} = \bar{v}^k + \alpha M^{-T} H(\bar{v}^k - \tilde{v}^k) \). (Backward Substitution)

Let \( k := k + 1 \) and go to Step 1.

Figure: I-ADMM-G Algorithm
Notation

- For any $h_i$ and $z_i \in \mathbb{R}^{n_i}$, we define its proximal mapping

$$\text{prox}_{h_i}(z_i) = \text{Arg min}_{u \in \mathbb{R}^{n_i}} h_i(u) + \frac{1}{2}\|z_i - u\|^2,$$

and define $\tilde{L}_i^k(u) := L_i^k(u) - h_i(u)$, where

$$L_i^k(u) := L^\rho(\bar{x}_1^k, \ldots, \bar{x}_{i-1}^k, u, \tilde{x}_i^k, \ldots, \tilde{x}_m^k; \lambda^k).$$

- We have

$$x_{i,k}^* = \text{Arg min}_{u \in \mathbb{R}^{n_i}} L_i^k(u)$$

if and only if

$$\|\text{prox}_{h_i}(x_{i,k}^* - \nabla \tilde{L}_i^k(x_{i,k}^*)) - x_{i,k}^*\| = 0.$$

- Define $\psi : \mathbb{R} \to \mathbb{R}^+$ to be any function having the property that

$$\lim_{t \to 0} \psi(t) = 0 \quad \text{and} \quad \psi(t) = 0 \text{ if and only if } t = 0.$$
Figure: Solve \( \min_{u \in \mathbb{R}^n_i} L_i^k(u) \) inexactness

Parameters: \( 0 \leq \sigma < 1, \ \epsilon^k > 0, \ \{\omega^\ell\} \subset \mathbb{R}^+ \) with \( \sum_{\ell=1}^\infty \omega^\ell < \infty \).

Let \( r_i^0 = 0, \ y_i^0 = u_i^0 = x_i^k, \ \alpha^1 = 1, \ \ell = 1 \) and \( \text{flag} = \text{true} \).

while (\( \text{flag} = \text{true} \))

Choose \( \delta^\ell > 0 \) and \( \alpha^\ell \in (0, 1) \) for \( \ell \geq 2 \) such that

\[
f_i(\bar{y}_i^\ell) + \langle \nabla f_i(\bar{y}_i^\ell), y_i^\ell - \bar{y}_i^\ell \rangle + \frac{(1-\sigma)\delta^\ell}{2\alpha^\ell} \|y_i^\ell - \bar{y}_i^\ell\|^2 + \pi^\ell \geq f_i(y_i^\ell), \tag{II}
\]

where \( \bar{y}_i^\ell = (1 - \alpha^\ell)y_i^{\ell-1} + \alpha^\ell u_i^{\ell-1}, \ y_i^\ell = (1 - \alpha^\ell)y_i^{\ell-1} + \alpha^\ell u_i^\ell \),

\[
u_i^\ell = \arg\min_{u \in \mathbb{R}^{n_i}} \langle \nabla f_i(\bar{y}_i^\ell), u \rangle + \frac{\delta^\ell}{2}\|u - u_i^{\ell-1}\|^2 + \frac{\rho}{2}\|A_i u - b_i^k + \lambda^k / \rho\|^2 + h_i(u),
\]

and \( \pi^\ell = \epsilon^k \omega^\ell / (2\gamma^\ell) \) with \( \gamma^1 = 1 / \delta^1 \) and \( \gamma^\ell = \gamma^{\ell-1} / (1 - \alpha^\ell) \) for \( \ell \geq 2 \).

Let \( r_i^\ell = r_i^{\ell-1} + \delta^\ell \alpha^\ell \gamma^\ell \|u_i^\ell - u_i^{\ell-1}\|^2 \).

If \( \gamma^\ell \geq \Gamma_i^{k-1} \) and \( \|\text{prox}_{h_i}(y_i^\ell - \nabla \tilde{L}_i^k(y_i^\ell)) - y_i^\ell\| \leq \psi(\epsilon^{k-1}) \),

set \( x_i^{k+1} = u_i^\ell, \ \Gamma_i^k = \gamma^\ell, \ \bar{x}_i^k = y_i^\ell, \ \tilde{r}_i^k = r_i^\ell / \Gamma_i^k \) and \( \text{flag} = \text{false} \);

else \( \ell := \ell + 1 \).
If the Lipschitz constant $\zeta_i$ of $\nabla f_i$ is known, we could simply set
\[
\delta^\ell = \frac{1}{(1 - \sigma)} \frac{2\zeta_i}{\ell} \quad \text{and} \quad \alpha^\ell = \frac{2}{\ell + 1} \in (0, 1].
\]

Comment:

* We have
\[
\frac{(1 - \sigma)\delta^\ell}{\alpha^\ell} = \frac{(l + 1)\zeta_i}{l} > \zeta_i.
\]
Hence, the condition (II) in the AG algorithm will be satisfied.

* In addition, we have
\[
\gamma^\ell = \frac{1}{\delta^1} \left[ \prod_{j=2}^{\ell} (1 - \alpha^j) \right]^{-1} = \frac{\ell(\ell + 1)}{2\delta^1} \quad \text{and} \quad \xi^\ell := \delta^\ell \alpha^\ell \gamma^\ell = 1.
\]
When the Lipschitz constant of $\nabla f_i$ is unknown, let $\tau^\ell = 1/(\delta_0^\ell \eta^j)$, where $\delta_0^\ell \in [\delta_{\text{min}}, \delta_{\text{max}}]$ and $j \geq 0$ is the smallest integer such that with

$$\delta^\ell = \frac{2}{\tau^\ell + \sqrt{(\tau^\ell)^2 + 4\tau^\ell \Lambda^{\ell-1}}} \quad \text{and} \quad \alpha^\ell = \frac{1}{1 + \delta^\ell \Lambda^{\ell-1}} \in (0, 1],$$

the condition (II) in the AG algorithm is satisfied, where $\Lambda^{\ell-1} = \sum_{i=1}^{\ell-1} \frac{1}{\delta^i}$.

Comment:

* Then, we have

$$\frac{(1 - \sigma)\delta^\ell}{\alpha^\ell} = \frac{1 - \sigma}{\theta^\ell} = (1 - \sigma)\delta_0^\ell \eta^j.$$ 

Hence, the condition (II) in the AG algorithm will be satisfied as $j \to \infty$.

* In addition, we have

$$\gamma^\ell = \Lambda^\ell \geq \frac{\ell^2}{4(\eta \zeta_i/(1 - \sigma) + \delta_{\text{max}})} \quad \text{and} \quad \xi^\ell := \delta^\ell \alpha^\ell \gamma^\ell = 1.$$
Lemma. Suppose the accelerated optimal gradient method with the previous line search rule is applied to solve the subproblem. Then we have

\[ \|y_i^\ell - x_{i,k}^*\|^2 \leq \frac{1}{\nu_i \rho} \left( \|u_i^0 - x_{i,k}^*\|^2 + \epsilon_k \sum_{j=1}^\ell \omega_j \right) \gamma^\ell, \]

where \( x_{i,k}^* = \text{Arg min } L_i^k(u) \) and \( \nu_i \) is the minimum eigenvalue of \( A_i^T A_i \).

Comment:

- We have optimal convergence rate

\[ \|y_i^\ell - x_{i,k}^*\|^2 \leq \mathcal{O}(\frac{1}{\ell^2}). \]

- The subproblem stopping condition

\[ \gamma^\ell \geq \Gamma_i^{k-1} \text{ and } \|\text{prox}_{\lambda_i}(y_i^\ell - \nabla \tilde{L}_i^k(y_i^\ell)) - y_i^\ell\| \leq \psi(e^{k-1}), \]

will be satisfied in finite number of steps.
Theorem. Let $\tilde{w}^k = (\tilde{x}_1^k, \tilde{x}_2^k, \ldots, \tilde{x}_n^k, \lambda^k)$, $\bar{w}^k = (\bar{x}_1^k, \bar{x}_2^k, \ldots, \bar{x}_n^k, \lambda^k)$ be the iterates generated by the I-ADMM-G algorithm. Suppose

- $\lim_{\ell \to \infty} \gamma^\ell = \infty$;
- $\xi^\ell := \delta^\ell \alpha^\ell \gamma^\ell$ is constant;

then,

$$\lim_{k \to \infty} \tilde{w}^k = \lim_{k \to \infty} \bar{w}^k = w^*,$$

where $w^* = (x_1^*, x_2^*, \ldots, x_n^*, \lambda^*)$ is an optimal primal-dual solution pair.

Proof. Denote $E_k = \rho \|\tilde{v}_e^k\|_G^2 + \frac{1}{\rho} \|\lambda^k\|_H^2 + \alpha \sum_{i=1}^m \frac{\|x_i^k\|_e^2}{\Gamma_i^k}$. We can show

$$E_k \geq E_{k+1} + \tau_k,$$

for large $k$, where

$$\tau_k = \tilde{c} \sum_{i=1}^m \frac{1}{\Gamma_i^k} \sum_{\ell=1}^{l_i^k} \|u_{i,k}^\ell - u_{i,k}^{\ell-1}\|^2 + \bar{c} (\rho \|\tilde{v}_e^k - \bar{v}_e^k\|_H^2 + \frac{1}{\rho} \|\lambda^k - \bar{\lambda}_e^k\|_H^2) - \alpha \sum_{i=1}^m \frac{\epsilon_k^i}{\Gamma_i^k} \sum_{\ell=1}^{l_i^k} \omega^\ell.$$
An image deblurring problem for the Cameraman image

- Optimize
  $$\min_u \frac{1}{2} \| Au - b \|^2 + \alpha \| u \|_{TV} + \beta \| \Phi^T u \|_1,$$
  which is equivalent to
  $$\min_u \frac{1}{2} \| Au - b \|^2 + \alpha \| w \|_{1,2} + \beta \| z \|_1, \text{ s. t. } Bu = w, \Phi^T u = z$$
  where $\Phi$ is the wavelet transform, $\alpha = 0.005$ and $\beta = 0.001$.
- size $256 \times 256$, $9 \times 9$ uniform blur and Gaussian noise $SNR = 40$
- Relative accuracy for the subproblem: $\psi(e) = \min\{0.1e, e^{1.1}\}$
- Plot the relative objective function error versus the CPU time ($\rho = 5 \times 10^{-4}$).
Plots of the relative objective error of the Cameraman image (size $256 \times 256$, $9 \times 9$ uniform blur and Gaussian noise $SNR = 40$)
Numerical Experiments

Three Partially Parallel Imaging (PPI) problems

- Optimize

\[
\min_u \frac{1}{2} \| Au - b \|^2 + \alpha \| u \|_{TV} + \beta \| \Phi^T u \|_1,
\]

which is equivalent to

\[
\min_u \frac{1}{2} \| Au - b \|^2 + \alpha \| w \|_{1,2} + \beta \| z \|_1, \text{ s. t. } Bu = w, \Phi^T u = z
\]

where \( \Phi \) is the wavelet transform, \( \alpha = 10^{-5} \) and \( \beta = 10^{-6} \).

- Relative accuracy for the subproblem: \( \psi(e) = \min\{0.1e, e^{1.1}\} \)

- Plot the relative objective function error versus the CPU time \( (\rho = 10^{-3}) \).
Figure: Plots of the relative objective error for data 1, 2 and 3.