Strong convergence and fast residual decay for monotone operator flows via Tikhonov regularization

Radu Ioan Boț

University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
1090 Vienna
Austria
www.mat.univie.ac.at/~rabot

ELLIIT Symposium on Optimization for Learning
Lund, Sweden, May 6-8, 2026

Der Wissenschaftsfonds.

Monotone equation

- Consider the monotone equation

$$V(x) = 0,$$

where

- $V : \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator, i. e.

$$\langle z - u, V(z) - V(u) \rangle \geq 0 \quad \forall z, u \in \mathcal{H};$$

- the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Monotone equation

- Consider the monotone equation

$$V(x) = 0,$$

where

- $V : \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator, i. e.

$$\langle z - u, V(z) - V(u) \rangle \geq 0 \quad \forall z, u \in \mathcal{H};$$

- the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of monotone operators

- Finding a global minimum of a convex and differentiable function $f : \mathcal{H} \to \mathbb{R}$ reduces to solving the monotone equation

$$\nabla f(x) = 0.$$

Monotone equation

- Consider the monotone equation

$$V(x) = 0,$$

where

- $V : \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator, i. e.

$$\langle z - u, V(z) - V(u) \rangle \geq 0 \quad \forall z, u \in \mathcal{H};$$

- the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of monotone operators

- Finding a global minimum of a convex and differentiable function $f : \mathcal{H} \to \mathbb{R}$ reduces to solving the monotone equation

$$\nabla f(x) = 0.$$

- Finding a saddle point of a convex-concave and differentiable function $\Phi : \mathcal{H} \times \mathcal{G} \to \mathbb{R}$ reduces to solving the monotone equation

$$\begin{pmatrix} \nabla_x \Phi(x, \lambda) \\ -\nabla_\lambda \Phi(x, \lambda) \end{pmatrix} = 0.$$

Monotone operator flows

- For $t_0 \geq 0$, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}\,(t) + V\,(x(t)) = 0 \\ x\,(t_0) = x^0. \end{cases}$$

- (Baillon-Brézis, 1976): The ergodic trajectory $t \mapsto \frac{1}{t}\int_0^t x(s)ds$ converges weakly to an element in $\operatorname{Zer} V$ as $t \to +\infty$.

Monotone operator flows

- For $t_0 \geq 0$, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0 \\ x(t_0) = x^0. \end{cases}$$

- (Baillon-Brézis, 1976): The ergodic trajectory $t \mapsto \frac{1}{t}\int_0^t x(s)ds$ converges weakly to an element in $\operatorname{Zer} V$ as $t \to +\infty$.
- Consider the counterclockwise $\pi/2$-rotation operator

$$V : \mathbb{R}^2 \to \mathbb{R}^2, \quad V(x,y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

and, for $t_0 := 0$, $(x(0), y(0)) = r_0(\cos(\theta_0), \sin(\theta_0))$. Then $\mathcal{S} = \{(0,0)\}$ and $(x(t), y(t)) = r_0(\cos(\theta_0 - t), \sin(\theta_0 - t))$, which is bounded, but does not have a limit as $t \to +\infty$. However, $t \mapsto \frac{1}{t}\int_0^t (x(s), y(s))ds$ converges to $(0,0)$ as $t \to +\infty$.

Monotone operator flows

- For $t_0 \geq 0$, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0 \\ x(t_0) = x^0. \end{cases}$$

- (Baillon-Brézis, 1976): The ergodic trajectory $t \mapsto \frac{1}{t}\int_0^t x(s)ds$ converges weakly to an element in $\operatorname{Zer} V$ as $t \to +\infty$.
- Consider the counterclockwise $\pi/2$-rotation operator

$$V : \mathbb{R}^2 \to \mathbb{R}^2, \quad V(x,y) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

and, for $t_0 := 0$, $(x(0), y(0)) = r_0(\cos(\theta_0), \sin(\theta_0))$. Then $\mathcal{S} = \{(0,0)\}$ and $(x(t), y(t)) = r_0(\cos(\theta_0 - t), \sin(\theta_0 - t))$, which is bounded, but does not have a limit as $t \to +\infty$. However, $t \mapsto \frac{1}{t}\int_0^t (x(s), y(s))ds$ converges to $(0,0)$ as $t \to +\infty$.

Gradient Descent Ascent (GDA) method (does not converge in general)

$$z^{k+1} := z^k - \theta V\left(z^k\right) \quad \forall k \geq 0.$$

- For the counterclockwise $\pi/2$-rotation operator it holds

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \end{pmatrix} := \begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix}\begin{pmatrix} x^k \\ y^k \end{pmatrix} \quad \forall k \geq 0.$$

Since $\rho\left(\begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix}\right) = \max\{|\lambda_1|, |\lambda_2|\} = \sqrt{1+\theta^2} > 1$, $(x^k, y^k)_{k\geq 0}$ does not converge to $(0,0)$ as $k \to +\infty$.

Monotone flows governed by cocoercive operators

- For $t_0 \geq 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0, \\ x(t_0) = x^0, \end{cases}$$

where

- $V : \mathcal{H} \to \mathcal{H}$ is a ρ-cocoercive operator for some $\rho > 0$, that is

$$\langle z - u, V(z) - V(u) \rangle \geq \rho \| V(z) - V(u) \|^2 \quad \forall z, u \in \mathcal{H};$$

- the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Monotone flows governed by cocoercive operators

- For $t_0 \geq 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0, \\ x(t_0) = x^0, \end{cases}$$

where

  - $V : \mathcal{H} \to \mathcal{H}$ is a ρ-cocoercive operator for some $\rho > 0$, that is

$$\langle z - u, V(z) - V(u) \rangle \geq \rho \| V(z) - V(u) \|^2 \quad \forall z, u \in \mathcal{H};$$

  - the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of cocoercive operators

- $\nabla f : \mathcal{H} \to \mathcal{H}$, for $f : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with ∇f Lipschitz continuous (Theorem of Baillon-Haddad);

Monotone flows governed by cocoercive operators

- For $t_0 \geq 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0, \\ x(t_0) = x^0, \end{cases}$$

where

- $V \colon \mathcal{H} \to \mathcal{H}$ is a ρ-cocoercive operator for some $\rho > 0$, that is

$$\langle z - u, V(z) - V(u) \rangle \geq \rho \|V(z) - V(u)\|^2 \quad \forall z, u \in \mathcal{H};$$

- the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of cocoercive operators

- $\nabla f : \mathcal{H} \to \mathcal{H}$, for $f : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with ∇f Lipschitz continuous (Theorem of Baillon-Haddad);
- the resolvent $J_M := (\mathrm{Id} + M)^{-1} : \mathcal{H} \to \mathcal{H}$ of a (set-valued) maximally monotone operator $M : \mathcal{H} \to 2^{\mathcal{H}}$;

Monotone flows governed by cocoercive operators

- For $t_0 \geq 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0, \\ x(t_0) = x^0, \end{cases}$$

where
  - $V : \mathcal{H} \to \mathcal{H}$ is a ρ-cocoercive operator for some $\rho > 0$, that is

$$\langle z - u, V(z) - V(u) \rangle \geq \rho \| V(z) - V(u) \|^2 \quad \forall z, u \in \mathcal{H};$$

  - the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of cocoercive operators

- $\nabla f : \mathcal{H} \to \mathcal{H}$, for $f : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with ∇f Lipschitz continuous (Theorem of Baillon-Haddad);
- the resolvent $J_M := (\mathrm{Id} + M)^{-1} : \mathcal{H} \to \mathcal{H}$ of a (set-valued) maximally monotone operator $M : \mathcal{H} \to 2^{\mathcal{H}}$;
- the Moreau-Yosida regularization $M^\rho := \frac{1}{\rho}(\mathrm{Id} - J_{\rho M}) : \mathcal{H} \to \mathcal{H}$ of a (set-valued) maximally monotone operator $M : \mathcal{H} \to 2^{\mathcal{H}}$.

Monotone flows governed by cocoercive operators

- For $t_0 \geq 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) = 0, \\ x(t_0) = x^0, \end{cases}$$

where

  - $V : \mathcal{H} \to \mathcal{H}$ is a ρ-cocoercive operator for some $\rho > 0$, that is

$$\langle z - u, V(z) - V(u) \rangle \geq \rho \|V(z) - V(u)\|^2 \quad \forall z, u \in \mathcal{H};$$

  - the solution set $\operatorname{Zer} V$ is assumed to be nonempty.

Examples of cocoercive operators

- $\nabla f : \mathcal{H} \to \mathcal{H}$, for $f : \mathcal{H} \to \mathbb{R}$ a convex and differentiable function with ∇f Lipschitz continuous (Theorem of Baillon-Haddad);
- the resolvent $J_M := (\operatorname{Id} + M)^{-1} : \mathcal{H} \to \mathcal{H}$ of a (set-valued) maximally monotone operator $M : \mathcal{H} \to 2^{\mathcal{H}}$;
- the Moreau-Yosida regularization $M^\rho := \frac{1}{\rho}(\operatorname{Id} - J_{\rho M}) : \mathcal{H} \to \mathcal{H}$ of a (set-valued) maximally monotone operator $M : \mathcal{H} \to 2^{\mathcal{H}}$.

Then it holds (Attouch-RIB-Nguyen, MOR, 2023):

  - $x(t)$ converges weakly to a zero of V as $t \to +\infty$;
  - $\|V(x(t))\| = o\left(\frac{1}{\sqrt{t}}\right)$ as $t \to +\infty$.

Monotone flows with corrector term for V monotone and L-Lipschitz continuous

- (Attouch-Svaiter, 2011): for $\mu, \gamma > 0$,

$$\begin{cases} \dot{x}(t) + \mu \frac{d}{dt} V(x(t)) + \gamma V(x(t)) = 0 \quad \forall t > 0, \\ x(0) = x^0. \end{cases}$$

- existence and uniqueness of the solution trajectory, convergence of $\|V(x(t))\|$ to 0 and weak convergece of $x(t)$ to a zero of V as $t \to +\infty$;

Monotone flows with corrector term for V monotone and L-Lipschitz continuous

- (Attouch-Svaiter, 2011): for $\mu, \gamma > 0$,

$$\begin{cases} \dot{x}(t) + \mu \frac{d}{dt} V(x(t)) + \gamma V(x(t)) = 0 \quad \forall t > 0, \\ x(0) = x^0. \end{cases}$$

- existence and uniqueness of the solution trajectory, convergence of $\|V(x(t))\|$ to 0 and weak convergece of $x(t)$ to a zero of V as $t \to +\infty$;
- (RIB-Nguyen, 2024): $\|V(x(t))\| = o\left(\frac{1}{\sqrt{t}}\right)$ as $t \to +\infty$.

Monotone flows with corrector term for V monotone and L-Lipschitz continuous

- (Attouch-Svaiter, 2011): for $\mu, \gamma > 0$,

$$\begin{cases} \dot{x}(t) + \mu \frac{d}{dt} V(x(t)) + \gamma V(x(t)) = 0 \quad \forall t > 0, \\ x(0) = x^0. \end{cases}$$

- existence and uniqueness of the solution trajectory, convergence of $\|V(x(t))\|$ to 0 and weak convergece of $x(t)$ to a zero of V as $t \to +\infty$;
- (RIB-Nguyen, 2024): $\|V(x(t))\| = o\left(\frac{1}{\sqrt{t}}\right)$ as $t \to +\infty$.

- (RIB-Schindler, 2024):

$$\begin{cases} d(X(t) + \mu(t)V(X(t))) = -(\gamma(t) - \dot{\mu}(t))V(X(t))dt + \sigma(t, X(t))dW(t) \quad \forall t > 0, \\ X(0) = X^0, \end{cases}$$

where

  - $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ is a filtered probability space;
  - the diffusion term $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ is measurable;
  - $X(\cdot)$ and $V(X(\cdot))$ are stochastic Itô processes with the same m-dimensional Brownian motion W.

- existence and uniqueness of the trajectory process, almost sure convergence of $\|V(X(t))\|$ to 0 and of $X(t)$ to a zero of V as $t \to +\infty$, ergodic convergence rates in expectation.

Numerical algorithms for V monotone and L-Lipschitz continuous

Extragradient (EG) method (Korpelevich, 1976; Antipin, 1976)

$$(\forall k \geq 0) \quad \begin{cases} y^{k+1} & := x^k - \theta V\left(x^k\right) \\ x^{k+1} & := x^k - \theta V\left(y^{k+1}\right). \end{cases}$$

- For $0 < \theta < \frac{1}{L}$, convergences weakly to a zero of V.
- Last iterate convergence rate (Gorbunov-Loizou-Gidel, 2022)

$$\left\|V\left(x^k\right)\right\| = O\left(\frac{1}{\sqrt{k}}\right) \text{ and } \mathtt{Gap}\left(x^k\right) = \sup_{u \in \mathbb{B}(x^*;\delta(x^0))} \left\langle x^k - u, V(u)\right\rangle = O\left(\frac{1}{\sqrt{k}}\right) \text{ as } k \to +\infty.$$

Optimistic Gradient Descent Ascent (OGDA) method (Popov, 1980)

$$(\forall k \geq 1) \quad x^{k+1} := x^k - 2\theta V\left(x^k\right) + \theta V\left(x^{k-1}\right)$$

- For $0 < \theta < \frac{1}{2L}$, convergences weakly to a zero of V.
- Last iterate convergence rate (Gorbunov-Taylor-Gidel, 2022; Cai-Oikonomou-Zheng, 2022)

$$\left\|V\left(x^k\right)\right\| = O\left(\frac{1}{\sqrt{k}}\right) \text{ and } \mathtt{Gap}\left(x^k\right) = \sup_{u \in \mathbb{B}(x^*;\delta(x^0))} \left\langle x^k - u, V(u)\right\rangle = O\left(\frac{1}{\sqrt{k}}\right) \text{ as } k \to +\infty.$$

Numerical algorithms for V monotone and L-Lipschitz continuous

Extra-Anchored-Gradient and Popov-type methods

$$(\forall k \geq 0) \quad \begin{cases} y^{k+1} & := x^k - \varepsilon_k (x^k - v) - \eta_k V(z^{k+1}) \\ x^{k+1} & := x^k - \varepsilon_k (x^k - v) - \theta_k V(y^{k+1}). \end{cases}$$

- EAG-C (Yoon-Ryu, 2021): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k = \eta \in \left(0, \frac{1}{8L}\right]$;
- EAG-V (Yoon-Ryu, 2021): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k$ fulfilling

$$\theta_0 \in \left(0, \frac{3}{4L}\right) \text{ and } \theta_{k+1} := \theta_k \left(1 - \frac{1}{(k+1)(k+3)} \frac{\theta_k^2 L^2}{1 - \theta_k^2 L^2}\right) \quad \forall k \geq 0;$$

- FEG (Lee-Kim, 2024): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+1}, \eta_k = \frac{1}{L}\left(1 - \frac{1}{k+1}\right), \theta_k = \frac{1}{L}$;
- APV (Tran-Dinh-Luo, 2024): $z^{k+1} = y^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k$ fulfilling

$$\theta_0 \in \left(0, \frac{1}{2L\sqrt{3}}\right) \text{ and } \theta_{k+1} := \frac{(1 - (\varepsilon_k)^2 - (2L\theta_k)^2)\varepsilon_{k+1}\theta_k}{(1 - (2L\theta_k)^2)(1 - \varepsilon_k)\varepsilon_k} \quad \forall k \geq 0.$$

Numerical algorithms for V monotone and L-Lipschitz continuous

Extra-Anchored-Gradient and Popov-type methods

$$(\forall k \geq 0) \quad \begin{cases} y^{k+1} & := x^k - \varepsilon_k (x^k - v) - \eta_k V(z^{k+1}) \\ x^{k+1} & := x^k - \varepsilon_k (x^k - v) - \theta_k V(y^{k+1}). \end{cases}$$

- EAG-C (Yoon-Ryu, 2021): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k = \eta \in \left(0, \frac{1}{8L}\right]$;
- EAG-V (Yoon-Ryu, 2021): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k$ fulfilling

$$\theta_0 \in \left(0, \frac{3}{4L}\right) \text{ and } \theta_{k+1} := \theta_k \left(1 - \frac{1}{(k+1)(k+3)} \frac{\theta_k^2 L^2}{1 - \theta_k^2 L^2}\right) \quad \forall k \geq 0;$$

- FEG (Lee-Kim, 2024): $z^{k+1} = x^k, \varepsilon_k = \frac{1}{k+1}, \eta_k = \frac{1}{L}\left(1 - \frac{1}{k+1}\right), \theta_k = \frac{1}{L}$;
- APV (Tran-Dinh-Luo, 2024): $z^{k+1} = y^k, \varepsilon_k = \frac{1}{k+2}$, and $\eta_k = \theta_k$ fulfilling

$$\theta_0 \in \left(0, \frac{1}{2L\sqrt{3}}\right) \text{ and } \theta_{k+1} := \frac{(1 - (\varepsilon_k)^2 - (2L\theta_k)^2)\varepsilon_{k+1}\theta_k}{(1 - (2L\theta_k)^2)(1 - \varepsilon_k)\varepsilon_k} \quad \forall k \geq 0.$$

The four methods exhibit a convergence rate of

$$\left\| V\left(x^k\right) \right\| = O\left(\frac{1}{k}\right) \quad \text{as } k \to +\infty.$$

Fast OGDA: continuous time approach

- For $t_0 > 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \ddot{x}(t) + \dfrac{\alpha}{t}\dot{x}(t) + \dfrac{d}{dt}V(x(t)) + \dfrac{\alpha}{2t}V(x(t)) = 0 \\ x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0, \end{cases}$$

for which

- $\alpha \geq 2$ and $\left(x^0, \dot{x}^0\right) \in \mathcal{H} \times \mathcal{H}$;
- we assume that admits a unique strong global solution $x : [t_0, +\infty) \to \mathcal{H}$ with the property that $t \mapsto V(x(t))$ is absolutely continuous on every compact interval.

Fast OGDA: continuous time approach

- For $t_0 > 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \ddot{x}(t) + \dfrac{\alpha}{t}\dot{x}(t) + \dfrac{d}{dt}V(x(t)) + \dfrac{\alpha}{2t}V(x(t)) = 0 \\ x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0, \end{cases}$$

for which

- $\alpha \geq 2$ and $\left(x^0, \dot{x}^0\right) \in \mathcal{H} \times \mathcal{H}$;
- we assume that admits a unique strong global solution $x : [t_0, +\infty) \to \mathcal{H}$ with the property that $t \mapsto V(x(t))$ is absolutely continuous on every compact interval.

- For $x^* \in \operatorname{Zer} V$ and $1 \leq \lambda \leq \alpha - 1$, let the energy function $\mathcal{E}_\lambda : [t_0, +\infty) \to [0, +\infty)$,

$$\mathcal{E}_\lambda(t) := \frac{1}{2}\left\|2\lambda\left(x(t) - x^*\right) + t\Big(2\dot{x}(t) + V(x(t))\Big)\right\|^2 + 2\lambda(\alpha - 1 - \lambda)\left\|x(t) - x^*\right\|^2$$
$$+2\lambda t\left\langle x(t) - x^*, V(x(t))\right\rangle + \frac{1}{2}t^2\left\|V(x(t))\right\|^2.$$

Fast OGDA: continuous time approach

- For $t_0 > 0$ we consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \ddot{x}(t) + \dfrac{\alpha}{t}\dot{x}(t) + \dfrac{d}{dt}V(x(t)) + \dfrac{\alpha}{2t}V(x(t)) = 0 \\ x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0, \end{cases}$$

for which

  - $\alpha \geq 2$ and $\left(x^0, \dot{x}^0\right) \in \mathcal{H} \times \mathcal{H}$;
  - we assume that admits a unique strong global solution $x : [t_0, +\infty) \to \mathcal{H}$ with the property that $t \mapsto V(x(t))$ is absolutely continuous on every compact interval.

- For $x^* \in \operatorname{Zer} V$ and $1 \leq \lambda \leq \alpha - 1$, let the energy function $\mathcal{E}_\lambda : [t_0, +\infty) \to [0, +\infty)$,

$$\mathcal{E}_\lambda(t) := \frac{1}{2}\left\|2\lambda\left(x(t) - x^*\right) + t\Big(2\dot{x}(t) + V(x(t))\Big)\right\|^2 + 2\lambda(\alpha - 1 - \lambda)\left\|x(t) - x^*\right\|^2$$
$$+2\lambda t\left\langle x(t) - x^*, V(x(t))\right\rangle + \frac{1}{2}t^2\left\|V(x(t))\right\|^2.$$

- $t \mapsto \mathcal{E}_{\alpha-1}(t)$ is nonincreasing on $[t_0, +\infty)$, thus

$$0 \leq \|V(x(t))\| \leq \sqrt{2\mathcal{E}_{\alpha-1}(t_0)}\,\frac{1}{t} \quad \text{and} \quad 0 \leq \left\langle x(t) - x^*, V(x(t))\right\rangle \leq \frac{\mathcal{E}_{\alpha-1}(t_0)}{2(\alpha-1)}\,\frac{1}{t}.$$

Convergence of trajectories and faster convergence rates

Theorem

If $\alpha > 2$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ be a solution of the dynamical system, and $x^* \in \operatorname{Zer} V$, then it holds

- $x(t)$ *converges weakly to a zero of* V *as* $t \to +\infty$*;*
- $$\|\dot{x}(t)\| = o\left(\frac{1}{t}\right) \quad \text{as } t \to +\infty,$$

  $$\left\langle x(t) - x^*, V(x(t)) \right\rangle = o\left(\frac{1}{t}\right) \text{ and } \|V(x(t))\| = o\left(\frac{1}{t}\right) \text{ as } t \to +\infty,$$

  and

  $$0 \leq \mathtt{Gap}(x(t)) = \sup_{u \in \mathbb{B}\left(x^*; \delta(x^0)\right)} \langle x(t) - u, V(u) \rangle = o\left(\frac{1}{t}\right) \text{ as } t \to +\infty.$$

Explicit temporal discretization/Explicit Fast OGDA method

Algorithm

Let $\alpha > 2, x^0, x^1, \bar{x}^0 \in \mathcal{H}$, *and* $0 < \theta < \frac{1}{4L}$. *For every* $k \geq 1$ *we set*

$$\bar{x}^k := x^k + \left(1 - \frac{\alpha}{k+\alpha}\right)\left(x^k - x^{k-1}\right) - \frac{\alpha\theta}{k+\alpha} V\left(\bar{x}^{k-1}\right)$$

$$x^{k+1} := \bar{x}^k - \theta\left(1 + \frac{k}{k+\alpha}\right)\left(V\left(\bar{x}^k\right) - V\left(\bar{x}^{k-1}\right)\right).$$

Convergence of the iterates and fast convergence rates

Theorem

Let $x^* \in \operatorname{Zer} V$ and $\left(x^k\right)_{k\geq 0}$ the sequence generated by the Explicit Fast OGDA method. Then it holds:

- *the sequence $\left(x^k\right)_{k\geq 0}$ converges weakly to a zero of V;*
- $$\left\| x^k - x^{k-1} \right\| = o\left(\frac{1}{k}\right), \quad \left\langle x^k - x^*, V\left(x^k\right)\right\rangle = o\left(\frac{1}{k}\right),$$
  $$\left\| V\left(x^k\right)\right\| = o\left(\frac{1}{k}\right), \quad \left\| V\left(\bar{x}^k\right)\right\| = o\left(\frac{1}{k}\right) \ as \ k \to +\infty.$$

  and
  $$0 \leq \mathtt{Gap}(x^k) = o\left(\frac{1}{k}\right) \ as \ k \to +\infty,$$

Convergence of the iterates and fast convergence rates

Theorem

Let $x^* \in \operatorname{Zer} V$ *and* $\left(x^k\right)_{k \geq 0}$ *the sequence generated by the Explicit Fast OGDA method. Then it holds:*

- *the sequence* $\left(x^k\right)_{k \geq 0}$ *converges weakly to a zero of* V*;*
-

$$\left\|x^k - x^{k-1}\right\| = o\left(\frac{1}{k}\right), \quad \left\langle x^k - x^*, V\left(x^k\right)\right\rangle = o\left(\frac{1}{k}\right),$$

$$\left\|V\left(x^k\right)\right\| = o\left(\frac{1}{k}\right), \quad \left\|V\left(\bar{x}^k\right)\right\| = o\left(\frac{1}{k}\right) \text{ as } k \to +\infty.$$

and

$$0 \leq \mathtt{Gap}(x^k) = o\left(\frac{1}{k}\right) \text{ as } k \to +\infty,$$

RIB, E.R. Csetnek, D.-K. Nguyen, *Fast Optimistic Gradient Descent Ascent (OGDA) method in continuous and discrete time*, Foundations of Computational Mathematics 25(1), 162–222, 2025

RIB, D.-K. Nguyen, *Fast Krasnosel'skiĭ-Mann algorithm with a convergence rate of the fixed point iteration of* $o\left(1/k\right)$, SIAM Journal on Numerical Analysis 61(6), 2813–2843, 2023

Numerical experiments

- We consider a minmax problem studied in (Ouyang-Xu, 2021)

$$\min_{u\in\mathbb{R}^n}\max_{v\in\mathbb{R}^n}\mathcal{L}(u,v) := \frac{1}{2}\langle u, Hu\rangle - \langle u, h\rangle - \langle v, Au-b\rangle,$$

where

$$A := \frac{1}{4}\begin{pmatrix} & & & -1 & 1 \\ & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ & -1 & 1 & & \\ -1 & 1 & & & \\ 1 & & & & \end{pmatrix} \in \mathbb{R}^{n\times n},$$

$$H := 2A^TA, \quad b := \frac{1}{4}\begin{pmatrix}1\\1\\\vdots\\1\\1\end{pmatrix} \in \mathbb{R}^n \text{ and } h := \frac{1}{4}\begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix} \in \mathbb{R}^n.$$

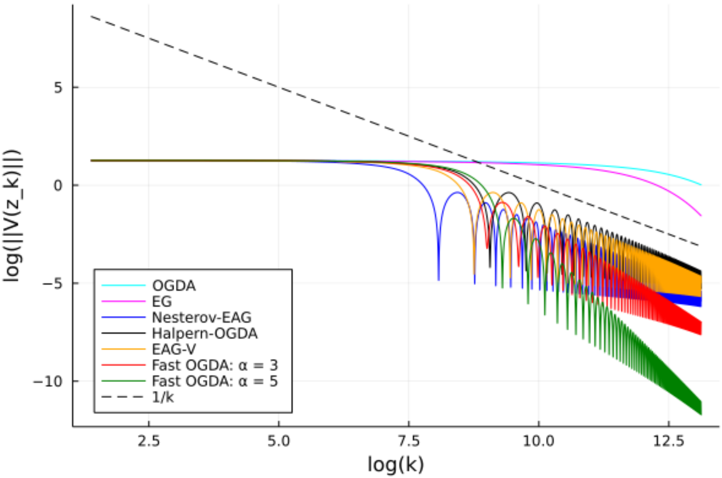
- Since $\|A\| \le \frac{1}{2}$, one can take for the monotone mapping

$$V(u,v) = \begin{pmatrix}\nabla_u\mathcal{L}(u,v)\\ -\nabla_v\mathcal{L}(u,v)\end{pmatrix}$$

as Lipschitz constant $L = 1$.

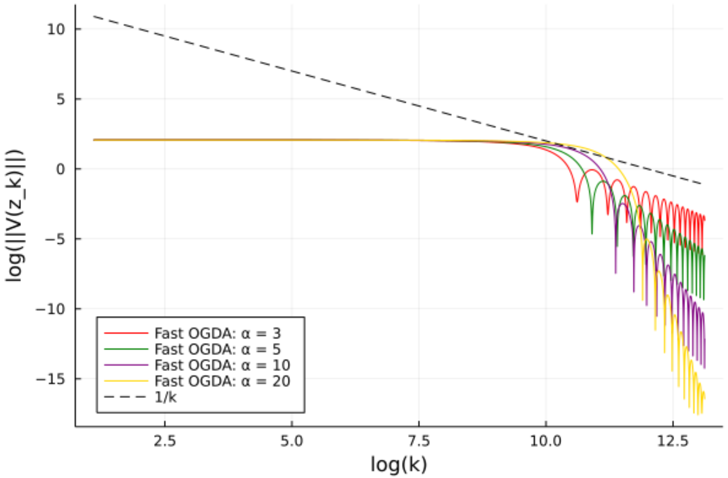
- The following numerical algorithms are used in the numerical experiments:
  1. Optimistic Gradient Descent Ascent (OGDA) method (Popov, 1980) with $s := \frac{0.48}{L}$;
  2. Extragradient (EG) method (Korpelevich, 1976; Antipin, 1976)) with $s := \frac{0.96}{L}$;
  3. Extra Anchored Gradient (EAG-V) method (Yoon-Ryu, 2021) with variable step sizes $(s_k)_{k\geq 0}$;
  4. Nesterov's accelerated variant (Nesterov-EAG) by (Tran-Dinh, 2022) of the Extra Anchored Gradient method;
  5. OGDA combined with Halpern anchoring scheme (Halpern-OGDA) by (Tran-Dinh-Luo, 2021);
  6. our explicit algorithm (Fast OGDA) with $s := \frac{0.24}{L}$ and various choices for α.

$n = 200$



log(||V(z_k)||)
log(k)
5
0
−5
−10
2.5
5.0
7.5
10.0
12.5
OGDA
EG
Nesterov-EAG
Halpern-OGDA
EAG-V
Fast OGDA: α = 3
Fast OGDA: α = 5
1/k

$n = 1000$: the influence of the parameter α



First order dynamical system with Tikhonov regularization

- Consider again the monotone equation

$$V(x) = 0,$$

where

- $V : \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator with $\operatorname{Zer} V$ nonempty.

First order dynamical system with Tikhonov regularization

- Consider again the monotone equation

$$V(x) = 0,$$

where

  - $V: \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator with $\operatorname{Zer} V$ nonempty.

- For $t_0 \geq 0$ and $\varepsilon: [t_0, +\infty) \to (0, +\infty)$, consider on $[t_0, +\infty)$ the Tikhonov regularized dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t) x(t) = 0, \\ x(t_0) = x^0. \end{cases}$$

First order dynamical system with Tikhonov regularization

- Consider again the monotone equation

$$V(x) = 0,$$

where

  - $V : \mathcal{H} \to \mathcal{H}$ is a continuous and monotone operator with $\operatorname{Zer} V$ nonempty.

- For $t_0 \geq 0$ and $\varepsilon : [t_0, +\infty) \to (0, +\infty)$, consider on $[t_0, +\infty)$ the Tikhonov regularized dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t) x(t) = 0, \\ x(t_0) = x^0. \end{cases}$$

  - The regularization function $\varepsilon : [t_0, +\infty) \to (0, +\infty)$ is required to be continuously differentiable and to satisfy

$$\lim_{t \to +\infty} \varepsilon(t) = 0 \quad \text{and} \quad \int_{t_0}^{+\infty} \varepsilon(t)\, dt = +\infty.$$

  - The dynamical system is assumed to admit a unique strong global solution $x : [t_0, +\infty) \to \mathcal{H}$ with the property that $t \mapsto V(x(t))$ is absolutely continuous on every compact interval.

A discussion of the underlying assumptions

- If $\varepsilon(t) \equiv \varepsilon_0 > 0$, then the trajectory $x(t)$ exhibits strong convergence towards the unique zero of the strongly monotone operator $V + \varepsilon_0 \operatorname{Id}$.
- If $V \equiv 0$, then the unique solution
$$x(t) = x^0 \exp\left(-\int_{t_0}^{t} \varepsilon(u)\, du\right)$$
of the system converges to 0 as $t \to +\infty$ if and only if $\int_{t_0}^{+\infty} \varepsilon(t) dt = +\infty$.
- If $V : \mathcal{H} \to \mathcal{H}$ is Lipschitz continuous, then the system has for every $x_0 \in \mathcal{H}$ a unique strong global solution $x : [t_0, +\infty) \to \mathcal{H}$. In this case, $t \mapsto V(x(t))$ is absolutely continuous on every compact interval and therefore almost everywhere differentiable on $[t_0, +\infty)$.

- We define
$$\gamma_\varepsilon : [t_0, +\infty) \to \mathbb{R}_{++}, \quad \gamma_\varepsilon(t) := \exp\left(\int_{t_0}^{t} \varepsilon(u)\, du\right).$$
- For every $t \geq t_0$ it holds $\dot{\gamma}_\varepsilon(t) = \varepsilon(t) \gamma_\varepsilon(t)$.
- In addition, γ_ε is strictly increasing, thus $\gamma_\varepsilon(t) \geq \gamma_\varepsilon(t_0) = 1$ for every $t \geq t_0$, and $\lim_{t \to +\infty} \gamma_\varepsilon(t) = +\infty$.
- For every $t \geq t_0$ it holds $\ddot{\gamma}_\varepsilon(t) = \left[\dot{\varepsilon}(t) + \varepsilon^2(t)\right] \gamma_\varepsilon(t)$.

▸ Given a trajectory solution $x\colon [t_0, +\infty) \to \mathcal{H}$ and $x^* \in \operatorname{Zer} V$, we consider the energy functions

$$\varphi\colon [t_0, +\infty) \to [0, +\infty) \quad \varphi(t) := \frac{1}{2}\left\|x(t) - x^*\right\|^2,$$

and

$$\psi\colon [t_0, +\infty) \to [0, +\infty), \quad \psi(t) := \frac{1}{2}\left\|V(x(t)) + \varepsilon(t)\, x(t)\right\|^2.$$

- Given a trajectory solution $x\colon [t_0, +\infty) \to \mathcal{H}$ and $x^* \in \operatorname{Zer} V$, we consider the energy functions

$$\varphi\colon [t_0, +\infty) \to [0, +\infty) \quad \varphi(t) := \frac{1}{2} \left\| x(t) - x^* \right\|^2,$$

and

$$\psi\colon [t_0, +\infty) \to [0, +\infty), \quad \psi(t) := \frac{1}{2} \left\| V(x(t)) + \varepsilon(t) x(t) \right\|^2.$$

- For a triple $\left(x^0, v, x^*\right) \in \mathcal{H} \times \mathcal{H} \times \mathcal{H}$, we denote

$$D_0\left(x^0, v, x^*\right) := \max\left\{\left\|x^0 - x^*\right\|, \left\|v - x^*\right\|\right\}.$$

- Given a trajectory solution $x\colon [t_0, +\infty) \to \mathcal{H}$ and $x^* \in \operatorname{Zer} V$, we consider the energy functions

$$\varphi\colon [t_0, +\infty) \to [0, +\infty) \quad \varphi(t) := \frac{1}{2} \left\| x(t) - x^* \right\|^2,$$

and

$$\psi\colon [t_0, +\infty) \to [0, +\infty), \quad \psi(t) := \frac{1}{2} \left\| V(x(t)) + \varepsilon(t) x(t) \right\|^2.$$

- For a triple $\left(x^0, v, x^*\right) \in \mathcal{H} \times \mathcal{H} \times \mathcal{H}$, we denote

$$D_0\left(x^0, v, x^*\right) := \max\left\{\left\|x^0 - x^*\right\|, \left\|v - x^*\right\|\right\}.$$

- The trajectory is bounded; for every $t \geq t_0$ it holds

$$\|x(t)\| \leq 2D_0\left(x^0, 0, x^*\right).$$

Proposition

If $\|V(x(t))\| \to 0$ as $t \to +\infty$, then $x(t)$ converges strongly to the minimum norm solution $x^* := \text{proj}_{\text{Zer}\, V}(0)$ as $t \to +\infty$.

Proposition

If $\|V(x(t))\| \to 0$ as $t \to +\infty$, then $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Choosing $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$, it holds for almost every $t \geq t_0$

$$\dot{\varphi}(t) + \varepsilon(t)\varphi(t) \leq \varepsilon(t)\left\langle x(t) - x^*, 0 - x^* \right\rangle .$$

Proposition

If $\|V(x(t))\| \to 0$ as $t \to +\infty$, then $x(t)$ converges strongly to the minimum norm solution $x^* := \text{proj}_{\text{Zer } V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Choosing $x^* := \text{proj}_{\text{Zer } V}(0)$, it holds for almost every $t \geq t_0$

$$\dot{\varphi}(t) + \varepsilon(t)\varphi(t) \leq \varepsilon(t)\left\langle x(t) - x^*, 0 - x^* \right\rangle.$$

Lemma (Gronwall's inequality in continuous time)

Let $h\colon [t_0, +\infty) \to \mathbb{R}$ be a locally absolutely continuous function, $g\colon [t_0, +\infty) \to \mathbb{R}$ a bounded function and $\varepsilon\colon [t_0, +\infty) \to [0, +\infty)$ a locally integrable function such that

$$\dot{h}(t) + \varepsilon(t) h(t) \leq \varepsilon(t) g(t) \quad \text{for almost every } t \geq t_0.$$

Then h is also bounded and, if $\int_{t_0}^{+\infty} \varepsilon(t)\, dt = +\infty$, then $\limsup_{t\to+\infty} h(t) \leq \limsup_{t\to+\infty} g(t)$.

Proposition

If $\|V(x(t))\| \to 0$ as $t \to +\infty$, then $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Choosing $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$, it holds for almost every $t \geq t_0$

$$\dot{\varphi}(t) + \varepsilon(t)\varphi(t) \leq \varepsilon(t)\left\langle x(t) - x^*, 0 - x^* \right\rangle .$$

Lemma (Gronwall's inequality in continuous time)

Let $h: [t_0, +\infty) \to \mathbb{R}$ be a locally absolutely continuous function, $g: [t_0, +\infty) \to \mathbb{R}$ a bounded function and $\varepsilon: [t_0, +\infty) \to [0, +\infty)$ a locally integrable function such that

$$\dot{h}(t) + \varepsilon(t) h(t) \leq \varepsilon(t) g(t) \quad \text{for almost every } t \geq t_0.$$

Then h is also bounded and, if $\int_{t_0}^{+\infty} \varepsilon(t)\, dt = +\infty$, then $\limsup_{t\to+\infty} h(t) \leq \limsup_{t\to+\infty} g(t)$.

- It holds

$$\limsup_{t\to+\infty} \varphi(t) \leq \limsup_{t\to+\infty} \left\langle x(t) - x_*, 0 - x_* \right\rangle .$$

The assumption $\|V(x(t))\| \to 0$ as $t \to +\infty$ yields that every weak cluster point of $x(\cdot)$ belongs to $\operatorname{Zer} V$. Thus, $\limsup_{t\to+\infty} \varphi(t) \leq 0$, which is nothing else than $x(t) \to x_*$ as $t \to +\infty$. □

Theorem (strong convergence of the trajectory)

If

$$\lim_{t \to +\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} = 0 \quad \text{or} \quad \int_{t_0}^{+\infty} |\dot{\varepsilon}(t)| \, dt < +\infty,$$

then $\|V(x(t))\| \to 0$ as $t \to +\infty$, and therefore the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Theorem (strong convergence of the trajectory)

If
$$\lim_{t \to +\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} = 0 \quad \text{or} \quad \int_{t_0}^{+\infty} |\dot{\varepsilon}(t)| \, dt < +\infty,$$

then $\|V(x(t))\| \to 0$ as $t \to +\infty$, and therefore the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Each of the two conditions imply $\psi(t) \to 0$, and further $\|V(x(t))\| \to 0$ as $t \to +\infty$.

Theorem (strong convergence of the trajectory)

If

$$\lim_{t \to +\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} = 0 \quad \text{or} \quad \int_{t_0}^{+\infty} |\dot{\varepsilon}(t)| \, dt < +\infty,$$

then $\|V(x(t))\| \to 0$ as $t \to +\infty$, and therefore the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Each of the two conditions imply $\psi(t) \to 0$, and further $\|V(x(t))\| \to 0$ as $t \to +\infty$.
- In the first scenario, it holds for almost every $t \geq t_0$

$$\dot{\psi}(t) + \varepsilon(t)\psi(t) \leq \varepsilon(t) \left(2D_0^2 \left(x^0, 0, x^* \right) \frac{|\dot{\varepsilon}(t)|^2}{\varepsilon(t)^2} \right).$$

Gronwall's inequality in continuous time yields $\lim_{t \to +\infty} \psi(t) = 0$.

Theorem (strong convergence of the trajectory)

If

$$\lim_{t\to+\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} = 0 \quad \text{or} \quad \int_{t_0}^{+\infty} |\dot{\varepsilon}(t)|\, dt < +\infty,$$

then $\|V(x(t))\| \to 0$ as $t \to +\infty$, and therefore the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Sketch of a proof.

- Each of the two conditions imply $\psi(t) \to 0$, and further $\|V(x(t))\| \to 0$ as $t \to +\infty$.
- In the first scenario, it holds for almost every $t \geq t_0$

$$\dot{\psi}(t) + \varepsilon(t)\psi(t) \leq \varepsilon(t)\left(2D_0^2\left(x^0, 0, x^*\right)\frac{|\dot{\varepsilon}(t)|^2}{\varepsilon(t)^2}\right).$$

Gronwall's inequality in continuous time yields $\lim_{t\to+\infty} \psi(t) = 0$.

- In the second scenario, for $s \geq t_0$ and every $t \geq s$, it holds

$$\gamma_\varepsilon^2(t)\psi(t) \leq \gamma_\varepsilon^2(s)\psi(s) + 2D_0^2\left(x^0, 0, x^*\right)\int_s^t |\dot{\varepsilon}(u)|\gamma_\varepsilon^2(u)du + \int_s^t |\dot{\varepsilon}(u)|\gamma_\varepsilon^2(u)\psi(u)du.$$

According to Gronwall's Lemma, we obtain for every $t \geq s$

$$\gamma_\varepsilon^2(t)\psi(t) \leq \left(\gamma_\varepsilon^2(s)\psi(s) + 2D_0^2\left(x^0, 0, x^*\right)\int_s^t |\dot{\varepsilon}(u)|\gamma_\varepsilon^2(u)du\right)\exp\left(\int_s^t |\dot{\varepsilon}(u)|\, du\right).$$

From here it yields for every $s \geq t_0$

$$\limsup_{t\to+\infty} \psi(t) \leq 2D_0^2\left(x^0, 0, x^*\right)\left(\int_s^{+\infty} |\dot{\varepsilon}(u)|\, du\right)\exp\left(\int_s^{+\infty} |\dot{\varepsilon}(u)|\, du\right).$$

Finally, letting $s \to +\infty$, we obtain $\lim_{t\to+\infty} \psi(t) = 0$. □

Remarks

- The condition $\int_{t_0}^{+\infty} |\dot{\varepsilon}(t)| \, dt < +\infty$ was introduced in (Cominetti, Peypouquet, Sorin, 2008).

Remarks

- The condition $\int_{t_0}^{+\infty} |\dot{\varepsilon}(t)| \, dt < +\infty$ was introduced in (Cominetti, Peypouquet, Sorin, 2008).
- The condition $\lim_{t \to +\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} = 0$ represents an enhancement of the condition (Israel, Reich, 1981; Cominetti, Peypouquet, Sorin, 2008)

$$t \mapsto \varepsilon(t) \text{ is decreasing} \quad \text{and} \quad \lim_{t \to +\infty} \frac{|\dot{\varepsilon}(t)|}{\varepsilon^2(t)} = 0.$$

For $\varepsilon(t) = \frac{\alpha}{t^q}$, where $0 < q \leq 1$ and $\alpha > 0$, it allows for the critical choice $q := 1$.

- We denote

$$\varepsilon_i := \inf_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \inf_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)} \quad \text{and} \quad \varepsilon_s := \sup_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \sup_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)}.$$

- We denote

$$\varepsilon_i := \inf_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \inf_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)} \quad \text{and} \quad \varepsilon_s := \sup_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \sup_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)}.$$

- Recall that

$$\gamma_\varepsilon(t) := \exp\left(\int_{t_0}^{t} \varepsilon(u)\, du\right).$$

- We denote

$$\varepsilon_i := \inf_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \inf_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)} \quad \text{and} \quad \varepsilon_s := \sup_{t \geq t_0} \frac{d}{dt}\left(\frac{1}{\varepsilon(t)}\right) = \sup_{t \geq t_0} \frac{-\dot{\varepsilon}(t)}{\varepsilon^2(t)}.$$

- Recall that

$$\gamma_\varepsilon(t) := \exp\left(\int_{t_0}^{t} \varepsilon(u)\, du\right).$$

Theorem (convegence rates)

Let $x\colon [t_0, +\infty) \to \mathcal{H}$ be the trajectory solution of the first-order dynamical system with Tikhonov regularization. Suppose that the regularization function $\varepsilon(\cdot)$ also satisfies

$$1 < \varepsilon_i \quad \text{or} \quad -1 < \varepsilon_i \leq \varepsilon_s < 1 \quad \text{or} \quad \varepsilon_i = \varepsilon_s = 1.$$

Then, the following statements are true:

1. as $t \to +\infty$ it holds

$$\|\dot{x}(t)\| = \begin{cases} \mathcal{O}\left(\frac{1}{\gamma_\varepsilon(t)}\right) & \text{if } 1 < \varepsilon_i \text{ or } \varepsilon_i = \varepsilon_s = 1, \\ \mathcal{O}(\varepsilon(t)) & \text{if } -1 < \varepsilon_i \leq \varepsilon_s < 1, \end{cases}$$

$$\|V(x(t))\| = \begin{cases} \mathcal{O}\left(\frac{1}{\gamma_\varepsilon(t)}\right) & \text{if } 1 < \varepsilon_i \text{ or } \varepsilon_i = \varepsilon_s = 1, \\ \mathcal{O}(\varepsilon(t)) & \text{if } -1 < \varepsilon_i \leq \varepsilon_s < 1; \end{cases}$$

2. the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \text{proj}_{\text{Zer}\, V}(0)$ as $t \to +\infty$.

Theorem (convergence rates in the case $\varepsilon(t) := \frac{\alpha}{t}$)

Let $t_0 > 0$, $\alpha > 0$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ be the trajectory solution of the dynamical system

$$\dot{x}(t) + V(x(t)) + \frac{\alpha}{t} x(t) = 0,$$

with initial condition $x(t_0) := x^0 \in \mathcal{H}$. Then, the following statements are true:

1. as $t \to +\infty$ it holds

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^{\min\{\alpha,1\}}}\right) \quad \text{and} \quad \|V(x(t))\| = \mathcal{O}\left(\frac{1}{t^{\min\{\alpha,1\}}}\right);$$

2. the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \mathrm{proj}_{\mathrm{Zer}\, V}(0)$ as $t \to +\infty$.

Theorem (convergence rates in the case $\varepsilon(t) := \frac{\alpha}{t}$)

Let $t_0 > 0$, $\alpha > 0$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ be the trajectory solution of the dynamical system

$$\dot{x}(t) + V(x(t)) + \frac{\alpha}{t} x(t) = 0,$$

with initial condition $x(t_0) := x^0 \in \mathcal{H}$. Then, the following statements are true:

1. as $t \to +\infty$ it holds

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^{\min\{\alpha,1\}}}\right) \quad \text{and} \quad \|V(x(t))\| = \mathcal{O}\left(\frac{1}{t^{\min\{\alpha,1\}}}\right);$$

2. the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Theorem (the case $\varepsilon(t) := \frac{\alpha}{t^q}, 0 < q < 1$)

Let $t_0 > 0, 0 < q < 1$, $\alpha > 0$ such that $\alpha t_0^{1-q} \geq 2q$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ be the trajectory solution of the dynamical system

$$\dot{x}(t) + V(x(t)) + \frac{\alpha}{t^q} x(t) = 0,$$

with initial condition $x(t_0) := x^0 \in \mathcal{H}$. Then, the following statements are true:

1. as $t \to +\infty$ it holds

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right) \quad \text{and} \quad \|V(x(t))\| = \mathcal{O}\left(\frac{1}{t^q}\right);$$

2. the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

Theorem (the case $\varepsilon(t) := \frac{1}{t\log(t)}$)

Let $t_0 > 1$ and $x : [t_0, +\infty) \to \mathcal{H}$ be the trajectory solution of the dynamical system

$$\dot{x}(t) + V(x(t)) + \frac{1}{t\log(t)} x(t) = 0,$$

with initial condition $x(t_0) := x^0 \in \mathcal{H}$. Then, the following statements are true:

1. as $t \to +\infty$ it holds

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{\log(t)}\right) \quad \text{and} \quad \|V(x(t))\| = \mathcal{O}\left(\frac{1}{\log(t)}\right).$$

2. the trajectory $x(t)$ converges strongly to the minimum norm solution $x^* := \operatorname{proj}_{\operatorname{Zer} V}(0)$ as $t \to +\infty$.

First order dynamical system with anchor point

- For $t_0 > 0$ and $v \in \mathcal{H}$ an anchor point, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t)(x(t) - v) = 0, \\ x(t_0) = x^0. \end{cases}$$

First order dynamical system with anchor point

- For $t_0 > 0$ and $v \in \mathcal{H}$ an anchor point, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t)(x(t) - v) = 0, \\ x(t_0) = x^0. \end{cases}$$

- $x(\cdot)$ is a trajectory solution if and only if

$$z: [t_0, +\infty) \to \mathcal{H}, \quad z(t) := x(t) - v \quad \forall t \geq t_0,$$

is the trajectory solution of

$$\begin{cases} \dot{z}(t) + V^v(z(t)) + \varepsilon(t) z(t) = 0, \\ z(t_0) = x^0 - v, \end{cases}$$

where

$$V^v: \mathcal{H} \to \mathcal{H}, \quad V^v(x) = V(x + v) \quad \forall x \in \mathcal{H}.$$

First order dynamical system with anchor point

- For $t_0 > 0$ and $v \in \mathcal{H}$ an anchor point, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t)(x(t) - v) = 0, \\ x(t_0) = x^0. \end{cases}$$

- $x(\cdot)$ is a trajectory solution if and only if

$$z\colon [t_0, +\infty) \to \mathcal{H}, \quad z(t) := x(t) - v \quad \forall t \geq t_0,$$

is the trajectory solution of

$$\begin{cases} \dot{z}(t) + V^v(z(t)) + \varepsilon(t) z(t) = 0, \\ z(t_0) = x^0 - v, \end{cases}$$

where

$$V^v\colon \mathcal{H} \to \mathcal{H}, \quad V^v(x) = V(x + v) \quad \forall x \in \mathcal{H}.$$

- The rates of convergence for the velocity and the operator norm follow from the results provided for the first order system with Tikhonov regularization, since $\dot{z}(t) = \dot{x}(t)$ and $V^v(z(t)) = V(x(t))$ for every $s \geq s_0$.

First order dynamical system with anchor point

- For $t_0 > 0$ and $v \in \mathcal{H}$ an anchor point, consider on $[t_0, +\infty)$ the dynamical system

$$\begin{cases} \dot{x}(t) + V(x(t)) + \varepsilon(t)(x(t) - v) = 0, \\ x(t_0) = x^0. \end{cases}$$

- $x(\cdot)$ is a trajectory solution if and only if

$$z\colon [t_0, +\infty) \to \mathcal{H}, \quad z(t) := x(t) - v \quad \forall t \geq t_0,$$

is the trajectory solution of

$$\begin{cases} \dot{z}(t) + V^v(z(t)) + \varepsilon(t) z(t) = 0, \\ z(t_0) = x^0 - v, \end{cases}$$

where

$$V^v\colon \mathcal{H} \to \mathcal{H}, \quad V^v(x) = V(x + v) \quad \forall x \in \mathcal{H}.$$

- The rates of convergence for the velocity and the operator norm follow from the results provided for the first order system with Tikhonov regularization, since $\dot{z}(t) = \dot{x}(t)$ and $V^v(z(t)) = V(x(t))$ for every $s \geq s_0$.

- The strong convergence of the trajectory $x(t) = z(t) + v$ to $\mathrm{proj}_{\mathrm{Zer}\, V}(v) = \mathrm{proj}_{\mathrm{Zer}\, V^v}(0) + v$ as $s \to +\infty$ follows also from the results provided for the first order system with Tikhonov regularization and time rescaling.

From first order dynamics with Tikhonov regularization to second order dynamics with vanishing damping term

- For $t_0 > 0$ and $\alpha > 1$, consider the second-order dynamical system

$$\begin{cases} \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0, \\ x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0. \end{cases}$$

From first order dynamics with Tikhonov regularization to second order dynamics with vanishing damping term

- For $t_0 > 0$ and $\alpha > 1$, consider the second-order dynamical system

$$\begin{cases} \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0, \\ x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0. \end{cases}$$

Proposition

Let $t_0 > 0$ and $\alpha > 1$. Then $x\colon [t_0, +\infty) \to \mathcal{H}$ is the trajectory solution of

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0$$

with initial conditions $x(t_0) = x^0 \text{ and } \dot{x}(t_0) = \dot{x}^0$ if and only if it is the trajectory solution of

$$\dot{x}(t) + V(x(t)) + \frac{\alpha - 1}{t}(x(t) - v) = 0,$$

with initial condition $x(t_0) = x^0$ and anchor point

$$v := x^0 + \frac{t_0}{\alpha - 1}\left(\dot{x}^0 + (t_0)V\left(x^0\right)\right).$$

Second-order dynamics with vanishing damping term: strong convergence and convergence rates

Theorem

Let $t_0 > 0$, $\alpha > 1$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ the trajectory solution of the second-order dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0$$

with initial conditions $x(t_0) = x^0$ and $\dot{x}(t_0) = \dot{x}^0$. Then, $x(t)$ converges strongly to $\operatorname{proj}_{\operatorname{Zer} V}\left(x^0 + \frac{t_0}{\alpha - 1}\left(\dot{x}^0 + V\left(x^0\right)\right)\right)$ as $t \to +\infty$, $\|\dot{x}(t)\| = O\left(\frac{1}{t^{\min\{\alpha-1,1\}}}\right)$, and $\|V(x(t))\| = O\left(\frac{1}{t^{\min\{\alpha-1,1\}}}\right)$ as $t \to +\infty$.

Second-order dynamics with vanishing damping term: strong convergence and convergence rates

Theorem

Let $t_0 > 0$, $\alpha > 1$, and $x\colon [t_0, +\infty) \to \mathcal{H}$ the trajectory solution of the second-order dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0$$

with initial conditions $x(t_0) = x^0$ and $\dot{x}(t_0) = \dot{x}^0$. Then, $x(t)$ converges strongly to $\mathrm{proj}_{\mathrm{Zer}\, V}\left(x^0 + \frac{t_0}{\alpha - 1}\left(\dot{x}^0 + V\left(x^0\right)\right)\right)$ as $t \to +\infty$, $\|\dot{x}(t)\| = O\left(\frac{1}{t^{\min\{\alpha-1,1\}}}\right)$, and $\|V(x(t))\| = O\left(\frac{1}{t^{\min\{\alpha-1,1\}}}\right)$ as $t \to +\infty$.

- If $\alpha \geq 2$, then $x(t)$ converges strongly to $\mathrm{proj}_{\mathrm{Zer}\, V}\left(x^0 + \frac{t_0}{\alpha - 1}\left(\dot{x}^0 + V\left(x^0\right)\right)\right)$, $\|\dot{x}(t)\| = O\left(\frac{1}{t}\right)$, and $\|V(x(t))\| = O\left(\frac{1}{t}\right)$ as $t \to +\infty$.

Remark

In case $\alpha = 2$, the Fast OGDA dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{\alpha}{2t}V(x(t)) = 0$$

is nothing else than

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0,$$

thus it inherits its asymptotic properties.

Remark

In case $\alpha = 2$, the Fast OGDA dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{\alpha}{2t}V(x(t)) = 0$$

is nothing else than

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \frac{d}{dt}V(x(t)) + \frac{1}{t}V(x(t)) = 0,$$

thus it inherits its asymptotic properties.

RIB, D.-K. Nguyen, *Tikhonov regularization of monotone operator flows not only ensures strong convergence of the trajectories but also speeds up the vanishing of the residuals*, Journal of Complexity 94, Article 102029, 2026

Generalized Extra-Anchored-Gradient method (G-EAG)

Let $x^0 \in \mathcal{H}$, $v \in \mathcal{H}$, $\theta \in (0, \frac{1}{L})$, and $(\varepsilon_k)_{k\geq 0}$ a positive sequence of parameters satisfying

$$\lim_{k\to+\infty} \varepsilon_k = 0\,, \quad \sum_{k=0}^{+\infty} \varepsilon_k = +\infty\,, \quad \text{and} \quad \lim_{k\to+\infty} \frac{|\varepsilon_{k+1} - \varepsilon_k|}{\varepsilon_k} = 0.$$

Consider the explicit iterative scheme

$$(\forall k \geq 0) \quad \begin{cases} y^{k+1} & := x^k - \theta\varepsilon_k(x^k - v) - \theta V(x^k) \\ x^{k+1} & := x^k - \theta\varepsilon_{k+1}(x^{k+1} - v) - \theta V(y^{k+1}). \end{cases}$$

Generalized Extra-Anchored-Gradient method

Generalized Extra-Anchored-Gradient method (G-EAG)

Let $x^0 \in \mathcal{H}$, $v \in \mathcal{H}$, $\theta \in (0, \frac{1}{L})$, and $(\varepsilon_k)_{k \geq 0}$ a positive sequence of parameters satisfying

$$\lim_{k \to +\infty} \varepsilon_k = 0\,, \quad \sum_{k=0}^{+\infty} \varepsilon_k = +\infty\,, \quad \text{and} \quad \lim_{k \to +\infty} \frac{|\varepsilon_{k+1} - \varepsilon_k|}{\varepsilon_k} = 0.$$

Consider the explicit iterative scheme

$$(\forall k \geq 0) \quad \begin{cases} y^{k+1} & := x^k - \theta \varepsilon_k (x^k - v) - \theta V(x^k) \\ x^{k+1} & := x^k - \theta \varepsilon_{k+1} (x^{k+1} - v) - \theta V(y^{k+1}). \end{cases}$$

Remark

Without loss of generality, we can assume $v = 0$. The second identity in (G-EAG) can be written as

$$\frac{x^{k+1} - x^k}{\theta} + V(y^{k+1}) + \varepsilon_{k+1} x^{k+1} = 0 \quad \text{for all } k \geq 0\,, \tag{1}$$

which shows that it can be understood as an explicit discretization of the monotone flow with Tikhonov regularization, y^{k+1} being an extrapolation point. If $\varepsilon_k = 0$ for all $k \geq 0$, the method reduces to EG.

Generalized Extra-Anchored-Gradient method: convergence analysis

- Given $x^* \in \operatorname{Zer} V$, we consider the discrete energy functions

$$\varphi^k := \frac{1}{2}\left\|x^k - x^*\right\|^2 \quad \text{for all } k \geq 0,$$

and

$$\psi^k := \frac{1}{2}\left\|V(x^k) + \varepsilon_k x^k\right\|^2 \quad \text{for all } k \geq 0.$$

Generalized Extra-Anchored-Gradient method: convergence analysis

- Given $x^* \in \operatorname{Zer} V$, we consider the discrete energy functions

$$\varphi^k := \frac{1}{2}\left\|x^k - x^*\right\|^2 \quad \text{for all } k \geq 0,$$

and

$$\psi^k := \frac{1}{2}\|V(x^k) + \varepsilon_k x^k\|^2 \quad \text{for all } k \geq 0.$$

Theorem (Gronwall's inequality in discrete time; Xu, 2002)

Let $(f^k)_k$ and $(h^k)_k$ be two sequences in $\mathbb{R}$ and $(\varepsilon_k)_k$ be a nonnegative sequence satisfying

$$f^{k+1} - f^k + \varepsilon_k f^k \leq \varepsilon_k h^k \quad \text{for all } k \geq 0\,. \tag{2}$$

Suppose further that $\varepsilon_k < 1$ for all $k \geq 0$. Then it holds

$$f^{k+1} \leq \frac{f^0}{\gamma_k} + \frac{1}{\gamma_k}\sum_{j=0}^{k}\gamma_j\varepsilon_j h^j \quad \text{for all } k \geq 0\,,$$

where $\gamma_k := \prod_{\ell=0}^{k}(1-\varepsilon_\ell)^{-1}$. In particular, if $\sum_{k=0}^{+\infty}\varepsilon_k = +\infty$, it holds

$$\limsup_{k\to+\infty} f^k \leq \limsup_{k\to+\infty} h^k\,.$$

Generalized Extra-Anchored-Gradient method: convergence analysis

- There exist a vanishing real sequence $(\eta_k)_k$ and $k_1 \geq 0$, such that

$$\varphi^{k+1} - \varphi^k + \frac{\theta\varepsilon_{k+1}}{2}\varphi^k \leq \frac{\theta\varepsilon_{k+1}}{2}\left(\|x^*\|^2 - \frac{1-\eta_k}{1+\theta\varepsilon_{k+1}}\|x^k\|^2\right) \quad \text{for all } k \geq k_1 \,.$$

Generalized Extra-Anchored-Gradient method: convergence analysis

- There exist a vanishing real sequence $(\eta_k)_k$ and $k_1 \geq 0$, such that

$$\varphi^{k+1} - \varphi^k + \frac{\theta\varepsilon_{k+1}}{2}\varphi^k \leq \frac{\theta\varepsilon_{k+1}}{2}\left(\|x^*\|^2 - \frac{1-\eta_k}{1+\theta\varepsilon_{k+1}}\|x^k\|^2\right) \quad \text{for all } k \geq k_1 .$$

- Assume that $(x^k)_{k\geq 0}$ is bounded. Then, for any $\delta \in (0,1)$, there exist $C_\delta \geq 0$ and $k_0 \geq 0$ such that

$$\psi^{k+1} - \psi^k + 2(1-\delta)\theta\varepsilon_k\psi^k \leq C_\delta\theta\varepsilon_k\left(\frac{\varepsilon_{k+1}-\varepsilon_k}{\theta\varepsilon_k}\right)^2 \quad \text{for all } k \geq k_0 .$$

Generalized Extra-Anchored-Gradient method: convergence analysis

- There exist a vanishing real sequence $(\eta_k)_k$ and $k_1 \geq 0$, such that

$$\varphi^{k+1} - \varphi^k + \frac{\theta\varepsilon_{k+1}}{2}\varphi^k \leq \frac{\theta\varepsilon_{k+1}}{2}\left(\|x^*\|^2 - \frac{1-\eta_k}{1+\theta\varepsilon_{k+1}}\|x^k\|^2\right) \quad \text{for all } k \geq k_1 .$$

- Assume that $(x^k)_{k\geq 0}$ is bounded. Then, for any $\delta \in (0,1)$, there exist $C_\delta \geq 0$ and $k_0 \geq 0$ such that

$$\psi^{k+1} - \psi^k + 2(1-\delta)\theta\varepsilon_k\psi^k \leq C_\delta\theta\varepsilon_k\left(\frac{\varepsilon_{k+1}-\varepsilon_k}{\theta\varepsilon_k}\right)^2 \quad \text{for all } k \geq k_0 .$$

Theorem (strong convergence of the iterates and convergence rates for the residual)

The sequence $(x^k)_{k\geq 0}$ converges strongly to the minimum norm solution $\operatorname{proj}_{\operatorname{Zer} V}(0)$ as $k \to +\infty$, and the convergence rate of the residual results from the following statement: For any $\delta \in (0,1)$, there exist $\widetilde{C}_\delta \geq 0$ and $k_0 \geq 0$ such that

$$\psi^{k+1} \leq \frac{\psi^{k_0}}{\gamma_k} + \frac{\widetilde{C}_\delta}{\gamma_k}\sum_{j=k_0}^{k}\gamma_j\varepsilon_j\left(\frac{\varepsilon_{j+1}-\varepsilon_j}{\theta\varepsilon_j}\right)^2 \quad \text{for all } k \geq k_0 ,$$

where $(\gamma_k)_{k\geq k_0}$ is the sequence defined by $\gamma_k := \prod_{l=k_0}^{k}(1-2(1-\delta)\theta\varepsilon_l)^{-1}$.

Generalized Extra-Anchored-Gradient method: convergence analysis

- There exist a vanishing real sequence $(\eta_k)_k$ and $k_1 \geq 0$, such that

$$\varphi^{k+1} - \varphi^k + \frac{\theta\varepsilon_{k+1}}{2}\varphi^k \leq \frac{\theta\varepsilon_{k+1}}{2}\left(\|x^*\|^2 - \frac{1-\eta_k}{1+\theta\varepsilon_{k+1}}\|x^k\|^2\right) \quad \text{for all } k \geq k_1 .$$

- Assume that $(x^k)_{k\geq 0}$ is bounded. Then, for any $\delta \in (0,1)$, there exist $C_\delta \geq 0$ and $k_0 \geq 0$ such that

$$\psi^{k+1} - \psi^k + 2(1-\delta)\theta\varepsilon_k\psi^k \leq C_\delta\theta\varepsilon_k\left(\frac{\varepsilon_{k+1}-\varepsilon_k}{\theta\varepsilon_k}\right)^2 \quad \text{for all } k \geq k_0 .$$

Theorem (strong convergence of the iterates and convergence rates for the residual)

The sequence $(x^k)_{k\geq 0}$ converges strongly to the minimum norm solution $\mathrm{proj}_{\mathrm{Zer}\,V}(0)$ as $k \to +\infty$, and the convergence rate of the residual results from the following statement: For any $\delta \in (0,1)$, there exist $\widetilde{C}_\delta \geq 0$ and $k_0 \geq 0$ such that

$$\psi^{k+1} \leq \frac{\psi^{k_0}}{\gamma_k} + \frac{\widetilde{C}_\delta}{\gamma_k}\sum_{j=k_0}^{k}\gamma_j\varepsilon_j\left(\frac{\varepsilon_{j+1}-\varepsilon_j}{\theta\varepsilon_j}\right)^2 \quad \text{for all } k \geq k_0 ,$$

where $(\gamma_k)_{k\geq k_0}$ is the sequence defined by $\gamma_k := \prod_{l=k_0}^{k}(1-2(1-\delta)\theta\varepsilon_l)^{-1}$.

Remark

Since $(x^k)_{k\geq 0}$ converges strongly to the minimum norm solution $\mathrm{proj}_{\mathrm{Zer}\,V}(0)$ as $k \to +\infty$ and $V(x^k) \to 0$ as $k \to +\infty$, $(y^{k+1})_{k\geq 0}$ converges also strongly to the same point.

Generalized Extra-Anchored-Gradient method: particular case

- Let
$$\varepsilon_k := \frac{\alpha}{\theta(k+\beta)} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0.$$

Generalized Extra-Anchored-Gradient method: particular case

- Let

$$\varepsilon_k := \frac{\alpha}{\theta(k+\beta)} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0.$$

Lemma (Chung, 1954; Polyak, 1987)

Let $(\psi^k)_{k\geq 1}$ be a non-negative sequence and let $p, c, C > 0$. Suppose that

$$\psi^{k+1} \leq \left(1 - \frac{c}{k}\right)\psi^k + \frac{C}{k^{p+1}} \quad \text{for all } k \geq 1\,.$$

Then the following holds true as $k \to +\infty$

$$\psi^k \leq \frac{C}{(c-p)k^p} + o\left(\frac{1}{k^p}\right), \qquad \text{if } c > p\,;$$

$$\psi^k = O\left(\frac{\log(k)}{k^c}\right), \qquad \text{if } c = p\,;$$

$$\psi^k = O\left(\frac{1}{k^c}\right), \qquad \text{if } c < p\,.$$

Generalized Extra-Anchored-Gradient method: particular case

- Let

$$\varepsilon_k := \frac{\alpha}{\theta(k+\beta)} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0.$$

Lemma (Chung, 1954; Polyak, 1987)

Let $(\psi^k)_{k \geq 1}$ be a non-negative sequence and let $p, c, C > 0$. Suppose that

$$\psi^{k+1} \leq \left(1 - \frac{c}{k}\right)\psi^k + \frac{C}{k^{p+1}} \quad \text{for all } k \geq 1\,.$$

Then the following holds true as $k \to +\infty$

$$\psi^k \leq \frac{C}{(c-p)k^p} + o\left(\frac{1}{k^p}\right), \qquad \text{if } c > p\,;$$

$$\psi^k = O\left(\frac{\log(k)}{k^c}\right), \qquad \text{if } c = p\,;$$

$$\psi^k = O\left(\frac{1}{k^c}\right), \qquad \text{if } c < p\,.$$

Theorem (strong convergence of the iterates and convergence rates for the residual)

The sequence $(x^k)_{k \geq 0}$ converges strongly to the minimum norm solution $\mathrm{proj}_{\mathrm{Zer}\, V}(0)$ as $k \to +\infty$, the residual showing the following convergence rate:

1. if $\alpha > 1$, then $\|V(x^k)\| = O(\frac{1}{k})$ as $k \to +\infty$;
2. if $\alpha \leq 1$, then $\|V(x^k)\| = O\left(\frac{1}{k^{\alpha(1-\delta)}}\right)$, for any $\delta \in (0,1)$, as $k \to +\infty$.

Generalized Extra-Anchored-Gradient method: particular case

- Let
$$\varepsilon_k := \frac{\alpha}{(k+\beta)^\eta} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0, \ \eta \in (0,1).$$

Generalized Extra-Anchored-Gradient method: particular case

- Let

$$\varepsilon_k := \frac{\alpha}{(k+\beta)^\eta} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0, \ \eta \in (0,1).$$

Lemma (Chung, 1954; Polyak, 1987)

Let $(\psi^k)_{k\geq 1}$ be a non-negative sequence and let $c, C > 0$, and $0 < s < 1$, $s < t$. Suppose that

$$\psi^{k+1} \leq \left(1 - \frac{c}{k^s}\right)\psi^k + \frac{C}{k^t} \quad \text{for all } k \geq 1.$$

Then the following holds true as $k \to +\infty$

$$\psi^k \leq \frac{C}{c}\frac{1}{k^{t-s}} + o\left(\frac{1}{k^{t-s}}\right).$$

Generalized Extra-Anchored-Gradient method: particular case

- Let
$$\varepsilon_k := \frac{\alpha}{(k+\beta)^\eta} \quad \text{for all } k \geq 0, \text{ and } \alpha, \beta > 0, \ \eta \in (0,1).$$

Lemma (Chung, 1954; Polyak, 1987)

Let $(\psi^k)_{k\geq 1}$ be a non-negative sequence and let $c, C > 0$, and $0 < s < 1$, $s < t$. Suppose that
$$\psi^{k+1} \leq \left(1 - \frac{c}{k^s}\right)\psi^k + \frac{C}{k^t} \quad \text{for all } k \geq 1.$$
Then the following holds true as $k \to +\infty$
$$\psi^k \leq \frac{C}{c}\frac{1}{k^{t-s}} + o\left(\frac{1}{k^{t-s}}\right).$$

Theorem (strong convergence of the iterates and convergence rates for the residual)

The sequence $(x^k)_{k\geq 0}$ converges strongly to the minimum norm solution $\operatorname{proj}_{\operatorname{Zer} V}(0)$ as $k \to +\infty$, the residual showing the following convergence rate:
$$\|V(x^k)\| = O(k^{-\eta}) \quad \text{as } k \to +\infty.$$

RIB, E. Chenchene, *Extra-Gradient Method with flexible anchoring: strong convergence and fast residual decay*, to appear in SIAM Journal on Optimization

Numerical experiments

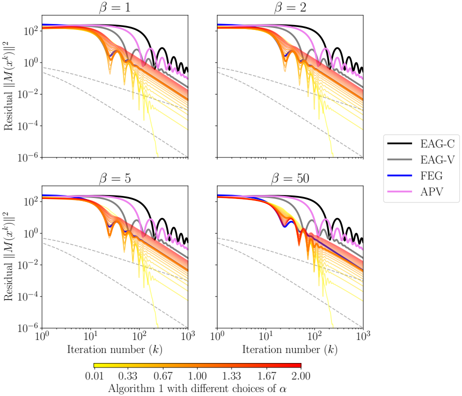
Let $\mathcal{H} = \mathbb{R}^{2d}$ for $d = 5$,

$$N := \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \quad \text{and} \quad x^* \in \mathbb{R}^{2d}, \quad b := Nx^*,$$

where x^* is randomly sampled. To induce some nonlinearity, consider $P\colon H \to H$ to be projection onto the ball with radius 1 and center x^*. Then, we define

$$V(x) := Nx - b + P(x) \quad \text{for all } x \in \mathbb{R}^{2d},$$

which is monotone and Lipschitz continuous with $V(x^*) = 0$. We choose $L := \|V\|_2 + 1$, and $\varepsilon_k = \frac{\alpha}{\theta(k+\beta)}$, for α in $(10^{-3}, 2]$ and $\beta = 1, 2, 5, 50$.

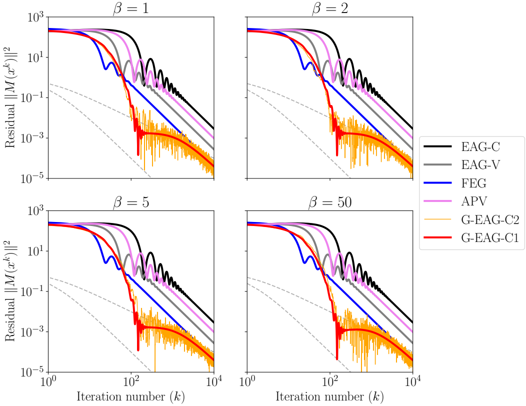


Numerical experiments: a practical enhancement

For the same problem, we choose

$$\varepsilon_k = \frac{\alpha_k}{\theta(k+\beta)} \quad \text{for all } k \geq 0\,, \tag{3}$$

where α_k is set to be either $\alpha_k := \frac{2}{\pi}\arctan(Mk)$ (falls in the setting of the main convergence theorem), with $M \in \mathbb{R}_+$ (we set $M = 10^{-3}$), or α_k to be equal to the same as before but further corrupted with Gaussian noise of variance $\sigma_k = \frac{1}{k}$ for all $k \geq 1$.



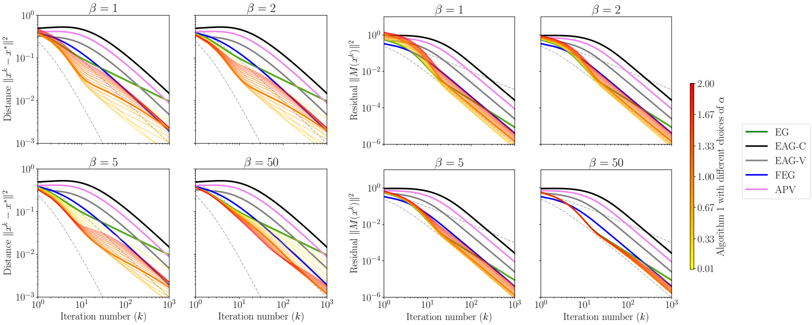
Numerical experiments: an infinite dimensional example

- On $\mathcal{H} := \ell^2$, the space of square summable sequences, we consider the operator

$$V(x) = x - Sx - b\,, \quad \text{where} \quad S(x_0, x_1, \dots) = (0, x_0, x_1, \dots)\,,$$

for all $x := (x_0, x_1, \dots) \in \mathcal{H}$, i.e. S is the right-shift operator and $b := (1, -1, 0, \dots) \in \mathcal{H}$. Since S is nonexpansive, V is monotone and 2-Lipschitz continuous. We see that $x^* := (1, 0, \dots)$ is the unique solution of $V(x) = 0$.

- We set a maximum number of iterations $K := 10^3$ and initialize $x^0 \in \mathrm{span}\{\delta^0, \delta^1\}$ randomly. In this way, $x_1, \dots, x_K$ will be always contained in $\mathrm{span}\{\delta^0, \dots, \delta^{2(K+1)}\} \simeq \mathbb{R}^{2(K+1)}$.



- G-EAG is particularly competitive also in terms of distance to solution. The example seems to suggest that in this case EG only converges weakly.

Thank you for your attention!

https://www.mat.univie.ac.at/~rabot/