FINITE-TIME GUARANTEES OF CONTRACTIVE STOCHASTIC APPROXIMATION: MEAN SQUARE AND TAIL BOUNDS

Siva Theja Maguluri

Industrial and Systems Engineering
Georgia Institute of Technology

JOINT WORK WITH

Zaiwei Chen Caltech

Martin Zubeldia, Univ of Minnesota

Karthikeyan Shanmugam Google Research

Sanjay Shakkottai UT Austin

BANACH FIXED POINT THEOREM

Want to find $\mathbf{x}^*$ that solves

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$$

A simple iteration

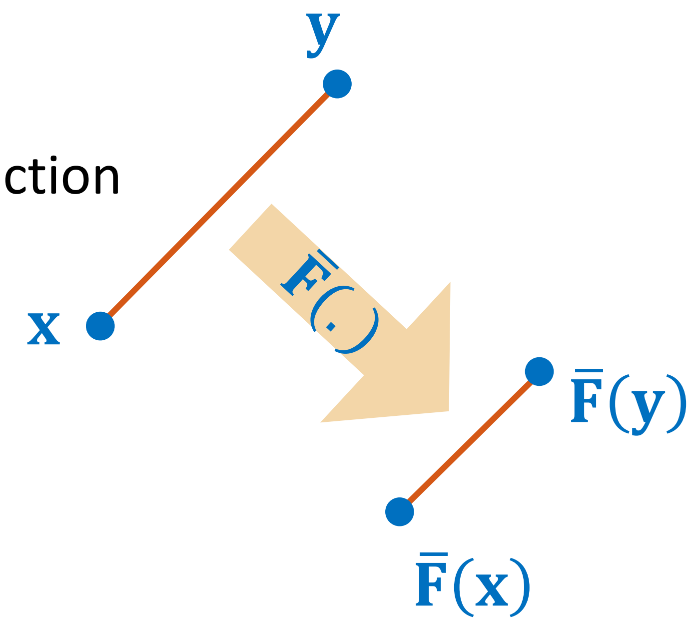
$$\mathbf{x}_{k+1} = \bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k$$

Noisy Oracle

Banach Fixed Point Theorem

$\mathbf{x}_k$ converges to $\mathbf{x}^*$ geometrically fast (linearly) if $\bar{\mathbf{F}}\,(.)$ is a contraction

Contraction: For all $\mathbf{x}$ and $\mathbf{y}$, $\left\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\right\| \leq \gamma \, \|\mathbf{x} - \mathbf{y}\|$



Works for any norm

STOCHASTIC APPROXIMATION

Want to find $\mathbf{x}^*$ that solves

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$$

A simple iteration

$$\mathbf{x}_{k+1} = \bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k$$

Noisy Oracle

Stochastic Approximation[Robbins, Monro '51]

$$\begin{aligned}\mathbf{x}_{k+1} &= (1-\alpha_k)\mathbf{x}_k + \alpha_k\left(\bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k\right) \\ &= \mathbf{x}_k + \alpha_k\left(\bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k - \mathbf{x}_k\right)\end{aligned}$$

Question: How well does this work?

OUTLINE

- Stochastic Approximation Introduction
- Finite Sample bounds on the mean-square error $\mathbb{E}[\|\mathbf{x}_k - \mathbf{x}^*\|^2]$
- High Probability (Tail) bounds on $\|\mathbf{x}_k - \mathbf{x}^*\|$
- Proof Sketch
  - Mean square - A Lyapunov function
  - Tail bounds - Exponential Supermartingale and Bootstrapping

STOCHASTIC APPROXIMATION

FIXED POINT PROBLEMS

Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

$$\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}_{\mathrm{k}} + \alpha_k\left(\bar{\mathbf{F}}(\mathbf{x}_{\mathrm{k}}) + \mathbf{w}_{\mathrm{k}} - \mathbf{x}_{\mathrm{k}}\right)$$

Optimization:

$$\min f(\mathbf{x})$$

$$-\eta\boldsymbol{\nabla} f(\mathbf{x}) + \mathbf{x} = \mathbf{x}$$

When f is smooth strongly convex, $\bar{\mathbf{F}}(\mathbf{x}) = -\eta\boldsymbol{\nabla} f(\mathbf{x}) + \mathbf{x}$ is contraction wrt ℓ_2-norm

SGD: $\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}_{\mathrm{k}} - \alpha_{\mathrm{k}}\left(\boldsymbol{\nabla} f(\mathbf{x}_{\mathrm{k}}) + \mathbf{w}_{\mathrm{k}}\right)$

FIXED POINT PROBLEMS

Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

$$\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}_{\mathrm{k}} + \alpha_k\left(\bar{\mathbf{F}}(\mathbf{x}_{\mathrm{k}}) + \mathbf{w}_{\mathrm{k}} - \mathbf{x}_{\mathrm{k}}\right)$$

Markov Decision Processes and RL:

$\bar{\mathbf{F}}\ (\cdot)$ is related to the Bellman operator.

TD learning, Q learning and their variants can be modeled as SA

The underlying norm is weighted ℓ_p (for TD) and ℓ_∞ (for Q learning)

FIXED POINT PROBLEMS

Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\left(\bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k - \mathbf{x}_k\right)$$

Linear Equations:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

$$(\mathbf{I} + \eta\mathbf{A})\mathbf{x} - \eta\mathbf{b} = \mathbf{x}$$

When $\mathbf{A}$ is Hurwitz ($\mathrm{Re}(\lambda_i) < 0$), $\bar{\mathbf{F}}(\mathbf{x}) = (\mathbf{I} + \eta\mathbf{A})\mathbf{x} - \eta\mathbf{b}$ is contraction wrt weighted ℓ_2-norm

Linear SA: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{A}\mathbf{x}_k - \mathbf{b}_k)$

MARKOVIAN STOCHASTIC APPROXIMATION

Want to find $\mathbf{x}^*$ that solves

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbb{E}_{\mathbf{Y}\sim\boldsymbol{\mu}}\left[\mathbf{F}(\mathbf{x},\mathbf{Y})\right] = \mathbf{x}$$

$$\mathbf{x}_{\mathrm{k}+1} = \mathbf{x}_{\mathrm{k}} + \alpha_{\mathrm{k}}(\mathbf{A}_{\mathrm{k}}\mathbf{x}_{\mathrm{k}} - \mathbf{b})$$

Markovian Stochastic Approximation

$$\mathbf{x}_{\mathrm{k}+1} = \mathbf{x}_{\mathrm{k}} + \alpha_{\mathrm{k}}(\mathbf{F}(\mathbf{x}_{\mathrm{k}},\mathbf{Y}_{\mathbf{k}}) + \mathbf{w}_{\mathrm{k}} - \mathbf{x}_{\mathrm{k}})$$

Multiplicative Noise

Additive Noise

(Main) Assumptions

- $\mathbf{Y}_{\mathbf{k}}$ is a finite state Ergodic Markov chain with stationary distribution $\boldsymbol{\mu}$
  - $\mathbf{Y}_{\mathbf{k}}$ is geometrically mixing
- Noise $\mathbf{w}_{\mathrm{k}}$ - iid or martingale difference, mean zero, $\|\mathbf{w}_{\mathrm{k}}\| \leq B(\|\mathbf{x}_{\mathrm{k}}\| + 1)$
- $\bar{\mathbf{F}}(.)$ is a contraction w.r.t arbitrary norm $\left\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\right\| \leq \gamma\,\|\mathbf{x} - \mathbf{y}\|$

MEAN SQUARE BOUNDS

FIXED STEP SIZE

Markovian Stochastic Approximation $\quad \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha\left(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k\right)$

$$\left\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\right\| \leq \gamma \left\|\mathbf{x} - \mathbf{y}\right\|$$

ℓ_∞-norm contraction

$\|\mathbf{x}_0 - \mathbf{x}^*\|_\infty^2$

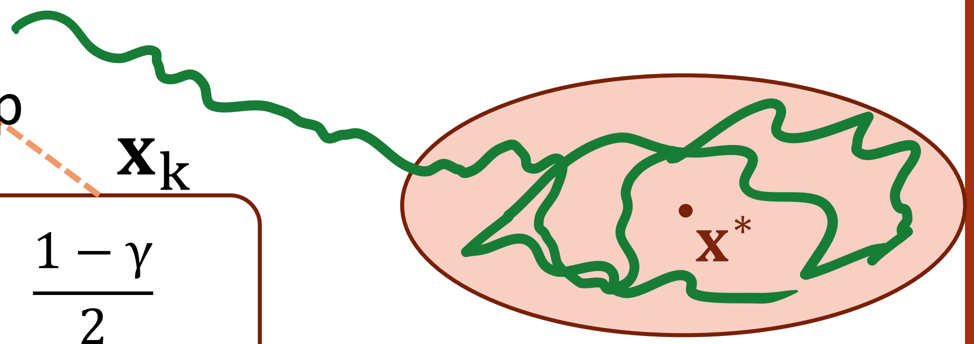
$\log d$

Theorem[Chen, M, Shakkottai, Shanmugam '21]: If the step-size α is small enough,

$$\mathbb{E}\left[\|\mathbf{x}_k - \mathbf{x}^*\|^2\right] \leq c_1(1 - c_2\alpha)^{k - \log\alpha^{-1}} + c_3\alpha\log\alpha^{-1}$$

- Given a target error ϵ, one can pick small enough step size so that eventually the mean square error is ϵ.
  - Mean Square sample complexity of $\tilde{O}\left(\frac{1}{\epsilon^2}\right)$

$\frac{1-\gamma}{2}$



DIMINISHING STEP SIZES

Markovian Stochastic Approximation $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$

$$\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\| \le \gamma \|\mathbf{x} - \mathbf{y}\|$$

$$\alpha_k \sim \frac{\alpha}{(k+h)^{\xi}}$$

Theorem[Chen, M, Shakkottai, Shanmugam '21]:

$$\mathbb{E}\left[\|\mathbf{x}_k - \mathbf{x}^*\|^2\right] \le \begin{cases} c_4 \dfrac{\ln k}{k^{\xi}} & \xi \in (0,1) \\ c_5 \dfrac{(\ln k)^2}{k^{\alpha c_2}} & \xi = 1, \alpha c_2 \le 1 \\ \hat{c}_6 \left(\dfrac{\log \mathrm{d}}{(1-\gamma)^3}\right) \dfrac{\ln k}{\mathrm{k}} & \xi = 1, \alpha c_2 > 1 \end{cases}$$

$\|\mathbf{x}_0 - \mathbf{x}^*\|_\infty^2$

$\dfrac{1-\gamma}{2}$

- This leads to a sample complexity of $\tilde{O}\left(\frac{1}{\epsilon^2}\right)$
  - With continual improvement beyond this.
  - Algorithm (choice of step-size) does not depend on ϵ

RELATED WORK

SA mode	Operator	Context	Literature
Additive noise	$\|.\|_2$-contraction	SGD	[Bottou et al 18]
Mult noise with boundedness	$\|.\|_\infty$-contraction	Q-learning	[Beck, Srikant 12,13] (poly d) (Need iterates to be bounded)
Linear	Hurwitz	TD-learning	[Srikant, Ying 19] (Markov Noise), [Lakshminarayanan and Szepesvari 18] (iid noise)
Markovian and Mult noise	Any norm contraction	SGD Q-learning TD-learning Off-policy TD	Our work Also recovers all prior

TAIL BOUNDS

TAIL BOUNDS

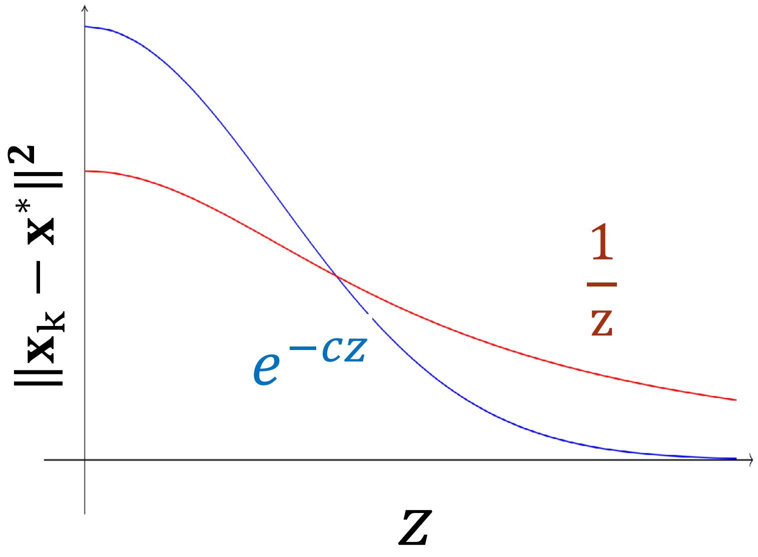
Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$$

Mean Square Bound:

$$\mathbb{E}[\|\mathbf{x}_k - \mathbf{x}^*\|^2] \leq O\left(\frac{1}{k}\right)$$

Using Markov Inequality, we get $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \geq O\left(\frac{1}{k}\right)z\right) \leq \frac{1}{z}$



Question: Can we get stronger tail bounds of the form

$$\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \geq O\left(\frac{1}{k}\right)z\right) \leq e^{-cz}?$$

YES in additive noise.

Not quite in multiplicative noise!

STOCHASTIC APPROXIMATION - ADDITIVE NOISE

Want to find $\mathbf{x}^*$ that solves

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbb{E}_{\mathbf{Y}\sim\boldsymbol{\mu}}[\mathbf{F}(\mathbf{x},\mathbf{Y})] = \mathbf{x}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{A}\mathbf{x}_k - \mathbf{b}_k)$$

Stochastic Approximation

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{F}(\mathbf{x}_k \quad) + \mathbf{w}_k - \mathbf{x}_k)$$

(Main) Assumptions

- Noise $\mathbf{w}_k$ - iid or martingale difference, mean zero, and is Sub Gaussian
- $\bar{\mathbf{F}}(.)$ is a contraction w.r.t arbitrary norm $\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\| \leq \gamma\,\|\mathbf{x} - \mathbf{y}\|$

ADDITIVE NOISE - EXPONENTIAL TAILS

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\alpha}{k+h}\left(\mathbf{F}(\mathbf{x}_k) + \mathbf{w}_k - \mathbf{x}_k\right)$$

Question: Can we get tail bounds of the form $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \geq O\left(\frac{1}{k}\right) z\right) \leq e^{-cz}$?

$$\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \geq O\left(\frac{1}{k}\right) O\left(\log\left(\frac{1}{\delta}\right)\right)\right) \leq \delta$$

Theorem[Zubeldia, Chen, Maguluri '23]: If α is large enough, for any $k \geq 0$, w.p. $(1-\delta)$,

$$\|\mathbf{x}_k - \mathbf{x}^*\|^2 \leq \frac{c}{k}\left(1 + \log\left(\frac{1}{\delta}\right)\right)$$

Sample complexity of $O\left(\frac{1}{\epsilon^2}\right)\log\left(\frac{1}{\delta}\right)$ to ensure $\|\mathbf{x}_k - \mathbf{x}^*\| \leq \epsilon$ w.p. $(1-\delta)$

This is a Gaussian like tail on the error $\|\mathbf{x}_k - \mathbf{x}^*\|$. $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\| \geq O\left(\frac{1}{\sqrt{k}}\right) z\right) \leq e^{-cz^2}$

MULTIPLICATIVE NOISE - THE CHALLENGE

- Linear SA to solve $\mathbf{Ax} = \mathbf{b}$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{A}_k\mathbf{x}_k - \mathbf{b}_k)$$

- Focus on multiplicative noise. Set $b_k = 0$, we get product of matrices

$$\mathbf{x}_{k+1} = \mathbf{x}_k(I + \alpha_k\mathbf{A}_k)$$

$\mathbb{E}[\mathbf{A}_k]$ is Hurwitz and
$\mathbb{E}[(I + \alpha_k\mathbf{A}_k)]$ is contraction

The matrix $(I + \alpha_k\mathbf{A}_k)$ is not a contraction. It is a contraction only in **expectation**.

- Mean Square bounds under constant step sizes: [Lakshminarayanan, Szepeswari '18] [Srikant, Ying '19]
- Tail Bounds under constant step sizes [Durmus et al '21]
  - Exponential tails if $\mathbf{A}_k$ is Hurwitz for all k. (i.e., assuming contraction at **all** times)
  - Polynomial tails otherwise.
  - Stationary distribution is heavy-tailed (Higher moments don't exist after a point) [Srikant, Ying '20]

We get exponential tails with diminishing step sizes and do it for general contractive SA

STOCHASTIC APPROXIMATION - MULTIPLICATIVE NOISE

Want to find $\mathbf{x}^*$ that solves

$$\bar{\mathbf{F}}(\mathbf{x}) = \mathbb{E}_{\mathbf{Y}\sim\boldsymbol{\mu}}[\mathbf{F}(\mathbf{x},\mathbf{Y})] = \mathbf{x}$$

$$\alpha_k = \frac{\alpha}{k+h}$$

Stochastic Approximation

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{F}(\mathbf{x}_k,\mathbf{Y_k}) + \quad - \mathbf{x}_k)$$

(Main) Assumptions

- $\mathbf{Y_k}$ is an iid process with stationary distribution $\boldsymbol{\mu}$
- With bounded support

- $\mathbf{Y}$
- $\|$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{A}_k\mathbf{x}_k - \mathbf{b})$$

If $\mathbf{A}_k$ is Gaussian, then, the MGF does not exist for $k \geq 3$

- $\bar{\mathbf{F}}(.)$ is a contraction w.r.t arbitrary norm $\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\| \leq \gamma\,\|\mathbf{x} - \mathbf{y}\|$

MULTIPLICATIVE NOISE – WEIBULLIAN TAILS

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\alpha}{k+h}(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) - \mathbf{x}$$

$\tilde{O}\left(\frac{1}{\epsilon^2}\right)\left(\log\left(\frac{1}{\delta}\right)\right)^M$ sample complexity

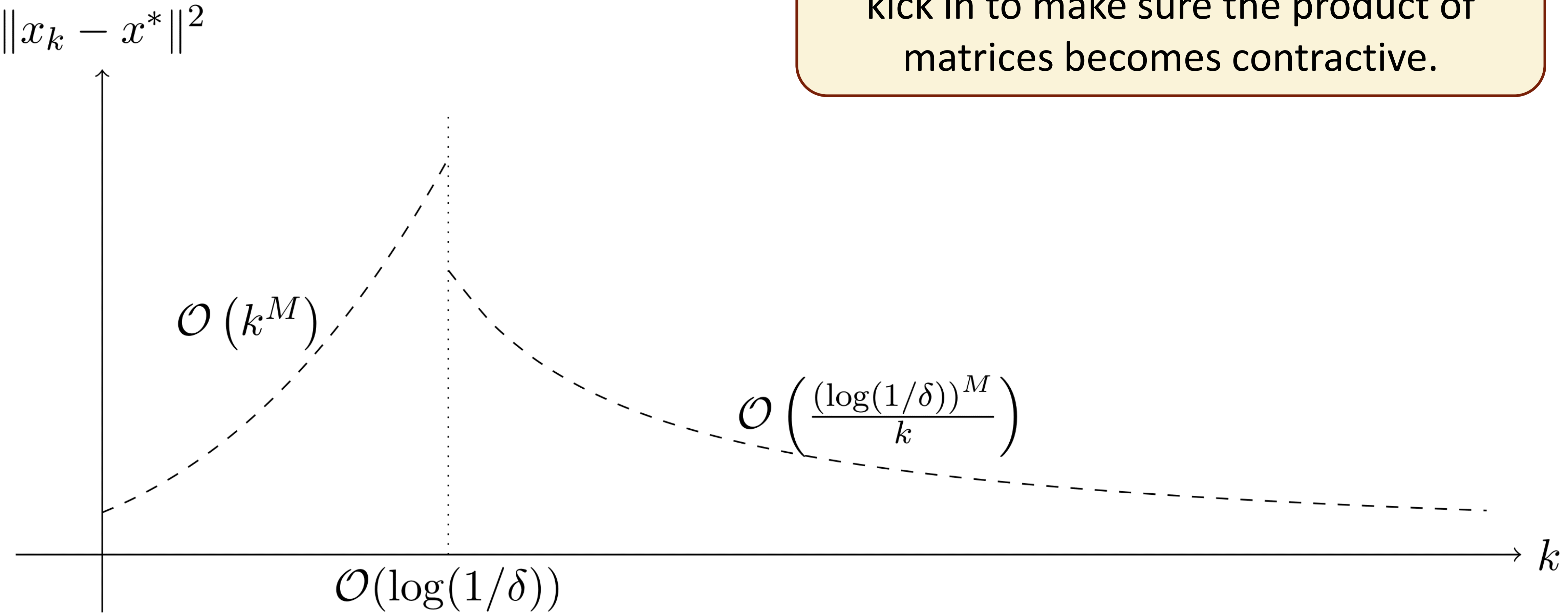
Theorem[Zubeldia, Chen, Maguluri '23]: For appropriate α, for a given k, w.p. $(1-\delta)$,

$$\|\mathbf{x}_k - \mathbf{x}^*\|^2 \le \begin{cases} \frac{c}{k}\left(1 + \left(\log\left(\frac{1}{\delta}\right)\right)^M\right) \\ \end{cases}$$

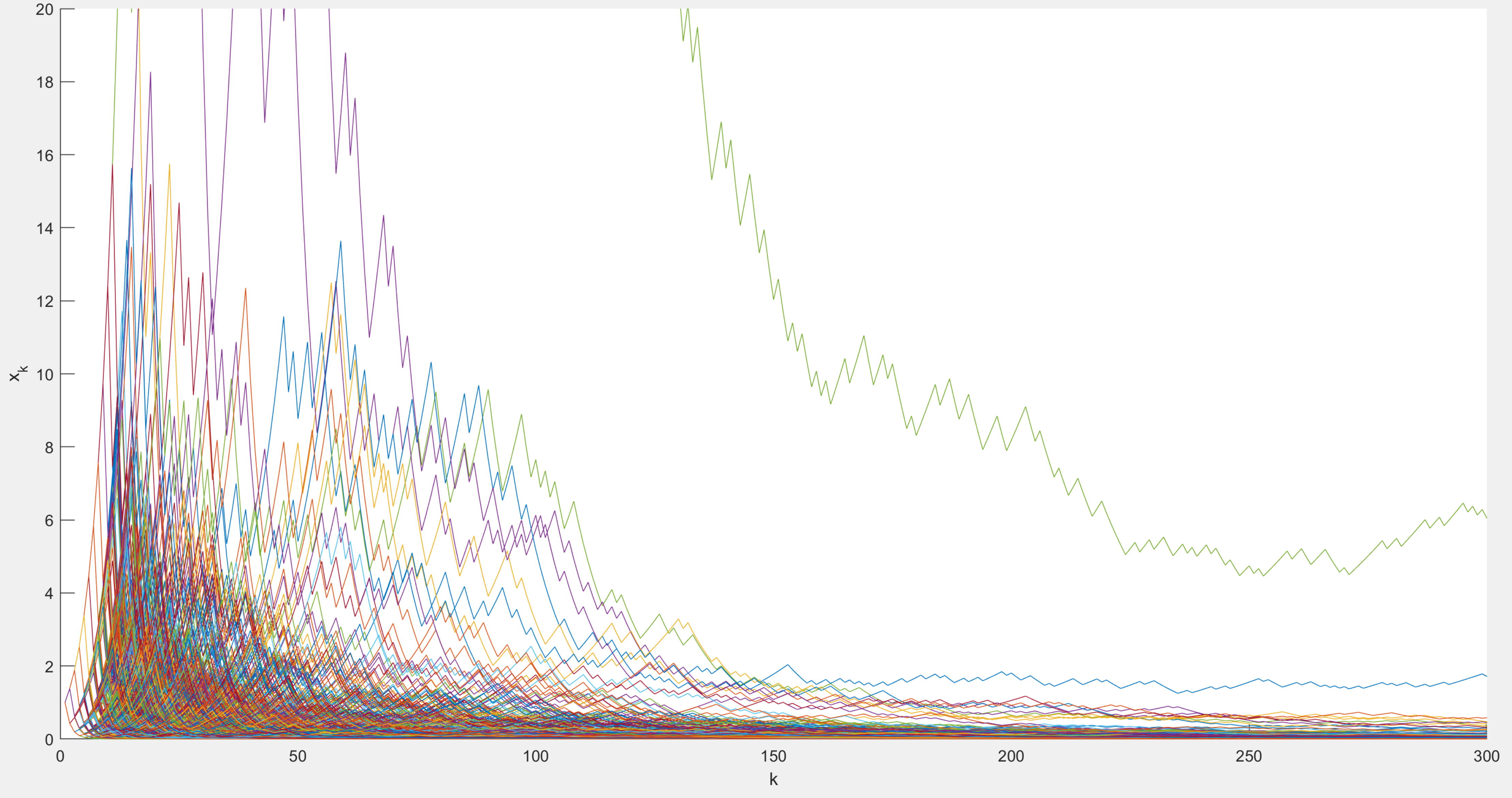
- M – integer ≥ 1 depends on how bad the bounded noise Y is (how expansive the operator can be)
- Corresponds to a tail of the form $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\| \geq O\left(\frac{1}{\sqrt{k}}\right) z\right) \leq e^{-cz^{\frac{2}{M}}}$
  - Weibullian tail (spans Gaussian, exponential and heavier – lighter than any ploynomial)
  - Counter example that (almost) matches this exponent.
- Why does the bound go up in the beginning?

WHY DOES THE ERROR GO UP?

Need enough samples for averaging to kick in to make sure the product of matrices becomes contractive.



ERROR GOES UP INDEED



ANY TIME CONCENTRATION

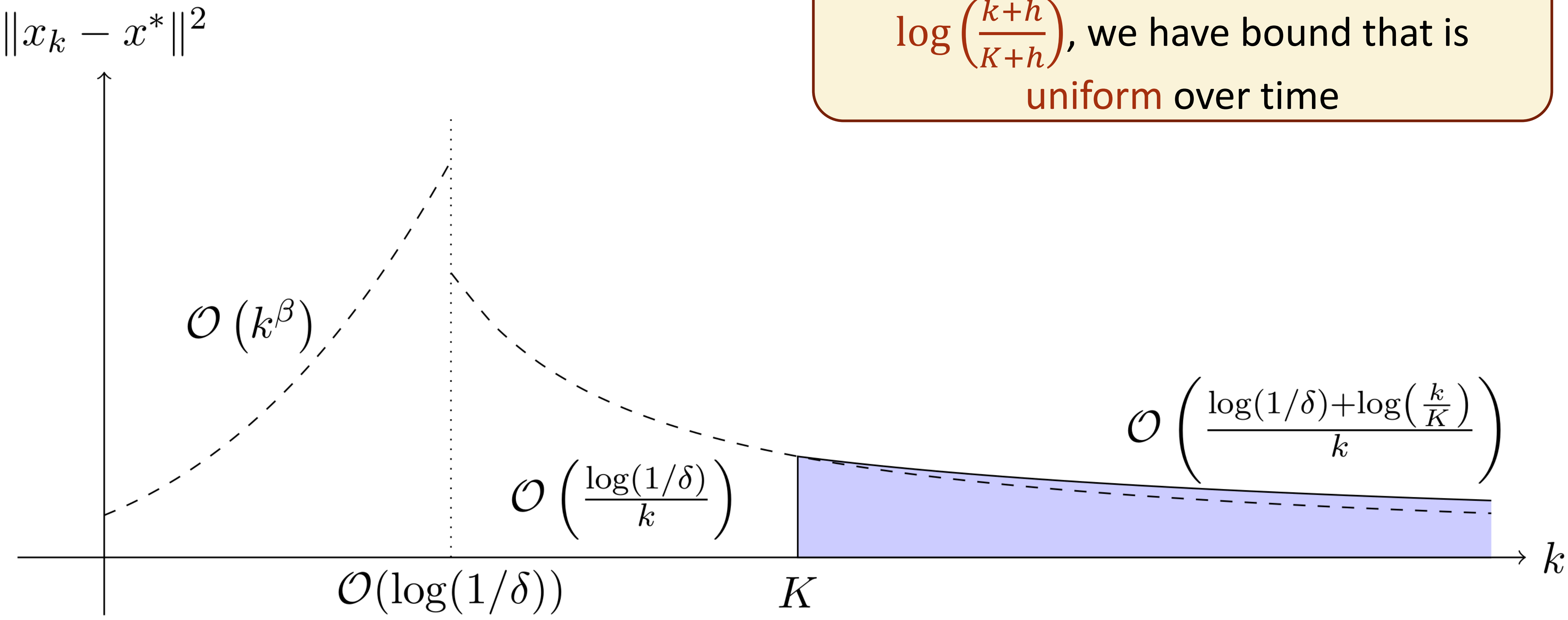
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\alpha}{k+h}\left(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k\right)$$

Theorem[Zubeldia, Chen, Maguluri '22]: For appropriate α, for a given K

$$\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \leq \begin{cases} \frac{c}{k}\left(1 + \left(\log\left(\frac{1}{\delta}\right)\right)^M\right) & if\ k \geq O\left(\log\left(\frac{1}{\delta}\right)\right) \\ k^\beta & \text{otherwise} \end{cases} \right) \geq (1-\delta)$$

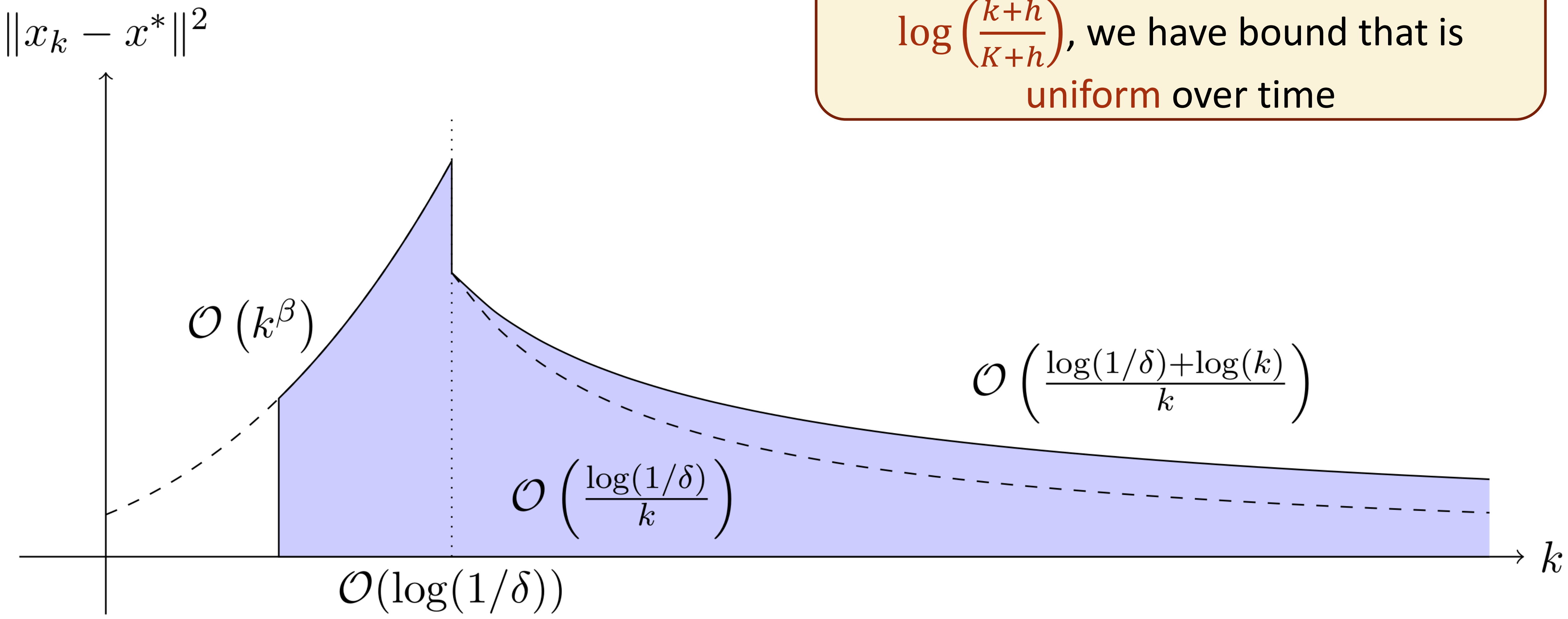
ANY TIME CONCENTRATION

With a small blowup factor of $\log\left(\frac{k+h}{K+h}\right)$, we have bound that is uniform over time



ANY TIME CONCENTRATION

With a small blowup factor of $\log\left(\frac{k+h}{K+h}\right)$, we have bound that is uniform over time



RELATED WORK

- Under boundedness
  - Either due to iterates being in compact set such as constrained optimization [Duchi et al ‘12], [Lan ‘20]
  - Or iterates are bounded due to other structural properties such as in Q Learning, [Evan-Dar et al ‘17], [Li et al ‘21], [Qu et al ‘20] or other related settings [Prashanth et al ‘21] [Thoppe et al ‘19], [Chandak ‘22]

- Constant Step Size that is picked as a function of ϵ and δ by obtaining a bound on just one point (or a window) of the tail
  - [Telgarsky ’22], [Mou et al ‘22], [Li et al ‘21]

- Result needs a bound on the iterates at some time n_0
  - [Thuppe et al ‘19], [Dalal ‘18]

- Our results in contrast, hold for potentially unbounded iterates, with diminishing step sizes and we bound the entire tail, without assuming any future bound.
  - Moreover, we allow for general norm contractions and we get anytime concentration.

PROOF SKETCH

MEAN SQUARE BOUNDS

STOCHASTIC APPROXIMATION: INTUITION

Stochastic Approximation $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$

Stochastic Approximation

$$\frac{\mathbf{x}_{k+1} - \mathbf{x}_k}{\alpha_k} = (\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$$

ODE

$$\dot{\mathbf{x}} = \left(\bar{\mathbf{F}}(\mathbf{x}) - \mathbf{x}\right)$$

- ODE Method [Borkar '09]:
  - Stochastic Approximation converges asymptotically if the ODE is globally asymptotically stable (gas)
  - Show gas using a Lyapunov function, $M(\mathbf{x}) = \|\mathbf{x}\|_\infty^2$: $\frac{dM(\mathbf{x}-\mathbf{x}^*)}{dt} \leq -\gamma M(\mathbf{x} - \mathbf{x}^*)$
- Want: Error bounds on original SA. We do not use the ODE method.
- Challenge: We need to handle error terms

Control the Errors

$$\underbrace{\mathbf{x}_{k+1} - \mathbf{x}_k}_{\text{Discretization Error}} = \alpha_k \Big(\underbrace{\bar{\mathbf{F}}(\mathbf{x}_k) - \mathbf{x}_k}_{\text{ODE Term}} + \underbrace{\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) - \bar{\mathbf{F}}(\mathbf{x}_k)}_{\text{Markovian Error}} + \underbrace{\mathbf{w}_k}_{\text{Additive Noise Error}} \Big)$$

ODE VS STOCHASTIC APPROXIMATION

Stochastic Approximation

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \alpha_k(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$$

ODE

$$\dot{\mathbf{x}} = \left(\bar{\mathbf{F}}(\mathbf{x}) - \mathbf{x}\right)$$

BAD NEWS

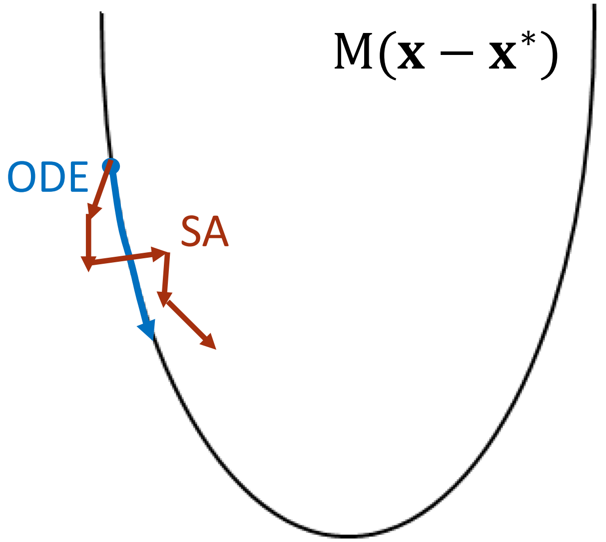
Lyapunov function $M(\mathbf{x}) = \|\mathbf{x}\|_\infty^2$ is not smooth

Smoothness: $M(\mathbf{y}) \leq M(\mathbf{x}) + \langle \nabla M(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2}\|\mathbf{y} - \mathbf{x}\|_\infty^2$

Approximation: $M(\mathbf{x}) \leq \|\mathbf{x}\|_\infty^2 \leq cM(\mathbf{x})$

$$\frac{dM(\mathbf{x} - \mathbf{x}^*)}{dt} \leq -\gamma M(\mathbf{x} - \mathbf{x}^*)$$

$$M(\mathbf{x}_{k+1} - \mathbf{x}^*) - M(\mathbf{x}_k - \mathbf{x}^*) \leq -\gamma\alpha_k M(\mathbf{x}_k - \mathbf{x}^*) + o(\alpha_k)$$



THE LYAPUNOV FUNCTION

Smoothness: $M(\mathbf{y}) \leq M(\mathbf{x}) + \langle \nabla M(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_\infty^{\mathbf{2}}$

Approximation: $M(\mathbf{x}) \leq \|\mathbf{x}\|_\infty^2 \leq cM(\mathbf{x})$

$$M(\mathbf{x}) = \|\mathbf{x}\|_\infty^2 \,\square\, \frac{1}{\mu} g(\mathbf{x}) = \min_{\mathbf{u}} \left\{ \|\mathbf{u}\|_\infty^2 + \frac{1}{\mu} g(\mathbf{x} - \mathbf{u}) \right\}$$

Moreau Envelope

$\|\mathbf{x}\|_\infty^2 \,\square\, \frac{1}{2\mu} \|\mathbf{x}\|_{\mathbf{2}}^2$

HANDLING THE ERRORS

Smoothness

$\|\mathbf{w}_k\| \le A(\|\mathbf{x}_k\| + 1)$

$$\underbrace{\mathbf{x}_{k+1} - \mathbf{x}_k}_{\text{Discretization Error}} = \alpha_k \left(\underbrace{\bar{\mathbf{F}}(\mathbf{x}_k) - \mathbf{x}_k}_{\text{ODE Term}} + \underbrace{\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) - \bar{\mathbf{F}}(\mathbf{x}_k)}_{\text{Markovian Error}} + \underbrace{\mathbf{w}_k}_{\text{Additive Noise Error}} \right)$$

- Due to smoothness, we are good, if we have a handle on Markovian Error
  - Exploit geometric mixing [Srikant, Ying '19] [Bertsikas, Tsitsiklis '96]

PROOF SKETCH

TAIL BOUNDS

PROOF SKETCH

- **Step 1 – Additive noise (or if iterates are bounded)**
  - Develop a proof framework based on Moreau envelope Lyapunov function to get exponential tails at a given time k (assuming the iterates are bounded).
- **Step 2 - Anytime concentration**
  - Generalize the result from Step 1 to get anytime concentration using Supermartingales and Ville's (Doob's) maximal inequality.
- **Step 3 - Bootstrapping**
  - Finally consider the real case of unbounded iterates, and use the previous two steps to inductively bootstrap from the worst case upper bound.

RECALL

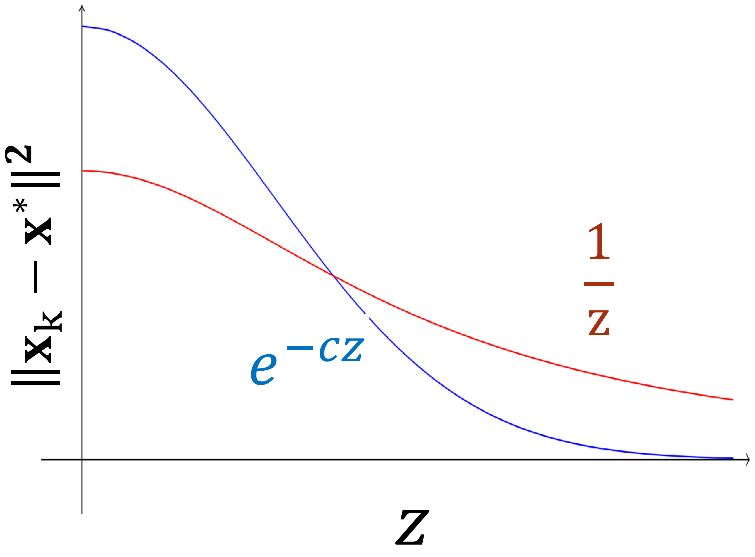
Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \;\; (\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k)$$

Mean Square Bound:

$$\mathbb{E}\left[\|\mathbf{x}_k - \mathbf{x}^*\|^2\right] \le O\left(\frac{1}{k}\right)$$

Using Markov Inequality, we get $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \ge O\left(\frac{1}{k}\right) z\right) \le \frac{1}{z}$



Question: Can we get stronger tail bounds of the form

$$\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2 \ge O\left(\frac{1}{k}\right) z\right) \le e^{-cz}?$$

YES in additive noise.

Not quite in multiplicative noise!

STEP 1: EXPONENTIAL TAIL BOUNDS

- Use $e^{\mathrm{M}(\mathbf{x})}$ as Lyapunov function to bound $\mathbb{E}\left[e^{\mathrm{M}(\mathbf{x}_\mathrm{k})}\right]$ and obtain tail bounds
  - Doesn't work – we don't get a recursion

Goal: $\mathbb{P}\left(k\|\mathbf{x}_\mathrm{k} - \mathbf{x}^*\|^2 \geq z\right) \leq e^{-cz}$

- Use $e^{\frac{k\mathrm{M}(\mathbf{x})}{\mathcal{B}}}$ as Lyapunov function to bound $\mathbb{E}\left[e^{\frac{k\mathrm{M}(\mathbf{x}_k)}{\mathcal{B}}}\right]$
  - $\mathcal{B}$ is the bound we assume on the iterates
  - Key trick: Incorporate the rate into the Lyapunov function
  - It works – We get a recursion (In the bounded case). Solving it, we get

$$\mathbb{E}\left[e^{k\mathrm{M}(\mathbf{x}_k)}\right] \leq c e^{o(1)\mathrm{M}(\mathbf{x}_0)}$$

- Applying Markov inequality, we get the exponential tail bounds.

STEP 2: ANY TIME CONCENTRATION

- Supermartingale - $\mathbb{E}[Z_{k+1}|\mathcal{F}_k] \leq Z_k$

$$\mathbb{P}\left(\sup_{k \geq K} Z_k > z\right) \leq \frac{\mathbb{E}[Z_K]}{z}$$

- Ville's (or Doob's) maximal inequality
- Lyapunov function, $e^{\frac{k\mathrm{M}(\mathbf{x}_k)}{\mathcal{B}}}$ is (almost) decreasing in expectation
  - because we incorporated the rate in it
  - Not quite – need to add a compensator term

$$e^{\frac{k\mathrm{M}(\mathbf{x}_k)}{\mathcal{B}} - c\log(k)} \text{ is a supermartingale}$$

- We get Anytime concentration (still assuming bounded iterates) using the maximal inequality
  - The compensator $\log\left(\frac{k}{K}\right)$ term gives the blowup factor of log in the result

STEP 3: BOOTSTRAPPING

$\mathbf{x}_k \leq \mathcal{B}$ for all k → **Step 1 and Step 2** → $\mathbf{x}_k \leq \tilde{O}\left(\frac{\mathcal{B}}{k}\right)$ for all k whp

When iterates $\mathbf{x}_k$ are not bounded, start with a worst case upper bound $\mathbf{x}_k \leq O\left(k^M\right)$ for all k

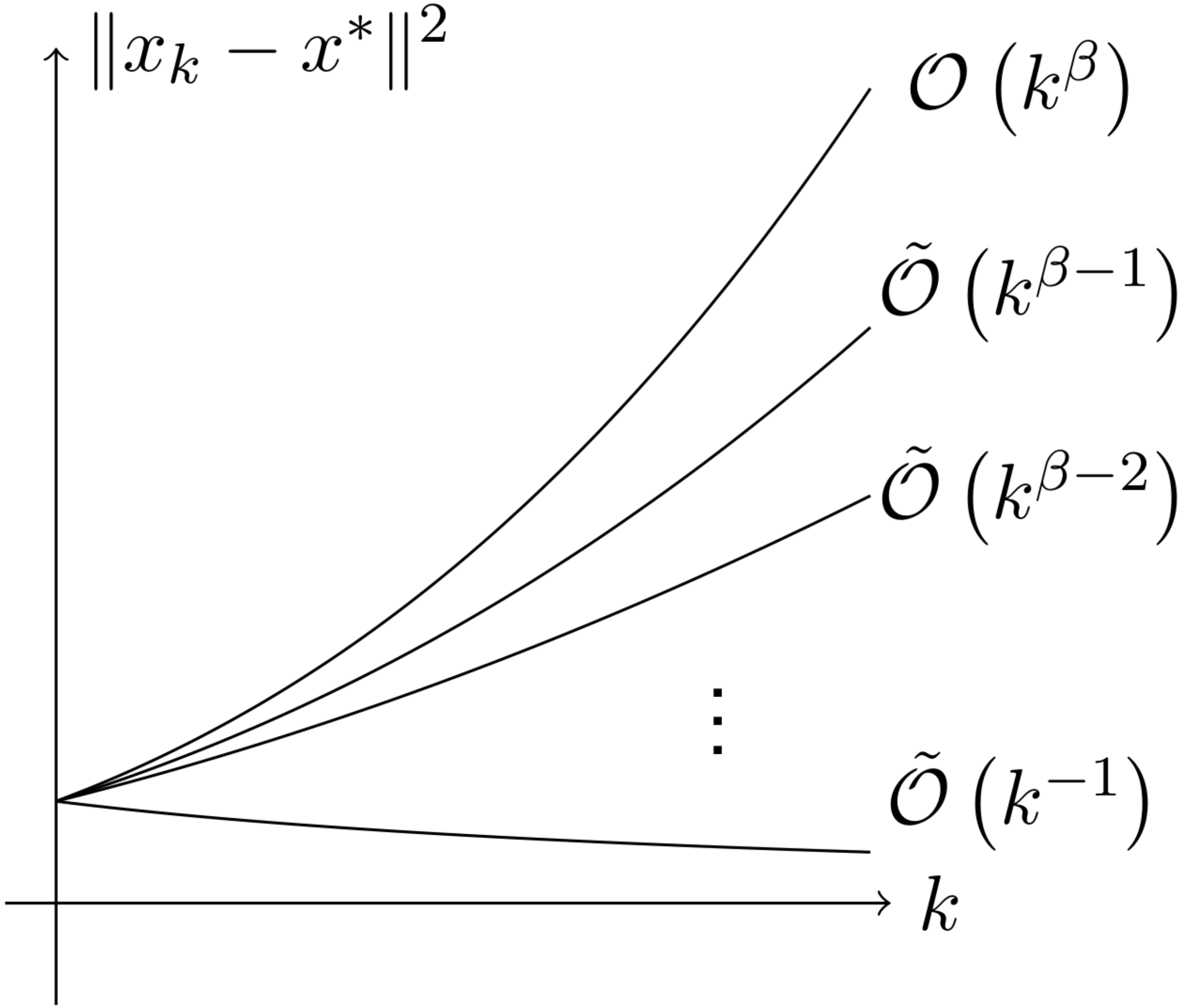
$\mathbf{x}_k \leq \tilde{O}\left(k^M\right)$ for all k whp → **Step 1 and Step 2** → $\mathbf{x}_k \leq \tilde{O}\left(k^{M-1}\right)$for all k whp

⋮

Bootstrap Inductively
Need Anytime Concentration

$\mathbf{x}_k \leq \tilde{O}(1)$ for all k whp → **Step 1 and Step 2** → $\mathbf{x}_k \leq \tilde{O}\left(\frac{1}{k}\right)$ for all k whp.

STEP 3: BOOTSTRAPPING



CONCLUSION

- Stochastic Approximation of a contractive operator under general norm
  - Both Additive and Multiplicative Noise
- Mean Square Convergence under Markovian Noise
  - $\tilde{O}\left(\frac{1}{k}\right)$ rate of convergence and $\tilde{O}\left(\frac{1}{\epsilon^2}\right)$ mean square sample complexity
  - Moreau Envelope of the norm square as the Lyapunov function
- Anytime Exponential Concentration under iid Noise
  - Additive noise: $O\left(\frac{1}{k}\right)$ rate Exponential tails and $O\left(\frac{1}{\epsilon^2}\right)\log\left(\frac{1}{\delta}\right)$ sample complexity
  - Multiplicative noise: $O\left(\frac{1}{k}\right)$ rate Weibullain tails and $O\left(\frac{1}{\epsilon^2}\right)\left(\log\left(\frac{1}{\delta}\right)\right)^M$ sample complexity
  - Proof based on Exponential supermartingales and Bootstrapping
  - Future work: Markovian noise (and a simpler proof?)

THANK YOU

Questions?

Theoretical Foundations of Reinforcement Learning

- Stochastic Approximation
- Off-Policy RL
- Actor-Critic
- Sample Complexity
- Average Reward
- Federated RL
- Multi Agent RL
- Concentration