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 Purdue

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

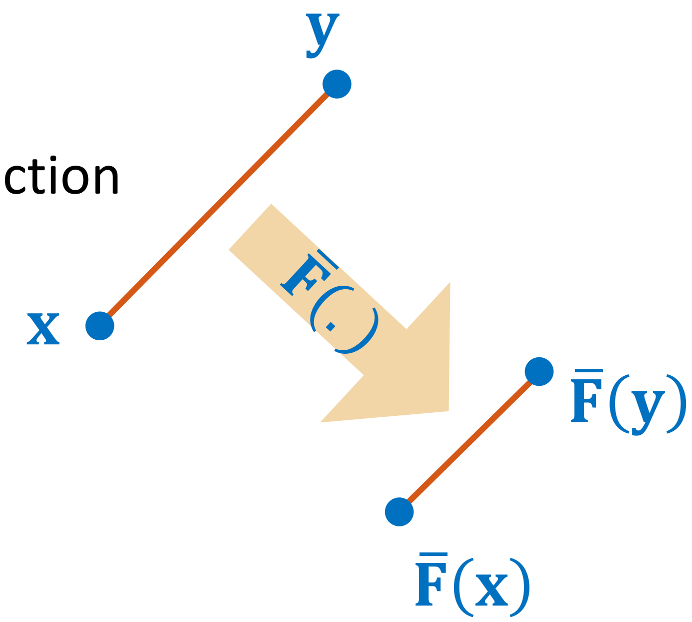
$$\mathbf{x}_{k+1} = \bar{\mathbf{F}}(\mathbf{x}_k) + \mathbf{w}_k$$

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



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]

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

Question: How well does this work?

OUTLINE

- Stochastic Approximation Introduction
  - Connection to Reinforcement Learning
- Finite Sample bounds on the mean-square error $\mathbb{E}[\|\mathbf{x}_k - \mathbf{x}^*\|^2]$
  - Illustration in Reinforcement Learning
  - Proof Sketch – A novel Lyapunov function
  - Other Applications and Generalizations
- High Probability (Tail) bounds on $\|\mathbf{x}_k - \mathbf{x}^*\|$
  - Proof Sketch - Exponential Supermartingale and Bootstrapping

STOCHASTIC APPROXIMATION

FIXED POINT PROBLEMS

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

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

Optimization:

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

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

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

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

FIXED POINT PROBLEMS

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

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

Linear Equations:

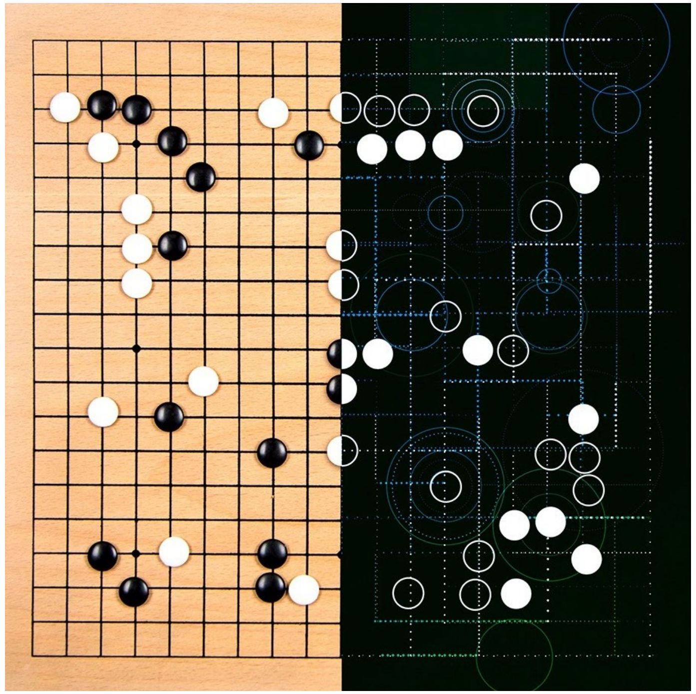
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

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

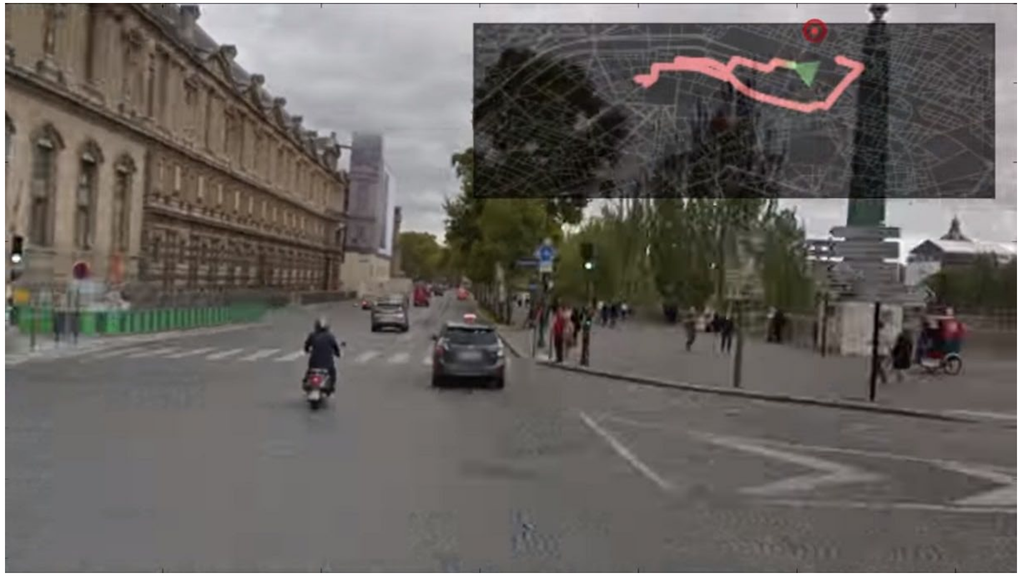
When $\mathbf{A}$ is Hurwitz ($\mathrm{Re}(\lambda_i) < 0$), $\bar{\mathbf{F}}(\mathbf{x}) = (\mathbf{I} + \eta\mathbf{A})\mathbf{x} - \eta\mathbf{b}$ is contraction wrt weighted ℓ_2-norm

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

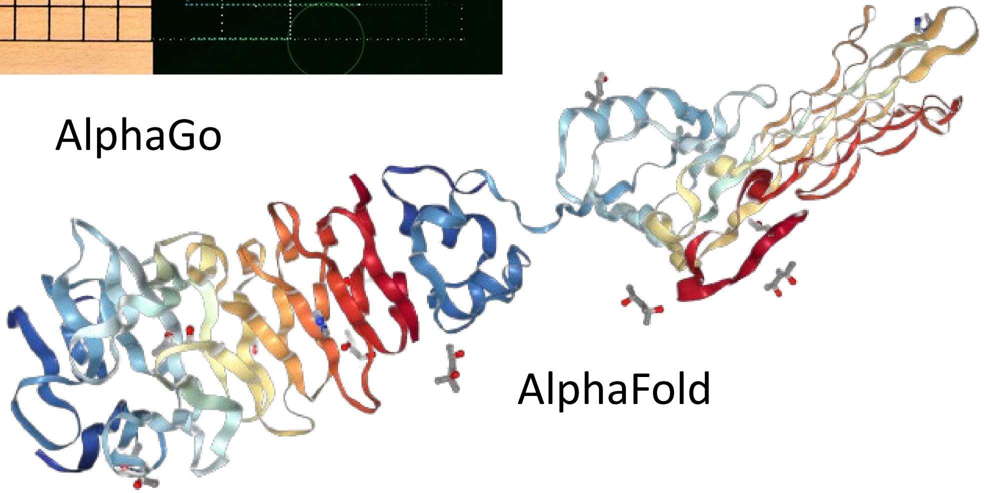
REINFORCEMENT LEARNING



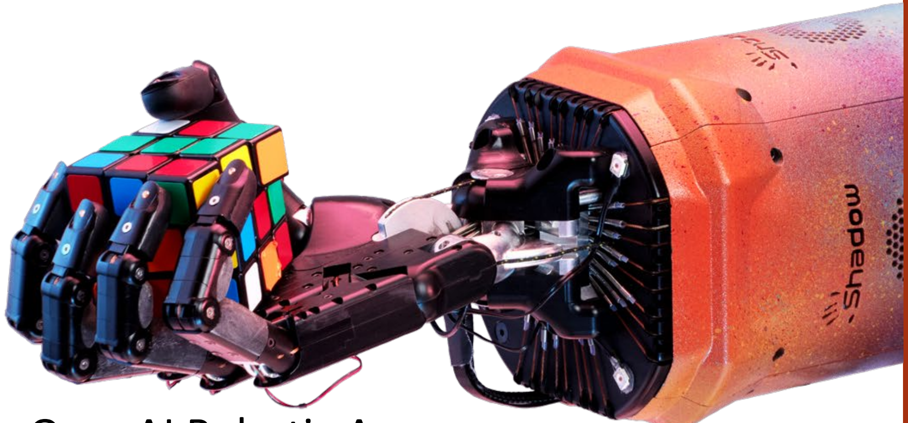
AlphaGo



Deepmind's StreetLearn Navigation



AlphaFold



OpenAI Robotic Arm

REINFORCEMENT LEARNING

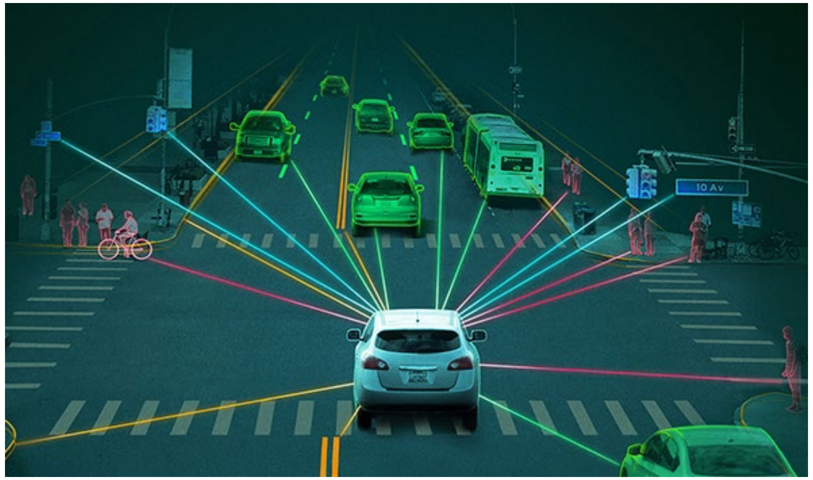
Policy π

Agent

Action A_k

State S_{k+1}

Reward R_k



Environment

Goal: Maximize long-term rewards

REINFORCEMENT LEARNING

State Value function

$$V_\pi(s) = \mathbb{E}\left[\sum_{k=0}^{\infty} \beta^k R(S_k, A_k) \middle| S_0 = s, A_k \sim \pi(S_k)\right]$$

State-Action Value function

$$Q_\pi(s, a) = \mathbb{E}\left[\sum_{k=0}^{\infty} \beta^k R(S_k, A_k) \middle| S_0 = s, A_0 = a, A_k \sim \pi(S_k)\right]$$

Goal - Control: Find the optimal policy π

Smaller Goal - Policy Evaluation: Evaluate the value function of policy π

BELLMAN EQUATION

Policy Evaluation

$$\underbrace{V_\pi(s) = \mathbb{E}[R(s,A) + \beta V_\pi(S_1) | S_0 = s, A \sim \pi(s)]}$$

$$\mathbf{V}_\pi = \mathbf{H}(\mathbf{V}_\pi)$$

Control Problem

$$\underbrace{Q^*(s,a) = \mathbb{E}\left[R(s,a) + \beta \max_{a'} Q^*(S_1,a') \middle| S_0 = s, A_0 = a\right]}$$

$$\mathbf{Q}^* = \mathbf{H}(\mathbf{Q}^*)$$

Need to solve the Fixed point Equation

When the operator $\mathbf{H}(.)$ not known

Q- LEARNING AND ASYNCHRONOUS SA

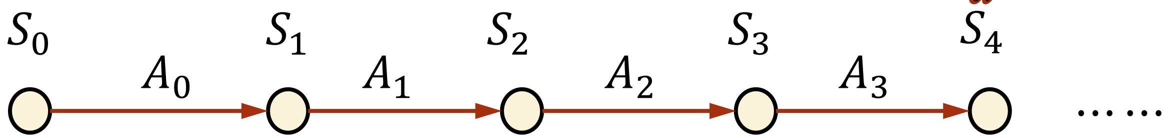
- Bellman Equation for the control problem:

$$Q^*(s,a) = \mathbb{E}\left[R(s,a) + \beta \max_{a'} Q^*(S_1,a') \,\middle|\, S_0 = s, A_0 = a\right]$$

$$\mathbf{Q}^* = \mathbf{H}(\mathbf{Q}^*)$$

- Q-Learning
  - Sample using a fixed policy π

$$Q_{k+1}(S_k,A_k) = Q_k(S_k,A_k) + \alpha_k\left[R(S_k,A_k) + \beta \max_a Q_k(S_{k+1},a) - Q_k(S_k,A_k)\right]$$



- Stochastic Approximation
  - Asynchronous SA: Update only one component. Other components are not updated

Markov Chain

$$\mathbf{Q}_{k+1} = \mathbf{Q}_k + \alpha_k(\mathbf{H}(\mathbf{Q}_k) + \mathbf{w}_k - \mathbf{Q}_k)$$

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

MARKOVIAN STOCHASTIC APPROXIMATION

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

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

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

Markovian Stochastic Approximation

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

Multiplicative Noise

Additive Noise

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 $\mathbf{w}_k$ - iid or martingale difference, mean zero, $\|\mathbf{w}_k\| \leq B(\|\mathbf{x}_k\| + 1)$
- $\bar{\mathbf{F}}(.)$ is a contraction w.r.t arbitrary norm $\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\| \leq \gamma\,\|\mathbf{x} - \mathbf{y}\|$
- $\mathbf{F}(.)$ is Lipschitz in $\mathbf{x}$ uniformly in $\mathbf{Y}$.

MEAN SQUARE BOUNDS

FIXED STEP SIZE

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

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

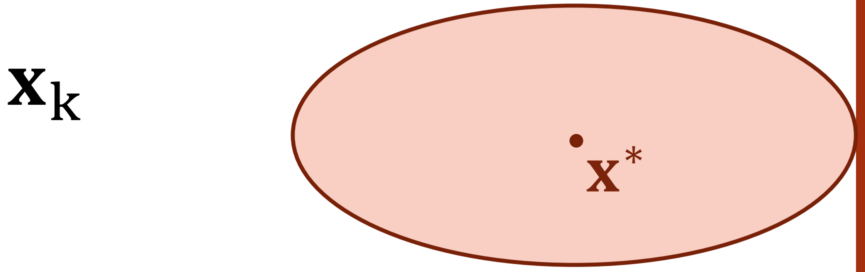
ℓ_∞-norm contraction

$\log d$

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

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

- 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 $\tilde{O}\left(\frac{1}{\epsilon^2}\right)$



DIMINISHING STEP SIZES

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

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

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

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

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

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

- 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 results

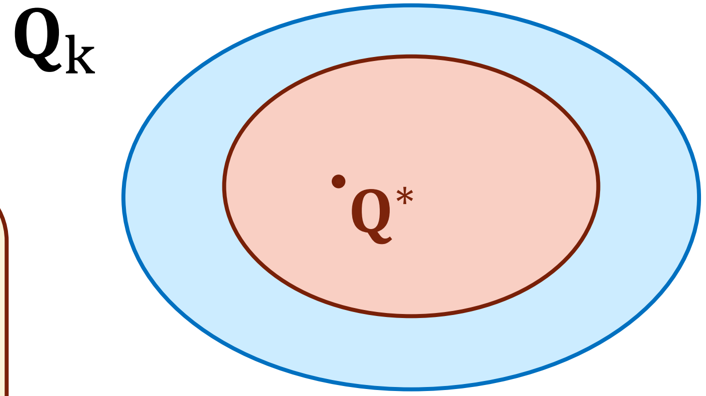
REINFORCEMENT LEARNING

Q-LEARNING: FINITE TIME BOUNDS

Q-Learning – using constant step size

$$Q_{k+1}(S_k, A_k) = Q_k(S_k, A_k) + \alpha \left[R(S_k, A_k) + \beta \max_{a'} Q_k(S_{k+1}, a') - Q_k(S_k, A_k)\right]$$

Corresponding $\bar{\mathbf{F}}(.)$ is a ℓ_∞-norm contraction



Theorem[Chen, M, Shakkottai, Shanmugam '21]: For appropriate choice of α,

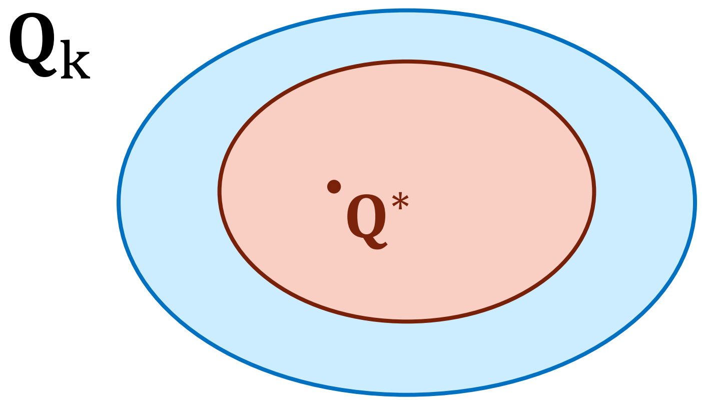
$$\mathbb{E}[\|\mathbf{Q}_\mathrm{k} - \mathbf{Q}^*\|_\infty^2] \leq \underbrace{\mathrm{c}_1(1 - \mathrm{c}_2\alpha)^{\mathrm{k} - \log \alpha^{-1}}}_{\frac{\epsilon}{2}} + \underbrace{\mathrm{c}_3 \alpha \log \alpha^{-1}}_{\frac{\epsilon}{2}}$$

Question: How many iteration do we need to get within ϵ of Q^*?

Q-LEARNING: SAMPLE COMPLEXITY

Q-Learning – using constant step size

$$Q_{k+1}(S_k, A_k) = Q_k(S_k, A_k) + \alpha_k \left[R(S_k, A_k) + \beta \max_{a'} Q_k(S_{k+1}, a') - Q_k(S_k, A_k) \right]$$



Sample Complexity: How many iteration do we need so that
$$\mathbb{E}\left[\|\mathbf{Q}_k - \mathbf{Q}^*\|_\infty\right] \leq \epsilon \text{ ?}$$

$(|\mathcal{S}||\mathcal{A}|)^3$

Theorem[Chen, M, Shakkottai, Shanmugam '21]: Sample complexity of Q learning is

$$\tilde{O}\left(\frac{1}{(1-\beta)^5}\right) \tilde{O}\left(\frac{1}{\mu_{\min}^3}\right) \tilde{O}\left(\frac{1}{\epsilon^2}\right)$$

RELATED WORK – Q LEARNING

- Constant step size [Beck, Srikant 12,13]
  - We improve by a factor of $(|\mathcal{S}||\mathcal{A}|)^2$ (best case scenario)
- Constant step size high probability bounds [Li, et al 20]
- Diminishing step sizes high probability bounds [Even-Dar etal ‘03], [Qu, Wierman ‘20]

CONTRACTIVE STOCHASTIC APPROXIMATION: APPLICATIONS

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

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

- Reinforcement Learning:
  - Q-Learning
  - TD-Learning – Policy Evaluation
    - N-step TD learning and TD(λ) – Efficiency of Bootstrapping
    - Off-policy TD learning – Bias-Variance Trade off [Chen, M, Shanmugam, Shakkottai '21]
  - Polyak-Ruppert Averaged Q learning [Li, Yang, Zhang, Jordan '21]
- Federated Reinforcement Learning [Khodadadian, Sharma, Joshi, M, '22]
- Robust Reinforcement Learning [Wang, Si, Blanchet, Zhou '23]
- Markov Chain Variance Estimation [Agrawal, LA, M '24]
- Stochastic Gradient Descent [Bottou et al '18]
- Linear Stochastic Approximation [Srikant, Ying '19]

Many more future applications!

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

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

- 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

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

Control the Errors

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

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

BAD NEWS

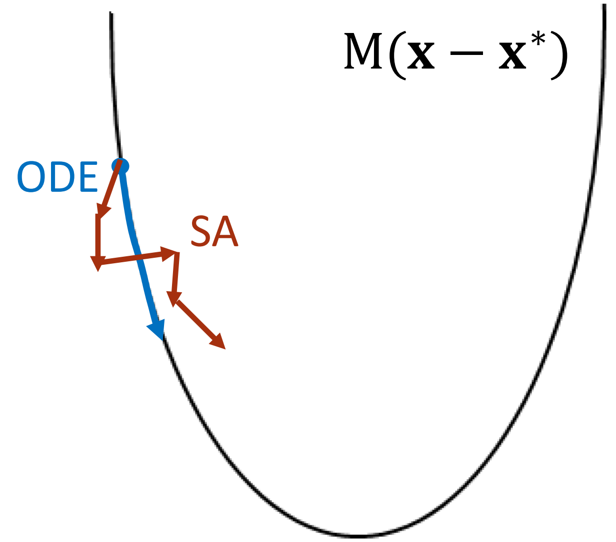
Lyapunov function $M(\mathbf{x}) = \|\mathbf{x}\|_\infty^2$ is not smooth

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

WISHLIST

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

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

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

Moreau Envelope

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

HANDLING THE ERRORS

Smoothness

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

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

- Due to smoothness, we are good, if we have a handle on Markovian Error
  - Exploit geometric mixing [Srikant, Ying '19]
  - Or use Poisson Equation Approach [Benveniste '90] [Haque, M '24]

CONTRACTIVE STOCHASTIC APPROXIMATION: GENERALIZATIONS

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

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

- Contraction wrt a seminorm [Zhang, Zhang, M ‘21]
  - Seminorm $p(\mathbf{x}) = 0$ does not imply $\mathbf{x} = 0$. Examples: $\|\mathbf{x}_{\parallel}\|$ or span($\mathbf{x}$)
  - Used to study Average Reward Reinforcement Learning
- Go beyond contraction [Nguyen, M ‘23]
  - If the ODE, $\dot{\mathbf{x}} = \left(\bar{\mathbf{F}}(\mathbf{x}) - \mathbf{x}\right)$ has a Lyapunov function with some rate
- Two Player Zero Sum Games
  - [Chen, Zhang, Mazumdar, Ozdaglar, Wierman ‘23]

STOCHASTIC APPROXIMATION: OPEN GENERALIZATIONS

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

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

- Contractive SA: Variance Reduction
  - [Mou, Khamaru, Wainwright, Bartlett, Jordan '22]
- Improving the State space dependance using Reverse Experience Replay
  - In Linear case [Agarwal, Chaudhuri, Jain, Nagaraj, Netrapalli '21]
- Nonexpansive operators $\|\bar{\mathbf{F}}(\mathbf{x}) - \mathbf{x}^*\| \leq \|\mathbf{x} - \mathbf{x}^*\|$
  - ℓ_2-norm is known [Chen, M, Shakkottai, Shanmugam '20]
- Two-time scale SA
  - Lot of work. Optimal rates are known only in special cases.
  - Polyak-Ruppert Averaging

TAIL BOUNDS

TAIL BOUNDS

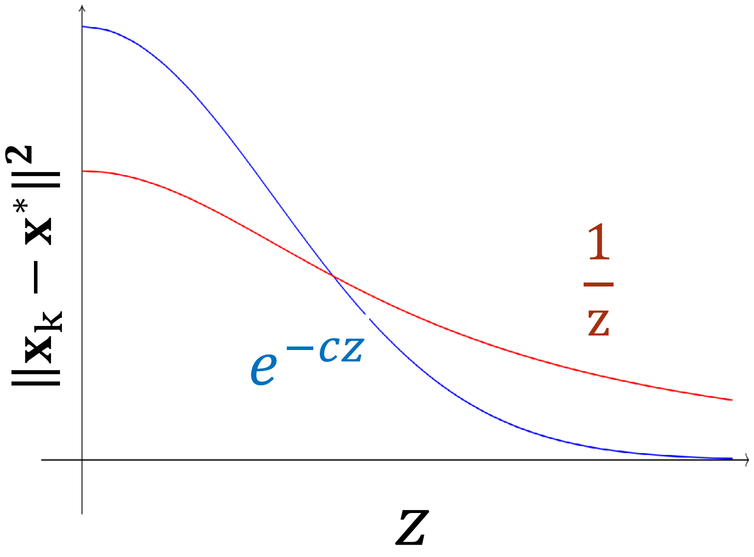
Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

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

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

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

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

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

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 $\|\bar{\mathbf{F}}(\mathbf{x}) - \bar{\mathbf{F}}(\mathbf{y})\| \leq \gamma\, \|\mathbf{x} - \mathbf{y}\|$
- $\mathbf{F}(.)$ is Lipschitz in $\mathbf{x}$ uniformly in $\mathbf{Y}$.

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. $(1-\delta)$

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

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

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

Stochastic Approximation

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

Assumptions

- $\mathbf{Y}_{\mathbf{k}}$ are iid with distribution $\boldsymbol{\mu}$ (or $\mathbf{F}(\mathbf{x}_{\mathrm{k}},\mathbf{Y}_{\mathbf{k}})$ is a martingale)
- With bounded support - More precisely, $\|\mathbf{F}(\mathbf{x}_{\mathrm{k}},\mathbf{Y}_{\mathbf{k}}) - \bar{\mathbf{F}}(\mathbf{y})\| \leq C(1+\|\mathbf{x}_{\mathrm{k}}\|)$

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

MULTIPLICATIVE NOISE – WEIBULLIAN TAILS

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

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

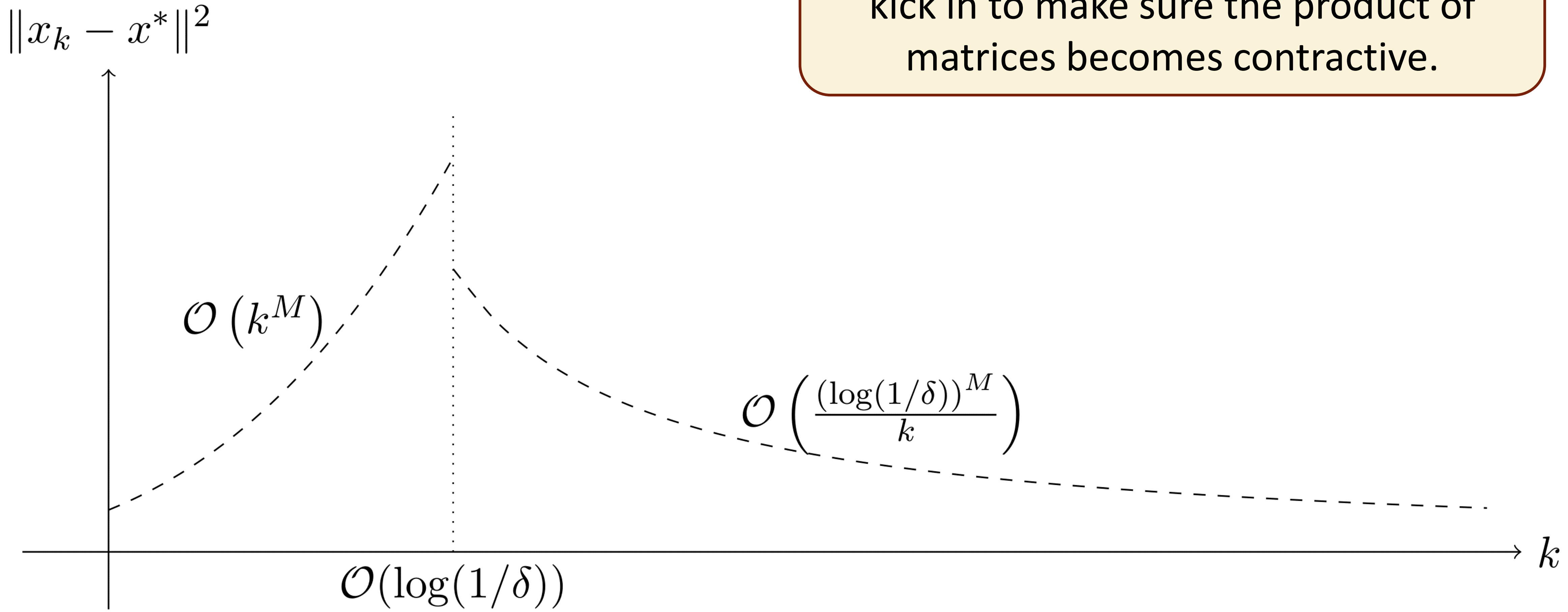
Theorem[Zubeldia, Chen, Maguluri '23]: For appropriate α, for a given k, w.p. $(1-\delta)$,

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

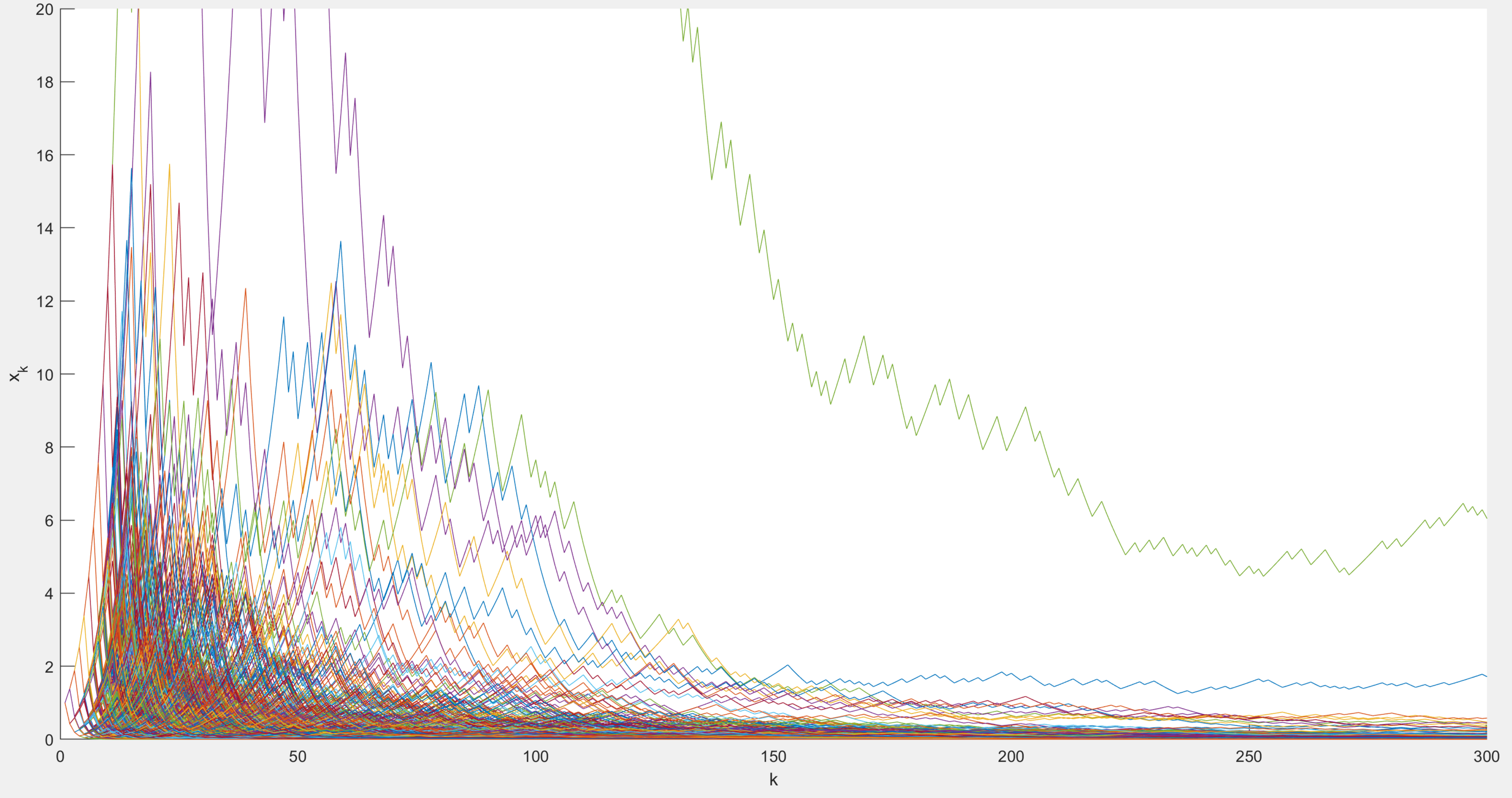
- M – integer ≥ 1 depends on how bad the bounded noise Y is (how expansive the operator can be)
- Corresponds to a tail of the form $\mathbb{P}\left(\|\mathbf{x}_k - \mathbf{x}^*\| \ge O\left(\frac{1}{\sqrt{k}}\right) z\right) \le e^{-cz^{\frac{2}{M}}}$
  - Weibullian tail (spans Gaussian, exponential 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

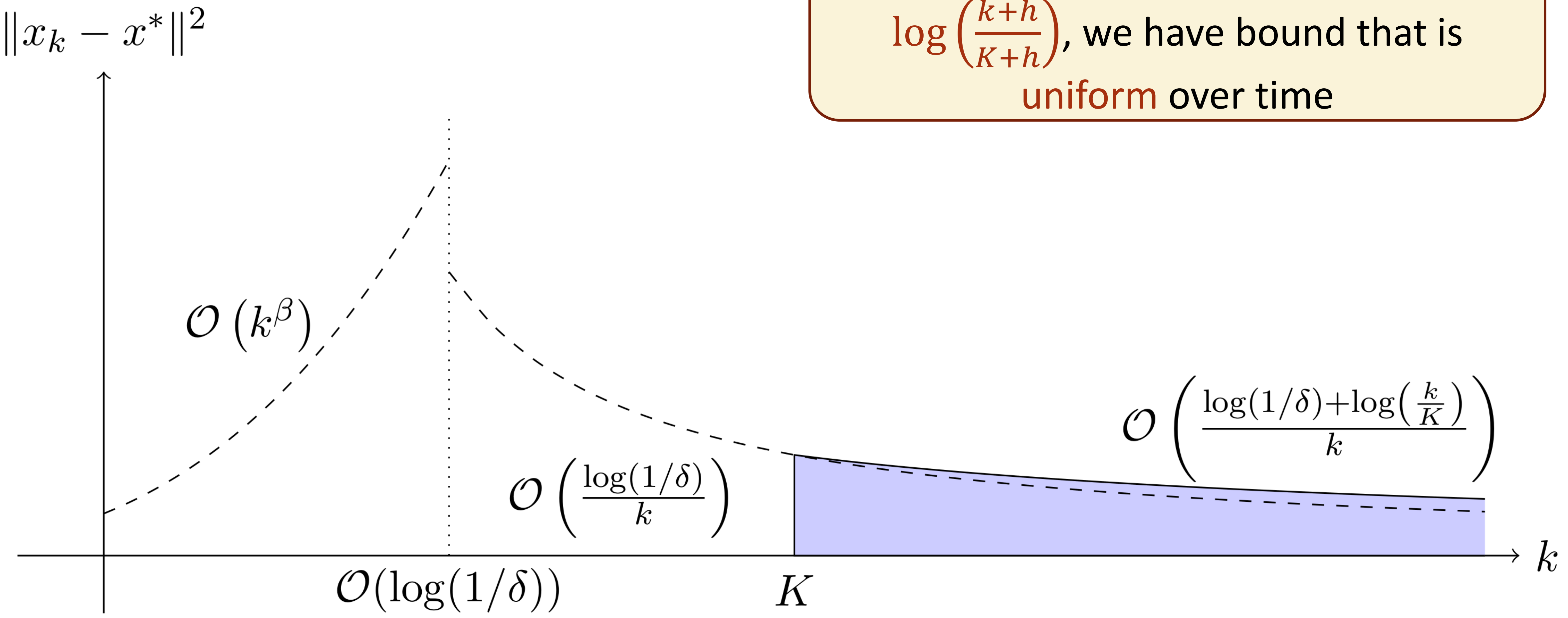
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\alpha}{k+h}\left(\mathbf{F}(\mathbf{x}_k, \mathbf{Y}_k) + \mathbf{w}_k - \mathbf{x}_k\right)$$

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

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

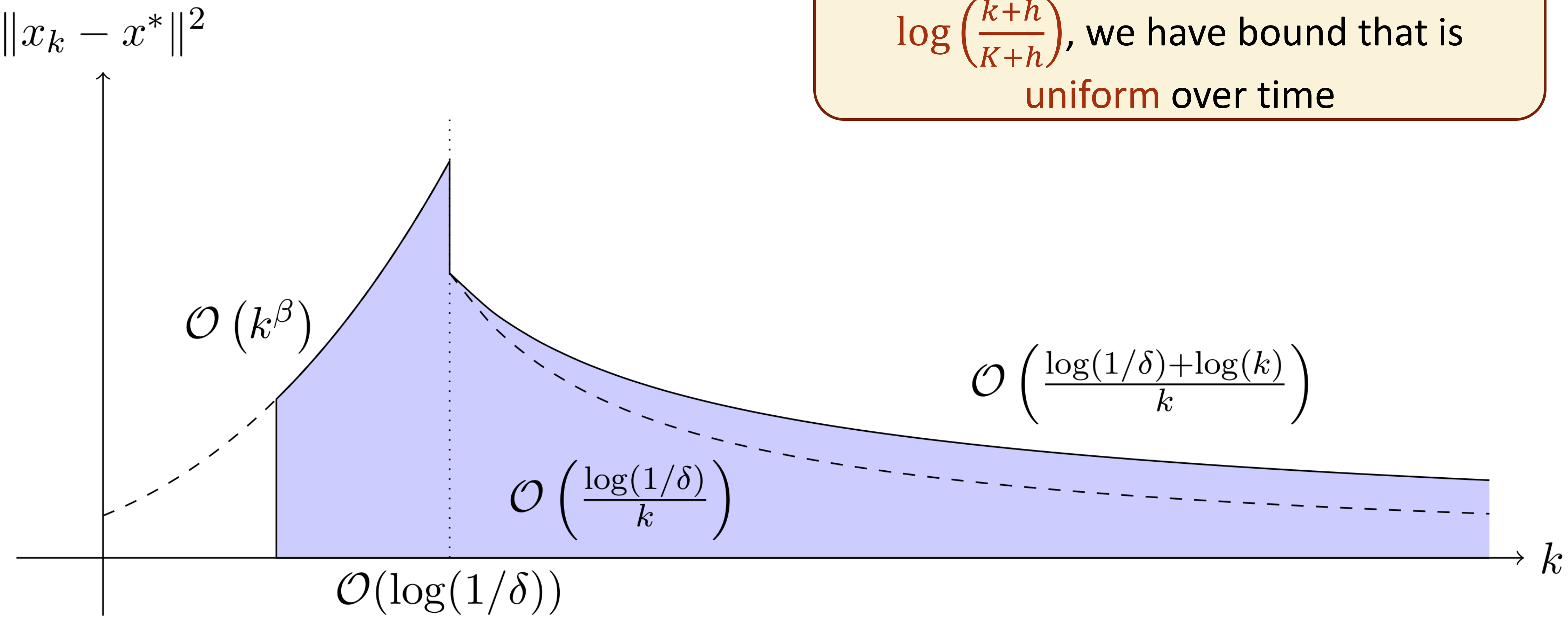
ANY TIME CONCENTRATION

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

TAIL BOUNDS

PROOF SKETCH

- **Step 1 – Additive noise case or if iterates are bounded**
  - Proof framework based on exponential Lyapunov function (and Moreau envelope)
- **Step 2 - Anytime concentration**
  - Using Ville’s (Doob’s) maximal inequality for supermartingales
- **Step 3 - Bootstrapping**
  - Inductively use the previous two cases

RECALL

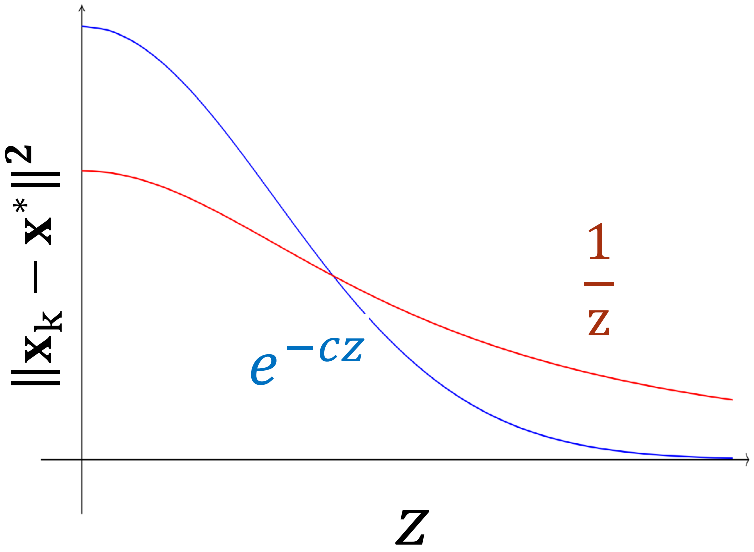
Stochastic Approximation to solve $\bar{\mathbf{F}}(\mathbf{x}) = \mathbf{x}$

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

Mean Square Bound:

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

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



Question: Can we get stronger tail bounds of the form

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

YES in additive noise.

Not quite in multiplicative noise!

STEP 1: EXPONENTIAL TAIL BOUNDS

- Use exponential Lyapunov function to bound MGF and obtain tail bounds.

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

- 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
  - We get a recursion (In the bounded case). Solving it, we get

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

- 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

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

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

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

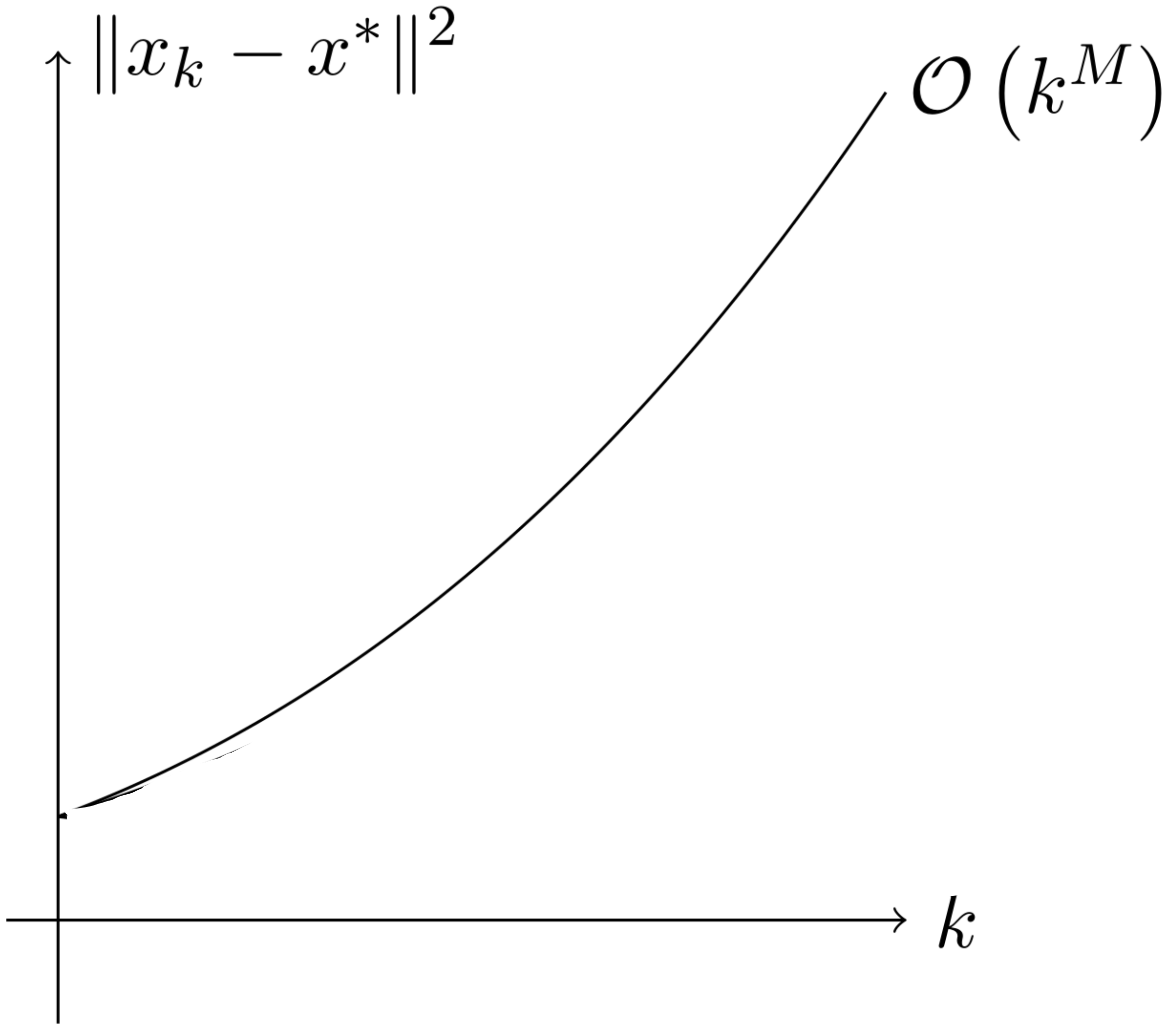
⋮

Bootstrap Inductively

Need Anytime Concentration

$\mathbf{x}_k \le \tilde{O}(1)$ for all k whp → **Step 1 and Step 2** → $\mathbf{x}_k \le \tilde{O}\left(\frac{1}{k}\right)$ 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 squared norm 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
  - Ongoing work: Both additive and multiplicative noise, Markovian noise (and a simpler proof?)

THANK YOU

Questions?