When ExtraGradient meets VR:
bridjing two giants to boost VIs

part II

Y. Kabikov, G. Chirkov
leads: D. Medyakov, G. Molodtsov
MIPT, 2025

a brief reminder

variational inequality:

$$\forall z \in \mathcal{Z} \hookrightarrow \langle F(z^*), z - z^* \rangle \geqslant 0$$

finite-sum approach:

$$F(z) = \frac{1}{n}\sum_{i=1}^{n} F_i(z)$$

Stochastic EG and variance problem:

$$z^{t+\frac{1}{2}} = z^t - \gamma F_{i_t}(z^t),$$
$$z^{t+1} = z^t - \gamma F_{j_t}(z^{t+\frac{1}{2}})$$

$$\mathbf{E}\{\|F_{i_t}(z^t) - F(z^t)\|^2\} \not\to 0, \quad z^t \to z^*$$

ExtraSAGA method

$$z^{t+1/2} = z^t - \gamma G^{t-1}$$
$$G^t = F_{i_t}(z^{t+1/2}) - y^t_{i_t} + \tfrac{1}{n}\sum_{j=1}^n y^t_j$$
$$y^{t+1}_i = F_{i_t}(z^{t+1/2}) \text{ for } i = i_t$$
$$z^{t+1/2} = z^t - \gamma G^t$$

ExtraSAGA method (updated)

$$z^{t+1/2} = z^t - \boxed{\frac{1}{n}\sum_{j=1}^{n} y_j^t}$$

$$G^t = F_{i_t}(z^{t+1/2}) - y_{i_t}^t + \frac{1}{n}\sum_{j=1}^{n} y_j^t$$

$$y_i^{t+1} = F_{i_t}(z^{t+1/2}) \text{ for } i = i_t$$

$$z^{t+1/2} = z^t - \gamma G^t$$

Positive result!

We've proved the ExtraSAGA convergence with the 1/6L step (gamma)

Experiments!

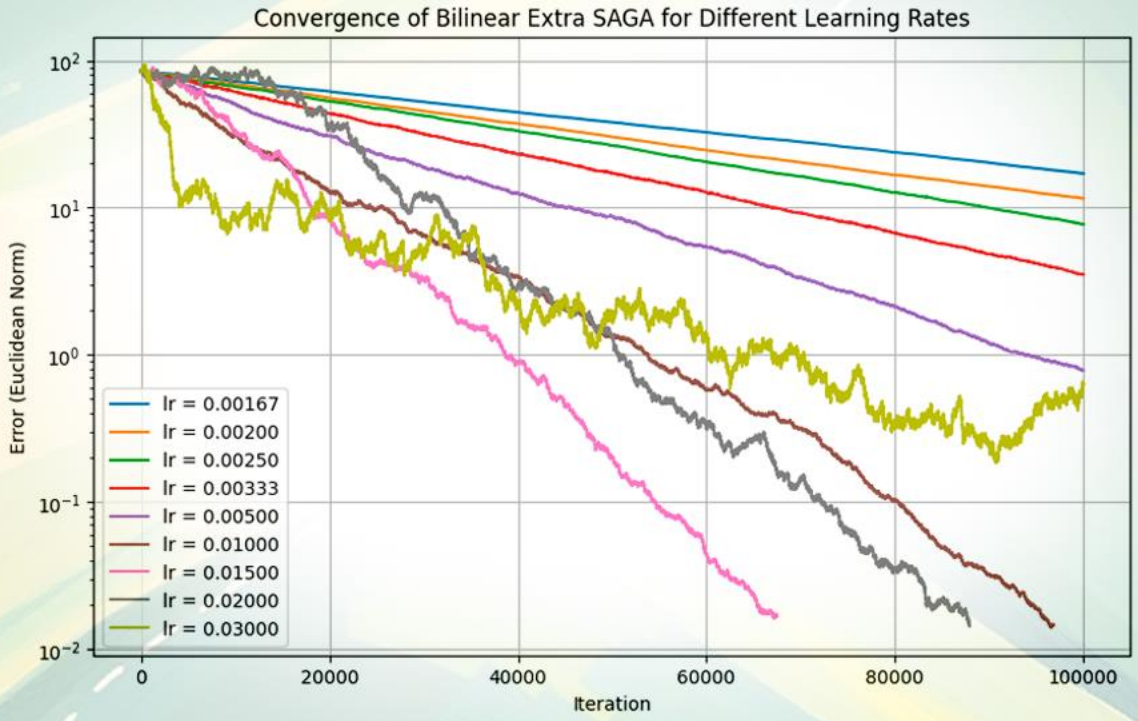
Experiments:

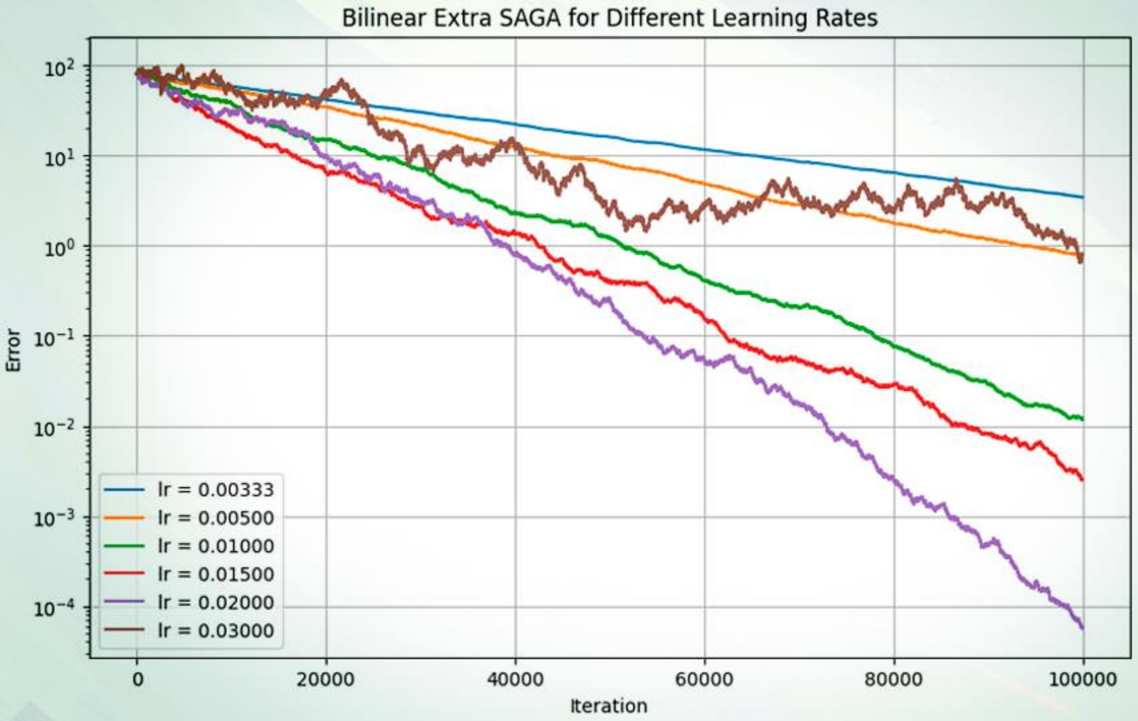
Bilinear convex-concave min-max (saddle points)

$$\min_{x \in \mathbb{R}^{d_x}} \max_{y \in \mathbb{R}^{d_y}} \left[f(x, y) \right]$$

$$f(x, y) = (x - b_x)^T A (y - b_y) + \frac{\lambda}{2} \|x - b_x\|^2 - \frac{\lambda}{2} \|y - b_y\|^2$$

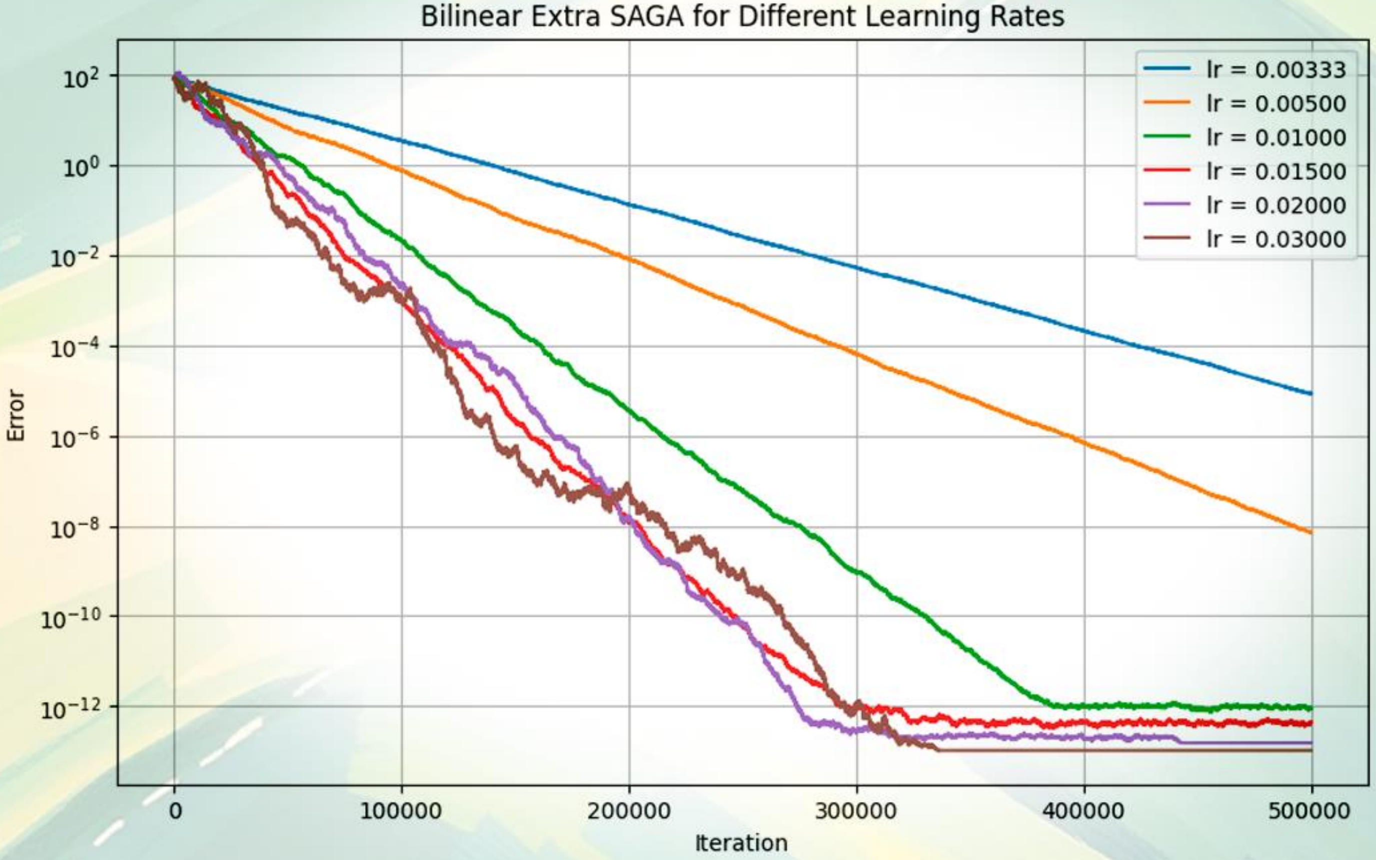
Experiments:
Bilinear convex-concave min-max (saddle points)





L = 100, mu = 0.1, n = 100, d = 100

Experiments:
Bilinear convex-concave min-max (saddle points)



L = 100, mu = 0.1, n = 100, d = 100

Denoising

$$\{(ih, jh) : 1 \leq i \leq M, 1 \leq j \leq N\}$$

picture = cartesian grid

$$(\nabla u)_{i,j} = \begin{pmatrix} (\nabla u)^1_{i,j} \\ (\nabla u)^2_{i,j} \end{pmatrix} \quad (\nabla u)^1_{i,j} = \begin{cases} \frac{u_{i+1,j} - u_{i,j}}{h} & \text{if } i < M \\ 0 & \text{if } i = M \end{cases}, \quad (\nabla u)^2_{i,j} = \begin{cases} \frac{u_{i,j+1} - u_{i,j}}{h} & \text{if } j < N \\ 0 & \text{if } j = N \end{cases}$$

$$\min_x \int_\Omega |Du| + \frac{\lambda}{2} \|u - g\|_2^2$$

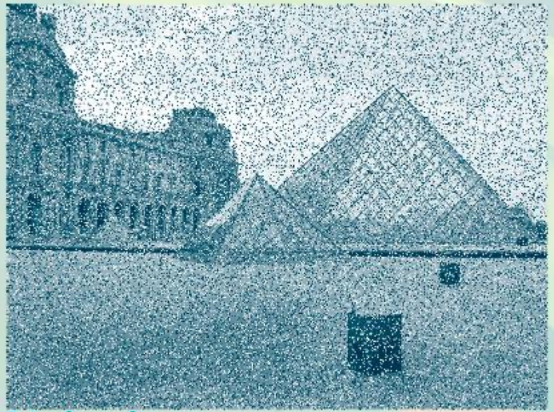
the ROF model

$$\min_{u \in X} \|\nabla u\|_1 + \lambda \|u - g\|_1$$

TV-L 1 model

Denoising

Clean image

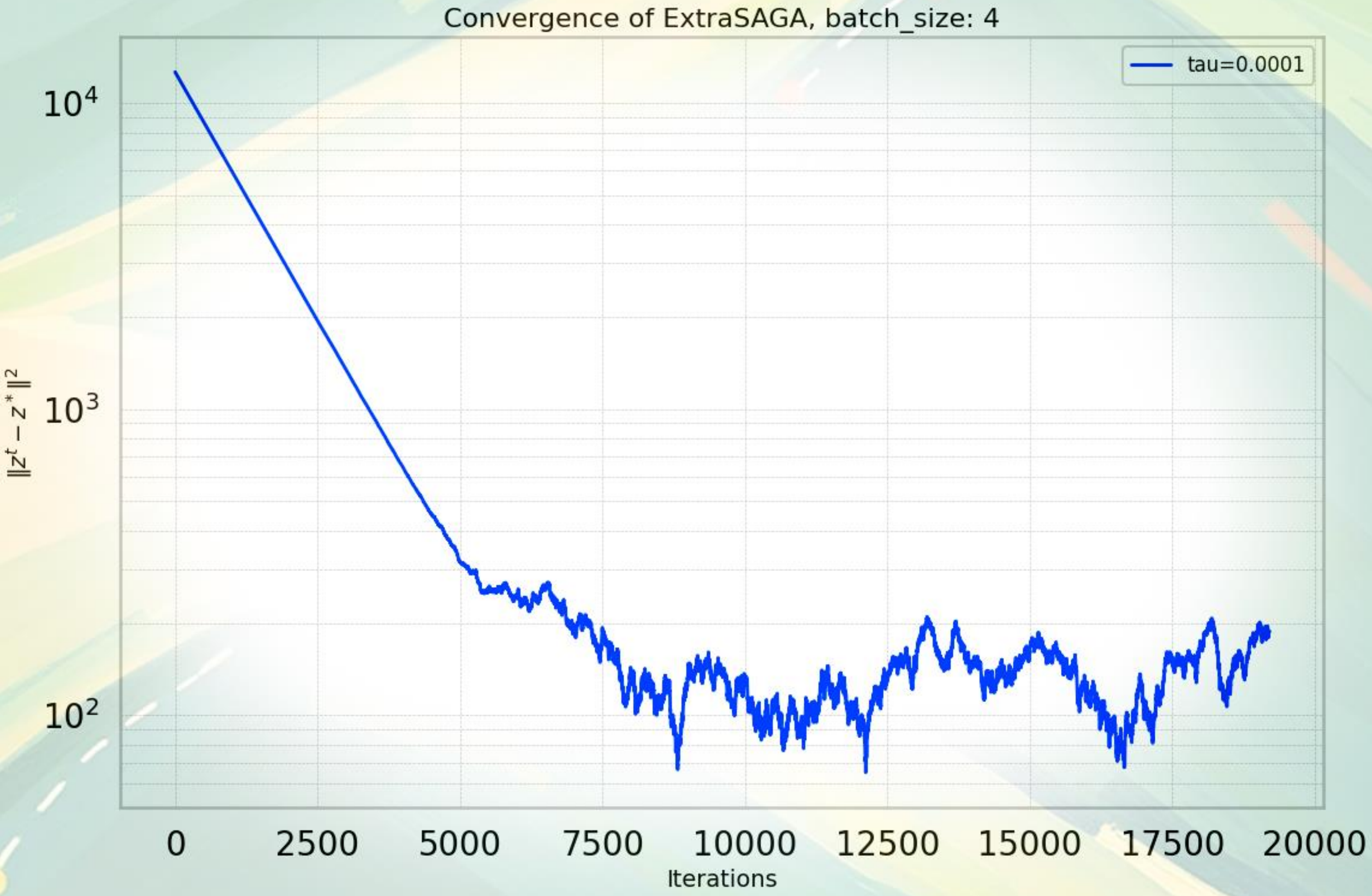


Noisy image

ROF (lambda = 8)

TV-L 1 (lambda = 1.5)

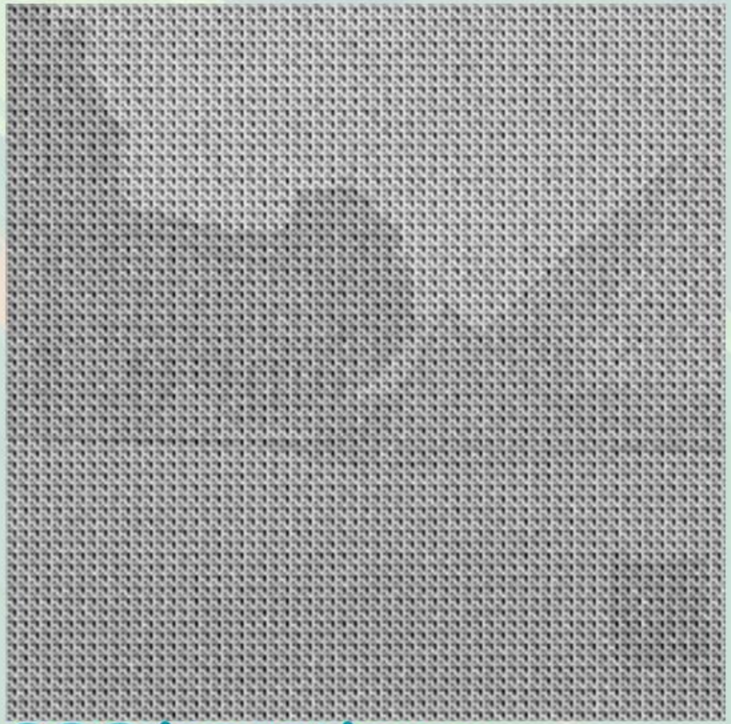
Experiments:
Denoising



Experiments:
Denoising

noisy input

300 iterations



600 iterations

Goals:

To fix Denoising experiments.
Experiments with different datasets.
GANs, Adversarial training.
Arcticle writing and layout.
To finish at the end of the April, 2025.

Thanks for listening!