Skoltech
Skolkovo Institute of Science and Technology

Accelerations of Kaczmarz method for solving linear systems

NLA/Optimization Final Project

Artem Sevastopolskiy
Sergey Samsonov
Mikhail Yakhlakov
Ekaterina Ivanova

Kaczmarz method for SLE

$$Ax = b,\ A \in \mathbb{R}^{n\times m},\ A = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$$

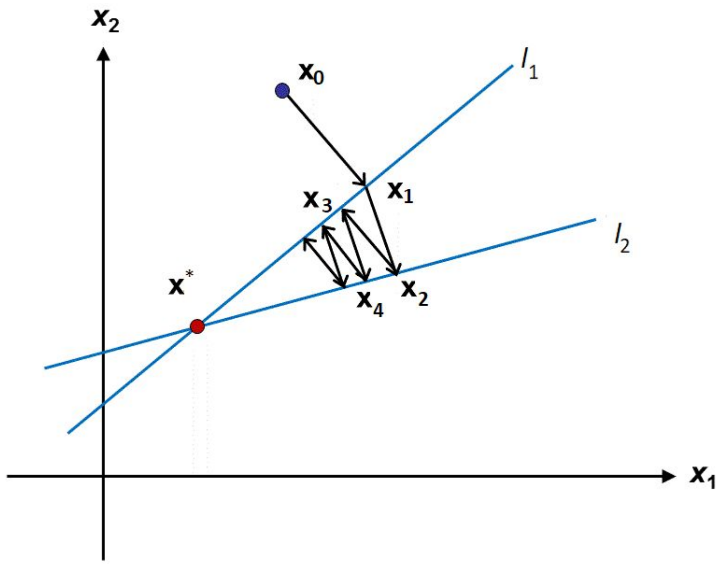
$$x^0$$

$$x_1^1, \dots, x_n^1,$$

$$\dots \quad \dots \quad \dots,$$

$$x_1^k, \dots, x_n^k \ :$$

$$x_i^k = x_{i-1}^k + \frac{b_i - \langle a_i, x_{i-1}^k \rangle}{\|a_i\|^2} a_i,\ i = \overline{1, n}$$



Example for n = m = 2

Randomized Kaczmarz method

[Strohmer, Vershynin 2009] Iterate hyperplanes in random order:

$$\mathbb{P}(\text{ choose } \quad \boldsymbol{k}) = \frac{\|a_k\|_2^2}{\|A\|_F^2}$$

$$x^k = x^{k-1} + \frac{b_{i_k} - \langle a_{i_k}, x^{k-1} \rangle}{\|a_{i_k}\|^2} a_{i_k}, \ \ k = 1, 2, \dots$$

For consistent, for over-determined system exponential convergence rate holds:

$$\mathbb{E}\|x_k - x\|_2^2 \leq (1 - \kappa(A)^{-2})^k \|x_0 - x\|_2^2$$

$$\kappa(A) = \|A\|_F \|A^{-1}\|_2$$

(assuming that left inverse exists)

Kaczmarz with Lyusternik acceleration

x^0

$x_1^1, \ldots, x_n^1,$ $\qquad x_i^k = x_{i-1}^k + \frac{b_i - \langle a_i, x_{i-1}^k \rangle}{\|a_i\|^2} a_i, \; i = \overline{1,n}$

$\ldots \; \ldots \; \ldots,$

$x_1^k, \ldots, x_n^k :$ $\qquad x^k := x_n^k$

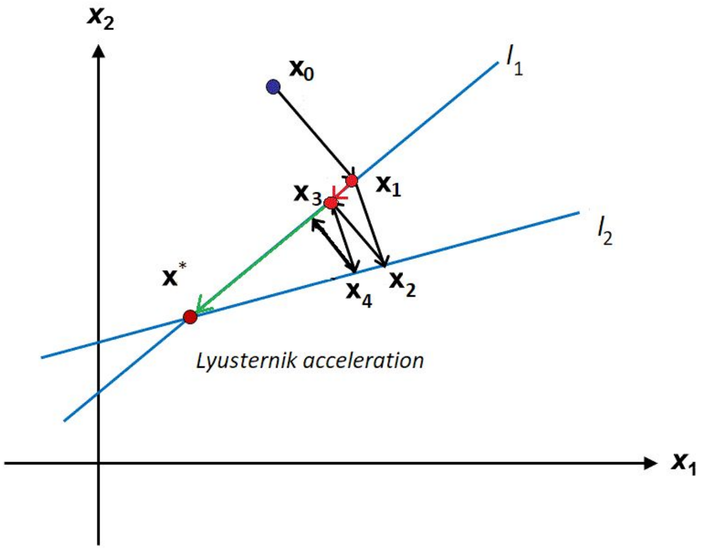
Idea: for large enough k,

- $x^k - x^{k-1} \approx \alpha v_1, \; v_1$ corresponds to λ_{max}
- $\|x^k - x^{k-1}\| \approx \mu \|x^{k-1} - x^{k-2}\|$

So we having the “shooting” rule:

when $\cos(x^k - x^{k-1}, x^{k-1} - x^{k-2}) \approx 1$,

shoot: $x^{k+1} := \frac{x^k - \mu x^{k-1}}{1 - \mu}$ (by sum of geometric series)



Where Kaczmarz is useful?

- overdetermined systems $(Ax = b,\ A \in \mathbb{R}^{n \times m},\ n > m)$
- ill-conditioned systems $(\|A\|\|A^{-1}\| \gg 1)$
- other special cases

Fourier matrix

$$f(t) = \sum_{l=-r}^{r} x_l \cos(2\pi i l t), x = \{x_l\}_{l=-r}^{l=r}$$

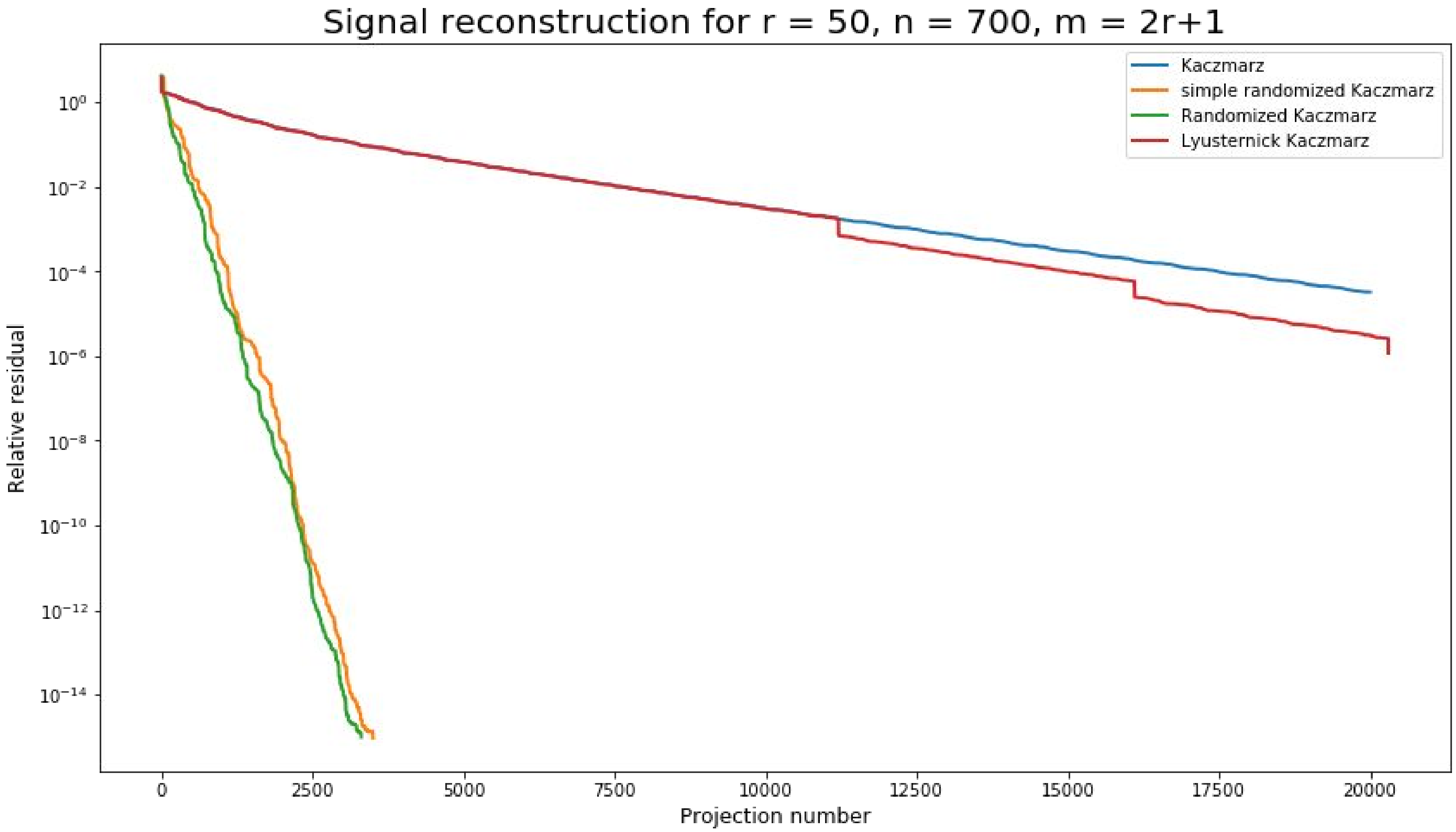
Assume that we are given non-uniformly sampled nodes and values in them:

$$\{t_k\}_{k=1}^{m}, \{f(t_k)\}_{k=1}^{m}$$

Recovering signal leads us to system

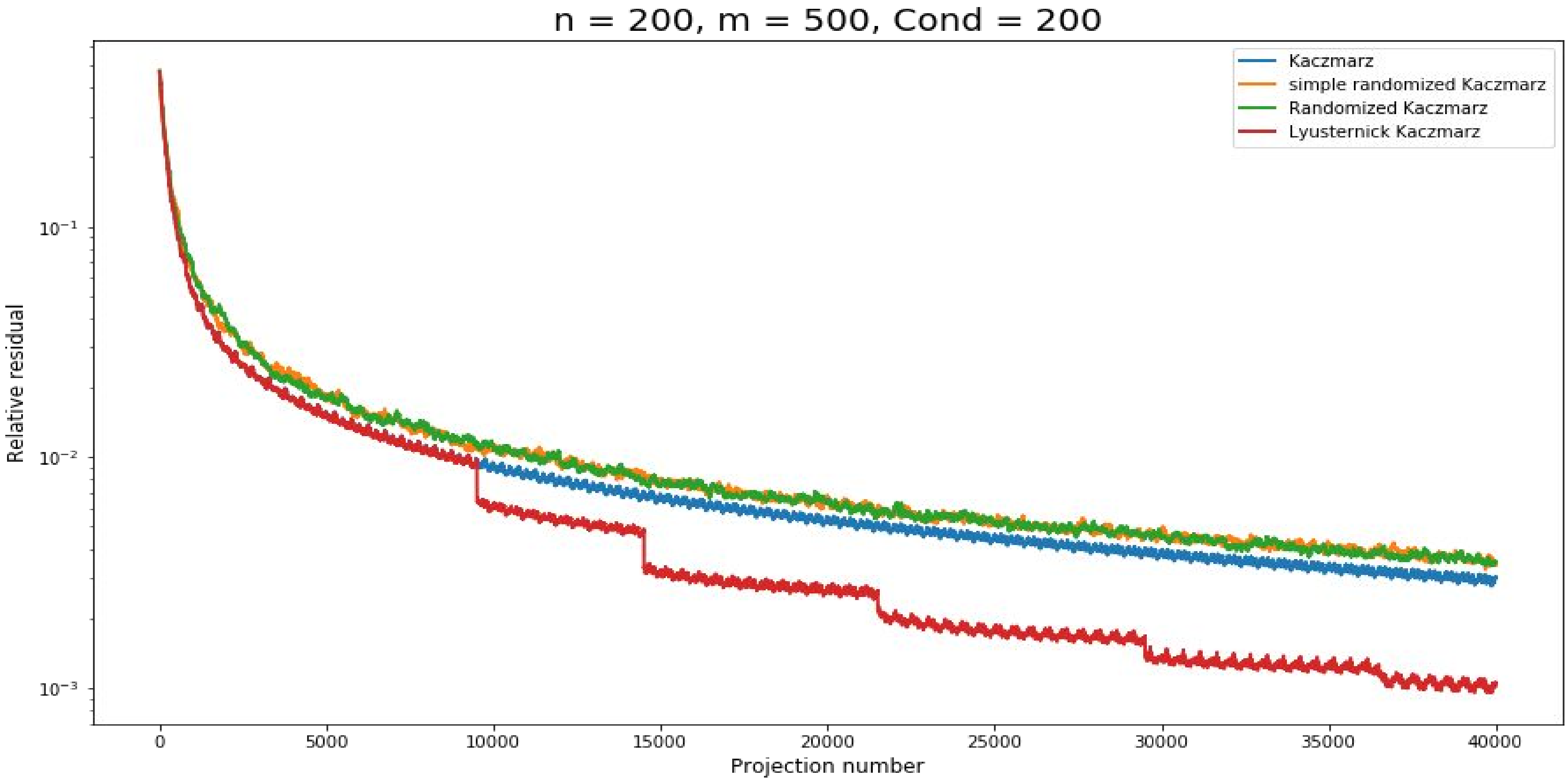
$$Ax = b, A_{j,k} = \sqrt{\omega_j}\cos(2\pi i k t_j), b_j = \sqrt{\omega_j} f(t_j), \omega_j = \frac{t_{j+1} - t_{j-1}}{2}$$

Here weights are supposed to compensate for varying density in the sample.

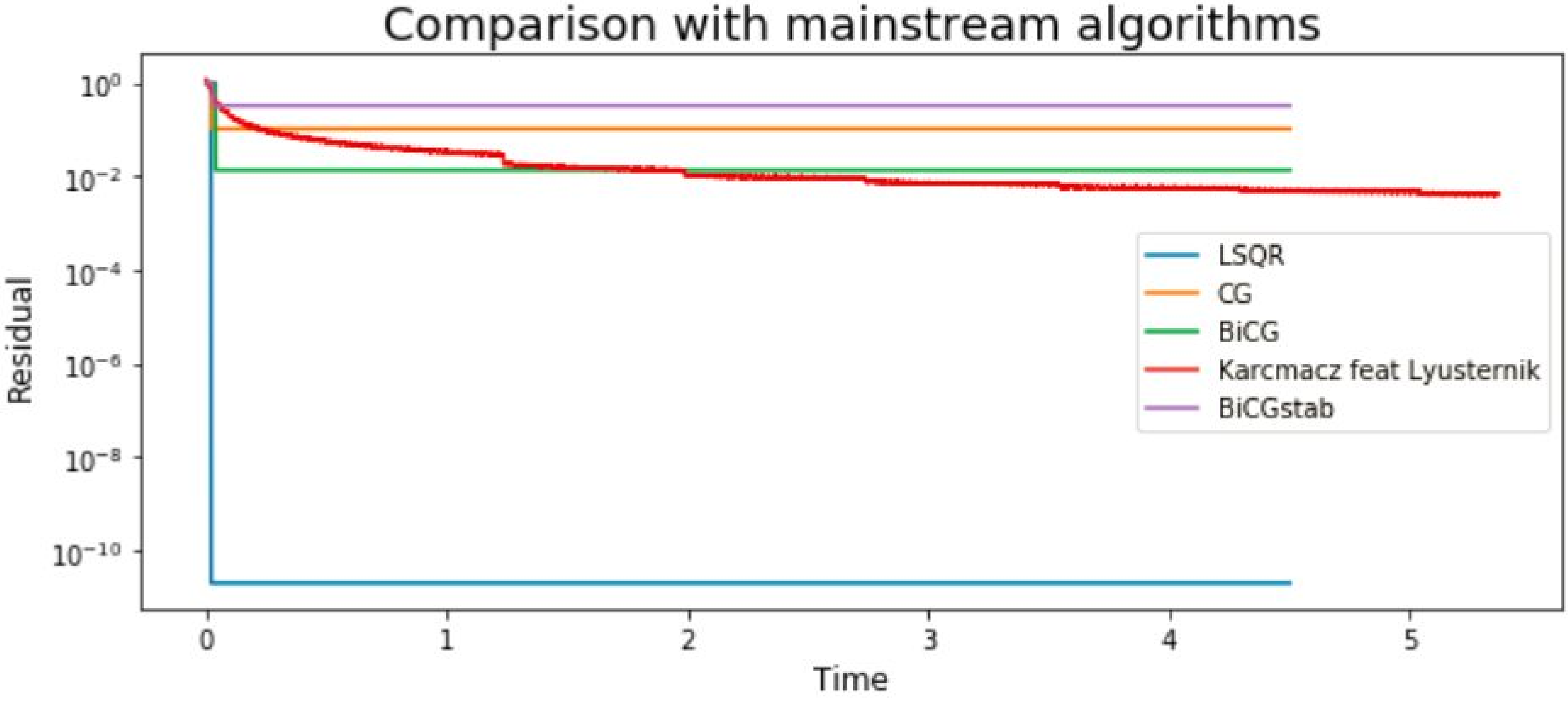


Signal reconstruction for r = 50, n = 700, m = 2r+1
Kaczmarz
simple randomized Kaczmarz
Randomized Kaczmarz
Lyusternick Kaczmarz
Relative residual
Projection number
0
2500
5000
7500
10000
12500
15000
17500
20000

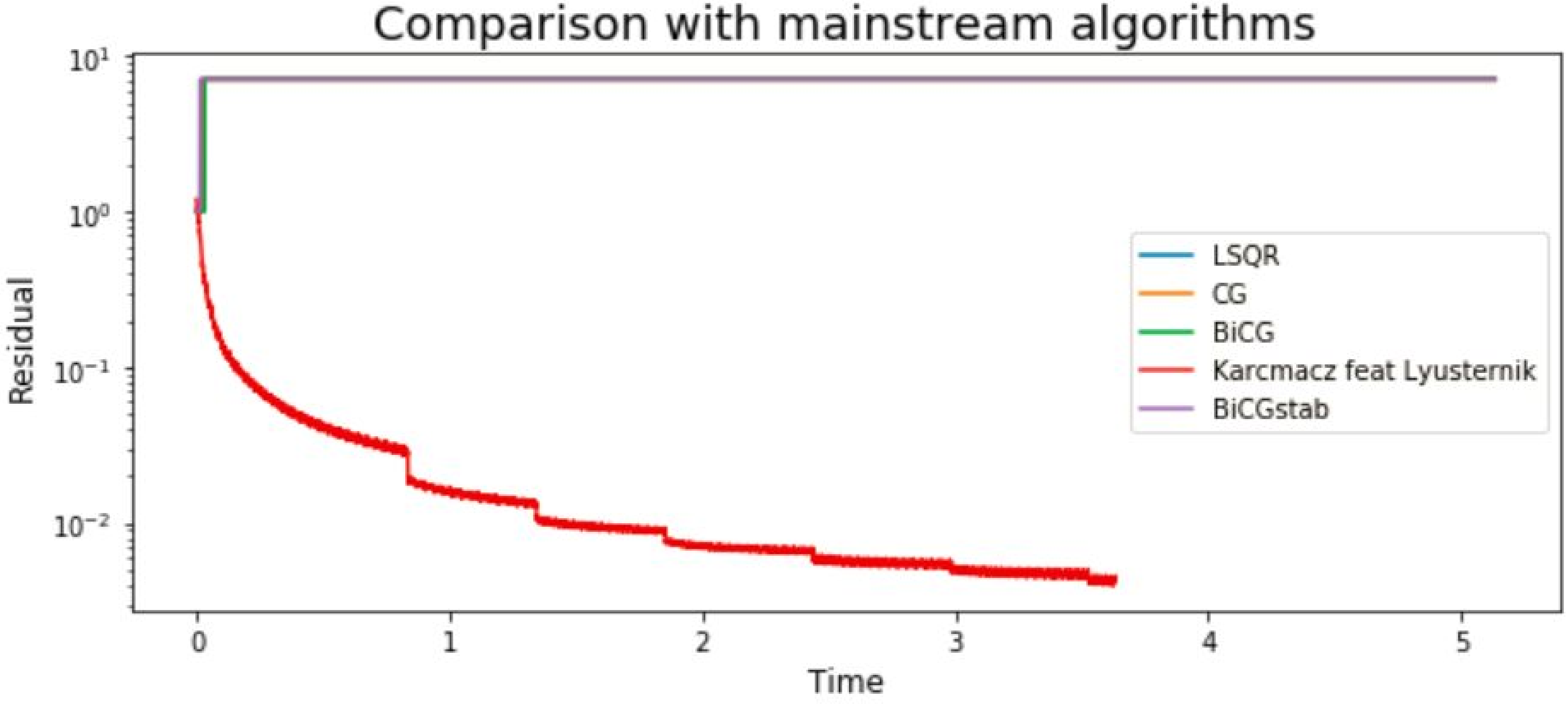
Over-determined systems



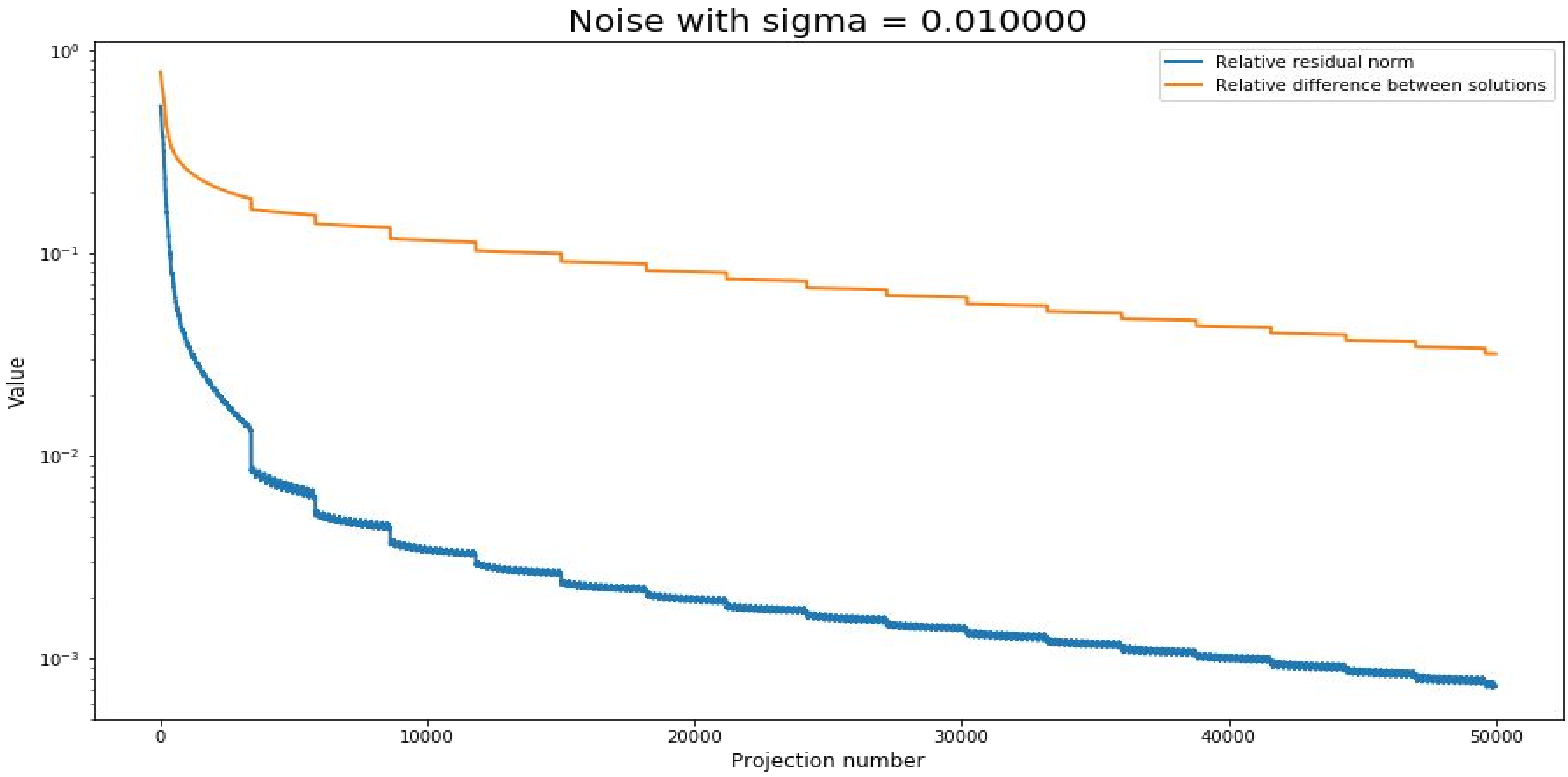
Over-determined systems



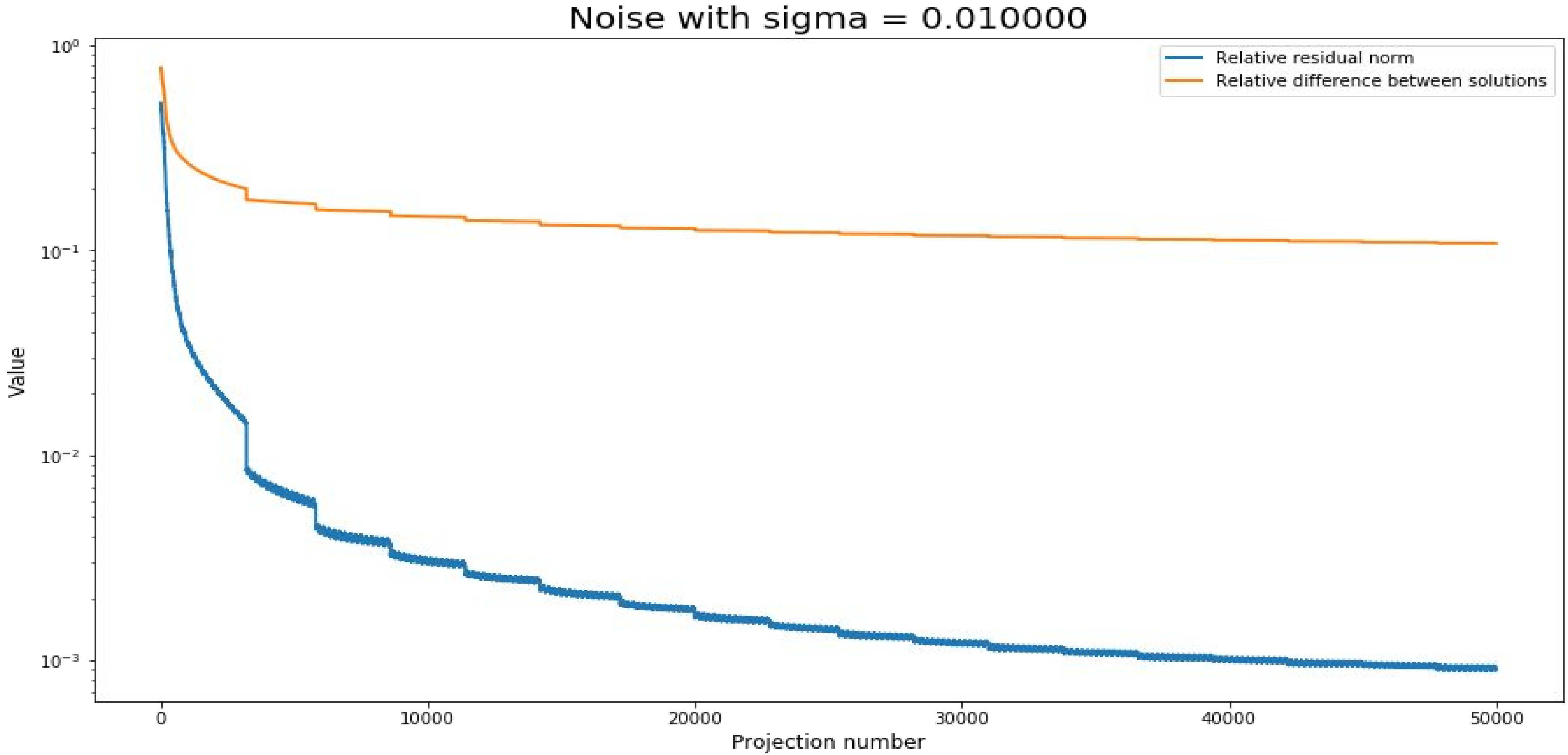
Ill-conditioned systems



Stability with moderate conditional number



Lack of stability for ill-conditioned matrix



Contribution:

Artem Sevastopolskiy	Sergey Samsonov	Mikhail Yakhlakov	Ekaterina Ivanova
team leader Kaczmarz with Lyusternik acceleration - programming presentation	presentation Kaczmarz - programming ill-conditioned systems	over-determined systems comparison with third-party methods	stability research comparison with third-party methods presentation

Thank you for your attention!

https://en.wikipedia.org/wiki/Lazar_Lyusternik