

A primal-dual extragradient method for nonlinear forward models in function spaces

Christian Clason¹ Tuomo Valkonen²

¹Faculty of Mathematics, Universität Duisburg-Essen

²Department of Mathematical Sciences, University of Liverpool

SIAM Conference on Imaging Science

Albuquerque, May 23, 2016

Primal-dual extragradient method:

- first-order algorithm for nonsmooth convex problems with linear operators [Chambolle/Pock 2011]
- *very* popular in imaging (TV denoising, deblurring, ...)
- **acceleration** (Nesterov, $\mathcal{O}(1/k^2)$ convergence)
- version for **nonlinear** operators [Valkonen 2014]

Here:

- application to **parameter identification for PDEs**
- \rightsquigarrow **function space** algorithm

Difficulty:

- convergence proof requires **set-valued analysis** in **infinite-dimensional spaces**

L^1 -fitting

$$\min_u \frac{1}{\alpha} \|S(u) - y^\delta\|_{L^1} + \frac{1}{2} \|u\|_{L^2}^2$$

L^∞ -fitting/Morozov

$$\min_u \frac{1}{2} \|u\|_{L^2}^2 \quad \text{s. t.} \quad |S(u)(x) - y^\delta(x)| \leq \delta \quad \text{a. e. in } \Omega$$

$S : U \subset L^2(\Omega) \rightarrow L^2(\Omega)$, $S(u) =: y$ satisfies

$$\begin{cases} -\Delta y + uy = f \\ \partial_\nu y = 0 \end{cases}$$

- 1 Overview
- 2 Algorithm
- 3 Set-valued analysis
- 4 Application to parameter identification
 - L^1 fitting
 - L^∞ fitting
- 5 Numerical examples

$$\min_{u \in X} F(K(u)) + G(u)$$

- $F : Y \rightarrow \overline{\mathbb{R}}, G : X \rightarrow \overline{\mathbb{R}}$ convex, lower semicontinuous
- X, Y Hilbert spaces
- $K \in C^2(X, Y)$ (here: $K(u) = S(u) - y^\delta$)
- saddle point formulation:

$$\min_{u \in X} \sup_{v \in Y^*} G(u) + \langle K(u), v \rangle - F^*(v)$$

- $F^* : Y^* \rightarrow \overline{\mathbb{R}}$ Fenchel conjugate

K linear:

$$\begin{cases} u^{k+1} = \text{prox}_{\tau G}(u^k - \tau K^* v^k) \\ \bar{u}^{k+1} = 2u^{k+1} - u^k \\ v^{k+1} = \text{prox}_{\sigma F^*}(v^k + \sigma K \bar{u}^{k+1}) \end{cases}$$

- σ, τ step sizes, $\sigma\tau < \|K\|^{-2}$
- $\text{prox}_{\sigma F}(v) = \arg \min_u \frac{1}{2\sigma} \|u - v\|^2 + F(u)$ proximal mapping

K nonlinear:

$$\begin{cases} u^{k+1} = \text{prox}_{\tau G}(u^k - \tau K'(u^k)^* v^k) \\ \bar{u}^{k+1} = 2u^{k+1} - u^k \\ v^{k+1} = \text{prox}_{\sigma F^*}(v^k + \sigma K(\bar{u}^{k+1})) \end{cases}$$

- σ, τ step sizes, $\sigma\tau < \sup_{u \in B_R} \|K'(u)\|^{-2}$
- $\text{prox}_{\sigma F}(v) = \arg \min_u \frac{1}{2\sigma} \|u - v\|^2 + F(u)$ proximal mapping
- $K'(u)$ Fréchet derivative, $K'(u)^*$ adjoint

K nonlinear, accelerated:

$$\begin{cases} u^{k+1} = \text{prox}_{\tau_k G}(u^k - \tau^k K'(u^k)^* v^k) \\ \omega_k = 1/\sqrt{1 + 2c\tau^k} \quad \tau^{k+1} = \omega_k \tau^k \quad \sigma_{k+1} = \sigma_k / \omega_k \\ \bar{u}^{k+1} = u^{k+1} + \omega_k (u^{k+1} - u^k) \\ v^{k+1} = \text{prox}_{\sigma_{k+1} F^*}(v^k + \sigma_{k+1} K(\bar{u}^{k+1})) \end{cases}$$

- σ, τ step sizes, $\sigma_0 \tau_0 < \sup_{u \in B_R} \|K'(u)\|^{-2}$
- $\text{prox}_{\sigma F}(v) = \arg \min_u \frac{1}{2\sigma} \|u - v\|^2 + F(u)$ proximal mapping
- $K'(u)$ Fréchet derivative, $K'(u)^*$ adjoint
- $c \geq 0$ acceleration parameter

Theorem

Iterates converge locally to saddle point (\bar{u}, \bar{v}) if

- 1 G is c_G -strongly convex (here: $c_G = 1$)
- 2 $c = c_n \in [0, c_G)$, $c_n = 0$ for $n > N \in \mathbb{N}$ (finite acceleration)
- 3 *metric regularity* around saddle point

(cf. [Valkonen 2014])

Difficulty:

- metric regularity in **function spaces**
- \rightsquigarrow requires **infinite-dimensional** set-valued analysis
- here: only rough outline, no details

- 1 Overview
- 2 Algorithm
- 3 Set-valued analysis**
- 4 Application to parameter identification
 - L^1 fitting
 - L^∞ fitting
- 5 Numerical examples

Saddle-point problem

$$\min_{u \in X} \sup_{v \in Y^*} G(u) + \langle K(u), v \rangle - F^*(v)$$

Primal-dual optimality conditions

$$\begin{cases} K(\bar{u}) \in \partial F^*(\bar{v}) \\ -K'(\bar{u})^* \bar{v} \in \partial G(\bar{u}) \end{cases}$$

Primal-dual optimality conditions

$$\begin{cases} K(\bar{u}) \in \partial F^*(\bar{v}) \\ -K'(\bar{u})^* \bar{v} \in \partial G(\bar{u}) \end{cases}$$

Set inclusion for $H : L^2(\Omega)^2 \rightrightarrows L^2(\Omega)^2$

$$0 \in H_{\bar{u}}(\bar{u}, \bar{v}) := \begin{pmatrix} \partial G(\bar{u}) + K'(\bar{u})^* \bar{v} \\ \partial F^*(\bar{v}) - K(\bar{u}) \end{pmatrix}$$

Set inclusion for $H : L^2(\Omega)^2 \rightrightarrows L^2(\Omega)^2$

$$0 \in H_{\bar{u}}(\bar{u}, \bar{v}) := \begin{pmatrix} \partial G(\bar{u}) + K'(\bar{u})^* \bar{v} \\ \partial F^*(\bar{v}) - K(\bar{u}) \end{pmatrix}$$

Metric regularity at (\bar{u}, \bar{v})

$$\inf_{q: w \in H_{\bar{u}}(q)} \|q - (\bar{u}, \bar{v})\| \leq L \|w\| \quad \text{for all } \|w\| \leq \rho$$

- interpretation: small perturbation w of 0
⇒ small perturbation q of saddle point (\bar{u}, \bar{v})
- Lipschitz property for set-valued $H_{\bar{u}}^{-1}$ at $((\bar{u}, \bar{v}), 0)$

Metric regularity at (\bar{u}, \bar{v})

$$\inf_{q: w \in H_{\bar{u}}(q)} \|q - (\bar{u}, \bar{v})\| \leq L \|w\| \quad \text{for all } \|w\| \leq \rho$$

Mordukhovich criterion

$$L_H = \inf_{t>0} \sup \left\{ \|\widehat{D}^* H(q'|w')\| \mid q' \in B((\bar{u}, \bar{v}), t), w' \in H(q') \cap B(w, t) \right\}$$

- Aubin constant L_H is minimal choice of L
- $\widehat{D}^* H$ regular coderivative of H (cf. $L = \|\nabla f\|$ for $f \in C^1$)
- \rightsquigarrow set-valued analysis in function spaces

Difficulties:

- multiple non-equivalent concepts (regular, limiting)
- calculus not tight

Here:

- set-valued mappings from subdifferentials of **pointwise** functionals
- \rightsquigarrow infinite-dimensional (regular) derivatives **pointwise** via nice finite-dimensional (regular, graphical) derivatives
- cf. pointwise Fenchel conjugates, subdifferentials [Ekeland]

$$0 \in H_{\bar{u}}(\bar{u}, \bar{v}) = \begin{pmatrix} \partial G(\bar{u}) + K'(\bar{u})^* \bar{v} \\ \partial F^*(\bar{v}) - K(\bar{u}) \end{pmatrix}$$

$$DH_{\bar{u}}(q|w)(\Delta q) = \begin{pmatrix} D[\partial G](u|\xi - K'(\bar{u})^* v)(\Delta u) + K'(\bar{u})^* \Delta v \\ D[\partial F^*](v|\eta + K'(\bar{u})u + c_{\bar{u}})(\Delta v) - K'(\bar{u})\Delta u \end{pmatrix}$$

$$\widetilde{DH}_{\bar{u}}(q|w)(\Delta q) = \begin{cases} T_q \Delta q + V(q|w)^\circ & \Delta q \in V(q|w) \\ \emptyset & \text{otherwise} \end{cases}$$

- $q = (u, v)$, $w = (\xi, \eta)$, $c_{\bar{u}} = K(\bar{u}) - K'(\bar{u})\bar{u}$
- T_q linear Operator (independent of w), $V(q|w)$ cone
- $\widehat{D}^* H_{\bar{u}}(q|w) = [DH_{\bar{u}}(q|w)]^{*+}$ upper adjoint of convexification

$$\widetilde{DH}_{\bar{u}}(q|w)(\Delta q) = \begin{cases} T_q \Delta q + V(q|w)^\circ & \Delta q \in V(q|w) \\ \emptyset & \text{otherwise} \end{cases}$$

Then: Aubin constant $L_H \leq c < \infty$ **iff**

$$\sup_{t>0} \inf_{\substack{(\Delta w, z) \in W^t(q|w), \\ \|\Delta w\| > 0}} \frac{\|T_q^* \Delta w - z\|}{\|\Delta w\|} \geq c^{-1} > 0$$

$$W^t(q|w) = \bigcup \left\{ V(q'|w') \times V(q'|w')^\circ \mid \begin{array}{l} w' \in H_{\bar{u}}(q'), \\ \|q' - q\| < t, \|w' - w\| < t \end{array} \right\}$$

- 1 Overview
- 2 Algorithm
- 3 Set-valued analysis
- 4 Application to parameter identification**
 - L^1 fitting
 - L^∞ fitting
- 5 Numerical examples

Primal-dual optimality conditions

$$\begin{cases} S(\bar{u}) - y^\delta \in \partial F^*(\bar{v}) \\ -S'(\bar{u})^* \bar{v} = \bar{u} \end{cases}$$

Metric regularity **around** (\bar{u}, \bar{v}) if either

- 1 $\sup_{t>0} \inf \left\{ \frac{\|S'(\bar{u})S'(\bar{u})^*z - v\|}{\|z\|} \mid (z, v) \in V_{\partial F^*}^t(\bar{v} | y^\delta - S(\bar{u})), z \neq 0 \right\} > 0$
- 2 Moreau–Yosida regularization: $F^* \mapsto F_\gamma^* := F^* + \frac{\gamma}{2} \|\cdot\|^2$
- 3 finite-dimensional data: $Y \mapsto Y_h$

In case 1: $\|S'(\bar{u})^*z\| \geq c\|z\|$ for $z \in V_{\partial F^*}^t(\bar{v} | y^\delta - S(\bar{u}))$ necessary!

$$\min_u \frac{1}{\alpha} \|S(u) - y^\delta\|_{L^1} + \frac{1}{2} \|u\|_{L^2}^2$$

■ $F(y) = \int_{\Omega} \alpha^{-1} |y(x)| dx \rightsquigarrow f^* = \delta_{[-\alpha^{-1}, \alpha^{-1}]}(z)$

■ $S : U \subset L^2(\Omega) \rightarrow L^2(\Omega), \quad S(u) =: y$ satisfies

$$\begin{cases} -\Delta y + uy = f \\ \partial_\nu y = 0 \end{cases}$$

Here: $z \in V_{\partial F^*}^t(v|\eta)$ if

$$z(x) \in \begin{cases} \{0\} & |v'(x)| = \alpha^{-1} \text{ and } \eta'(x) \neq 0 \\ -\text{sign } v'(x)[0, \infty) & |v'(x)| = \alpha^{-1} \text{ and } \eta'(x) = 0 \\ \mathbb{R} & |v'(x)| < \alpha^{-1} \text{ and } \eta'(x) = 0 \end{cases}$$

for some $\|v' - \bar{v}\| \leq t, \|\eta' - \bar{\eta}\| \leq t$

- S compact operator: $\|S'(\bar{u})^* z\| \geq c\|z\|$ only holds for $z = 0$
- $\bar{\eta} = S(\bar{u}) - y^\delta, \quad \alpha \bar{v} \in \text{sign } \bar{\eta}$
- \rightsquigarrow in general **not satisfied!**

$$\left\{ \begin{array}{l} z^{k+1} = S'(u^k)^* v^k \\ u^{k+1} = \frac{1}{1 + \tau_k} (u^k - \tau_k z^{k+1}) \\ \omega_k = 1/\sqrt{1 + 2c\tau_k} \quad \tau_{k+1} = \omega_k \tau_k \quad \sigma_{k+1} = \sigma_k / \omega_k \\ \bar{u}^{k+1} = u^{k+1} + \omega_k (u^{k+1} - u^k) \\ v^{k+1} = \text{proj}_{[-\alpha^{-1}, \alpha^{-1}]} \left(\frac{1}{1 + \sigma_{k+1}\gamma} (v^k + \sigma_{k+1} (S(\bar{u}^{k+1}) - y^\delta)) \right) \end{array} \right.$$

- $S'(u^k)^* v^k$ solution of adjoint equation
- proj_C pointwise projection on convex set $C \subset \mathbb{R}$
- Moreau–Yosida parameter $\gamma \geq 0$
- local convergence if $\gamma > 0$ or finite-dimensional

$$\min_u \frac{1}{2} \|u\|_{L^2}^2 \quad \text{s. t.} \quad |S(u)(x) - y^\delta(x)| \leq \delta \quad \text{a. e. in } \Omega$$

■ $F(y) = \delta_{\{|y(x)| \leq \delta\}}(y) \rightsquigarrow f^* = \delta|z|$

■ $S : U \subset L^2(\Omega) \rightarrow L^2(\Omega), S(u) =: y$ satisfies

$$\begin{cases} -\Delta y + uy = f \\ \partial_\nu y = 0 \end{cases}$$

■ for $t = 0$: $z(x) = 0$ if $|S(\bar{u})(x) - y^\delta(x)| < \delta \rightsquigarrow$ estimate unlikely

$$\left\{ \begin{array}{l} z^{k+1} = S'(u^k) * v^k \\ u^{k+1} = \frac{1}{1 + \tau_k} (u^k - \tau_k z^{k+1}) \\ \omega_k = 1/\sqrt{1 + 2c\tau_k} \quad \tau_{k+1} = \omega_k \tau_k \quad \sigma_{k+1} = \sigma_k / \omega_k \\ \bar{u}^{k+1} = u^{k+1} + \omega_k (u^{k+1} - u^k) \\ \bar{v}^{k+1} = v^k + \sigma_{k+1} (S(\bar{u}^{k+1}) - y^\delta) \\ v^{k+1} = \frac{1}{1 + \sigma_{k+1}\gamma} (|\bar{v}^{k+1}| - \delta\sigma)^+ \text{sign}(\bar{v}^{k+1}) \end{array} \right.$$

- local convergence if $\gamma > 0$ or finite-dimensional

- 1 Overview
- 2 Algorithm
- 3 Set-valued analysis
- 4 Application to parameter identification
 - L^1 fitting
 - L^∞ fitting
- 5 Numerical examples

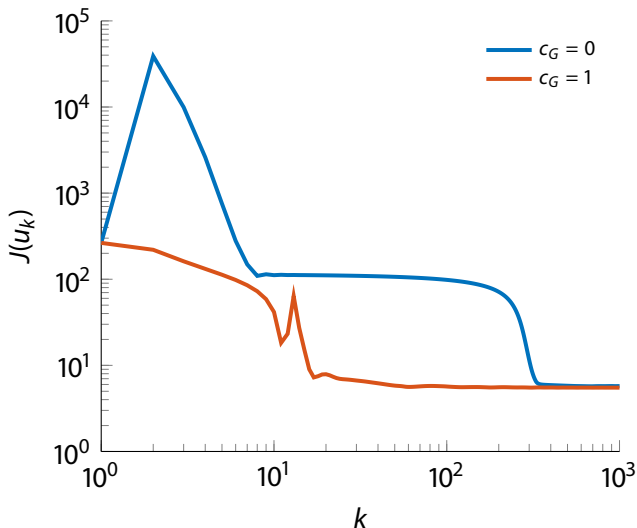
- $\Omega = [-1, 1]$, FE (P_1 - P_0) discretization of (y, u)
- random impulsive noise:

$$y^\delta(x) = \begin{cases} y^\dagger(x) + \|y^\dagger\|_\infty \xi(x) & \text{with probability 0.3} \\ y^\dagger(x) & \text{else} \end{cases}$$

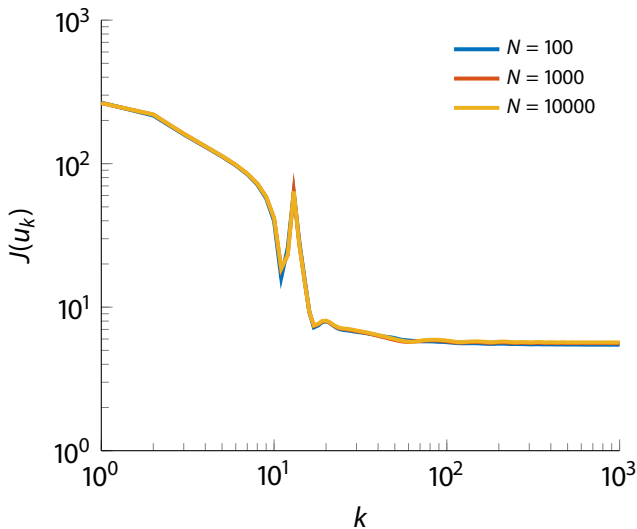
$$y^\dagger = S(u^\dagger), \quad \xi(x) \in \mathcal{N}(0, 0.1), \quad \rightsquigarrow \alpha = 10^{-2}$$

- $\sigma_0 = \tilde{L}^{-1}, \tau_0 = 0.99\tilde{L}^{-1}, \tilde{L} = \|S(u^0)\|/\|u^0\|$
- $\gamma = 10^{-12}, u^0 \equiv 1, v^0 = 0$ (no warmstart!)
- compare $c \in \{0, 1\}, N \in \{100, 1000, 10000\}$

L^1 fitting: acceleration (same data)



L^1 fitting: discretization (avg. of 10)



- $\Omega = [-1, 1]$, FE (P_1 - P_0) discretization of (y, u)

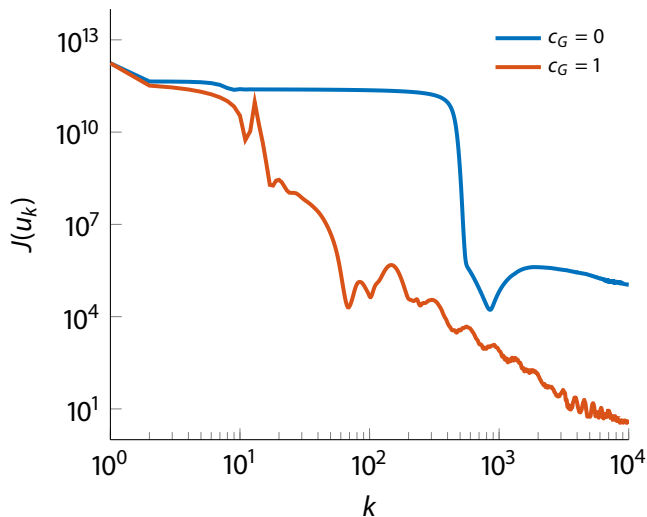
- quantization noise:

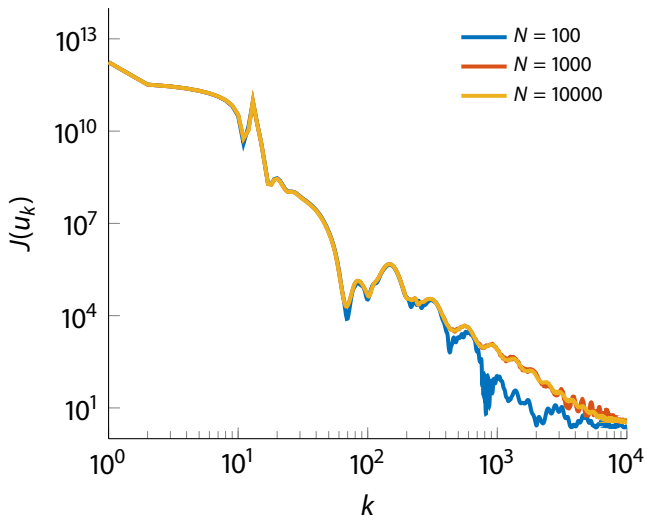
round y^\dagger to $n_b = 11$ equidistant values

- $\gamma = 10^{-12}$, $\sigma_0 = \tilde{L}^{-1}$, $\tau_0 = 0.99\tilde{L}^{-1}$, $\tilde{L} = \|S(u^0)\|/\|u^0\|$

- $u^0 \equiv 1, v^0 = 0$ (no warmstart!)

- compare $c \in \{0, 1\}$, $N \in \{100, 1000, 10000\}$





Primal-dual extragradient methods in function space:

- can be accelerated
- analyzed using set-valued analysis in function space
- requires Moreau–Yosida regularization
- \rightsquigarrow no norm gap, continuation needed; mesh-independence

Outlook:

- full acceleration
- partial stability (w.r.t. primal variable only)
- other PDE-constrained optimization problems

Preprints/Code (soon):

http://www.uni-due.de/mathematik/agclason/clason_pub.php