Universal Reduced-Operator Method and High-Order Global Curvature Bounds

Nikita Doikov

Cornell University, ORIE

Based on joint work with Yurii Nesterov (**arXiv:2511.07341**)

INFORMS Annual Meeting, San Francisco

November 3, 2026

Motivation: Smooth Convex Optimization

Problem: $\min_{x \in \mathbb{R}^n} f(x)$, where

- f is convex, differentiable and smooth:

$$\|\nabla f(y) - \nabla f(x)\| \quad \leq \quad L\|y - x\|, \qquad x, y \in \mathbb{R}^n$$

- Characterization of the global convergence of gradient methods. E.g., for the **gradient descent**: $x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k)$, the rate is

$$f(x_k) - f^\star \quad \leq \quad \frac{2L\|x_0 - x^\star\|^2}{k+4}.$$

However, we can assume that the gradient is Hölder, for a certain degree $\nu \in [0, 1]$:

$$\|\nabla f(y) - \nabla f(x)\| \quad \leq \quad H_\nu \|y - x\|^\nu, \qquad x, y \in \mathbb{R}^n$$

- The rate of the gradient descent becomes

$$f(x_k) - f^\star \quad = \quad O\left(1/k^{\frac{1+\nu}{2}}\right).$$

- E.g., for $\nu = 0$ (non-smooth optimization), $f(x_k) - f^\star = O(1/\sqrt{k})$

Main Question

$$\|\nabla f(y) - \nabla f(x)\| \quad \leq \quad H_\nu \|y - x\|^\nu$$

Which ν to choose?

- In practice, we do not know the "right" $0 \leq \nu \leq 1 \quad \Rightarrow \quad$ *universal algorithms*?
- The same objective f can belong to several "problem classes" simultaneously
- Consider $f_1(x) = \frac{1}{2}x^2$ and $f_2(x) = \frac{2}{3}|x|^{3/2}$: [Nesterov, 2025]
  - f_1 corresponds to $\nu = 1$
  - f_2 corresponds to $\nu = 1/2$
  - **However**, their sum

$$f(x) \quad = \quad f_1(x) + f_2(x)$$

does not belong to any problem class with any $0 \leq \nu \leq 1$!

Our work: **Global Curvature Bound** — new parameter-free characterization of smoothness

$\Rightarrow$ works for high-order and universal algorithms

$\Rightarrow$ works for metric spaces and for variational inequalities

Second-Order Methods

At each point $x \in \mathbb{R}^n$ we have an access to second-order oracle:

$$x \quad \mapsto \quad f(x), \nabla f(x), \nabla^2 f(x)$$

Assume that the Hessian is Hölder continuous, for some $0 \leq \nu \leq 1$:

$$\|\nabla^2 f(y) - \nabla^2 f(x)\| \quad \leq \quad H_\nu \|y - x\|^\nu$$

$\Rightarrow$ a natural method minimizes the global second-order upper model of the function:

$$x_{k+1} = \arg\min_{y \in \mathbb{R}^n} \left\{ f(x_k) + \langle \nabla f(x_k), y - x_k \rangle + \tfrac{1}{2} \langle \nabla^2 f(x_k)(y - x_k), y - x_k \rangle + \frac{H_\nu \|y - x_k\|^{2+\nu}}{(1+\nu)(2+\nu)} \right\}$$

- Can be implemented as efficient as the classic Newton step
- For $\nu = 1$: cubic regularization of Newton's method
- Global rate:

$$f(x_k) - f^\star \quad = \quad O\Big(1/k^{(1+\nu)}\Big)$$

[Nesterov-Polyak, 2006; Cartis-Gould-Toint, 2011; Grapiglia-Nesterov, 2017]

Super-Universal Newton Method

- A simple algorithm that converges globally for all problem classes:

Super-Universal Newton

$$x_{k+1} \quad = \quad x_k - \left(\nabla^2 f(x_k) + M\|\nabla f(x_k)\| I\right)^{-1} \nabla f(x_k), \qquad k \geq 0$$

[**D**-Mishchenko-Nesterov, 2024; **SIOPT**]

- M is a regularization constant (either predefined or using an adaptive search)
- ν-Hölder Hessian: $f(x_k) - f^\star = O\big(1/k^{(1+\nu)}\big)$
- ν-Hölder third derivative: $f(x_k) - f^\star = O\big(1/k^{(2+\nu)}\big)$
- Quasi-Self-Concordant f: $f(x_k) - f^\star = O\big(\exp(-\frac{k}{MD})\big)$ [**D**, 2025; **Math. Prog.**]
- Inexact Hessians: $H_k \approx \nabla^2 f(x_k)$ [Semenov-Jaggi-**D**, 2026; **ICLR**]

The goal is to avoid the presence of parameter ν from the analysis.

Parameter-free Problem Class?

1. Modulus of uniform convexity. [Vladimirov-Nesterov-Chekanov, 1978; Azé-Penot, 1995]
   - **Strongly convex** functions, for some $\sigma > 0$:

$$\alpha f(x) + (1-\alpha) f(y) \quad \geq \quad f(\alpha x + (1-\alpha) y) + \frac{\alpha(1-\alpha)\sigma}{2} \|y - x\|^2, \qquad x, y \in \mathbb{R}^n, \alpha \in (0,1)$$

   - **Uniformly convex** functions: replace $2 \mapsto q$.
   - **Exact modulus of uniform convexity**:

$$\mu(t) \quad := \quad \inf_{\substack{x,y \in \mathrm{dom}\, f, \\ \alpha \in (0,1)}} \left\{ \frac{\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)}{\alpha(1-\alpha)} \ : \ \|y - x\| \geq t \right\}$$

   Then, it holds: $\alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y) + \alpha(1-\alpha) \cdot \mu(\|y - x\|)$

2. Global Curvature Bound for a function f:

$$\hat{\mu}(t) \quad := \quad \sup_{\substack{x,y \in \mathrm{dom}\, f, \\ \alpha \in (0,1)}} \left\{ \frac{|\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)|}{\alpha(1-\alpha)} \ : \ \|y - x\| \leq t \right\}$$

   - **Parameter-free analysis** of gradient methods [Nesterov, 2025]

Operators Between Metric Spaces

Let M and N be two metric spaces and let $A : M \to N$ be a mapping

Example. Let $M = \mathbb{E}$ be Euclidean space and $f : \mathbb{E} \to \mathbb{R}$ be our objective. Set

$$A(x) \quad := \quad \nabla f(x) \quad \in \quad \mathbb{E}^* \quad = \quad N$$

- $d(x, y) := \|x - y\|$

Example. M can be a Riemannian manifold.

For any two points $x, y \in M$ we assume to have a curve $\gamma : [0, 1] \to M$ called geodesic:

$$d(\gamma(s), \gamma(t)) \quad = \quad |s - t| d(x, y), \qquad 0 \leq s, t \leq 1.$$

NB: for the normed vector space $\mathbb{E}$ geodesics are straight lines:

$$\gamma(t) \quad = \quad x + t(y - x), \qquad 0 \leq t \leq 1.$$

Global Curvature Bound

Let $A : M \to N$. How to measure the curvature of A?

- Let $x, y \in M \quad \Rightarrow \quad$ geodesic $\gamma : [0,1] \to M$ s.t. $\gamma(0) = x$, $\gamma(1) = y$
- We have $A(x), A(y) \in N \quad \Rightarrow \quad$ geodesic $\xi : [0,1] \to N$ in the target space, $\xi(0) = A(x)$ and $\xi(1) = A(y)$
- **Define** the point-wise measure of curvature:

$$\Delta_A(x; y) \;:=\; \sup_{0<t<1} \tfrac{1}{t(1-t)} d\big(\, A(\gamma(t)),\; \xi(t)\,\big)$$

  shows how far the image of A is from the geodesic
- **The Global Curvature Bound** of operator A:

$$\kappa_A(r) \;:=\; \sup_{x,y} \Big\{\; \Delta_A(x,y) \;:\; d(x,y) \le r \;\Big\}$$

- **Main Theorem.** Let N have non-positive curvature. Then, for any $\beta \in [0,1]$:

$$\kappa_A(\beta r) \;\ge\; \beta^2 \kappa_A(r) \qquad \text{(local quadratic growth)}$$

Example: Variational Inequalities in Euclidean Space

Problem (Variational Inequality):

- Convex set $Q \subseteq \mathbb{R}^n$ in Euclidean space
- $V : Q \to \mathbb{R}^n$ be a monotone differentiable mapping
- **Goal:** find a point $x^\star \in Q$ s.t.

$$\langle V(x^\star), x - x^\star \rangle \quad \geq \quad 0, \qquad x \in Q$$

- **Application:** min-max optimization: $\min_x \max_y f(x, y)$: set $V = [\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}]^T$

Classic algorithm: **Extragradient method** [Korpelevich, 1976; Nemirovski, 2004]

- first-order algorithm with $O(1/k)$ rate, when V is Lipschitz

New universal second-order algorithm (using the Jacobian DV of V).

$$\kappa_V(r) \quad := \quad \sup_{\substack{x,y \in Q \\ 0<t<1}} \left\{ \frac{1}{t(1-t)} \left\| tV(y) + (1-t)V(x) - V(x + t(y-x)) \right\| \ : \ \|y - x\| \leq r \right\}$$

- Our main assumption: $\kappa_V(\cdot)$ is invertible

New Algorithm: Reduced-Operator Method

Initialization: $x_0 \in Q$, $M_0 > 0$, required tolerance $\delta, \varepsilon > 0$. Set $v_0 = x_0$ and $A_0 = 0$

Iterate, for $k \geq 0$**:**

1. Find the smallest integer $i_k \geq 0$ s.t. for $M^+ := 2^{i_k} M_k$ and for the computed x^+

$$\langle V(v_k) + DV(v_k)[x^+ - v_k] + \left(\sqrt{\tfrac{2}{5}\delta M^+} + M^+ r\right)(x^+ - v_k), x - x^+ \rangle \quad \geq \quad 0, \quad \forall x \in Q$$
$$r \;=\; \|x^+ - v_k\|$$

it holds

$$\langle V(x^+), v_k - x^+ \rangle \quad \geq \quad \tfrac{1}{3\sqrt{M^+}} \|V(x^+)\|^{3/2}$$

2. Set $x_{k+1} = x^+$ and $M_{k+1} = M^+$
3. Perform the reduced-operator step:

$$v_{k+1} \quad = \quad \mathrm{proj}_Q\big(v_k - a_{k+1} V(x_{k+1})\big), \quad \text{where} \quad a_{k+1} \;=\; \tfrac{\langle V(x_{k+1}), v_k - x_{k+1} \rangle}{\|V(x_{k+1})\|^2}$$

Convergence Analysis

The goal: find a point $\bar{x}_k$ such that

$$\max_{x \in Q} \langle V(x), \bar{x}_k - x \rangle \quad \leq \quad \varepsilon \qquad\qquad (\star)$$

- We did not fix a problem class yet!
- Denote by $\kappa_V(\cdot)$ the **global curvature bound** of V:

$$\kappa_V(r) \quad := \quad \sup_{\substack{x,y \in Q \\ 0<t<1}} \left\{ \tfrac{1}{t(1-t)} \left\| tV(y) + (1-t)V(x) - V(x + t(y-x)) \right\| \ : \ \|y - x\| \leq r \right\}$$

- The only assumption: $\kappa_V(\cdot)$ is invertible.

Main result. To achieve $(\star)$ it is enough to perform

$$k \quad = \quad O\left(\left[\frac{D}{\kappa_V^{-1}(\varepsilon/(5D))} \right]^{4/3} \right)$$

iterations (oracle calls to V and DV), where $D := \operatorname{diam}(Q)$

Comparison Example

Our result:

$$k \;=\; O\Big(\Big[\frac{D}{\kappa_V^{-1}(\varepsilon/(5D))}\Big]^{4/3}\Big)$$

- No specific assumption on the behavior of V!

Let DV be Hölder continuous, for some $0 \le \nu \le 1$:

$$\|DV(x) - DV(y)\| \;\le\; H_\nu \|x-y\|^\nu$$

Then,

$$\kappa_V(r) \;\le\; \frac{2^{1-\nu}}{1+\nu} \cdot H_\nu r^{1+\nu} \qquad \Rightarrow \qquad \kappa_V^{-1}(s) \;\ge\; \Big[\frac{(1+\nu)s}{2^{1-\nu}H_\nu}\Big]^{\frac{1}{1+\nu}}.$$

The complexity:

$$k \;=\; O\Big(\Big[\frac{H_\nu D^{2+\nu}}{\varepsilon}\Big]^{\frac{4}{3(1+\nu)}}\Big)$$

matches the optimal results from [Adil-Bullins-Jambulapati-Sachdeva, 2022; Lin-Jordan, 2025]

Conclusions

- Parameter-free analysis of optimization and numerical methods based on the notion of **Global Curvature Bound**
- Works for abstract metric spaces (e.g. Riemannian manifolds)
- High-order methods for variational inequalities: **optimal and universal algorithms**

Thank you for your attention!