V.V. Ulyanov (HSE University and Lomonosov MSU). CLT for algorithms of stochastic approximation

In the talk, we consider two recurrent sequenses of random vectors $\{\theta_k\}$ and $\{w_k\}$:

$$\theta_{k+1} = \theta_k + \beta_k(b_1 - A_{11}\theta_k - A_{12}w_k + V_{k+1}),$$
$$w_{k+1} = w_k + \gamma_k(b_2 - A_{21}\theta_k - A_{22}w_k + W_{k+1}),$$

where b_i and A_{ij} are some vectors and some matrices resp., V_{k+1}, W_{k+1} are the noise terms, the scalars β_k, γ_k are step sizes. We consider $\beta_k = c_{0,\beta}(k+k_0)^{-b}$ and $\gamma_k = c_{0,\gamma}(k+k_0)^{-a}$ with exponents a and b satisfying $1/2 < a < b < 1$. These type of sequences $\{\theta_k\}$ and $\{w_k\}$ arise in two-timescale stochastic approximation (TTSA) algorithms as stochastic approximations for the uniqe solutions $(\theta^\star, w^\star)$ for the system of linear equations

$$A_{11}\theta + A_{12}w = b_1, \quad A_{21}\theta + A_{22}w = b_2.$$

Theorem. *Under the martingale noise setting, let some assumptions are met (see details https://doi.org/10.48550/arXiv.2508.07928). Then for all large enough n, one has*

$$dist_K\big(\beta_n^{-1/2}\tilde{\theta}_{n+1}, \mathcal{N}(0, \Sigma_\infty^{\mathrm{last}})\big) \lesssim_{\log n} n^{-1/4},$$

where $dist_K$ is the Kolmogorov distance with respect to all convex sets, $\tilde{\theta}_{n+1}$ is the last iterate.

Similar results are obtained for the sequence $\{w_k\}$, for Polyak-Ruppert averaged TTSA and for Markov noise, although with worse order of approximation error.

The talk is based on the joint resuts with Bogdan Butyrin, Artemy Rubtsov, Alexey Naumov, Sergey Samsonov, see details https://doi.org/10.48550/arXiv.2508.07928