# Optimization with Markovian Noise

## First Order Methods

Boris Prokhorov

Scientific adviser - Aleksandr Beznosikov

May 5, 2024

# Table of Contents

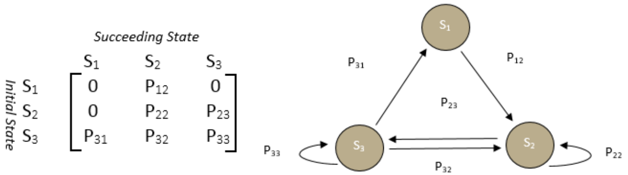
# Table of Contents

We study the minimization problem

$$\min_{x \in \mathbb{R}^d} f(x) := \mathbb{E}_{Z \sim \pi}[F(x, Z)] \tag{1}$$

Where $Z$ is a Markov Chain.

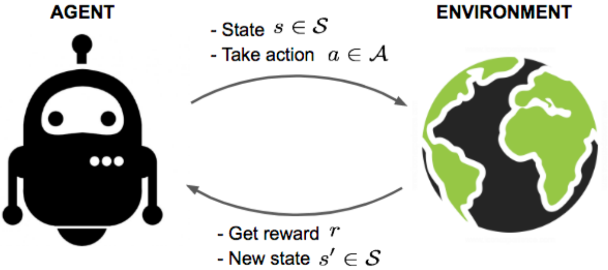
# Introduction
## Quick intro to Markov Chains




**Transition Matrix** **Transition Diagram**

# Introduction
## Motivation

*...Reinforcement Learning...*



Our research aims to expand the results of [1].
Following the article, we consider additional assumptions:

- (A1,A2) $f$ is $(\mu,\ L)$-smooth.
- (A3) $Z$ is uniformly geometrically ergodic with mixing time $\tau$.
- (A4) $F$ is $(\sigma,\ \delta)$-bounded:

$$\|\nabla F(x, Z) - \nabla f(x)\|^2 < \sigma^2 + \delta^2 \|\nabla f(x)\|^2$$

The main result of [1] is the following

**Theorem.** Under A1-A4 the problem (1) can be solved (in terms of $\mathbb{E}\left[\|x^{(k)} - x^*\|^2\right] \leq \varepsilon$) in

$$\tilde{\mathcal{O}}\left(\tau\left[(1+\delta^2)\sqrt{\frac{L}{\mu}}\log\frac{1}{\varepsilon} + \frac{\sigma^2}{\mu^2\varepsilon}\right]\right) \quad \text{oracle calls}\,. \tag{2}$$

And the bound is tight for $\delta = 1$.

# Table of Contents

Our main focus is on the optimality of the bound (2).

Summary of the interim results:

1. For $\delta = 1$ the bound is tight. (**Theorem**, [1])
2. For $\delta > 1$ the bound is tight. (**New**)
   The proof generalizes the previous result but keeps the main idea.
3. For $\delta < \frac{\mu}{4L}$ and $\sigma = 0$ the bound is **not** tight. (**New**)
   For now, we have an upper bound of $\tilde{\mathcal{O}}\left(\frac{L}{\mu}\log\frac{1}{\varepsilon}\right)$ oracle calls which is better than (2) if $\tau$ is large.

# Our Contribution

## Future work

Our main goal is to determine the algorithmic complexity of the problem. To do so, we need to eliminate the gaps between the lower and upper bounds:

1. For $\delta < \frac{\mu}{4L}$ we believe that the upper bound can be improved using Nesterov acceleration [2].
2. For $\frac{\mu}{4L} < \delta < 1$ current methods are not working and more research is needed.

# Table of Contents

# Bibliography I

[1] Aleksandr Beznosikov et al. *First Order Methods with Markovian Noise: from Acceleration to Variational Inequalities*. 2023. arXiv: 2305.15938 [math.OC].

[2] Yurii Nesterov. “A method for solving the convex programming problem with convergence rate $O(1/k^2)$”. In: *Proceedings of the USSR Academy of Sciences* 269 (1983), pp. 543–547.