Primal Dual Algorithms for Convex Models and Applications to Image Restoration, Registration and Nonlocal Inpainting

Ernie Esser

UCLA
2-16-10
Summary of Contributions

- General framework for a class of primal-dual algorithms
  - Connections between PDHG, AMA, PFBS, ADMM, Douglas Rachford splitting and split inexact Uzawa (Ch. 3, Fig. 3.1)
- Convergence of split Bregman by connection to ADMM (Sec. 2.3.2.3)
- Convergence of modified PDHG by connection to split inexact Uzawa (Sec. 3.4.2.3)
- Operator splitting techniques for extending application to a large class of convex models (Sec. 2.1 and 3.6.1)
- Clarification of dual interpretations and general shrinkage formulas (Sec. 2.4.2)
- Proposed convex model for image registration (Ch. 4)
- Proposed convex model for nonlocal patch-based image inpainting (Ch. 5)
- Demonstrated successful application of ADMM and PDHG variants to image restoration, multiphase segmentation and the proposed registration and nonlocal inpainting models
Key References


Status of Papers


From ATC to Dissertation

Main ATC Topics:

- Fourier inpainting with compressive sensing applications
- Adapting regularizer to different applications (TV, Mumford Shah, $H^{-1}$)
- Fast numerical implementations combining operator splitting, quadratic penalty methods and alternating minimization

Focus shifted away from Fourier inpainting and studying different regularizers

Focused instead on studying efficient primal-dual algorithms for solving convex programs and their application to image restoration
Motivation to Study Primal Dual Algorithms

- Split Bregman and PDHG, both primal-dual methods, demonstrated clear potential to be significantly more efficient than previous methods used to solve convex models in image processing, but convergence properties were initially unclear.
- Overwhelming number of seemingly related methods, but with often unclear connections.
- Need for methods with simple, explicit iterations capable of solving large scale, often non-differentiable convex models.
A Framework for Relating Algorithms

\[ \min_u F_P(u) = J(Au) + H(u) \]

\[ \max_p F_D(p) = -J^*(p) - H^*(-A^T p) \]

Legend:
- **(P):** Primal
- **(D):** Dual
- **(PD):** Primal-Dual
- **(SPP):** Split Primal
- **(SPP):** Split Dual

AMA: Alternating Minimization Algorithm (4.2.1)
PFBS: Proximal Forward Backward Splitting (4.2.1)
ADMM: Alternating Direction Method of Multipliers (4.2.2)
PDHG: Primal Dual Hybrid Gradient (4.2)
PDHGM: Modified PDHG (4.2.3)
Bold: Well Understood Convergence Properties
\[
\min_{z \in \mathbb{R}^n, u \in \mathbb{R}^m} F(z) + H(u) \quad (P0)
\]

\[
Bz + Au = b
\]

\[
z^{k+1} = \arg \min_z F(z) - F(z^k) - \langle p_z^k, z - z^k \rangle + \frac{\alpha}{2} \| b - Au^k - Bz^k \|^2
\]

\[
u^{k+1} = \arg \min_u H(u) - H(u^k) - \langle p_u^k, u - u^k \rangle + \frac{\alpha}{2} \| b - Au - Bz^{k+1} \|^2
\]

\[
p_z^{k+1} = p_z^k + \alpha B^T (b - Au^{k+1} - Bz^{k+1})
\]

\[
p_u^{k+1} = p_u^k + \alpha A^T (b - Au^{k+1} - Bz^{k+1})
\]

\[
p_z^0 = 0 \quad p_u^0 = 0 \quad p_u^k \in \partial H(u^k) \quad p_z^k \in \partial F(z^k)
\]


UCLA CAM Report [08-29], April 2008.
ADMM

\[ L_\alpha(z, u, \lambda^k) = F(z) + H(u) + \langle \lambda^k, b - Au - Bz \rangle + \frac{\alpha}{2} \| b - Au - Bz \|^2 \]

\[ z^{k+1} = \arg \min_z L_\alpha(z, u^k, \lambda^k) \]

\[ u^{k+1} = \arg \min_u L_\alpha(z^{k+1}, u, \lambda^k) \]

\[ \lambda^{k+1} = \lambda^k + \alpha (b - A u^{k+1} - B z^{k+1}) \]

Equivalence to Split Bregman with

\[ p_z^k = B^T \lambda^k \]
\[ p_u^k = A^T \lambda^k \]
\[ \lambda^0 = 0 \]


Dual Interpretations

\[ L(z, u, \lambda) = F(z) + H(u) + \langle \lambda, b - Au - Bz \rangle \]

\[ q(\lambda) = \inf_{u \in \mathbb{R}^m, z \in \mathbb{R}^n} L(z, u, \lambda) = -F^*(B^T \lambda) - H^*(A^T \lambda) + \langle \lambda, b \rangle \]

Dual Problem: \[ \max_{\lambda \in \mathbb{R}^d} q(\lambda) \quad (Q0) \]

Strong Duality: \[ F(z^*) + H(u^*) = q(\lambda^*) \]

Saddle Point Characterization: \((z^*, u^*)\) solves (P0) and \(\lambda^*\) solves (Q0) iff

\[ L(z^*, u^*, \lambda) \leq L(z^*, u^*, \lambda^*) \leq L(z, u, \lambda^*) \quad \forall \; z, u, \lambda \]

Optimality Conditions \[ Au^* + Bz^* = b \]

\[ B^T \lambda^* \in \partial F(z^*) \quad z^* \in \partial F^*(B^T \lambda^*) \]

\[ A^T \lambda^* \in \partial H(u^*) \quad u^* \in \partial H^*(A^T \lambda^*) \]
Douglas Rachford Splitting

Define: $\Psi(\lambda) = B\partial F^*(B^T\lambda) - b$ \quad $\phi(\lambda) = A\partial H^*(A^T\lambda)$

An Approach for solving dual: Find $0 \in \Psi(\lambda) + \phi(\lambda)$

Formal Douglas Rachford Splitting:

\[
0 \in \frac{\hat{\lambda}^k - \lambda^k}{\alpha} + \Psi(\hat{\lambda}^k) + \phi(\lambda^k),
\]

\[
0 \in \frac{\lambda^{k+1} - \lambda^k}{\alpha} + \Psi(\hat{\lambda}^k) + \phi(\lambda^{k+1}).
\]

ADMM Equivalent Version:

\[
\hat{\lambda}^k = \arg \min_{\hat{\lambda}} F^*(B^T\hat{\lambda}) - \langle \hat{\lambda}, b \rangle + \frac{1}{2\alpha} \| \hat{\lambda} - (2\lambda^k - y^k) \|^2
\]

\[
\lambda^{k+1} = \arg \min_{\lambda} H^*(A^T\lambda) + \frac{1}{2\alpha} \| \lambda - (y^k - \lambda^k + \hat{\lambda}^k) \|^2
\]

\[
y^{k+1} = y^k + \hat{\lambda}^k - \lambda^k
\]

General Moreau Decomposition

Let $J$ be a closed proper convex function on $\mathbb{R}^n$, $f \in \mathbb{R}^m$ and $A \in \mathbb{R}^{n \times m}$. Then

$$f = \arg \min_u J(Au) + \frac{1}{2\alpha} \|u - f\|^2 + \alpha A^T \arg \min_p J^*(p) + \frac{\alpha}{2} \|A^T p - \frac{f}{\alpha}\|^2$$

Can use this to directly show equivalence between ADMM applied to (P0) and Douglas Rachford Splitting applied to (Q0)
ADMM Convergence

**Theorem 1** (Eckstein, Bertsekas) Consider the problem (P0) where $F$ and $H$ are closed proper convex functions, $F(z) + \|Bz\|^2$ is strictly convex and $H(u) + \|Au\|^2$ is strictly convex. Let $\lambda^0 \in \mathbb{R}^d$ and $u^0 \in \mathbb{R}^m$ be arbitrary and let $\alpha > 0$. Suppose we are also given sequences $\{\mu_k\}$ and $\{\nu_k\}$ such that $\mu_k \geq 0$, $\nu_k \geq 0$, $\sum_{k=0}^{\infty} \mu_k < \infty$ and $\sum_{k=0}^{\infty} \nu_k < \infty$. Suppose that

\[
\|z^{k+1} - \arg\min_{z \in \mathbb{R}^n} F(z) + \langle \lambda^k, -Bz \rangle + \frac{\alpha}{2} \|b - Au^k - Bz\|^2\| \leq \mu_k
\]

\[
\|u^{k+1} - \arg\min_{u \in \mathbb{R}^m} H(u) + \langle \lambda^k, -Au \rangle + \frac{\alpha}{2} \|b - Au - Bz^{k+1}\|^2\| \leq \nu_k
\]

\[
\lambda^{k+1} = \lambda^k + \alpha(b - Au^{k+1} - Bz^{k+1}).
\]

If there exists a saddle point of $L(z, u, \lambda)$, then $z^k \to z^*$, $u^k \to u^*$ and $\lambda^k \to \lambda^*$, where $(z^*, u^*, \lambda^*)$ is such a saddle point. On the other hand, if no such saddle point exists, then at least one of the sequences $\{u^k\}$ or $\{\lambda^k\}$ must be unbounded.

Generalize Application by Operator Splitting

Convert
\[ \min_{u \in \mathbb{R}^m} \ J(u) \] to
\[ \min_{z \in \mathbb{R}^n, u \in \mathbb{R}^m} \ F(z) + H(u) \]

\[ Ku = f \]
\[ Bz + Au = b \]

where \( J(u) \) can be written as a sum of closed proper convex functions \( H \) and \( G_i \) composed with linear operators,

\[ J(u) = H(u) + \sum_{i=1}^{N} G_i(A_i u + b_i) \]

by letting \( F(z) = \sum_{i=1}^{N} G_i(z_i) \)

\[ z = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix} \]

\[ B = \begin{bmatrix} -I \\ 0 \end{bmatrix} \]

\[ A = \begin{bmatrix} A_1 \\ \vdots \\ A_N \\ K \end{bmatrix} \]

\[ b = \begin{bmatrix} -b_1 \\ \vdots \\ -b_N \\ f \end{bmatrix} \]
**TV-$l_1$ Formulation**

$$\min_{u \in \mathbb{R}^m} \|Du\|_E + \beta\|Ku - f\|_1,$$

Where $D$ is discrete gradient and $\| \cdot \|_E$ is norm defined so that $\|Du\|_E = \|u\|_{TV}$.

Assume $\ker(D) \cap \ker(K) = \{0\}$ and let

$$z = \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} Du \\ Ku - f \end{bmatrix} \quad B = -I \quad A = \begin{bmatrix} D \\ K \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ f \end{bmatrix}.$$  

Write Lagrange multiplier for constraint $Au + Bz = b$ as $\lambda = \begin{bmatrix} p \\ q \end{bmatrix}$

Augmented Lagrangian: $L_\alpha(w, v, u, p, q) =$

$$\|w\|_E + \beta\|v\|_1 + \langle p, w - Du \rangle + \langle q, v - Ku + f \rangle + \frac{\alpha}{2}\|w - Du\|^2 + \frac{\alpha}{2}\|v - Ku + f\|^2.$$
TV-$l_1$ Iterations

\[ w^{k+1} = \arg\min_w \|w\|_E + \frac{\alpha}{2} \|w - Du^k + \frac{p^k}{\alpha}\|^2 \]

\[ v^{k+1} = \arg\min_v \beta\|v\|_1 + \frac{\alpha}{2} \|v - Ku^k + f + \frac{q^k}{\alpha}\|^2 \]

\[ u^{k+1} = \arg\min_u \frac{\alpha}{2} \|Du - w^{k+1} - \frac{p^k}{\alpha}\|^2 + \frac{\alpha}{2} \|Ku - v^{k+1} - f - \frac{q^k}{\alpha}\|^2 \]

\[ p^{k+1} = p^k + \alpha(w^{k+1} - Du^{k+1}) \]

\[ q^{k+1} = q^k + \alpha(v^{k+1} - Ku^{k+1} + f), \]

where $p^0 = q^0 = 0$, $u^0$ is arbitrary and $\alpha > 0$. 
TV-$l_1$ Numerical Result

<table>
<thead>
<tr>
<th>Image Size</th>
<th>Iterations</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 × 64</td>
<td>40</td>
<td>1s</td>
</tr>
<tr>
<td>128 × 128</td>
<td>51</td>
<td>5s</td>
</tr>
<tr>
<td>256 × 256</td>
<td>136</td>
<td>78s</td>
</tr>
<tr>
<td>512 × 512</td>
<td>359</td>
<td>836s</td>
</tr>
</tbody>
</table>

Iterations until $\|u^k - u^{k-1}\|_\infty \leq .5$, $\|Du^k - w^k\|_\infty \leq .5$ and $\|v^k - u^k + f\|_\infty \leq .5$

$\beta = .6, .3, .15$ and .075, $\alpha = .02, .01, .005$ and .0025
Better Understanding of PDHG

\[
\min_u J(Au) + H(u) \quad (P)
\]

\[
J(Au) = J^{**}(Au) = \sup_p \langle p, Au \rangle - J^*(p)
\]

Saddle Point Formulation:

\[
\min_u \sup_p -J^*(p) + \langle p, Au \rangle + H(u) \quad (PD)
\]

Interpret PDHG as primal-dual proximal point method:

\[
p^{k+1} = \arg\max_{p \in \mathbb{R}^n} -J^*(p) + \langle p, Au^k \rangle - \frac{1}{2\delta_k} \|p - p^k\|^2
\]

\[
u^{k+1} = \arg\min_{u \in \mathbb{R}^m} H(u) + \langle A^T p^{k+1}, u \rangle + \frac{1}{2\alpha_k} \|u - u^k\|^2
\]

Connections to Other Methods

\[
\begin{align*}
\min_u F_P(u) & \quad = J(Au) + H(u) \\
\max_p F_D(p) & \quad = -J^*(p) - H^*(-A^T p)
\end{align*}
\]

\[
\begin{align*}
\min_u \sup_{u,p} L_{PD}(u,p) & \quad = \langle p, Au \rangle - J^*(p) - H^*(u) \\
\min_u \sup_{u,p} L_{PD}(u,p) & \quad = \langle p, Au \rangle - J^*(p) - H^*(u) + \langle u, -A^T p - y \rangle
\end{align*}
\]

Legend:
- (P): Primal
- (D): Dual
- (PD): Primal-Dual
- (SP\_P): Split Primal
- (SP\_D): Split Dual
PFBS, AMA and PDHG

PFBS on (P)
(additional PDHG proximal penalty in red)

\[ p^{k+1} = \arg \min_{p \in \mathbb{R}^n} J^*(p) + \langle -Au^k, p \rangle + \frac{1}{2\delta_k} \| p - p^k \|_2^2 \]

\[ u^{k+1} = \arg \min_{u \in \mathbb{R}^m} H(u) + \langle u, A^T p^{k+1} \rangle + \frac{1}{2\alpha_k} \| u - u^k \|_2^2 \]

Equivalent (by Moreau decomposition) to AMA applied to (SPD)
(again with additional PDHG penalty in red)

\[ p^{k+1} = \arg \min_{p \in \mathbb{R}^m} J^*(p) + \langle -Au^k, p \rangle + \frac{1}{2\delta_k} \| p - p^k \|_2^2 \]

\[ y^{k+1} = \arg \min_{y \in \mathbb{R}^m} H^*(y) - \langle u^k, y \rangle + \frac{\alpha_k}{2} \| y + A^T p^{k+1} \|_2^2 \]

\[ u^{k+1} = u^k + \alpha_k ( -A^T p^{k+1} - y^{k+1} ) \]

- PFBS on (D) and AMA on (SPP) are connected to PDHG in the analogous way.
PDHG and ADMM

Add augmented Lagrangian penalty to first step of AMA to get ADMM.

ADMM on (SPP)

\[ u^{k+1} = \arg \min_{u \in \mathbb{R}^m} H(u) + \langle A^T p^k, u \rangle + \frac{\delta_k}{2} \|Au - w^k\|^2 \]

\[ w^{k+1} = \arg \min_{w \in \mathbb{R}^n} J(w) - \langle p^k, w \rangle + \frac{\delta_k}{2} \|w - Au^{k+1}\|^2 \]

\[ p^{k+1} = p^k + \delta_k (Au^{k+1} - w^{k+1}) \]

ADMM on (SPD)

\[ p^{k+1} = \arg \min_{p \in \mathbb{R}^m} J^*(p) + \langle -Au^k, p \rangle + \frac{\alpha_k}{2} \|A^T p + y^k\|^2 \]

\[ y^{k+1} = \arg \min_{y \in \mathbb{R}^m} H^*(y) - \langle u^k, y \rangle + \frac{\alpha_k}{2} \|y + A^T p^{k+1}\|^2 \]

\[ u^{k+1} = u^k + \alpha_k (-A^T p^{k+1} - y^{k+1}) \]
Decoupling Variables

One can add additional proximal-like penalties to ADMM iterations and obtain a more explicit algorithm that still converges.

Given a step of the ADMM algorithm of the form

\[ u^{k+1} = \arg \min_u J(u) + \frac{\delta}{2} \| Ku - f^k \|^2, \]

modify the objective functional by adding

\[ \frac{1}{2} \left\langle u - u^k, \left( \frac{1}{\alpha} - \delta K^T K \right) (u - u^k) \right\rangle, \]

where \( \alpha \) is chosen such that \( 0 < \alpha < \frac{1}{\delta \| K^T K \|} \).

Modified update is given by

\[ u^{k+1} = \arg \min_u J(u) + \frac{1}{2\alpha} \| u - u^k + \alpha \delta K^T (Ku^k - f^k) \|^2. \]

Split Inexact Uzawa (special case)

Consider adding $\frac{1}{2} \langle p - p^k, (\frac{1}{\delta_k} - \alpha_k A A^T)(p - p^k) \rangle$ to the first step of ADMM on (SPD), with $0 < \delta_k < \frac{1}{\alpha_k \|A\|^2}$.

Split Inexact Uzawa on (SPD)

$$p^{k+1} = \arg \min_{p \in \mathbb{R}^n} J^*(p) + \frac{1}{2\delta_k} \| p - p^k - \delta_k A u^k + \alpha_k \delta_k A (A^T p^k + y^k) \|^2_2$$

$$y^{k+1} = \arg \min_{y \in \mathbb{R}^m} H^*(y) - \langle u^k, y \rangle + \frac{\alpha_k}{2} \| y + A^T p^{k+1} \|^2_2$$

$$u^{k+1} = u^k + \alpha_k (-A^T p^{k+1} - y^{k+1})$$

Note: In general we could similarly modify both minimization steps in ADMM, but by only modifying the first step we can obtain an interesting PDHG-like interpretation.

Convergence of Split Inexact Uzawa

Convergence requires

- $\alpha_k = \alpha > 0$
- $\delta_k = \delta > 0$
- $0 < \delta < \frac{1}{\alpha \|A\|^2}$

Then if $p^*$ is optimal for (D) and $y^* = -A^T p^*$,

- $\|A^T p^k + y^k\|_2 \rightarrow 0$
- $J^*(p^k) \rightarrow J^*(p^*)$
- $H^*(y^k) \rightarrow H^*(y^*)$
- All convergent subsequences of $(p^k, y^k, u^k)$ converge to a saddle point of $L_D$

Modified PDHG

Split Inexact Uzawa applied to (SPP) with fixed step size parameters:

\[ u^{k+1} = \arg \min_{u \in \mathbb{R}^m} H(u) + \frac{1}{2\alpha} \| u - u^k + \alpha A^T p^k + \delta \alpha A^T (Au^k - w^k) \|^2_2 \]

\[ w^{k+1} = \arg \min_{w \in \mathbb{R}^n} J(w) - \langle p^k, w \rangle + \frac{\delta}{2} \| Au^{k+1} - w \|^2_2 \]

\[ p^{k+1} = p^k + \delta (Au^{k+1} - w^{k+1}) \]

PDHGMp:

\[ u^{k+1} = \arg \min_{u \in \mathbb{R}^m} H(u) + \langle A^T (2p^k - p^{k-1}) , u \rangle + \frac{1}{2\alpha} \| u - u^k \|^2_2 \]

\[ p^{k+1} = \arg \min_{p \in \mathbb{R}^n} J^*(p) - \langle p, Au^{k+1} \rangle + \frac{1}{2\delta} \| p - p^k \|^2_2 \]

Generalize Application of PDHG

Many seemingly more complicated problems can be written in the form (P).

Example:

\[ \sum_{i=1}^{N} \phi_i(B_i A_i u + b_i) + H(u) = \sum_{i=1}^{N} J_i(A_i u) + H(u) = J(A u) + H(u), \]

where \( A = \begin{bmatrix} A_1 \\ \vdots \\ A_N \end{bmatrix} \) and \( J_i(z_i) = \phi_i(B_i z_i + b_i) \). Let \( p = \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} \). Then

\[ J^*(p) = \sum_{i=1}^{N} J_i^*(p_i) \]
Applying PDHGMP to \( \min_u \sum_{i=1}^{N} \phi_i(B_iA_iu + b_i) + H(u) \) yields:

\[
 u^{k+1} = \arg \min_u H(u) + \frac{1}{2\alpha} \left\| u - \left( u^k - \alpha \sum_{i=1}^{N} A_i^T (2p_i^k - p_i^{k-1}) \right) \right\|_2^2 \\
 p_i^{k+1} = \arg \min_{p_i} J_i^*(p_i) + \frac{1}{2\delta} \left\| p_i - (p_i^k + \delta A_i u^{k+1}) \right\|_2^2 \quad \text{for } i = 1, \ldots, N
\]

- Need \( 0 < \alpha \delta < \frac{1}{\|A\|^2} \) for stability
Especially Efficient for Certain Functionals

Want simple, explicit solutions to minimization subproblems.

Convex constraint $u \in S$ can be handled by adding convex indicator function

$$g_S(u) = \begin{cases} 0 & \text{if } u \in S \\ \infty & \text{otherwise.} \end{cases}$$

Simple update when orthogonal projection

$$\Pi_S(f) = \arg \min_u g_S(u) + \|u - f\|^2$$

is easy to compute

$$S = \{z : \|z - f\|_2 \leq \epsilon\} \quad \Pi_S(z) = f + \frac{z - f}{\max\left(\frac{\|z-f\|_2}{\epsilon}, 1\right)}$$
A Few Other Examples

- $J(z) = \|z\|_2 \Rightarrow J^*(p) = g\{p: \|p\|_2 \leq 1\}$
- $J(z) = \frac{1}{2\alpha} \|z\|_2^2 \Rightarrow J^*(p) = \frac{\alpha}{2} \|p\|_2^2$
- $J(z) = \|z\|_1 \Rightarrow J^*(p) = g\{p: \|p\|_\infty \leq 1\}$
- $J(z) = \|z\|_E \Rightarrow J^*(p) = g\{p: \|p\|_{E^*} \leq 1\}$
- $J(z) = \|z\|_\infty \Rightarrow J^*(p) = g\{p: \|p\|_1 \leq 1\}$
- $J(z) = \max(z) \Rightarrow J^*(p) = g\{p: p \geq 0 \text{ and } \|p\|_1 = 1\}$

Note: Although there’s no simple formula for projecting a vector onto the $l_1$ unit ball (or its positive face) in $\mathbb{R}^n$, this can be computed with $O(n \log n)$ complexity.
General Shrinkage Formula (1)

Define $J(w) = \sum_i \|w_i\|$ and compute the Legendre transform of $J$.

$$J^*(p) = \sup_w \langle p, w \rangle - J(w)$$

$$= \sum_i \sup_{w_i} \langle p_i, w_i \rangle - \|w_i\|$$

$$= \sum_i \sup_{w_i} \|w_i\|(\|p_i\| - 1)$$

$$= \begin{cases} 0 & \text{if } \max_i \|p_i\| \leq 1 \\ \infty & \text{otherwise.} \end{cases}$$

Consider $$\min_w J(w) + \frac{1}{2\mu} \|w - f\|^2$$
General Shrinkage Formula (2)

Apply Moreau decomposition:

$$\arg\min_w J(w) + \frac{1}{2\mu}||w - f||^2 = f - \mu \arg\min_p J^*(p) + \frac{\mu}{2}||p - \frac{f}{\mu}||^2$$

$$= f - \mu \arg\min_{\{p:\max_i ||p_i||\leq 1\}} \frac{\mu}{2}||p - \frac{f}{\mu}||^2$$

$$= f - \mu \Pi_{\{p:\max_i ||p_i||\leq 1\}}\left(\frac{f}{\mu}\right)$$

$$= f - \Pi_{\{p:\max_i ||p_i||\leq \mu\}}(f)$$

Reduces to $l_1$-$l_2$ scalar shrinkage formula when $w_i$ is scalar.

$$\arg\min_w \|w\|_1 + \frac{1}{2\mu}||w - f||^2 = f - \Pi_{\{p:\|p_i\|_\infty \leq \mu\}}(f)$$
Constrained TV Deblurring Example

\[
\min_{\| Ku - f \|_2 \leq \epsilon} \| u \|_{TV}
\]

can be rewritten as

\[
\min_u \| Du \|_E + g_T(Ku),
\]

where \( g_T \) is the indicator function for \( T = \{ z : \| z - f \|_2 \leq \epsilon \} \)

In order to treat both \( D \) and \( K \) explicitly, let

\[
H(u) = 0 \quad \text{and} \quad J(Au) = J_1(Du) + J_2(Ku),
\]

where \( A = \begin{bmatrix} D \\ K \end{bmatrix} \).

Write the dual variable as \( p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \) and apply PDHGMP.
Deblurring Numerical Result

$K$ convolution operator for normalized Gaussian blur with Std. dev. 3

$h$ clean image

$f = Kh + \eta$

$\eta$ zero mean Gaussian noise Std. dev. 1

$\epsilon = 256$

$\alpha = .2, \delta = .55$

Original, blurry/noisy and image recovered from 300 PDHGMp iterations
Convex Image Registration Model

Given images \( u \) and \( \phi \), minimize

\[
\frac{1}{2} \| \phi(x + v(x)) - u(x) \|^2 + \frac{\gamma}{2} \| Dv_1 \|^2 + \frac{\gamma}{2} \| Dv_2 \|^2
\]

with respect to displacement field \( v \).

Obtain convex relaxation by adding edges with unknown weights \( c_{i,j} \) such that

\[
(v^i_1, v^i_2) = \left( \sum_{j \sim i} c_{i,j} y^j_1 - x^i_1, \sum_{j \sim i} c_{i,j} y^j_2 - x^i_2 \right) = ((A y_1 c - x_1)_i, (A y_2 c - x_2)_i)
\]

Replace \( \phi(x + v(x))_i \) with \( (A \phi c)_i = \sum_{j \sim i} c_{i,j} \phi_j \).

\[
F(c) = \frac{1}{2} \| A \phi c - u \|^2 + \frac{\eta}{2} \| D(A y_1 c - x_1) \|^2 + \frac{\eta}{2} \| D(A y_2 c - x_2) \|^2 + g_C(c)
\]

Minimize over \( c \in C \), \( C = \{ c : c_{i,j} \geq 0 \text{ and } \sum_{j \sim i} c_{i,j} = 1 \} \).
Encouraging Localized Weights

Reindex \( c \in \mathbb{R}^{M \times W} = \begin{bmatrix} c_1 & \ldots & c_W \end{bmatrix}, c_w \in \mathbb{R}^M \) assuming \( W \) weights for each of the \( M \) pixels. Define \( C = \{ c : \sum_w c_w = 1, c_{m,w} \geq 0 \} \).

Encourage spatial smoothness of \( c_w \) by adding \( \sum_{w=1}^{W} \| D\mathcal{X}_w c \|_2 \), where \( \mathcal{X}_w c = c_w \).

To control local error: \( \| \frac{A_{\phi} c - u}{\tau} \|_{\infty} \leq 1 \) for some data dependent \( \tau \in \mathbb{R}^M \)

To control average error: \( \| A_{\phi} c - u \|_2 \leq \epsilon \) for some \( \epsilon \geq 0 \).

Define \( T_2 = \{ z : \| z - u \|_2 \leq \epsilon \} \quad T_{\infty} = \{ z : \| \frac{z - u}{\tau} \|_{\infty} \leq 1 \} \).

Proposed functional, with \( g \) denoting indicator function:

\[
F(c) = g_C(c) + \sum_{w=1}^{W} \| (D\mathcal{X}_w c) \|_2 + g_{T_2}(A_{\phi} c) + g_{T_{\infty}}(A_{\phi} c) + \frac{\eta}{2} \| D(A_{y_1} c - x_1) \|_2^2 + \frac{\eta}{2} \| D(A_{y_2} c - x_2) \|_2^2
\]
Multiscale Approach

Effect of downsampling on resolution and search window size

Interpolate low resolution displacement estimate \((A_{y_1} c - x_1, A_{y_2} c - x_2)\) to use as initial guess at next finer resolution.

Apply PDHGMp to minimize \(F(c)\) at each scale, from coarse to fine.
Pencil Registration Example

Registration of Low Resolution Photo of Two Pencils

Iterations required for $\|c^{k+1} - c^k\|_\infty \leq \frac{0.002}{W}$

<table>
<thead>
<tr>
<th>Scale</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>14276</td>
</tr>
<tr>
<td>1</td>
<td>6560</td>
</tr>
<tr>
<td>0</td>
<td>10403</td>
</tr>
</tbody>
</table>
Convex Model for Patch-Based Inpainting

Regions Defined for Nonlocal Inpainting Model

\[ h_0 \]
Proposed Inpainting Model

\[ G(c, u) = g_C(c) + g_S(u) + \frac{\mu}{2} \| A c - B u \|_F^2 + \| D c \|_1, \]

where

\[ S = \{ u : u(v) = h(v) \text{ for } v \in \Omega_o \} \]

the set of \( u \) that agree with boundary condition.

\[ C = \{ c : c(p, m) \geq 0 \text{ and } \sum_p c(p, m) = 1 \forall m \} \]

the set of \( c \) satisfying normalization constraint.

\[ \| A c - B u \|_F^2 \] ensures unknown patches agree with \( u \).

\[ \| D c \|_1 \] encourages spatial correspondence of patches.

Weighted Average Image Update

Solve for $u$ by projecting weighted average onto $S$.

$$
 u(v) = \begin{cases} 
 h(v) & v \in \Omega_o \\
 ((B^* B)^{-1} B^* A c)(v) & v \in \Omega 
\end{cases}.
$$

Split up data fidelity term into parts on $\Omega_o$ and $\Omega$.

$$
 x_{\Omega_o}(q, v) = \begin{cases} 
 1 & v \in \Omega_o \\
 0 & \text{otherwise}
\end{cases}, \quad x_{\Omega}(q, v) = \begin{cases} 
 1 & v \in \Omega \\
 0 & \text{otherwise}
\end{cases}.
$$

$$
 h_0(v) = \begin{cases} 
 h(v) & v \in \Omega_o \\
 0 & \text{otherwise},
\end{cases} \quad f = x_{\Omega_o} \cdot B h_0,
$$

$$
 A_{\Omega_o} = x_{\Omega_o} \cdot A \quad A_{\Omega} = x_{\Omega} \cdot (I - B (B^* B)^{-1} B^*) A.
$$

Now we can define a convex functional just in terms of $c$, 

$$
 F(c) = g_C(c) + \frac{\mu_{\Omega_o}}{2} \| A_{\Omega_o} c - f \|_F^2 + \frac{\mu_{\Omega}}{2} \| A_{\Omega} c \|_F^2 + \| D c \|_1.
$$

Brick Wall Inpainting Example (1)

Inpainting brick wall using $15 \times 15$ patches but without $\|Dc\|_1$ correspondence term

Using PDHGMp to minimize $F(c)$
Brick Wall Inpainting Example (2)

Inpainting brick wall using $15 \times 15$ patches and including correspondence term
Grass Inpainting Example

Inpainting grass using $15 \times 15$ patches and including correspondence term
Nonconvex Modification for Binary Weights

To encourage binary weights, add nonconvex term analogous to use of double well potential in phase field approach to segmentation.

\[ \gamma \sum_{p,m} c_{p,m}(1 - c_{p,m}). \]

\[ F_{nc}(c) = g_C(c) + \gamma \langle c, 1 \rangle - \gamma \|c\|_F^2 + \frac{\mu_{\Omega_o}}{2} \|A_{\Omega_o} c - f\|_F^2 + \frac{\mu_{\Omega}}{2} \|A_{\Omega} c\|_F^2 + \|Dc\|_1 \]

Since objective functional for PDHGMp \( c^{k+1} \) update adds \( \frac{1}{2\alpha} \|c - c^k\|_F^2 \), the PDHGMp subproblems remain strictly convex if \( 0 \leq \gamma < \frac{1}{2\alpha} \).

Convergence theory no longer applies, but method works empirically to yield good inpainting solutions with binary weights.

For a slight variant of the problem (Replace \( \|Dc\|_1 \) with smooth \( \xi(Dc) \) where \( \xi(d) = \min_z \|z\|_1 + \frac{1}{2\varepsilon} \|z - d\|_2^2 \), \( \nabla(\xi(Dc)) = \frac{1}{\varepsilon} D^T \Pi\{z: \|z\|_\infty \leq \varepsilon\}(Dc) \))

PFBS will converge to a local minimum.
Grass Inpainting with Nonconvex Model

Inpainting grass using $15 \times 15$ patches, correspondence term and nonconvex modification.
Brick Inpainting with Nonconvex Model

Inpainting brick wall using $45 \times 45$ patches, correspondence term and nonconvex modification
Conclusions

• The general framework for a class of primal-dual algorithms greatly reduces the dauntingly large space of potential methods.

• Operator splitting techniques extend application to a large class of convex models, including many of practical interest in image processing.

• The work on PDHG variants for minimizing non-differentiable convex functionals by a sequence of simple, explicit iterations makes it possible to tackle large scale problems that might otherwise be considered too computationally intensive.

• Effectiveness of the modified PDHG method is demonstrated for convex models of image restoration, registration, patch-based inpainting and a relaxation of multiphase segmentation.