Proceedings of Machine Learning Research vol 125:1–51, 2020 33rd Annual Conference on Learning Theory

On Linear Stochastic Approximation: Fine-grained Polyak-Ruppert and Non-Asymptotic Concentration

Wenlong Mou WMOU@BERKELEY.EDU
Chris Junchi Li JUNCHILI@BERKELEY.EDU
Martin J. Wainwright WAINWRIG@EECS.BERKELEY.EDU
Peter L. Bartlett BARTLETT@CS.BERKELEY.EDU
Michael I. Jordan JORDAN@CS.BERKELEY.EDU
Department of EECS and Department of Statistics, UC Berkeley

Editors: Jacob Abernethy and Shivani Agarwal

Abstract

We undertake a precise study of the asymptotic and non-asymptotic properties of stochastic approximation procedures with Polyak-Ruppert averaging for solving a linear system $\bar{A}\theta = \bar{b}$. When the matrix $\bar{A}$ is Hurwitz, we prove a central limit theorem (CLT) for the averaged iterates with fixed step size and number of iterations going to infinity. The CLT characterizes the exact asymptotic covariance matrix, which is the sum of the classical Polyak-Ruppert covariance and a correction term that scales with the step size. Under assumptions on the tail of the noise distribution, we prove a non-asymptotic concentration inequality whose main term matches the covariance in CLT in any direction, up to universal constants. When the matrix $\bar{A}$ is not Hurwitz but only has non-negative real parts in its eigenvalues, we prove that the averaged LSA procedure actually achieves an $O(1/T)$ rate in mean-squared error. Our results provide a more refined understanding of linear stochastic approximation in both the asymptotic and non-asymptotic settings. We also show various applications of the main results, including the study of momentum-based stochastic gradient methods as well as temporal difference algorithms in reinforcement learning.

Keywords: Linear stochastic approximation, Polyak-Ruppert iteration averaging, TD learning, Momentum SGD, constant step size.

1. Introduction

Fixed-point algorithms based on stochastic approximation (SA) play a central role in a wide variety of disciplines (Robbins and Monro, 1951; Bertsekas and Tsitsiklis, 1989; Bottou et al., 2016; Lai, 2003). In general, given the goal of solving an underlying deterministic equation, SA methods perform updates based on randomized approximations to the current residual. An important special case is provided by stochastic gradient methods for optimization, which play an increasingly important role in large-scale machine learning and statistics (Nemirovski et al., 2009; Moulines and Bach, 2011).

Moving beyond the setting of optimization, there are many other kinds of problems in which stochastic approximation is a workhorse. For example, many problems in reinforcement learning involve the solution of fixed-point equations, and algorithms like TD (Sutton, 1988) and Q-learning (Watkins and Dayan, 1992) solve them via stochastic approximation. Moreover, even for stochastic optimization, accelerated methods that include momentum terms in their updates involve non-symmetric operators, and so require more general SA techniques for their analysis.

The celebrated Polyak-Ruppert averaging procedure (Polyak and Juditsky, 1992; Ruppert, 1988) stabilizes and accelerates stochastic approximation algorithms by taking an average over iterates. It is known that for suitably decaying step sizes, a central limit theorem (CLT) can be established for the averaged iterates. Moreover, Polyak-Ruppert averaging can achieve an optimal covariance, in the sense of local asymptotic minimaxity. Asymptotic results of this kind have provided the underpinnings for the development of online statistical inference methods. Recently, numerous non-asymptotic results have also been established in the settings of stochastic optimization (see Section 1.1). Notably, the work of Nemirovski et al. (2009); Moulines and Bach (2011) and Jain et al. (2017) gives non-asymptotic bounds for stochastic gradient methods as applied to (strongly) convex objectives; here the main term depends on the trace of the optimal covariance matrix.

There remains, however, a major mismatch between the classical CLTs and the non-asymptotic rates. Although the non-asymptotic results are valid for a finite number of iterations and are more reliable, they do lose some of the quantitative aspects of the CLT results. In particular, bounds on mean square error give much less information than the optimal covariance matrix, and the lack of high-probability bounds make them inapplicable in important applications such as policy evaluation. On the other hand, many important effects can vanish when the asymptotic limit is taken. In general, the trade-off between asymptotic limits and the rate of approach to asymptotic limits can be crucial. Such trade-offs should reflect the effect of the step size, and provide guidance for step-size selection.

In this paper, we consider the problem of linear stochastic approximation, where the goal is to solve a system, $\bar{A}\theta = \bar{b}$, of linear equations from noisy observations $(A_t, b_t)_{t=1}^{\infty}$. This problem is not only of intrinsic interest, within areas such as linear regression and TD learning, but it provides leverage on nonlinear SA problems, where analysis generally proceeds via local linearization.

In this paper, we make three primary contributions. First, we characterize the asymptotic covariance for the averaged iterates in the Polyak-Ruppert procedure for constant step size linear stochastic approximation. In addition to the classical $\bar{A}^{-1}\Sigma(\bar{A}^{-1})^\top$ term, we find a correction term that depends on the step size. A central limit theorem is shown for the averaged constant step-size procedure. Second, under stronger tail assumptions, we show a non-asymptotic concentration inequality for the averaged iterates in any direction, the leading term of which is the asymptotic covariance in this direction, while other terms keep the optimal rates. Thus, we achieve the best of both worlds. Finally, we show that even if the matrix A is not Hurwitz, as long as the real part of eigenvalues are non-negative, a non-asymptotic second moment bound is still valid for the Polyak-Ruppert procedure, again yielding a $1/\sqrt{T}$ rate. This goes beyond the regime of stable dynamical systems, and completes the picture of possibilities and impossibilities for linear stochastic approximation. When applied to momentum-based stochastic gradient descent (SGD) and temporal difference (TD) learning for value function estimation, our results capture many interesting phenomena, including the acceleration effect of momentum-based SGD, instance-dependent ℓ_∞-bounds for policy evaluation with near-optimal rates, and gap-independent results for the average-reward TD algorithm.

Technical overview: Similar to past work (Polyak and Juditsky, 1992; Ruppert, 1988), our analysis is based on representing the term $\bar{A}(\bar{\theta}_T - \theta^*)$ using a martingale to account for the noise at each step, where $\bar{\theta}_T$ denotes the averaged iterates. Our setting involves additional noise terms, due to the stochasticity in our observations of the matrix $\bar{A}$. As a consequence, the conditional covariance of the martingale difference terms at each step are dependent on the current iterate θ_t. Handling this issue requires the ergodicity of $\{\theta_t\}_{t\geq 0}$ as a Markov chain. Having established ergodicity, we can then prove an asymptotic result by combining Lindeberg-type CLTs with ergodic theorems.

In order to move from the asymptotic to the non-asymptotic setting, we study the projection of the iterate θ_T, for each time T, in some fixed but arbitrary direction. We can then apply the Burkholder-Davis-Gundy inequality to the higher moments of the supremum of a martingale, which separates the leading variance term and other terms that vanish at faster rates in T. Similar to the asymptotic case, the concentration results require a non-asymptotic bound on the deviation of the empirical averages of a function along a Markov chain, when compared to an expectation under the stationary distribution. In order to obtain such a bound, we exploit metric ergodic concentration inequalities (Joulin and Ollivier, 2010) combined with a coupling estimate.

In the case when the matrix $\bar{A}$ is not Hurwitz but has non-negative real parts in its eigenvalues, the process $\{\theta_t\}_{t\geq 0}$ does not generally approach θ^*. In the critical case, the dynamics is governed by a pure rotation with stochastic terms diffusing in all directions. However, when averaging is applied, both the effect of rotation and the random noise can be controlled. The step size is chosen to decay at the faster rate $1/\sqrt{T}$ in order to prevent an exponenential blowup.

Notation: For a matrix $W \in \mathbb{C}^{d\times d}$, we use $\{\lambda_i(W)\}_{i=1}^d$ to denote its eigenvalues. The spectral radius is given by $\rho(W) := \max_{i\in[d]} |\lambda_i(W)|$. For an invertible matrix W, we define the condition number $\kappa(W) = \|W\|_{\mathrm{op}} \cdot \|W^{-1}\|_{\mathrm{op}}$, where the operator norm is given by $\|W\|_{\mathrm{op}} := \sup_{\|x\|_2=1} \|Wx\|_2$. We use $a_T \lesssim b_T$ to denote $\forall T \geq 1, a_T \leq Cb_T$ for a universal constant $C > 0$. And we use $a_T \precsim b_T$ to denote $a_T \leq b_T \cdot \log^c(T/\delta)$ for a universal constant $c > 0$.

1.1. Related work

In the past decade, the growth of interest in stochastic gradient descent (SGD) has revived both theoretical and applied interest in stochastic approximation. There is a long line of work on the asymptotic regime of stochastic approximation algorithms (Ruppert, 1988; Polyak and Juditsky, 1992; Kushner and Yin, 2003; Borkar, 2008; Benveniste et al., 2012; Li et al., 2018). One core idea is that of averaging iterates along the path, which can be shown to have favorable statistical properties in the asymptotic setting (Ruppert, 1988; Polyak and Juditsky, 1992). (See, for instance, Theorem 1 in Ruppert (1988).) More recent papers (Chen et al., 2020; Su and Zhu, 2018; Liang and Su, 2019; Li et al., 2018) have developed iterative algorithms for constructing asymptotically valid confidence intervals for statistical problems.

In addition to asymptotic results, there are also a wide range of non-asymptotic results for stochastic approximation algorithms (see, e.g., Nemirovski et al. (2009); Rakhlin et al. (2012); Wang and Bertsekas (2016); Dieuleveut et al. (2017a,b); Jain et al. (2017, 2018, 2019); Lakshminarayanan and Szepesvari (2018)). Perhaps most closely related to our work is the analysis of Lakshminarayanan and Szepesvari (2018), who study linear stochastic approximation with constant step sizes combined with Polyak-Ruppert averaging. Relative to the analysis given here, their bound has a dependency on the Hurwitz parameter and condition number for eigenvector matrix in the leading term, which are sub-optimal. Moreover, the effect of the step size choice on the estimation error is not fully captured by the MSE bound. For more discussion about related works in stochastic optimization and reinforcement learning, see Appendix A.

2. Background and problem formulation

We begin by introducing the stochastic approximation algorithm to be analyzed in this paper, along with discussion of some of its applications.

2.1. Linear stochastic approximation

In this paper, we study stochastic approximation procedures for solving a linear system of the form $\bar{A}\theta = \bar{b}$, where the deterministic quantities $\bar{A} \in \mathbb{R}^{d\times d}$ and $\bar{b} \in \mathbb{R}^d$ are parameters of the problem. Throughout the paper, we assume that the matrix $\bar{A}$ is invertible, so that the solution θ^* to the equation exists and is unique. Suppose that we can observe a sequence of random variables of the form $\{(A_t, b_t)\}_{t\geq 1}$, assumed to be independent and identically distributed (i.i.d.), and exhibiting an unbiasedness property:

$$\mathbb{E}(A_t \mid \mathcal{F}_{t-1}) = \bar{A}, \quad \text{and} \quad \mathbb{E}(b_t \mid \mathcal{F}_{t-1}) = \bar{b}, \tag{1}$$

where $\mathcal{F}_{t-1}$ denotes the σ-field generated by $\{(A_k, b_k)\}_{k=1}^{t-1}$. Given observations of this form, our goal is to form an estimate $\widehat{\theta}$ of the solution vector θ^*. For some given initial vector θ_0, we consider the following linear stochastic approximation (LSA) procedure:

$$\theta_{t+1} = \theta_t - \eta(A_{t+1}\theta_t - b_{t+1}), \quad \text{for } t = 0, 1, 2, \ldots, \tag{2}$$

where $\eta > 0$ is a pre-specified step size. Our focus will be the Polyak-Ruppert averaged sequence $\{\bar{\theta}_T\}_{T\geq 1}$ given by

$$\bar{\theta}_T := \frac{1}{T}\sum_{t=0}^{T-1} \theta_t. \tag{3}$$

In particular, our goals are to establish guarantees for the renormalized error sequence $\sqrt{T}(\bar{\theta}_T - \theta^*)$, both in an asymptotic (i.e., $T \to \infty$) and non-asymptotic (i.e., finite T) setting.

2.2. Some motivating examples

Let us consider some applications that motivate the analysis of this paper. We begin with the simple example of stochastic gradient methods for linear regression:

Example 1 (Stochastic gradient methods for linear regression) Let $X \in \mathbb{R}^d$ be a vector of features, and let $Y \in \mathbb{R}$ be a scalar response. A linear predictor of Y based on X takes the form $\langle X,\, \theta\rangle = \sum_{j=1}^d X_j\theta_j$ for some weight vector $\theta \in \mathbb{R}^d$. If we view the pair (X, Y) as random, we can consider a vector θ^* that is optimal in the sense of minimizing the mean-squared error of the prediction—that is,

$$\theta^* \in \arg\min_{\theta\in\mathbb{R}^d} \mathbb{E}\Big[\Big(Y - \langle X,\, \theta\rangle\Big)^2\Big], \tag{4}$$

where $\mathbb{E}$ denotes an expectation over the joint distribution of (X, Y). A straightforward computation yields that θ^* must be a solution of the linear system $A\theta = b$, where $\bar{A} := \mathbb{E}[XX^\top] \in \mathbb{R}^{d\times d}$ and $\bar{b} := \mathbb{E}[XY] \in \mathbb{R}^d$. Note that θ^* exists and is unique whenever $\bar{A}$ is strictly positive definite.

In practice, we do not know the joint distribution of (X, Y), but might have access to a sequence of paired observations, say $\{(X_t, Y_t)\}_{t\geq 1}$, i.i.d. across different time instances t. The standard SGD algorithm computes an estimate of θ^* via the recursive update

$$\theta_{t+1} = \theta_t - \eta X_{t+1}\big(\langle X_{t+1},\, \theta_t\rangle - Y_{t+1}\big) \quad \text{for } t = 0, 1, 2 \ldots. \tag{5}$$

Note that this update is a special case of Eq (2), with the choices $A_t = X_t X_t^T$ and $b_t = X_t Y_t$. ♣

As a continuation of the previous example, let us consider a more sophisticated algorithm for online linear regression, one based on the introduction of an additional momentum component.

Example 2 (Stochastic gradient with momentum) For this particular example, let us adopt the shorthand $A_t = X_t X_t^T$ and $b_t = X_t Y_t$. Given a step size $\eta > 0$ and a momentum term $\alpha > 0$, consider a recursion over a pair $(\theta_t, v_t) \in \mathbb{R}^d \times \mathbb{R}^d$, of the following form:

$$\begin{cases} \theta_{t+1} = \theta_t - \eta v_t \\ v_{t+1} = v_t - \eta \alpha v_t + \eta (A_{t+1} \theta_{t+1} - b_{t+1}). \end{cases} \tag{6}$$

Let us reformulate these updates in the form (2), where we lift the problem to dimension $2d$ and use a tilde to denote lifted quantities. We find that the algorithm can be formulated as an update of the $2d$-dimensional vector $\widetilde{\theta}_t := \begin{bmatrix} \theta_t & v_t \end{bmatrix}^T \in \mathbb{R}^{2d}$ according to the recursion (2), where

$$\tilde{A}_t := \begin{bmatrix} 0 & I_d \\ -A_t & \alpha I_d + \eta A_t \end{bmatrix}, \quad \text{and} \quad \tilde{b}_t := \begin{bmatrix} 0 \\ -b_t \end{bmatrix}.$$

The underlying deterministic problem is to solve the $2d$-dimensional linear system $\tilde{A}\widetilde{\theta} = \tilde{b}$, where $\tilde{A} = \mathbb{E}[\tilde{A}_t]$ and $\tilde{b} = \mathbb{E}[\tilde{b}_t]$. It can be seen that $\theta^* \in \mathbb{R}^d$ is a solution to the original problem if and only if the vector $\widetilde{\theta}^* := \begin{bmatrix} \theta^* & 0 \end{bmatrix}^T$ is a solution to the lifted problem. In the sequel, we will use our general theoretical results to show why the addition of the momentum term can be beneficial. ♣

The area of stochastic control and reinforcement learning is another fertile source of stochastic approximation algorithms, and we devote our next two examples to the problems of exact and approximate policy evaluation.

Example 3 (TD algorithms in reinforcement learning) We now describe how the TD(0)-algorithm in reinforcement learning can be seen as an instance of the update (2). In this example, we discuss the TD algorithm for exact policy evaluation; in Example 4 to follow, we discuss the extension to TD with linear function approximation.

We begin by reviewing the background on Markov reward processes necessary to describe the problem; see Bertsekas (1995); Puterman (2005); Sutton and Barto (2018) for more details. We focus on a discrete Markov reward process (MRP) with D states; any such MRP is specified by a pair $(P, r) \in \mathbb{R}^{D \times D} \times \mathbb{R}^D$. The matrix $P \in \mathbb{R}^{D \times D}$ is row-stochastic, with entry $P_{ij} \in [0, 1]$ representing the probability of transitioning to state j from state i. The vector $r \in \mathbb{R}^D$ is the reward vector, with r_i denoting the reward received when in state i.

Discounted case: If future rewards are discounted with a factor $\gamma \in (0, 1)$, then the value function of the Markov reward process is a vector θ^* that solves the Bellman equation $\theta^* = r + \gamma P \theta^*$. This linear equation can be seen as a special case of our general set-up with

$$\bar{A} := I_D - \gamma P, \quad \text{and} \quad \bar{b} := r,$$

where I_D denotes the D-dimensional identity matrix.

There are various observation models in reinforcement learning, with one of the simpler ones being the *generative model*. In this setting, at each time t, we observe the following quantities:

- for each state $i \in [D]$, a random reward $R_{t,i}$ satisfying $\mathbb{E}[R_{t,i}] = r_i$. For simplicity, from now on, we assume that $R_{t,i} \in [-1, 1]$ almost surely, for any $i \in [D]$ and $t \geq 0$.
- for each state $i \in [D]$, a next state J is drawn randomly according to the transition vector $P_{i,\cdot}$.

We place this model in our general LSA framework by setting $b_t = R_t$ for each time t, and defining a random matrix $A_t \in \{0, 1\}^{D \times D}$ with a single one in each row; in particular, row i contains a one in position J, where J is the randomly drawn next state for i.

Average-reward case: Average-reward TD algorithm solves the fixed-point equation $\theta^* = r + P\theta^*$ via stochastic approximation. We assume the same generative model as in the discounted case. However, the matrix $\bar{A} = I - P$ is not invertible, with $\lambda_1(P) = 1$. In such case, the algorithm can be seen as LSA within the quotient space $\mathbb{R}^S/\mathrm{Ker}(\bar{A})$ (assuming the Markov chain is irreducible and consequently no multiplicity of eigenvalue 1, and $\dim(\mathrm{Ker}(\bar{A})) = 1$), by subtracting the mean. See Tsitsiklis and Van Roy (2002) for more details. ♣

Our framework can also be applied to TD with linear function approximation and stochastic quadratic minimax optimization. See Appendix B for detailed discussion with these examples.

3. Main results and their consequences

We now turn to the statements of our main results. We begin with the easier case when the matrix $\bar{A}$ is Hurwitz (meaning that all its eigenvalues have a positive real part), and provide both asymptotic and non-asymptotic guarantees for the Polyak-Ruppert sequence. Targeting Example 3, we also extend the non-asymptotic guarantees to the ℓ_∞ case with mild dimension dependency. We then turn to the more challenging critical case, in which the Hurwitz condition is violated (or the eigengap is too small to be quantitatively useful), and prove bounds on the mean-squared error. For all our results, we impose an i.i.d. condition:

Assumption 1 *The sequences $\{A_t\}_{t\geq 1}$ and $\{b_t\}_{t\geq 1}$ have* i.i.d. *entries.*

3.1. Asymptotic and Non-asymptotic Guarantees for Hurwitz Matrices

This section is devoted to guarantees that hold for a Hurwitz matrix.

Assumption 2 *The matrix $\bar{A} \in \mathbb{R}^{d\times d}$ is Hurwitz, meaning that*

$$\lambda := \min_{i\in[d]} \mathrm{Re}\left(\lambda_i(\bar{A})\right) > 0. \tag{7}$$

Our non-asymptotic statement involves various factors that pertain to properties that are implied by the Hurwitz condition. In particular, it is known (Perko, 2013) that any Hurwitz matrix is similar to a complex matrix D such that $D + D^{\mathsf{H}}$ is positive definite. Formally, we have:

Lemma 1 *For any Hurwitz matrix $\bar{A}$, there exists a non-degenerate matrix $U \in \mathbb{C}^{d\times d}$ such that $\bar{A} = UDU^{-1}$ for some matrix $D \in \mathbb{C}^{d\times d}$ that satisfies $D + D^{\mathsf{H}} \succeq \min_{i\in[d]} \mathrm{Re}(\lambda_i(\bar{A}))I_d$.*

For completeness, we provide a proof of this known result in Appendix K.1. From now on, we will use this decomposition for the Hurwitz matrix $\bar{A}$.

3.1.1. AN ASYMPTOTIC GUARANTEE

We begin with the asymptotic guarantee. In addition to Hurwitz condition on $\bar{A}$ and the i.i.d. assumption stated previously, this result requires second-moment control on the noise sequences $\Xi_t = A_t - \bar{A}$ and $\xi_t = b_t - \bar{b}$. (We denote by Ξ_A and ξ_b a generic random variable following the same distributions as Ξ_t and ξ_t.)

Assumption 3 *There exist finite scalars v_A^2 and v_b^2 such that*

$$\mathbb{E}\,\|\Xi_A u\|_2^2 \le v_A^2, \quad \textit{and} \quad \mathbb{E}|\xi_b^\top u|^2 \le v_b^2,$$

for any fixed vector u in the sphere $\mathbb{S}^{d-1}$. Moreover, the random elements Ξ_t and ξ_t are uncorrelated.

With these assumptions in place, we are now ready to state our first result, which is an asymptotic guarantee. We let Ξ_A denote a random matrix following the same distribution as each Ξ_t variable, and similarly, let ξ_b denote a random vector following the distribution of each ξ_t vector. Given these quantities, we define the following covariance matrix:

$$\Sigma^* := \mathrm{cov}(\xi_b + \Xi_A \theta^*) \;=\; \mathrm{cov}(\xi_b) + \mathrm{cov}(\Xi_A \theta^*). \tag{8}$$

Note that Σ^* is the sum of the covariances of the two kinds of noise involved in the stochastic approximation scheme. Given Σ^* and $\bar{A}$, we define a linear equation in a matrix variable Λ:

$$\bar{A}\Lambda + \Lambda\bar{A}^\top - \eta\bar{A}\Lambda\bar{A}^\top - \eta\mathbb{E}(\Xi_A\Lambda\Xi_A^\top) = \eta\Sigma^*. \tag{9}$$

As shown in the sequel (cf. Lemma 11), this matrix equation always has a unique PSD solution, which we denote by Λ_η^*. In fact, the matrix Λ_η^* corresponds to the covariance matrix of the stationary distribution of the Markov process $(\theta_t)_{t\ge 0}$.

Theorem 2 *Suppose that the matrix $\bar{A}$ is Hurwitz (Assumption 2), the i.i.d. condition (Assumption 1) and the second-moment condition (Assumption 3) hold, and the random elements A_t and b_t both have finite $(2+\delta)$-order moments for some $\delta > 0$. Then there exists a constant $\eta_0 > 0$ such that for any $\eta \in \big(0, \eta_0\big)$, we have*

$$\sqrt{T}(\bar{\theta}_T - \theta^*) \xrightarrow{d} \mathcal{N}\left(0, \bar{A}^{-1}\big(\mathbb{E}[\Xi_A\Lambda_\eta^*\Xi_A^\top] + \Sigma^*\big)(\bar{A}^{-1})^\top\right),$$

where the d-dimensional matrix Λ_η^* is the unique solution to equation (9).

See Appendix E for the proof of this theorem.

Note that when $\eta \to 0$, then equation (9) becomes a rescaled version of the classical Lyapunov equation $\bar{A}\Lambda + \Lambda\bar{A}^T = \eta\Sigma$, the solution of which specifies the stationary covariance matrix of a stochastic linear system. For suitably decaying step sizes, a minor extension[1] of arguments due to Polyak and Juditsky (1992) give an asymptotic statement involving the solution to the classic Lyapunov equation. On the other hand, for the constant step-size setting studied here, our result includes an additional correction term corresponding to the lingering effect of the non-zero step size. Theorem 2 specifies the asymptotic covariance matrix in this more general setting.

1. Such an extension is required to handle the randomness in A_t in addition to that in b_t.

When η is small, the matrix Λ_η^* scales linearly with η. The main term $\bar{A}^{-1}\Sigma^*(\bar{A}^{-1})^\top$ corresponds to the asymptotic limit of the classical Polyak-Ruppert averaging procedure. However, the effect of step size is not fully captured by the classical CLT. This additional term precisely characterizes the effect of step size on the asymptotic behavior of the averaged iterates.

As an important application of the general result in Theorem 2, we study SGD with momentum in Example 2. The momentum does not change the leading term in the asymptotic covariance matrix. On the other hand, compared to vanilla SGD, the momentum improves both the mixing time of the process and the correction term in the asymptotic covariance, by a factor of $\sqrt{\lambda_{min}(\bar{A})}$. See Section C.1 for more discussions.

3.1.2. NON-ASYMPTOTIC CONCENTRATION

As highlighted in classical Le Cam theory (cf. Van der Vaart (2000)), the asymptotic guarantee in Theorem 2 leads to asymptotic risk bounds in any fixed direction, and under any bowl-shaped loss function. It is natural to expect non-asymptotic concentration bounds that relate the error of $\bar{\theta}_T$ with that of a Gaussian random variable, up to some high-order terms, in any direction and under any gauge norm. This section gives an affirmative answer to the question.

For non-asymptotic concentration results, additional tail conditions need to be imposed on the noise distribution. In particular, we replace the second-moment bounds in Assumption 3 with the following stronger conditions:

Assumption 3′ *For some $p \geq 2$, there exist positive scalars $\sigma_A, \sigma_b, \alpha, \beta > 0$ such that for any u in the Euclidean sphere $\mathbb{S}^{d-1}$, we have*

$$\left(\mathbb{E}\,\|\Xi_A u\|_2^p\right)^{\frac{1}{p}} \overset{(i)}{\leq} p^\alpha \sigma_A, \quad \left(\mathbb{E}\left|\xi_b^\top u\right|^p\right)^{\frac{1}{p}} \overset{(ii)}{\leq} p^\beta \sigma_b. \tag{10}$$

Moreover, the noise components (Ξ_t and ξ_t) are uncorrelated.

The p-moment condition (10) with the parameters (α, β) provides a natural generalization of the notions of sub-Gaussian and sub-exponential tails (cf. Chap. 2, Wainwright (2019a)). Focusing on the inequality (ii) in the condition (10), the setting $\beta = \frac{1}{2}$ corresponds to a vector with sub-Gaussian tails, whereas the case $\beta = 1$ corresponds to the sub-exponential case. Generally, if we take the p-th power of a sub-Gaussian random variable, then it satisfies the condition (10) with exponent $p/2$.

Under these conditions, we can prove a result that gives a concentration guarantee at a given (finite) iteration T. The guarantee depends on the matrix U from Assumption 2 and Lemma 1 via its condition number, $\kappa(U) = \|U\|_{\mathrm{op}} \cdot \|U^{-1}\|_{\mathrm{op}}$. For a given iteration T and tolerance parameter $\delta \in (0, 1)$, we require a positive step size η that satisfies the bound

$$\eta < \frac{\lambda}{\rho^2(\bar{A}) + \kappa^2(U)\sigma_A^2 \log^{2\alpha+1}(T/\delta)}, \tag{11a}$$

where $\lambda = \min_{i \in [d]} \mathrm{Re}(\lambda_i(\bar{A})) > 0$ is the Hurwitz constant of $\bar{A}$, and $\rho(\bar{A})$ is its spectral radius.

Our result also involves the asymptotic covariance matrix from Theorem 2, namely the quantity

$$\Gamma^*(\eta) := \bar{A}^{-1}\left(\Sigma^* + \mathbb{E}(\Xi_A \Lambda_\eta^* \Xi_A^\top)\right)(\bar{A}^{-1})^\top. \tag{11b}$$

We bound the deviations of the rescaled process $\sqrt{T}(\bar{\theta}_T - \theta^*)$ in terms of the error term

$$\begin{aligned} \Delta(T, \delta) &:= V(\theta^*) \left(\frac{\sigma_A + \sigma_b}{T^{1/4}} + \frac{1 + \sqrt{\sigma_A / \lambda}}{\eta \sqrt{T}} \right) \log^{2\max(\alpha,\beta)+2}\left(\frac{T}{\delta}\right), \quad \text{where} \\ V(\theta^*) &:= \frac{\kappa^2(U)}{\min_{i \in [d]} |\lambda_i(\bar{A})|} \left\{ \|\theta^* - \theta_0\|_2 + \|\theta^*\|_2 + \sqrt{\tfrac{\eta}{\lambda}} \big(\sigma_A \|\theta^*\|_2 + \sigma_b \sqrt{d} \big) \right\}. \end{aligned} \tag{11c}$$

Given these definitions, we have the following non-asymptotic bound:

Theorem 3 *Fix an iteration number T and a tolerance $\delta \in (0, 1/T)$, and suppose that the i.i.d. condition (Assumption 1), higher-order moment condition (Assumption 3′), and Hurwitz condition all hold (Assumption 2). Then there exists a constant $c > 0$ such that for any step size $\eta > 0$ satisfying the bound (11a) and for any $v \in \mathbb{S}^{d-1}$, we have*

$$\mathbb{P}\left[\sqrt{T} \big| v^\top (\bar{\theta}_T - \theta^*) \big| \leq c \sqrt{\log(\tfrac{1}{\delta})} \Big\{ \sqrt{v^\top \Gamma^*(\eta) v} + \Delta(T, \delta) \Big\} \right] \geq 1 - \delta, \tag{12}$$

where the matrix $\Gamma^*(\eta)$ and deviation term $\Delta(T, \delta)$ are defined in Eq (11b), (11c), respectively.

See Appendix F for the proof of this theorem.

Remarks: A few comments are in order: first, we note that the leading term of $\sqrt{v^T \Gamma^*(\eta) v}$ of this non-asymptotic bound matches the term arising from the asymptotic covariance in Theorem 2, up to universal constants and the $\sqrt{\log(1/\delta)}$ term. This matches the behavior of a Gaussian random vector following the asymptotic distribution in Theorem 2 up to universal constants. Second, although the step size is required to belong to an interval depending on T and δ, the dependence is only logarithmic. In fact, our step-size condition (11a) differs only by these logarithmic factors from the stability threshold $\frac{\lambda}{\rho^2(\bar{A}) + \kappa^2(U) v_A^2}$, assuming σ_A and v_A are of the same order.

Second, in the definition of $\Delta(T, \delta)$, observe that the $\frac{1}{\sqrt{T}}$ term is accompanied by a $\frac{1}{\eta}$ dependence, while the $T^{-\frac{1}{4}}$ term does not diverge as $\eta \to 0^+$. This behavior is natural, because the former comes from the ergodicity of the process $\{\theta_t\}_{t=0}^{\infty}$, while the latter comes from the concentration.

Finally, let us consider the issue of how to set the step size η as a function of T so as to achieve an optimal bound for this pre-specfied T. Note that the step-size-dependent term from the matrix $\Gamma^*(\eta)$ scales linearly in η. Collecting the terms from $V(\theta^*)$ and $\Delta(T, \delta)$ that depend on the pair (T, η), we arrive at a bound that scales as

$$\underbrace{\eta}_{\text{From } \Gamma^*(\eta)} + \underbrace{\sqrt{\eta} \left\{ \frac{1}{T^{1/4}} + \frac{1}{\eta \sqrt{T}} \right\}}_{\text{From } \Delta(T, \delta)}.$$

In order to minimize this bound, the optimal choice is to set $\eta = T^{-1/3}$, which leads to the overall error scaling as $T^{-1/3}$. Thus, with this scaling, we can conclude that Theorem 3 guarantees a high-probability bound of the form

$$\sqrt{T} \big| v^\top (\bar{\theta}_T - \theta^*) \big| \precsim \sqrt{v^\top \bar{A}^{-1} (\Sigma^*) (\bar{A}^{-1})^\top v} + \mathcal{O}\left(T^{-1/3}\right),$$

where the notation $\precsim$ denotes inequality up to constants and logarithmic factors in (T, δ).

In addition, we note that Theorem 3 is useful for TD learning with linear function approximation. See Example 4 and Appendix C.2 for more details.

Constructing non-asymptotic confidence sets: The classical Polyak-Ruppert procedure gives a locally asymptotically-optimal covariance matrix, which can also be used for the construction of asymptotic confidence sets. Theorem 3 has analogous consequences for purposes of non-asymptotic inference. When going from asymptotically valid inference methods to the non-asymptotic counterparts, Berry-Esseen-type estimates are often used. But the sizes of confidence sets constructed in this way have polynomial dependence on the confidence level δ, even if the data themselves are not heavy-tailed. When a large number of confidence sets or tests are needed to be constructed, the size of each confidence set can expand in a rapid way. In contrast to this undesirable behavior, we now show how Theorem 3 yields a confidence set with better dependence on the confidence level.

Using the notation of Theorem 3, we define the positive definite matrix

$$B(T, \delta) := \Gamma^*(\eta) \log(\tfrac{d}{\delta}) + \Delta(T, \tfrac{\delta}{d}) I_d, \tag{13}$$

and the associated weighted Euclidean norm $\|v\|_{B(T,\delta)} = \sqrt{v^\top B(T, \delta) v}$. Using this weighted norm, we then define an ellipse that yields a confidence set that has coverage $1 - \delta$.

Corollary 4 *Under the conditions of Theorem 3, there is a universal known constant $c > 0$ such that the ellipse $\mathcal{E}(T, \delta) = \left\{ \theta \in \mathbb{R}^d \mid \|\theta - \bar{\theta}_T\|_{B(T,\delta)} \leq c \sqrt{d/T} \right\}$, centered at the averaged iterate $\bar{\theta}_T$, has the coverage guarantee $\mathbb{P}\left[\mathcal{E}(T, \delta) \ni \theta^*\right] \geq 1 - \delta$.*

From the definition (13) of the ellipse parameters (recalling the definition of $\Delta(T, \delta)$ from equation (11c), it can be seen that the size of our confidence set depends only logarithmically (as opposed to polynomially) on $1/\delta$. In terms of computing the confidence ellipse $\mathcal{E}(T, \delta)$, an obstacle is the fact that the the matrix $\Gamma^*(\eta)$ is unknown (depending on both the unknown $\bar{A}$, and other aspects of the noise distribution). However, we believe that it should be possible to estimate $\Gamma^*(\eta)$ based on the sample path of the algorithm itself. Notably, in their study of stochastic gradient methods, Chen et al. (2020) construct an online estimator for the asymptotic covariance. An interesting direction for future work is to extend estimators of this type to the class of stochastic approximation procedures considered here.

3.2. Some extensions beyond the basic setting

We now turn to some extensions that move beyond the basic setting of ℓ_2-bounds when the matrix $\bar{A}$ is Hurwitz. We begin in Section 3.2.1 by deriving some ℓ_∞-bounds that are useful in the analysis of the TD algorithm. In Section 3.2.2 to follow, we develop a relaxation of the Hurwitz condition.

3.2.1. BOUNDS IN THE ℓ_∞-NORM

In this section, we extend the analysis framework of Theorem 3 to the ℓ_∞-setting. Under somewhat stronger assumption on the linear operator and the noise distribution, we establish an ℓ_∞-bound in which leading term matches the ℓ_∞-norm of the asymptotic distribution in Theorem 2. Notably, the correction term and concentration error bounds has only logarithmic dependence on the dimensionality of the problem, as opposed to the polynomial dependence in Theorem 3. This much milder dimension dependence is important in applications, such as TD algorithms in reinforcement learning, where the dimension may be very large. See Appendix C.2 for the implication of this general theorem to TD learning with Example 3.

In order to obtain the tight dimension dependence, we impose the following stronger condition:

Assumption 4 *The stochastic oracles satisfy* $\|b_t\|_\infty \leq 1$ *and* $\forall u \in \mathbb{R}^d,\ \|A_t u\|_\infty \leq \|u\|_\infty$ a.s.

In addition, we replace the Hurwitz condition with the following stronger contraction condition:

Assumption 5 *There is a constant* $\bar{\lambda} > 0$ *such that the random matrix* $I - A_t$ *is a* $(1-\bar{\lambda})$*-contractive with respect to the* ℓ_∞*-norm, almost surely, meaning that*

$$\|(I - A_t)v\|_\infty \leq (1 - \bar{\lambda})\, \|v\|_\infty \qquad \textit{for all } v \in \mathbb{R}^d.$$

Under Assumption 4, we are able to establish an upper bound on each coordinate direction $\mathbf{e}_j$, leading to a high-probability upper bound on $\left\|\bar{\theta}_T - \theta^*\right\|_\infty$. Naturally, this bound involves the *maximal coordinate-wise variance*: $\sigma^2_{\max} := \max_{j=1,\dots,d} \mathbf{e}_j^T \Gamma^*(\eta)\mathbf{e}_j$.

Theorem 5 *Fix an iteration number* T *and a tolerance* $\delta \in (0, 1/T)$*, and suppose that the i.i.d. condition (Assumption 1), the almost-sure* ℓ_∞ *bound condition (Assumption 4), and the almost-sure* ℓ_∞ *contraction condition (Assumption 5) all hold. Then there exists a constant* $c > 0$ *such that for any step size* $\eta > 0$ *satisfying the bound* (11a)*, we have*

$$\mathbb{P}\left[\sqrt{T}\left\|\bar{\theta}_T - \theta^*\right\|_\infty \leq c\sqrt{\sigma^2_{max}\log(d/\delta)} + c\frac{\bar{\lambda}^{-2}\eta + \bar{\lambda}^{-1}}{T^{\frac{1}{4}}}\sqrt{\log\frac{d}{\delta}} + c\frac{\bar{\lambda}^{-\frac{5}{2}}}{\eta\sqrt{T}}\right] \geq 1 - \delta.$$

See Appendix G for the proof of this theorem.

We note that the theorem can actually be slightly refined by replacing the term $\sigma^2_{\max}\log(d/\delta)$ with the quantity $Q\left((\mathbf{e}_j^\top \Gamma^*(\eta)\mathbf{e}_j)_{j=1}^d; \delta\right)$, where for a vector $v = (v_1, v_2, \cdots, v_d) \in \mathbb{R}^d$, we define

$$Q(v; \delta) := \inf\left\{q \ \middle|\ e^{-q/v_1} + e^{-q/v_2} + \cdots + e^{-q/v_d} \leq \delta\right\}. \tag{14}$$

For example, if the maximal variance σ^2_{max} is much larger than second largest term $(\sigma')^2$ in the diagonal of $\Gamma^*(\eta)$, the quantity Q is upper bounded by $\sigma^2_{max}\log(1/\delta) + (\sigma')^2\log(d/\delta)$. For TD learning, this slightly improves the instance-dependent bound of Pananjady and Wainwright (2019).

3.2.2. CRITICAL CASE

In many real-world situations, the Hurwitz assumption may be violated, or the eigengap can be too small to be useful. At the population level, solving the deterministic equation $\bar{A}\theta = b$ is possible as long as the eigenvalues of $\bar{A}$ are bounded away from zero. Thus, it is natural to wonder whether the linear stochastic approximation scheme (2) still behaves well without this assumption. Furthermore, when the Hurwitz constant λ is positive but extremely small, does one necessarily obtain a slow convergence rate? In this section, we show that the non-asymptotic rates for LSA remain valid even in the critical case with no contraction at all.

In this section, we prove a non-asymptotic convergence rate for LSA in the critical case. We replace the Hurwitz condition on $\bar{A}$ (stated as Assumption 2) with the following assumption:

Assumption 2′ *The matrix* $\bar{A}$ *is diagonalizable with* $\bar{A} = UDU^{-1}$*, and* $\min_{i\in[d]} \mathrm{Re}\left(\lambda_i(\bar{A})\right) \geq 0$.

The reader might wonder why Assumption 2′ includes a diagonalizability condition, which was not needed before. Unfortunately, unlike the Hurwitz case, the diagonalizability assumption is unavoidable in the critical case. In particular, the Polyak-Ruppert procedure is not even consistent when A has purely imaginary eigenvalues and is non-diagonalizable at the same time, even in the noiseless case. We show this with an explicit construction in Appendix K.2.

Theorem 6 *Suppose that the* i.i.d. *condition (Assumption 1), the eigenvalue condition (Assumption $2'$), and the second-moment bounds (Assumption 3) all hold. Then, given a total number of iterations T, for the step size $\eta = \frac{1}{(\rho(\bar{A})+3\kappa(U)v_A)\sqrt{T}}$, there is a universal constant c such that*

$$\mathbb{E}\left\|\bar{A}\bar{\theta}_T - \bar{b}\right\|_2^2 \leq c\,\frac{\kappa^2(U)(\rho^2(\bar{A}) + \kappa^2(U)v_A^2)\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2 + v_b^2 d + v_A^2\left\|\theta^*\right\|_2^2}{T}. \tag{15}$$

See Appendix H for the proof of this theorem.

Theorem 6 is particularly useful in the asymmetric case, where the eigenvalues of $\bar{A}$ can be complex though the matrix itself is real. Even if the matrix $\bar{A}$ has an eigenvalue whose real part is exactly zero but with imaginary part being non-zero, which is beyond the classical regime of stable dynamical systems, the $1/T$ rate in mean-squared error is still guaranteed by averaging. More precisely, we have

$$\mathbb{E}\left\|\bar{\theta}_T - \theta^*\right\|_2^2 \leq c\,\kappa^2(U)\frac{\kappa^2(U)(\rho^2(\bar{A}) + \kappa^2(U)v_A^2)\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2 + v_b^2 d + v_A^2\left\|\theta^*\right\|_2^2}{\min_{i\in[d]}|\lambda_i(\bar{A})|^2 T}.$$

Although Theorem 6 achieves the correct $O(1/T)$ rate for mean-squared error, the problem-dependent pre-factor is not optimal in general. Indeed, a superior problem-dependent rate $\frac{v_A^2\|\theta^*\|_2^2 + v_b^2 d}{T}$ can be achieved by a plug-in estimator solving $\bar{A}_n\hat{\theta} = \bar{b}_n$, where $\bar{A}_n$ and b_n are empirical averages. In comparison, the initial distance $\mathbb{E}\left\|\theta^* - \theta_0\right\|_2^2$ appears in Theorem 6. Intuitively, one can view this term as the counterpart of the correction term in Theorem 2 when the dynamics itself fails to converge. It is also worth noticing that the step size choice $O(1/\sqrt{T})$ is crucial in this case: a larger step size makes the dynamical system exponentially blow up, and a smaller step size leads to a suboptimal rate. It is an interesting open question how to achieve the optimal problem-dependent constant using stochastic approximation.

That being said, Theorem 6 does exhibit the general effectiveness of LSA as it achieves the optimal $O(1/T)$ rate in the critical case, with completely online update and $O(d)$ storage. This is the first time that a stochastic approximation procedure has been shown to achieve the correct rate *without* the Hurwitz assumption, and demonstrates the additional advantage of averaging in such settings. Note that the quantity $\min_{i\in[d]}|\lambda_i(\bar{A})|$ can be much larger than the smallest real part of eigenvalues in many applications. An important application of Theorem 6 is average-reward TD learning in Example 3, which is further discussed in Appendix C.2.

4. Discussion

In this paper, we established several new results for constant step-size linear stochastic approximation combined with Polyak-Ruppert averaging. In the case where $\bar{A}$ is a Hurwitz matrix, we establish a central limit theorem, with asymptotic covariance characterizing the effect of the constant step size. Non-asymptotically, we derive high-probability concentration bounds for the averaged iterates in any direction, whose leading term matches the non-asymtotic behavior of a Gaussian random variable with the limiting distribution, and has poly-logarithmic dependence on the failure probability. We also study the critical case where the real part of eigenvalues are only guaranteed to be non-negative, and establish a gap-independent $\mathcal{O}(1/T)$ rate in mean-squared error. We illustrate the effectiveness of our abstract results by considering momentum SGD for linear regression and TD learning, and uncover new aspects of the LSA approach to these problems.

Acknowledgements

This research was partially supported by Office of Naval Research Grant DOD ONR-N00014-18-1-2640 and National Science Foundation Grant DMS-1612948 to MJW, and by the Army Research Office (ARO) under contract W911NF-17-1-0304 under the Multidisciplinary University Research Initiative (MURI) to MIJ.

References

Alekh Agarwal, Peter L Bartlett, Pradeep Ravikumar, and Martin J Wainwright. Information-theoretic lower bounds on the oracle complexity of stochastic convex optimization. *IEEE Transactions on Information Theory*, 58(5):3235–3249, 2012.

Andreas Anastasiou, Krishnakumar Balasubramanian, and Murat A Erdogdu. Normal approximation for stochastic gradient descent via non-asymptotic rates of martingale CLT. *arXiv preprint arXiv:1904.02130*, 2019.

Albert Benveniste, Michel Métivier, and Pierre Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer, 2012.

Dimitris P. Bertsekas. *Dynamic Programming and Stochastic Control*, volume 1. Athena Scientific, Belmont, MA, 1995.

Dimitris P Bertsekas and John N Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1989.

Jalaj Bhandari, Daniel Russo, and Raghav Singal. A finite time analysis of temporal difference learning with linear function approximation. *arXiv preprint arXiv:1806.02450*, 2018.

Vivek S Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *arXiv preprint arXiv:1606.04838*, 2016.

Donald L Burkholder, Burgess J Davis, and Richard F Gundy. Integral inequalities for convex functions of operators on martingales. In *Proc. Sixth Berkeley Symp. Math. Statist. Prob*, volume 2, pages 223–240, 1972.

Xi Chen, Jason D Lee, Xin T Tong, Yichen Zhang, et al. Statistical inference for model parameters in stochastic gradient descent. *The Annals of Statistics*, 48(1):251–273, 2020.

Aaron Defazio, Francis Bach, and Simon Lacoste-Julien. Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. In *Advances in Neural Information Processing Systems*, pages 1646–1654, 2014.

Aymeric Dieuleveut, Alain Durmus, and Francis Bach. Bridging the gap between constant step size stochastic gradient descent and Markov chains. *arXiv preprint arXiv:1707.06386*, 2017a.

Aymeric Dieuleveut, Nicolas Flammarion, and Francis Bach. Harder, better, faster, stronger convergence rates for least-squares regression. *The Journal of Machine Learning Research*, 18(1): 3520–3570, 2017b.

Jianqing Fan, Wenyan Gong, Chris Junchi Li, and Qiang Sun. Statistical sparse online regression: A diffusion approximation perspective. In *International Conference on Artificial Intelligence and Statistics*, pages 1017–1026, 2018.

David A Freedman. On tail probabilities for martingales. *The Annals of Probability*, 3(1):100–118, 1975.

Peter Hall and Christopher C Heyde. *Martingale Limit Theory and Its Application*. Academic Press, 1980.

Prateek Jain, Praneeth Netrapalli, Sham M Kakade, Rahul Kidambi, and Aaron Sidford. Parallelizing stochastic gradient descent for least squares regression: mini-batching, averaging, and model misspecification. *The Journal of Machine Learning Research*, 18(1):8258–8299, 2017.

Prateek Jain, Sham M Kakade, Rahul Kidambi, Praneeth Netrapalli, and Aaron Sidford. Accelerating stochastic gradient descent for least squares regression. In *Conference On Learning Theory*, pages 545–604, 2018.

Prateek Jain, Dheeraj Nagaraj, and Praneeth Netrapalli. Making the last iterate of SGD information theoretically optimal. *arXiv preprint arXiv:1904.12443*, 2019.

Rie Johnson and Tong Zhang. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in Neural Information Processing Systems*, pages 315–323, 2013.

Aldéric Joulin and Yann Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *The Annals of Probability*, 38(6):2418–2442, 2010.

Olav Kallenberg. *Foundations of Modern Probability*. Springer Science & Business Media, 2006.

Belhal Karimi, Blazej Miasojedow, Éric Moulines, and Hoi-To Wai. Non-asymptotic analysis of biased stochastic approximation scheme. *arXiv preprint arXiv:1902.00629*, 2019.

Koulik Khamaru, Ashwin Pananjady, Feng Ruan, Martin J Wainwright, and Michael I Jordan. Is temporal difference learning optimal? An instance-dependent analysis. *arXiv preprint arXiv:2003.07337*, 2020.

Harold Kushner and G George Yin. *Stochastic Approximation and Recursive Algorithms and Applications*, volume 35. Springer, 2003.

Tze Leung Lai. Stochastic approximation. *The Annals of Statistics*, 31(2):391–406, 2003.

Chandrashekar Lakshminarayanan and Csaba Szepesvari. Linear stochastic approximation: How far does constant step-size and iterate averaging go? In *International Conference on Artificial Intelligence and Statistics*, pages 1347–1355, 2018.

Tianyang Li, Liu Liu, Anastasios Kyrillidis, and Constantine Caramanis. Statistical inference using SGD. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Tengyuan Liang and Weijie J Su. Statistical inference for the population landscape via moment-adjusted stochastic gradients. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(2):431–456, 2019.

Stephan Mandt, Matthew D Hoffman, and David M Blei. Stochastic gradient descent as approximate Bayesian inference. *The Journal of Machine Learning Research*, 18(1):4873–4907, 2017.

Eric Moulines and Francis R Bach. Non-asymptotic analysis of stochastic approximation algorithms for machine learning. In *Advances in Neural Information Processing Systems*, pages 451–459, 2011.

Arkadi Nemirovski, Anatoli Juditsky, Guanghui Lan, and Alexander Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on optimization*, 19(4):1574–1609, 2009.

Arkadii Nemirovskii and David Borisovich Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley, 1983.

Balamurugan Palaniappan and Francis Bach. Stochastic variance reduction methods for saddle-point problems. In *Advances in Neural Information Processing Systems*, pages 1416–1424, 2016.

Ashwin Pananjady and Martin J Wainwright. Value function estimation in Markov reward processes: Instance-dependent ℓ_∞-bounds for policy evaluation. *arXiv preprint arXiv:1909.08749*, 2019.

Bob Pepin. Time averages of stochastic processes: a martingale approach. *arXiv preprint arXiv:1810.10945*, 2018.

Lawrence Perko. *Differential Equations and Dynamical Systems*, volume 7. Springer Science & Business Media, 2013.

Boris T Polyak and Anatoli B Juditsky. Acceleration of stochastic approximation by averaging. *SIAM Journal on Control and Optimization*, 30(4):838–855, 1992.

LA Prashanth, Nathaniel Korda, and Rémi Munos. Stochastic approximation for speeding up LSTD (and LSPI). *arXiv preprint arXiv:1306.2557*, 2013.

Mark L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.

Alexander Rakhlin, Ohad Shamir, and Karthik Sridharan. Making gradient descent optimal for strongly convex stochastic optimization. In *Proceedings of the 29th International Conference on Machine Learning*, pages 449–456, 2012.

Herbert Robbins and Sutton Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, pages 400–407, 1951.

R Tyrrell Rockafellar. Monotone operators associated with saddle-functions and minimax problems. *Nonlinear Functional Analysis*, 18(part 1):397–407, 1970.

Nicolas L Roux, Mark Schmidt, and Francis R Bach. A stochastic gradient method with an exponential convergence rate for finite training sets. In *Advances in Neural Information Processing Systems*, pages 2663–2671, 2012.

David Ruppert. Efficient estimations from a slowly convergent Robbins-Monro process. *Technical Report, Cornell University Operations Research and Industrial Engineering*, 1988.

Ohad Shamir and Tong Zhang. Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes. In *Proceedings of the 30th International Conference on Machine Learning*, pages 71–79, 2013.

Aaron Sidford, Mengdi Wang, Xian Wu, and Yinyu Ye. Variance reduced value iteration and faster algorithms for solving markov decision processes. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 770–787. SIAM, 2018.

Weijie J Su and Yuancheng Zhu. Uncertainty quantification for online learning and stochastic approximation via hierarchical incremental gradient descent. *arXiv preprint arXiv:1802.04876*, 2018.

Richard S Sutton. Learning to predict by the methods of temporal differences. *Machine learning*, 3 (1):9–44, 1988.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2nd edition, 2018.

John N Tsitsiklis and Benjamin Van Roy. On average versus discounted reward temporal-difference learning. *Machine Learning*, 49(2-3):179–191, 2002.

Aad W Van der Vaart. *Asymptotic Statistics*, volume 3. Cambridge university press, 2000.

Cédric Villani. *Optimal Transport: Old and New*, volume 338. Springer Science & Business Media, 2008.

Martin J Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, volume 48. Cambridge University Press, 2019a.

Martin J Wainwright. Stochastic approximation with cone-contractive operators: Sharp ℓ_∞-bounds for Q-learning. *arXiv preprint arXiv:1905.06265*, 2019b.

Martin J Wainwright. Variance-reduced Q-learning is minimax optimal. *arXiv preprint arXiv:1906.04697*, 2019c.

Mengdi Wang and Dimitri P Bertsekas. Stochastic first-order methods with random constraint projection. *SIAM Journal on Optimization*, 26(1):681–717, 2016.

Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.

Appendix A. Additional Related Works

Several bounds have been established on function values in stochastic optimization. After processing N samples, the averaged iterate enjoys an $O(1/N)$ and $O(1/\sqrt{N})$ optimization error bounds for strongly convex and convex objectives (Nemirovski et al., 2009; Rakhlin et al., 2012; Shamir and Zhang, 2013). Such optimization error bounds are optimal in the sense that they match the statistical lower bounds under a stochastic first-order oracle (Agarwal et al., 2012; Nemirovskii and Yudin, 1983). Dieuleveut et al. (2017b) studied a momentum accelerated stochastic gradient scheme with appropriate regularization, proving its optimality in the critical case. Nevertheless when applied to (often high-dimensional) statistical models with specific distributional assumptions, the aforementioned sharp results often lose essential statistical information due to their coarse-grained nature.

Stochastic approximation methods have also been widely applied in reinforcement learning; in particular, TD learning (Sutton, 1988) and Q-learning (Watkins and Dayan, 1992) are based on linear and nonlinear stochastic approximation updates for policy evaluation and Q-function learning, respectively. It should be noted that the various Bellman-type operators arising in RL do not correspond to gradients of functions, so that the analysis requires different techniques from stochastic optimization. A recent line of work has focused on the non-asymptotic analysis of TD learning and Q-learning algorithms. Prashanth et al. (2013) studies TD algorithms with linear function approximation using Polyak-Ruppert average, but their rate is slower than the optimal $O(1/\sqrt{T})$ one. Bhandari et al. (2018) studied TD with linear function approximation and established bounds with the optimal rate on the mean-squared error. Wainwright (2019b,c) analyzed Q-learning as a special case of a cone-contractive operator, and established sharp ℓ_∞-norm bounds, both for ordinary Q-learning and a variance-reduced version thereof. Variance-reduced Q-learning under the generative model is also analyzed in Sidford et al. (2018). Concurrent to our work, Khamaru et al. (2020) studies the local asymptotic minimax complexity of the value function estimation problem, and obtain a non-asymptotic upper bound that matches the leading terms using variance reduced TD algorithms. Karimi et al. (2019) studied general biased stochastic approximation procedures, in particular proving convergence of online EM and policy gradient methods.

Additional perspectives and variations on stochastic approximation appear in the literature, with improved non-asymptotic convergence properties in particular cases. Recent work also studies tail averaging with parallelization (Jain et al., 2017), momentum-based schemes (Jain et al., 2018; Dieuleveut et al., 2017b), Markov chain perspectives (Dieuleveut et al., 2017a), variational Bayesian perspectives (Mandt et al., 2017) and diffusion approximation perspectives (Fan et al., 2018). Pepin (2018) studies ergodict concentration inequalities of averaged Markov processes, with applications to a special case of Polyak-Ruppert procedure. Berry-Esseen bounds are also obtained for the normal approximation in Polyak-Ruppert CLT (Anastasiou et al., 2019). There is also significant work on last-iterate SGD (Jain et al., 2019) and variance-reduced estimators (see, e.g., Roux et al. (2012); Johnson and Zhang (2013); Defazio et al. (2014)). Our discussion of these variants is limited in this paper; it will be interesting to study whether these variants can be shown to have the desirable statistical properties that we uncover here under a similar set of assumptions.

Appendix B. Additional Examples

In this section, we describe two more examples for the general LSA procedure (2) in addition to Example 1, 2 and 3.

We first consider a generalization of Example 3, which allows a linearly parametrized family to represent the value function.

Example 4 (TD Algorithm with linear function approximation) In practice, the state space $\mathcal{X}$ can be extremely large or possibly infinite. In such settings, the exact approach to policy evaluation, as described in the previous example, becomes both computationally infeasible and statistically inefficient. In practice, it is typical to combine TD algorithms with a linear function approximation step. Suppose that we are given a feature map $\phi : \mathcal{X} \to \mathbb{R}^d$. We consider the set of value functions $V : \mathcal{X} \to \mathbb{R}$ that have a linear parameterization of the form $V_\theta(x) = \langle \theta, \phi(x) \rangle = \sum_{j=1}^d \theta_j \phi_j(x)$ for some vector of weights $\theta \in \mathbb{R}^d$. We use $\mathcal{L}_\phi$ to denote the collection of all such linearly parameterized value functions.

In this more general context, the TD(0) algorithm seeks to compute a particular approximation to the original value function, as we now describe. Suppose that the Markov process $(X_t)_{t\geq 0}$ has a unique stationary distribution μ, and let $\Pi_{\mathcal{L}_\phi,\mu} : \mathcal{X} \to \mathcal{L}_\phi$ denote the $L^2(\mu)$-projection onto the linear space $\mathcal{L}_\phi$—that is $\Pi_{\mathcal{L}_\phi,\mu}(V) := \arg\min_{V_\theta \in \mathcal{L}_\phi} \|V - V_\theta\|_{L^2(\mu)}$. We can then define the *projected Bellman equation* as

$$V = \Pi_{\mathcal{L}_\phi,\mu}\Big(r + \gamma P V\Big), \tag{16}$$

where $r : \mathcal{X} \to \mathbb{R}$ is the reward function of the Markov reward process. It can be shown that this equation has a unique fixed point V^*, known as the TD approximation. Since V^* must belong to $\mathcal{L}_\phi$, we can write $V^*(x) = \langle \theta^*, \phi(x) \rangle$ for some $\theta^* \in \mathbb{R}^d$.

With this set-up, we can now describe the more general instantiation of the TD(0) algorithm, which uses linear stochastic approximation to solve the projected Bellman equation (16). Using the optimality conditions for projection, it can be shown that the vector θ^*, which characterizes the projected Bellman fixed point V^*, must satisfy the linear equation

$$\mathbb{E}(\phi(X)\phi(X)^\top)\theta^* = \mathbb{E}(R(X)\phi(X)) + \gamma \mathbb{E}(\phi(X)\phi(X^+)^\top)\theta^*.$$

Here the expectations are taken over the joint distribution of a pair (X, X^+), where X is distributed according to the stationary distribution μ, and X^+ is drawn from the transition kernel P (conditioned on the previous state being X). Thus, we see that the fixed point θ^* must satisfy an equation of the form $\bar{A}\theta^* = \bar{b}$, where

$$\bar{A} := \mathbb{E}(\phi(X)\phi(X)^\top) - \gamma \mathbb{E}(\phi(X)\phi(X^+)^\top), \quad \text{and} \quad \bar{b} = \mathbb{E}(R(X)\phi(X)).$$

The TD(0) algorithm corresponds to linear stochastic approximation for solving this equation. At time t, if we are given a triplet (X_t, X_t^+, R_t), where X_t is distributed according to μ; the next state X_t^+ is drawn from P conditioned on the previous state X_t, and R_t is a random reward. We can then run linear stochastic approximation using the quantities

$$A_t = \phi(X_t)\phi(X_t)^T - \gamma \phi(X_t)\phi(X_t^+)^T \quad \text{and} \quad b_t = R_t \phi(X_t). \tag{17}$$

We return to analyze this algorithm in Section C.2.2. ♣

Finally, we turn to an example of a minimax saddle-point problem (Rockafellar, 1970), which has broad application in computational game theory, machine learning and robust statistics (see Palaniappan and Bach (2016) and references therein).

Example 5 (Minimax games) We consider a minimax saddle-point problem of the following form:

$$\min_{x\in\mathbb{R}^n}\max_{y\in\mathbb{R}^m}\frac{1}{2}\begin{bmatrix}x\\y\\1\end{bmatrix}^\top\cdot\begin{bmatrix}P_{xx} & P_{xy} & c_x\\P_{xy}^\top & P_{yy} & c_y\\c_x^\top & c_y^\top & 0\end{bmatrix}\cdot\begin{bmatrix}x\\y\\1\end{bmatrix}. \tag{18}$$

In a computational game theory setting, for example, the vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ represent the actions of the two players. The payoff matrix $P \in \mathbb{R}^{(n+m)\times(n+m)}$ satisfies the PSD conditions $P_{xx} \succeq 0$ and $P_{yy} \preceq 0$, so that the game is of the convex-concave type. The matrix game (18) is a type of saddle-point problem, and its solution reduces to solving the linear system

$$\begin{bmatrix}P_{xx} & P_{xy}\\-P_{xy}^\top & -P_{22}\end{bmatrix}\cdot\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-c_x\\c_y\end{bmatrix}. \tag{19}$$

Thus, this problem fits into our general set-up with $\bar{A} = P$ and $\bar{b} = \begin{bmatrix}-c_x & c_y\end{bmatrix}^T$, so that $d = n+m$. Note that the conditions $P_{xx} \succ 0$ and $P_{yy} \prec 0$ imply that $A = P$ is Hurwitz. The setting of $P_{xx} = 0$ and $P_{yy} = 0$ corresponds to the so-called critical case. ♣

Appendix C. Applications

In this section, we illustrate the usefulness of our four main theorems by applying them to some concrete problems, namely the momentum SGD algorithm discussed in Example 2 and the temporal difference (TD) algorithm discussed in Example 3.

C.1. Stochastic gradient method with momentum

Recall the SGD with momentum algorithm for linear regression that was previously introduced in Example 2. In this section, we use our general theory to analyze it. As defined in Example 2, at the population level the algorithm involves a matrix $\tilde{A} \in \mathbb{R}^{d\times d}$ and vector $\tilde{b} \in \mathbb{R}^{2d}$. For the linear regression setting, we can assume without loss of generality that $\theta^* = 0$, by the translation invariance. At each time t, the algorithm makes use of a pair $(\tilde{A}_t, \tilde{b}_t)$ that are unbiased estimates of these population quantities. The momentum SGD update rule takes the form

$$\tilde{\theta}_{t+1} = \tilde{\theta}_t - \eta(\tilde{A}_{t+1}\tilde{\theta}_t - \tilde{b}_{t+1}). \tag{20}$$

Consider the noise variables $\tilde{\Xi}_t = \tilde{A}_t - \tilde{A}$ and $\tilde{\xi}_t = \tilde{b}_t - \tilde{b}$. It can be seen that they satisfy the same second moment or higher moment assumptions as Ξ_t and ξ_t do, with the constants $(\sqrt{1+\eta^2}\sigma_A, \sigma_b)$ or $(\sqrt{1+\eta^2}v_A, v_b)$.

The addition of momentum to SGD has two effects: it changes the mixing time of the process $(\theta_t)_{t\geq 0}$, and it alters the structure of the asymptotic covariance matrix $\Gamma^*(\eta)$. The spectrum of $\tilde{A}$ plays a central role in these effects; accordingly, let us investigate the structure of this spectrum. Suppose that the matrix $\bar{A}$ is positive definite, and let $\{\lambda_i\}_{i=1}^d$ denote its eigenvalues.

We claim that for any $\alpha \in \mathbb{R}_+ \setminus \{2\sqrt{\lambda_i} - \eta\lambda_i\}_{i=1}^d$, the matrix $\tilde{A} \in \mathbb{R}^{2d\times 2d}$ is diagonalizable, with paired (complex) eigenvalues

$$\left(\frac{(\alpha+\eta\lambda_i)+\sqrt{(\alpha+\eta\lambda_i)^2-4\lambda_i}}{2}, \frac{(\alpha+\eta\lambda_i)+\sqrt{(\alpha+\eta\lambda_i)^2+4\lambda_i}}{2}\right) \quad \text{for } i = 1, \ldots, d. \tag{21}$$

See Appendix K.3 for the proof of this claim.

Let us now consider the consequences of the spectrum (21) for the mixing time of the process $(\theta_t)_{t\geq 0}$. We claim that when the parameter α is suitably chosen, the mixing rate of the momentum-based method is faster by a factor of $1/\sqrt{\lambda_{\min}(\bar{A})}$. Introduce the shorthand

$$\nu_i := \frac{(\alpha+\eta\lambda_i)+\sqrt{(\alpha+\eta\lambda_i)^2-4\lambda_i}}{2}, \quad \text{for } i=1,\ldots,d.$$

For an index i such that $\alpha > 2\sqrt{\lambda_i}-\eta\lambda_i$, we have $\nu_i \in \mathbb{R}$, and for index i such that $\alpha < 2\sqrt{\lambda_i}-\eta\lambda_i$, we have $\mathrm{Re}(\nu_i) = \alpha+\eta\lambda_i$. Therefore, for $\lambda = \lambda_{\min}(\bar{A})$, we have:

$$\min_i \mathrm{Re}(\lambda_i(\tilde{A})) = \begin{cases} \alpha+\eta\lambda - \sqrt{(\alpha+\eta\lambda)^2-4\lambda} \geq \frac{2\lambda}{\alpha+\eta\lambda}, & \alpha \geq 2\sqrt{\lambda}-\eta\lambda \\ \alpha+\eta\lambda, & \alpha < 2\sqrt{\lambda}-\eta\lambda. \end{cases}$$

When we take $\alpha \asymp \sqrt{\lambda_{\min}(\bar{A})}$, we have $\min_i \mathrm{Re}(\lambda_i(\tilde{A})) \asymp \sqrt{\lambda_{\min}(\bar{A})}$.

Now Lemma 12 implies that for given step size $\eta > 0$, the mixing time is upper bounded by

$$\frac{1}{\eta \min \mathrm{Re}(\lambda_i(\tilde{A}))} \asymp \frac{1}{\eta\sqrt{\lambda_{\min}(\bar{A})}}.$$

Consequently, the use of momentum speeds up the mixing time by a factor of $(1/\sqrt{\lambda_{\min}(\bar{A})})$, which is significant in the regime $\lambda_{\min}(\bar{A}) \ll 1$.

Furthermore, we study the effect of momentum on the asymptotic covariance. We make the following claim:

Claim 1 *For the momentum SGD update (6) with $\alpha \asymp \sqrt{\lambda_{min}(A)}$, the asymptotic covariance in Theorem 2 restricted to θ-components is of the form $\bar{A}^{-1}\Sigma^*\bar{A}^{-1}+L_\eta$, where the matrix L_η satisfies the following upper bound:*

$$\mathrm{Tr}(L_\eta) \lesssim \eta\frac{v_A^2\kappa^2(U)v_b^2 d}{\lambda_{\min}(A)^{3/2}},$$

where the matrix is written as $\tilde{A} = UDU^{-1}$ in the decomposition in Lemma 1.

A similar analysis can be carried out to show that SGD with averaging achieves a covariance at stationarity that has a larger correction term $O(\eta\lambda_{min}(\bar{A})^{-3})$ than momentum with SGD. However, whether momentum SGD can exceed SGD in correction term involves computing $\kappa(U)$ and choosing η. We leave this as future work.

A straightforward calculation shows that the leading term $\tilde{A}^{-1}\begin{bmatrix}0 & 0\\ 0 & \Sigma^*\end{bmatrix}(\tilde{A}^{-1})^\top$ in the θ-component is the same as $\bar{A}^{-1}\Sigma^*\bar{A}^{-1}$. Now we consider the correction term $\tilde{A}^{-1}\mathbb{E}(\tilde{\Xi}_A\Lambda_\eta^*(\tilde{\Xi}_A)^\top)(\tilde{A}^{-1})^\top$. Note that Λ_η^* is the stationary covariance of $(\theta_t)_{t\geq 0}$. Simple calculation leads to the upper bound:

$$\mathrm{Tr}(\tilde{A}^{-1}\mathbb{E}(\tilde{\Xi}_A\Lambda_\eta^*(\tilde{\Xi})_A^\top)(\tilde{A}^{-1})^\top) \leq (\min_i|\lambda_i(\tilde{A})|)^{-2}(1+\eta^2)v_A^2\mathbb{E}_{\pi_\eta}\|\theta_t-\theta^*\|_2^2.$$

As we will see in Lemma 11 in Appendix D.1.2, the stationary covariance satisfies the following upper bound:

$$\mathbb{E}_{\pi_\eta}\|\theta-\theta^*\|_2^2 \leq \kappa^2(U)\frac{\eta}{\min_i \mathrm{Re}\lambda_i(\tilde{A})}v_b^2 d.$$

Noting that we have $\min_i \mathrm{Re}(\lambda_i(\tilde{A})) \asymp \sqrt{\lambda_{\min}(A)}$, plugging into the above upper bound proves the trace bound in Claim 1.

C.2. Temporal difference learning

We discuss the applications of our main theorems in TD learning, in both exact (Example 3) and linear function approximation (Example 4) settings. We consider both the discounted case ($\gamma < 1$) as well as the undiscounted case ($\gamma = 1$). Theorem 3, 5 and 6 turn out to have nontrivial implications to the TD algorithm in these cases.

C.2.1. ANALYSIS OF TD WITHOUT FUNCTION APPROXIMATION

We start with the case of exact TD(0). We follow the model definition and assumptions in Example 3.

Non-asymptotic bounds in the Hurwitz case Recall that in the generative model, the one-step observation P_t satisfies $\|Pv\|_\infty \leq \|v\|_\infty$ for any vector v. For discount factor $\gamma \in [0, 1)$, the matrix γP_t is γ-contractive under the $\|\cdot\|_\infty$ norm. Consequently, Assumption 5 is satisfied by the observation model, and we can apply Theorem 5.

In order to state the result, we require a few additional pieces of notation. Define the D-dimensional vector $\sigma^* \in \mathbb{R}^D$ of standard deviations, with

$$\sigma^*_j := \sqrt{\mathrm{var}(R(j)) + \mathrm{var}(Z(j,:)\theta^*)}, \quad \text{for } j = 1, \ldots, D.$$

Since the rows of Z_t and entries of R_t are independent, the matrix Σ_* in the main term is actually $\mathrm{diag}(\sigma^*(j)^2)_{j\in[D]}$. It is easy to see that the structure of the stochastic oracles (A_t, b_t) satisfies Assumption 4 and Assumption 5. Thus, we can apply Theorem 5. Doing so yields a result that involves the matrix

$$\Gamma^*(\eta) := (I - \gamma P)^{-1}(\mathrm{diag}(\sigma^*(j)^2)_{j\in[D]} + \Lambda^*_\eta)(I - \gamma P^\top)^{-1}, \tag{22}$$

where the matrix Λ^*_η was defined in equation (9). The result also involves the function Q defined in equation (14).

Corollary 7 *Consider the* i.i.d. *observational model for Markov reward processes defined above. Given a discount factor $\gamma \in (0, 1)$ and a failure probability $\delta > 0$, the averaged TD(0) algorithm based on step size $\eta \in (0, 1)$ satisfies the bound*

$$\sqrt{T}\left\|\hat{\theta}_T - \theta^*\right\|_\infty \lesssim \sqrt{Q(\mathrm{diag}(\Gamma^*(\eta)); \delta)} + T^{-\frac{1}{4}}\left(\frac{\eta}{(1-\gamma)^2} + \frac{1}{1-\gamma}\right)\sqrt{\log\frac{d}{\delta}} + \frac{T^{-\frac{1}{2}}}{\eta(1-\gamma)^{-\frac{5}{2}}},$$

with probability at least $1 - \delta$.

When the step size is chosen to be of order $\eta = O(T^{-\frac{1}{3}})$, the leading term of Corollary 7 is an instance-dependent term that slightly improves upon that of the offline plug-in estimator in Pananjady and Wainwright (2019), which was shown to be minimax optimal.

Critical case: Application of Theorem 6. While most of existing results in policy evaluation require the discount factor to be bounded away from one, our second result certifies that, even if there is no discount at all (i.e., when $\gamma = 1$, corresponding to the average reward RL setting), the linear stochastic approximation achieves a $O(1/\sqrt{T})$ error decay, as long as the error is measured in terms of Bellman error (i.e., the deficiency in the fixed point relation). Furthermore, for discounted problems, the results show that the Bellman error can be bounded independently of the $(1 - \gamma)$ factor:

Corollary 8 *Suppose the transition matrix P is diagonalizable with $P = UD_PU^{-1}$, for $\eta = \frac{1}{(1+3\kappa(U)v(P))\sqrt{T}}$, for any $\gamma \in [0,1]$, we have*

$$\mathbb{E}\left\|\bar{\theta}_T - (\gamma P\bar{\theta}_T + r)\right\|_2^2 \lesssim \frac{\kappa^2(U)(1+\kappa^2(U)v(P)^2)\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2 + v(r)^2 D + v(P)^2\left\|\theta^*\right\|_2^2}{T}.$$

In the setting of average reward TD learning, by subtracting the stationary average reward (see discussion in Example 3), we can still translate the bound in Bellman error to the parameter estimation error. Corollary 8 implies that:

$$\mathbb{E}\left\|\bar{\theta}_T - \theta^*\right\|_2^2 = O\left(\kappa^2(U)\frac{v(r)^2 D + v(P)^2\left\|\theta^*\right\|_2^2 + \kappa^2(U)(1+\kappa^2(U)v(P)^2)\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2}{T\cdot \min_{i\geq 2}|1-\lambda_i(P)|^2}\right),$$

where the problem-dependent complexity term is $\min_{i\geq 2}|1-\lambda_i(P)|$, as opposed to the real-part of the eigengap $\min_{i\geq 2}(1-\mathrm{Re}(\lambda_i(P)))$ in the Hurwitz case. In particular, suppose that the transition matrix P has a complex eigenvalue of the form $e^{i\alpha}$ for some $\alpha \ll 1$.[2] In this case, we have $\min_{i\geq 2}|1-\lambda_i(P)| \asymp \alpha$ but $\min_{i\geq 2}(1-\mathrm{Re}(\lambda_i(P))) \asymp \alpha^2$. The dependency on α in the critical case bound can even be better than the bound we get by treating the matrix as Hurwitz. Specifically, Corollary 8 yields a bound of order $O(1/\alpha\sqrt{T})$; on the other hand, although the leading term in Theorem 3 is near-optimal, due to the presence of a $\frac{1}{\eta \min_{i\geq 2}|1-\lambda_i(P)|T}$ term in the bound, it leads to an $O(1/\alpha^3 T)$ term, as the step size has to be chosen such that $\eta \lesssim \alpha^2$. Corollary 8 leads to a better $O(\frac{1}{\alpha^2\varepsilon^2})$ sample complexity, compared with the $O(\frac{1}{\alpha^2\varepsilon^2} + \frac{1}{\alpha^3\varepsilon})$ complexity guaranteed by the theorem in the Hurwitz case. This is mainly because the step size choice $\eta \lesssim \alpha^2$ suggested by Theorem 3 is too conservative, compared to the gap-independent $O(1/\sqrt{T})$ choice implied by Theorem 6.

C.2.2. TD WITH LINEAR FUNCTION APPROXIMATION

We now consider an application of Theorem 3 and Theorem 6 to the use of the TD algorithm in conjunction with linear function approximation; recall Example 4. Note that for any vector $v \in \mathbb{S}^{d-1}$, by the Cauchy-Schwartz inequality, we have

$$v^\top \mathbb{E}(\phi(X)\phi(X^+))v \leq (v^\top \mathbb{E}(\phi(X)\phi(X))v)^{\frac{1}{2}}(v^\top \mathbb{E}(\phi(X^+)\phi(X^+))v)^{\frac{1}{2}} = v^\top \mathbb{E}(\phi(X)\phi(X))v.$$

So we have $\min_i \mathrm{Re}(\lambda_i(A)) \geq (1-\gamma)\min_i \lambda_i(\mathbb{E}\phi(X)\phi(X)^\top) > 0$ and Theorem 3 is applicable in this case. For the following results, we make two assumptions on the tail behavior:

- The feature vector $\phi(X)$ is a centered and σ_ϕ-sub-Gaussian random vector when X follows the stationary distribution μ, namely:

$$\mathbb{E}_\mu \phi(X) = 0, \quad \forall v \in \mathbb{R}^d, p \geq 2,\ (\mathbb{E}_\mu|\langle v,\, \phi(X)\rangle|^p)^{\frac{1}{p}} \leq \sigma_\phi\sqrt{p}.$$

- The random reward R_t satisfies the following moment bound:

$$\forall p \geq 2,\ (\mathbb{E}|R_t|^p)^{\frac{1}{p}} \leq \sigma_r\sqrt{p}.$$

2. This can happen, for example, in an N-state Markov chain where the transition from state i is deterministically to the state $(i+1) \mod N$. In such case the eigenvalues are $e^{\frac{2\pi k}{N}i}$.

In stating the resulting corollary, we let μ denote the stationary distribution of the Markov reward process; define the covariance matrix $M = \mathbb{E}_\mu \phi(X)\phi(X)^\top$, and the quantity

$$V(\theta^*) := \kappa(U)(\|\theta^* - \theta_0\|_2 + \|\theta^*\|_2 + \sqrt{\eta(1-\gamma)^{-1}}(\sqrt{d}\sigma_\phi \|\theta^*\|_2 + \sigma_r \sqrt{d})) \log^4 \frac{T}{\delta}.$$

Corollary 9 *Suppose that the model assumptions in Example 4 hold, we are given a discount factor $\gamma \in (0,1)$ and a failure probability $\delta > 0$, and we run the LSA algorithm using a step size $\eta \in \left(0, \frac{1-\gamma}{1+\kappa^2(U)\sigma_\phi^2 d \log^3 \frac{T}{\delta}}\right)$. Then for any vector $v \in \mathbb{S}^{d-1}$, the quantity $\sqrt{T}\left|v^\top(\hat{\theta}_T - \theta^*)\right|$ is upper bounded, up to a universal pre-factor, by*

$$\sqrt{v^\top \Gamma^*(\eta) v \log \frac{1}{\delta}} + \frac{\kappa(U)V(\theta^*)}{1-\gamma}\left(\frac{\sigma_\phi\sqrt{d} + \sigma_r}{T^{\frac{1}{4}}} + \frac{1 + \sqrt{\sigma_r/(1-\gamma)}}{\eta T}\right). \tag{23}$$

As a consequence of the bound (23), we are guaranteed that the rescaled error $\sqrt{T}\left\|\widehat{\theta}_T - \theta^*\right\|_{L^2(\mu)}$ is upper bounded as

$$\sqrt{\mathrm{Tr}\left(\Gamma^*(\eta) \cdot M\right) \log \frac{d}{\delta}} + \frac{\kappa(U)V(\theta^*)\sqrt{\|M\|_{\mathrm{op}} d} \log^4 \frac{dT}{\delta}}{1-\gamma}\left(\frac{\sigma_\phi\sqrt{d} + \sigma_r}{T^{\frac{1}{4}}} + \frac{1 + \sqrt{\sigma_r/(1-\gamma)}}{\eta T}\right),$$

with probability $1 - \delta$.

The proof of this bound simply follows by applying Corollary 9 on all of the eigenvectors of M, and using a union bound. Using a more refined ε-net argument (cf. Wainwright (2019a), Chapter 5), it is possible to reduce the log factor in the leading term, and match the behavior of a Gaussian random variable up to a constant factor and high-order terms. We omit the details.

Appendix D. Preliminary Steps in the Proofs

We now turn the proofs of our three main theorems, along with the various corollaries. Before proceeding to the arguments themselves, in this section, let us summarize some notation, and introduce the common initial steps used in the proofs of all the theorems.

Summary of notation: For an L^2-integrable quasi-martingale $\{X_t\}_{t \geq 1}$ adapted to the filtration $\{\mathcal{F}_{t \geq 0}\}$, we define

$$[X]_T := \sum_{t=0}^{T-1} \mathrm{var}\left(X_{t+1} | \mathcal{F}_t\right), \quad \text{and} \quad \langle X \rangle_T := \sum_{t=0}^{T-1} \left(X_{t+1} - \mathbb{E}(X_{t+1}|\mathcal{F}_t)\right)^2.$$

For two matrices A, B, we use $A \otimes B$ to denote their Kronecker product and $A \oplus B$ to denote their Kronecker sum. When it is clear from the context, we slightly overload the notation to let $A \otimes B$ denote the 4-th-order tensor produced by taking the tensor product of A and B. Note that Kronecker product is just a flattened version of the tensor. For any matrix A, we use $\mathrm{vec}(A)$ to denote the vector obtained by flattening A. For a k-th order tensor T, matrix M and vector v, we use $T[M]$ to denote the $(k-2)$-th order tensor obtained by applying T to matrix M, and similarly, we use $T[v]$ to denote the $(k-1)$-th order tensor obtained by applying T to vector v.

D.1. Preliminaries

We now state a few preliminary facts and auxiliary results that play an important role in the proof.

D.1.1. TELESCOPE IDENTITY

The proofs of all theorems make use of a basic telescope identity. In particular, we define the noise term

$$e_t(\theta) := \underbrace{(A_t - \bar{A})}_{\Xi_t}\theta - \underbrace{(b_t - b)}_{\xi_t}. \tag{24}$$

With this shorthand, some straightforward algebra shows that the Polyak-Ruppert averaged iterate $\bar{\theta}_T$ satisfies the *telescope relation*

$$\bar{A}(\bar{\theta}_T - \theta^*) = \frac{\theta_0 - \theta_T}{\eta T} - \frac{1}{T}\sum_{t=0}^{T-1} e_{t+1}(\theta_t), \tag{25}$$

involving the non-averaged sequence $\{\theta_t\}_{t\geq 1}$.

D.1.2. PROPERTIES OF THE PROCESS $\{\theta_t\}_{t\geq 0}$

We make repeated use of a number of basic properties of the Markov process $\{\theta_t\}_{t\geq 0}$, which we state here for future reference. All of these claims are proved in Appendix I.

Lemma 10 *Under Assumptions 1, 3, and 2, for any step size $\eta \in \left(0, \frac{\lambda}{\rho^2(\bar{A})+\kappa^2(U)v_A^2}\right)$ and any $t \geq 1$, we have the moment bounds*

$$\mathbb{E}\left\|\theta_t - \theta^*\right\|_2^2 \leq \kappa^2(U)\left(\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2 + \frac{\eta}{\lambda}(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d)\right). \tag{26a}$$

If we assume furthermore that $(2+\alpha)$-moments of the noises Ξ_A and ξ_b are finite, there exists a constant η_0, such that for $\eta < \eta_0$ we have:

$$\mathbb{E}\left\|\theta_t - \theta^*\right\|_2^{2+\alpha} \leq M \quad \text{for some } M < \infty. \tag{26b}$$

See Appendix I.1 for the proof of this claim.

For future use, we also state a foundational lemma on the stationary distribution of the Markov chain.

Lemma 11 *Under Assumptions 1, 3, and 2, for any choice of step size $\eta \in \left(0, \frac{\lambda}{\rho^2(\bar{A})+\kappa^2(U)v_A^2}\right)$, the Markov process $(\theta_t)_{t=0}^{+\infty}$ satisfies the following properties: (i) it has a unique stationary distribution π_η; and (ii) the stationary distribution has finite second moments, and concretely we have*

$$\mathbb{E}_{\pi_\eta}(\theta) = \theta^*, \quad \text{and} \quad \operatorname{cov}_{\pi_\eta}(\theta) = \Lambda_\eta^*, \tag{27a}$$

where Λ_η^* is the unique solution to equation (9). Finally, we have the moment bound

$$\mathbb{E}_{\pi_\eta}\left\|\theta - \theta^*\right\|_2^2 \leq \kappa^2(U)\frac{\eta}{\lambda}(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d). \tag{27b}$$

See Appendix I.2 for the proof of this claim.

In the following, we state a coupling result that allows us to prove existence of the stationary distribution, and to control the rate of convergence to stationarity. We first observe that using standard properties of the Kronecker product, the matrix equation (9) can be re-written in the following equivalent but vectorized form:

$$(A \oplus A - \eta A \otimes A - \eta \mathbb{E}(\Xi_A \otimes \Xi_A)) \operatorname{vec}(\Lambda) = \eta \operatorname{vec}(\Sigma^*). \tag{28}$$

Moreover, since we have $A \oplus A \succeq 2\lambda$ under Assumption 2, the minimal requirement (up to constant factors) on the step size η for equation (9) to have a PSD solution is:

$$A \oplus A - \eta A \otimes A - \eta \mathbb{E}(\Xi_A \otimes \Xi_A) \succeq \lambda I_{d \times d}. \tag{29}$$

With this definition, we have

Lemma 12 *Suppose that Assumptions 1, 3 and 2 all hold, and consider the Markov chain $(\theta_t)_{t \geq 0}$ with any step size $\eta > 0$ satisfying equation (29). Then for any two starting points $\theta_0^{(1)}$ and $\theta_0^{(2)}$, we have:*

$$\mathcal{W}_2(\mathcal{L}(\theta_T^{(1)}), \mathcal{L}(\theta_T^{(2)})) \leq e^{-\lambda \eta T/2} \kappa(U) \left\| \theta_0^{(1)} - \theta_0^{(2)} \right\|_2. \tag{30}$$

In particular, any $\eta \leq \frac{\lambda}{\rho(A)^2 + \kappa^2(U) v_A^2}$ satisfies equation (29) and makes the above claim true.

See Appendix I.3 for the proof of claim.

An elementary consequence of Lemma 12 is the following bound on the Wasserstein-2 distance:

$$\mathcal{W}_2\left(\mathcal{L}(\theta_T), \pi_\eta\right) \leq e^{-\frac{\eta \lambda T}{2}} \kappa(U) \mathcal{W}_2(\mu, \pi_\eta). \tag{31}$$

The proof of this claim is straightforward: we simply take the optimal coupling between the initial laws μ_0 and π_η, apply Lemma 12 conditionally on the starting points, and then take expectations.

Finally, we give control on the support size and coupling estimates on the process in the ℓ_∞ setting, which is used in the proof of Theorem 5.

Lemma 13 *Under Assumption 1, 4 and 5, for $\eta \leq 1$, given $\theta_0 \in [-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$, we have $\|\theta_t\|_\infty \leq \bar{\lambda}^{-1}$ for any $t \geq 0$. Furthermore, for any two starting points $\theta_0^{(1)}, \theta_0^{(2)} \in [-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$, we have:*

$$\mathcal{W}_{\|\cdot\|_\infty, \infty}(\mathcal{L}(\theta_1^{(1)}), \mathcal{L}(\theta_1^{(2)})) \leq (1 - \eta \bar{\lambda}) \left\| \theta_0^{(1)} - \theta_0^{(2)} \right\|_\infty.$$

See Appendix I.4 for the proof of this lemma.

Appendix E. Proof of Theorem 2

We are now equipped to prove Theorem 2. First, by the telescope identity (25), we have

$$\frac{\theta_T - \theta_0}{\eta \sqrt{T}} = -\bar{A} \left[\frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} (\theta_t - \theta^*) \right] - \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} e_{t+1}(\theta_t).$$

From its definition, it can be seen that the sequence $\{e_{t+1}(\theta_t)\}_{t\geq 0}$ is a vector martingale difference sequence with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ (for notational consistency, we let $\mathcal{F}_{-1}$ denote the trivial σ-field). Accordingly, we can apply a martingale CLT en route to establishing the claim. In order to do so, we begin by computing the relevant conditional second moments.

We let $r_t := \theta_t - \theta^*$ denote the error in the non-averaged sequence at time t. Observe that we have the relation $e_{t+1}(\theta_t) = e_{t+1}^{(1)} + e_{t+1}^{(2)}$, where

$$e_{t+1}^{(1)} := \Xi_{t+1} r_t, \quad \text{and} \quad e_{t+1}^{(2)} := -\xi_{t+1} + \Xi_{t+1}\theta^*.$$

Based on this decomposition, we can expand the conditional covariance of $e_{t+1}(\theta_t)$ as a sum of four terms:

$$\mathbb{E}\left[e_{t+1}(\theta_t)e_{t+1}(\theta_t)^\top \mid \mathcal{F}_t\right] = \mathbb{E}\left[e_{t+1}^{(1)}(e_{t+1}^{(1)})^\top + e_{t+1}^{(2)}(e_{t+1}^{(2)})^\top + e_{t+1}^{(1)}(e_{t+1}^{(2)})^\top + e_{t+1}^{(2)}(e_{t+1}^{(1)})^\top \mid \mathcal{F}_t\right].$$

We treat each of these four terms in turn. For the first term, we note that:

$$\begin{aligned}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[e_{t+1}^{(1)}(e_{t+1}^{(1)})^\top \mid \mathcal{F}_t\right] &= \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\Xi_{t+1}r_t r_t^\top \Xi_{t+1}^\top \mid \mathcal{F}_t\right] \\ &= \mathbb{E}(\Xi_A \otimes \Xi_A)\left[\frac{1}{T}\sum_{t=0}^{T-1} r_t r_t^\top\right]. \qquad (32a)\end{aligned}$$

Here $\mathbb{E}(\Xi_A \otimes \Xi_A)$ is a fourth-order tensor. As noted in Section D, the square brackets denote the tensor applying to a matrix $\frac{1}{T}\sum_{t=0}^{T-1} r_t r_t^\top$, resulting in a $d \times d$ matrix.

For the second term, by Assumption 1, the noises Ξ_t and ξ_t are uncorrelated, so we have:

$$\begin{aligned}\mathbb{E}\left[e_{t+1}^{(2)}(e_{t+1}^{(2)})^\top \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\left(-\xi_{t+1} + \Xi_{t+1}\theta^*\right)\left(-\xi_{t+1} + \Xi_{t+1}\theta^*\right)^\top \mid \mathcal{F}_t\right] \\ &= \mathbb{E}(\xi\xi^\top) + \mathbb{E}\left((\Xi_A\theta^*)(\Xi_A\theta^*)^\top\right). \qquad (32b)\end{aligned}$$

For the third term, we note that:

$$\begin{aligned}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[e_{t+1}^{(1)}(e_{t+1}^{(2)})^\top \mid \mathcal{F}_t\right] &= \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\Xi_{t+1}r_t(\Xi_{t+1}\theta^*)^\top \mid \mathcal{F}_t\right] \\ &= \mathbb{E}(\Xi_A \otimes \Xi_A)\left[\frac{1}{T}\sum_{t=0}^{T-1} r_t \theta^{*\top}\right]. \qquad (32c)\end{aligned}$$

Similarly, for the fourth term, we have

$$\begin{aligned}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[e_{t+1}^{(2)}(e_{t+1}^{(2)})^\top \mid \mathcal{F}_t\right] &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[e_{t+1}^{(2)}(e_{t+1}^{(1)})^\top \mid \mathcal{F}_t\right] \\ &= \mathbb{E}(\Xi_A \otimes \Xi_A)\left[\frac{1}{T}\sum_{t=0}^{T-1} \theta^* r_t^\top\right]. \qquad (32d)\end{aligned}$$

The second conditional expectation term is a deterministic quantity, while other three terms depend on the random variable r_t. When taking the quadratic variation of the martingale M_t, we get the partial sum of functions of a Markov chain $(\theta_t)_{t\geq 0}$. Accordingly, we now use Lemma 11, which guarantees the existence of a unique stationary measure π_η, in order to study the limits of the first three terms.

Note that for any vectors $u, v \in \mathbb{S}^{d-1}$, the functions $(u, v) \mapsto (u^\top \theta)(v^\top \theta)$ and $v \mapsto (v^\top \theta)(v^\top \theta^*)$ are L^1 integrable under the stationary measure π_η. Consequently, by Birkhoff's ergodic theorem (cf. Kallenberg (2006), Theorem 9.6), we have:

$$\frac{1}{T}\sum_{t=0}^{T-1} u^\top r_t r_t^\top v \to u^\top \mathbb{E}_{\pi_\eta}(\theta-\theta^*)(\theta-\theta^*)^\top v = u^\top \Lambda_\eta^* v, \quad \text{a.s.}$$
$$\frac{1}{T}\sum_{t=0}^{T-1} u^\top r_t {\theta^*}^\top v \to u^\top (\mathbb{E}_{\pi_\eta}\theta - \theta^*){\theta^*}^\top v = 0, \quad \text{a.s.}$$

Thus, the ergodic averages converge to the corresponding limits, which implies that

$$\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left[e_{t+1}^{(1)}(e_{t+1}^{(1)})^\top \mid \mathcal{F}_t\right] = \mathbb{E}(\Xi_A \otimes \Xi_A)\left[\frac{1}{T}\sum_{t=0}^{T-1} r_t r_t^\top\right] \to \mathbb{E}\left(\Xi_A \Lambda_\eta^* \Xi_A^\top\right), \quad \text{a.s.,} \quad \text{and}$$
$$\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left[e_{t+1}^{(1)}(e_{t+1}^{(2)})^\top \mid \mathcal{F}_t\right] = \mathbb{E}(\Xi_A \otimes \Xi_A)\left[\frac{1}{T}\sum_{t=0}^{T-1} r_t {\theta^*}^\top\right] \to 0, \quad \text{a.s.}$$

Combining the pieces yields

$$\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left[e_{t+1}(\theta_t)(e_{t+1}(\theta_t))^\top \mid \mathcal{F}_t\right] \to \mathbb{E}(\xi_b \xi_b^\top) + \mathbb{E}\left((\Xi_A\theta^*)(\Xi_A\theta^*)^\top\right) + \mathbb{E}\left(\Xi_A \Lambda_\eta^* \Xi_A^\top\right), \quad \text{a.s.}$$

In order to prove the martingale CLT, it remains to verify that the process $e_t(\theta_{t-1})$ satisfies a Lindeberg-type condition when projected in an arbitrary direction $u \in \mathbb{S}^{d-1}$. (Doing so is sufficient since Markov's inequality allows us to translate it to a Lyapunov-type condition.) Accordingly, we seek to bound a $(2+\alpha)$-moment of the martingale differences, which furthermore requires a uniform bound on the $(2+\alpha)$-moment for the process $(\theta_t)_{t\geq 0}$.

Using the $(2+\alpha)$-moment bound (26b) from Lemma 11, we have

$$\begin{aligned}
\mathbb{E}|u^\top e_{t+1}(\theta_t)|^{2+\alpha} &\leq \mathbb{E}\left|2u^\top e_{t+1}^{(1)}\right|^{2+\alpha} + \mathbb{E}\left|2u^\top e_{t+1}^{(2)}\right|^{2+\alpha} \\
&\leq 2^{2+\alpha}\mathbb{E}\left(\|\!|\Xi_{t+1}|\!\|_{\text{op}} \|r_t\|_2\right)^{2+\alpha} + 2^{2+\alpha}\mathbb{E}\left\|\Xi_A\theta^* - \xi_b\right\|_2^{2+\alpha} \\
&\leq 2^{2+\alpha}\mathbb{E}\|\!|\Xi_A|\!\|_{\text{op}}^{2+\alpha} \cdot M + 4^{2+\alpha}\left(\mathbb{E}\left\|\Xi_A\theta^*\right\|_2^{2+\alpha} + \mathbb{E}\left\|\xi_b\right\|_2^{2+\alpha}\right) := Q < +\infty.
\end{aligned}$$

Notably, the quantity Q is independent of t.

Therefore, for a fixed $\epsilon > 0$, the quantity $E := \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left[\left(u^\top e_{t+1}(\theta_t)\right)^2 1\left(\left|u^\top e_{t+1}(\theta_t)\right| > \epsilon\sqrt{T}\right)\right]$ is upper bounded as

$$\begin{aligned}
E &\leq \frac{1}{T}\sum_{t=0}^{T-1} \frac{1}{\epsilon^\alpha T^{\alpha/2}}\mathbb{E}\left[\left|u^\top e_{t+1}(\theta_t)\right|^{2+\alpha} 1\left(\left|u^\top e_{t+1}(\theta_t)\right| > \epsilon\sqrt{T}\right)\right] \\
&\leq \frac{1}{\epsilon^\alpha T^{\alpha/2}} \cdot \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\left|u^\top e_{t+1}(\theta_t)\right|^{2+\alpha} \leq \frac{1}{\epsilon^\alpha T^{\alpha/2}} \cdot Q.
\end{aligned}$$

Note that this bound converges to zero as $T \to \infty$.

Applying the one-dimensional martingale central limit theorem (cf. Corollary 3.1 in the book Hall and Heyde (1980)), we have the convergence of $\frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} u^\top e_{t+1}(\theta_t)$. Combined with the Cramér-Wold device, we conclude that $\frac{1}{\sqrt{T}} \sum_{t=0}^{T} e_{t+1}(\theta_t)$ converges in distribution to a zero-mean Gaussian with covariance $\mathbb{E}(\Xi_A \Lambda_\eta^* \Xi_A^\top) + \Sigma^*$. By Lemma 10, we have $\sqrt{T} \cdot \frac{1}{\eta T}(\theta_T - \theta^*) \to 0$ almost surely. Therefore, by the telescoping equation (25), we have:

$$A \left[\frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} (\theta_t - \theta^*) \right] \xrightarrow{d} \mathcal{N}\left(0, \mathbb{E}(\Xi_A \Lambda \Xi_A^\top) + \Sigma^*\right).$$

Taking the inverse of A completes the proof.

Appendix F. Proof of Theorem 3

In this section, we provide a proof for Theorem 3, the non-asymptotic concentration result. In order to prove this theorem, we require an auxiliary result that provides bounds on higher-order moments of the process.

Lemma 14 *Suppose that Assumptions 1, 3′ and 2 all hold. Given some $p \geq 2 \log T$, consider any step size $\eta \in \left(0, \frac{\lambda}{\rho^2(\bar{A}) + C p^{2\alpha+1} \kappa^2(U) \sigma_A^2}\right)$. Then there is a universal constant c such that*

$$(\mathbb{E} \|\theta_t - \theta^*\|_2^p)^{\frac{2}{p}} \leq c\, \kappa^2(U) \left((\mathbb{E} \|\theta_0 - \theta^*\|_2^p)^{\frac{2}{p}} + \frac{\eta}{\lambda} (p^{2\beta+1} \sigma_b^2 d + p^{2\alpha+1} \sigma_A^2 \|\theta^*\|_2^2) \right). \tag{33}$$

See Appendix F.1 for the proof of this claim. Recall that the matrix U is defined in Lemma 1, which guarantees that $\bar{A} = U D U^{-1}$. We will use this notation throughout the proof.

Equipped with this lemma, we now turn to the proof of the theorem. We consider the martingale term $M_t := \sum_{s=0}^{t-1} e_{s+1}(\theta_s)$. By the telescope equation (25), we need to bound in any direction the variation of $\frac{1}{T\eta} \bar{A}^{-1}(\theta_0 - \theta_T)$ and $\frac{1}{T} \sum_{t=0}^{T-1} \bar{A}^{-1} e_{t+1}(\theta_t)$, respectively. For any vector $v \in \mathbb{S}^{d-1}$, define $M_t^{(v)} := \sum_{s=0}^{t-1} \bar{A}^{-1} v^\top e_{s+1}(\theta_s)$. Since $M_t^{(v)}$ is a martingale, we can apply the discrete-time Burkholder-Davis-Gundy (BDG) inequality (Burkholder et al., 1972): it guarantees the existence of a finite constant C such that for any $p \geq 4$, we have

$$\mathbb{E} \sup_{0 \leq t \leq T} |M_t^{(v)}|^p \leq (Cp)^{\frac{p}{2}} \mathbb{E} \langle M^{(v)} \rangle_T^{\frac{p}{2}} = (Cp)^{\frac{p}{2}} \mathbb{E} \left(\sum_{t=0}^{T-1} (v^\top e_{t+1}(\theta_t))^2 \right)^{\frac{p}{2}}.$$

Moreover, we have

$$\begin{aligned} \mathbb{E} \left(\sum_{t=0}^{T-1} (v^\top e_{t+1}(\theta_t))^2 \right)^{\frac{p}{2}} &= \mathbb{E} \left(\sum_{t=0}^{T-1} \left((v^\top \Xi_{t+1} \theta_t)^2 + (\xi_{t+1}^\top v)^2 - 2(v^\top \Xi_{t+1} \theta_t)(v^\top \xi_{t+1}) \right) \right)^{\frac{p}{2}} \\ &\leq 6^{p/2} \sum_{j=1}^{3} I_j, \end{aligned}$$

where $I_1 := \mathbb{E}\left(\sum_{t=0}^{T-1}(v^\top \Xi_{t+1}\theta_t)^2\right)^{\frac{p}{2}}$, along with

$$I_2 := \mathbb{E}\left(\sum_{t=0}^{T-1}(v^\top \xi_{t+1})^2\right)^{\frac{p}{2}}, \quad \text{and} \quad I_3 := \mathbb{E}\left|\sum_{t=0}^{T-1}(v^\top \Xi_{t+1}\theta_t)(v^\top \xi_{t+1})\right|^{\frac{p}{2}}.$$

By the Cauchy-Schwartz inequality, we have:

$$\begin{aligned} I_3 &\leq \mathbb{E}\left(\sqrt{\sum_{t=0}^{T-1}(v^\top \Xi_{t+1}\theta_t)^2} \cdot \sqrt{\sum_{t=0}^{T-1}(v^\top \xi_{t+1})^2}\right)^{\frac{p}{2}} \\ &\leq \sqrt{\mathbb{E}\left(\sum_{t=0}^{T-1}(v^\top \Xi_{t+1}\theta_t)^2\right)^{\frac{p}{2}} \cdot \mathbb{E}\left(\sum_{t=0}^{T-1}(v^\top \xi_{t+1})^2\right)^{\frac{p}{2}}} = \sqrt{I_1 \cdot I_2} \leq (I_1 + I_2)/2. \end{aligned}$$

So we only need to bound the terms I_1 and I_2.

Denote the following quantity:

$$B_p := \|\theta_0 - \theta^*\|_2 + \frac{\eta}{\lambda}(\sigma_b\sqrt{d}(p\log T)^{\beta+1/2} + \sigma_A \|\theta^*\|_2 (p\log T)^{\alpha+1/2}). \tag{34}$$

According to Lemma 14, intuitively, for large p, the quantity $\kappa(U)B_p$ can be used as a uniform high-probability upper bound on the distances $\|\theta_t - \theta^*\|_2$, for $t = 0, 1, \cdots, T$. This quantity involves in the upper bounds of I_1. We also denote the matrix $\Sigma_\xi := \mathbb{E}(\xi_b\xi_b^\top)$.

We now state an auxiliary result that bounds each of these terms:

Lemma 15 *We have the bounds*

$$I_2 \leq (2v^\top \Sigma_\xi v T)^{\frac{p}{2}} + C_\beta^p \sigma_b^p \left((pT)^{\frac{p}{4}} + (p\log T)^{\frac{p}{2}(1+2\beta)}\right), \tag{35a}$$

and

$$\begin{aligned} (I_1)^{\frac{2}{p}} \leq\ & 3Tv^\top \mathbb{E}(\Xi_A(\Lambda_\eta^* + \theta^*\theta^{*\top})\Xi_A^\top)v + \frac{12v_A^2\kappa^2(U)}{\lambda\eta}\left(\mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2 + \|\theta_0 - \theta^*\|_2^2\right) \\ &+ C\|\mathbb{E}(\Xi_A vv^\top \Xi_A^\top)\|_{op}\frac{\kappa^2(U)}{\lambda}B_p\left(\sigma_A(B_p + \|\theta^*\|_2)(p\log T)^\alpha + \sigma_b\sqrt{d}(p\log T)^\beta\right)\sqrt{pT\log T} \\ &+ \sqrt{CpT}\sigma_A^2 p^{2\alpha}\kappa^2(U)B_p^2. \end{aligned} \tag{35b}$$

See Section F.2 for the proof of this claim.

Combining the results for I_1, I_2, I_3, we obtain the main moment bound on the supremum of martingale $M_t^{(v)}$. Denote the matrix $\tilde{\Sigma} := \mathbb{E}(\Xi_A\Lambda_\eta^*\Xi_A)$, and denote $Z_p := \sigma_A \|\theta^*\|_2 (p\log T)^\alpha +$

$\sigma_b\sqrt{d}(p\log T)^\beta$. We obtain:

$$\frac{1}{\sqrt{T}}\left(\mathbb{E}\sup_{0\le t\le T}|M_t^{(v)}|^p\right)^{\frac{1}{p}} \lesssim \sqrt{pv^\top(\Sigma^*+\tilde{\Sigma})v}+\sqrt{p}\sigma_b\left((\frac{p}{T})^{\frac{1}{4}}+\frac{(p\log T)^{\beta+1/2}}{\sqrt{T}}\right)$$
$$+p\log T\cdot T^{-\frac{1}{4}}\frac{\kappa(U)\sqrt{\|\tilde{\Sigma}\|_{\text{op}}}}{\sqrt{\lambda\eta}}(\sqrt{\frac{\eta}{\lambda}}Z_p+\|\theta^*-\theta_0\|_2)+\frac{v_A\kappa(U)}{\sqrt{T\lambda\eta}}(\|\theta^*\|_2+\|\theta_0\|_2+\sqrt{\text{trace}(\Lambda_\eta^*)})$$
$$+\sqrt{p}T^{-\frac{1}{4}}p^{\alpha+\beta}\sqrt{\sigma_A\sigma_b}\kappa(U)(\|\theta_0-\theta^*\|_2+\sqrt{\frac{\eta}{\lambda}}Z_p),$$

for $p>2\log T$ and η satisfying the assumption in the theorem.

For the bias term, we note that:

$$(\mathbb{E}\|\theta_T-\theta^*\|_2^p)^{\frac{2}{p}}\le\kappa^2(U)\left(\|\theta_0-\theta^*\|_2+\frac{\eta}{\lambda}Z_p\right).$$

Finally, putting together the previous results and merging the terms, we obtain the upper bound

$$\sqrt{T}\left(\mathbb{E}|v^\top A(\bar{\theta}_T-\theta^*)|^p\right)^{\frac{1}{p}}\lesssim\sqrt{pv^\top(\Sigma^*+\tilde{\Sigma})v}$$
$$+\kappa(U)(p\log T)^{2\max(\alpha,\beta)+2}\left(\frac{\sigma_A+\sigma_b}{T^{\frac{1}{4}}}+\frac{1+\sqrt{\sigma_A/\lambda}}{\eta\sqrt{T}}\right)\left(\|\theta^*\|_2+\|\theta_0-\theta^*\|_2+\sqrt{\frac{\eta}{\lambda}}(\sigma_A\|\theta^*\|_2+\sigma_b\sqrt{d})\right).$$

Applying Markov's inequality yields the claimed high-probability bound.

F.1. Proof of Lemma 14

We decompose $\bar{A}$ in the form $\bar{A}=UDU^{-1}$ that is guaranteed by Lemma 1. We study the dynamics of $\left\|U^{-1}(\theta_t-\theta^*)\right\|_2$. Defining the residual term $r_t:=\theta_t-\theta^*$, we observe that

$$\begin{aligned}
&\left\|U^{-1}r_{t+1}\right\|_2^2\\
&=(r_t-\eta(A+\Xi_{t+1})(r_t+\theta^*)-\eta\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}U^{-1}(r_t-\eta(A+\Xi_{t+1})(r_t+\theta^*)-\eta\xi_{t+1})\\
&=(U^{-1}r_t)^{\mathsf{H}}(I-\eta(D+D^{\mathsf{H}})+\eta^2D^{\mathsf{H}}D)(U^{-1}r_t)-2\eta\text{Re}\left((\Xi_{t+1}(r_t+\theta^*)+\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}(I-\eta D)U^{-1}r_t\right)\\
&\qquad+\eta^2\left\|U^{-1}(\Xi_{t+1}r_t+\Xi_{t+1}\theta^*+\xi_{t+1})\right\|_2^2\\
&\le(1-\eta\lambda+\eta^2\rho^2(\bar{A}))\left\|U^{-1}r_t\right\|_2^2-2\eta\text{Re}\left((\Xi_{t+1}(r_t+\theta^*)+\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}(I-\eta D)U^{-1}r_t\right)\\
&\qquad+3\eta^2\|U^{-1}\|_{\text{op}}^2\left(\|\Xi_{t+1}r_t\|_2^2+\|\Xi_{t+1}\theta^*\|_2^2+\|\xi_{t+1})\|_2^2\right).
\end{aligned}$$

Telescoping this expression, for $\eta \in \left(0, \frac{\lambda}{\rho^2(A)}\right)$, we have:

$$e^{\eta\lambda T}\left\|U^{-1}r_T\right\|_2^2 \leq \left\|U^{-1}r_0\right\|_2^2 - 2\eta \underbrace{\sum_{t=0}^{T-1} e^{\eta\lambda t}\mathrm{Re}\left((\Xi_{t+1}(r_t+\theta^*)+\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}(I-\eta D)U^{-1}r_t\right)}_{:=S_1(T)}$$
$$+ 3\eta^2 \underbrace{\sum_{t=0}^{T-1} e^{\eta\lambda t}\|U^{-1}\|_{\text{op}}^2\left(\|\Xi_{t+1}r_t\|_2^2 + \|\Xi_{t+1}\theta^*\|_2^2 + \|\xi_{t+1})\|_2^2\right)}_{:=S_2(T)}.$$

Note that the process $\{S_1(T)\}$ is a martingale and the process $\{S_2(T)\}$ is non-decreasing.

Let us adopt $\mathbb{E}\sup_{0\leq t\leq T}\left(e^{\lambda\eta t}\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{p}{2}}$ as a Lyapunov function. By Young's inequality we obtain:

$$\mathbb{E}\sup_{0\leq t\leq T}\left(e^{\lambda\eta t}\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{p}{2}} \leq 3^{\frac{p}{2}}\mathbb{E}\left\|U^{-1}r_0\right\|_2^p + 6^{\frac{p}{2}}\eta^{\frac{p}{2}}\mathbb{E}\sup_{1\leq t\leq T}|S_1(t)|^{\frac{p}{2}} + 9^{\frac{p}{2}}\eta^p\mathbb{E}(S_2(T))^{\frac{p}{2}}.$$

We upper bound the two terms respectively.

Upper bound for $|S_1|$: Note that:

$$\begin{aligned}
&\left|(\Xi_{t+1}(r_t+\theta^*)+\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}(I-\eta D)U^{-1}r_t\right| \\
&\quad\leq \left\|(U^{-1}\Xi_{t+1}r_t)+(U^{-1}\xi_{t+1})+U^{-1}\Xi_{t+1}\theta^*\right\|_2\cdot\|I-\eta D\|_{\text{op}}\cdot\left\|U^{-1}r_t\right\|_2 \\
&\quad\quad\leq 2\|U^{-1}\|_{\text{op}}\left(\|\Xi_{t+1}r_t\|_2+\|\xi_{t+1}\|_2+\|\Xi_{t+1}r_t\|_2\right)\left\|U^{-1}r_t\right\|_2.
\end{aligned}$$

Applying the Burkholder-Davis-Gundy inequality to the martingale $S_1(t)$, we have:

$$\begin{aligned}
&\mathbb{E}\sup_{1\leq t\leq T}|S_1(t)|^{\frac{p}{2}} \leq (Cp)^{\frac{p}{4}}\,\mathbb{E}\langle S_1\rangle_T^{\frac{p}{4}} \\
&= (Cp)^{\frac{p}{4}}\,\mathbb{E}\left(\sum_{t=0}^{T-1}e^{2\eta\lambda t}\left|(\Xi_{t+1}(r_t+\theta^*)+\xi_{t+1})^{\mathsf{H}}(U^{-1})^{\mathsf{H}}(I-\eta D)U^{-1}r_t\right|^2\right)^{\frac{p}{4}} \\
&\leq (Cp)^{\frac{p}{4}}\,\|U^{-1}\|_{\text{op}}^{\frac{p}{2}}\mathbb{E}\left(\sum_{t=0}^{T-1}e^{2\eta\lambda t}\left(\|\Xi_{t+1}r_t\|_2^2\left\|U^{-1}r_t\right\|_2^2 + (\|\xi_{t+1}\|_2^2+\|\Xi_{t+1}\theta^*\|_2^2)\left\|U^{-1}r_t\right\|_2^2\right)\right)^{\frac{p}{4}}.
\end{aligned}$$

By Hölder's inequality, we have:

$$\begin{aligned}
&\left(\sum_{t=0}^{T-1}e^{2\eta\lambda t}\left((\|\Xi_{t+1}r_t\|_2^2+\|\xi_{t+1}\|_2^2+\|\Xi_{t+1}\theta^*\|_2^2)\left\|U^{-1}r_t\right\|_2^2\right)\right)^{\frac{p}{4}} \\
&\quad\leq\left(\sum_{t=0}^{T-1}e^{\frac{2p}{p-4}\eta\lambda t}\right)^{\frac{p}{4}-1}\left(3\sum_{t=0}^{T-1}(\|\Xi_{t+1}r_t\|_2^{\frac{p}{2}}+\|\xi_{t+1}\|_2^{\frac{p}{2}}+\|\Xi_{t+1}\theta^*\|_2^{\frac{p}{2}})\left\|U^{-1}r_t\right\|_2^{\frac{p}{2}}\right).
\end{aligned}$$

For the geometric series, we have $\left(\sum_{t=0}^{T-1} e^{\frac{2p}{p-4}\eta\lambda t}\right)^{\frac{p}{4}-1} \leq \frac{1}{(\eta\lambda)^{\frac{p}{4}-1}} e^{\eta\lambda p T}$.

By Assumption 3′, we have:

$$\mathbb{E}\left\|\xi_{t+1}\right\|_2^{\frac{p}{2}} \leq p^{p\beta/2}(\sigma_b\sqrt{d})^{p/2}, \quad \mathbb{E}\left\|\Xi_{t+1}v\right\|_2^{\frac{p}{2}} \leq p^{p\alpha/2}\sigma_A^{p/2}\left\|v\right\|_2^{p/2}.$$

Putting together the pieces, we obtain:

$$\mathbb{E}\sup_{1\leq t\leq T}|S_1(t)|^{\frac{p}{2}} \leq \frac{(Cp)^{\frac{p}{4}}e^{\eta\lambda pT/2}}{(\lambda\eta)^{\frac{p}{4}}}\sum_{t=0}^{T-1}\Big(p^{\frac{p\beta}{2}}(\sigma_b\sqrt{d})^{\frac{p}{2}}\|U^{-1}\|_{\text{op}}^{\frac{p}{2}}\mathbb{E}\left\|U^{-1}r_t\right\|_2^{\frac{p}{2}} + p^{\frac{p\alpha}{2}}\sigma_A^{\frac{p}{2}}\kappa(U)^{\frac{p}{2}}\mathbb{E}\left\|U^{-1}r_t\right\|_2^{p} + p^{\frac{p\alpha}{2}}\sigma_A^{\frac{p}{2}}\|U^{-1}\|_{\text{op}}^{\frac{p}{2}}\mathbb{E}\left\|\theta^*\right\|_2^{p}\Big).$$

Upper bounds on S_2: By Young's inequality, we have:

$$\begin{aligned}(S_2(T))^{\frac{p}{2}} &= \left(\sum_{t=0}^{T-1}e^{\eta\lambda t}\|U^{-1}\|_{\text{op}}^2\left(\|\Xi_{t+1}r_t\|_2^2+\|\xi_{t+1}\|_2^2+\|\Xi_{t+1}\theta^*\|_2^2\right)\right)^{\frac{p}{2}}\\ &\leq \|U^{-1}\|_{\text{op}}^{p}\left[\left(3\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\xi_{t+1}\right\|_2^2\right)^{\frac{p}{2}}+\left(3\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\Xi_{t+1}r_t\right\|_2^2\right)^{\frac{p}{2}}+\left(3\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\Xi_{t+1}\theta^*\right\|_2^2\right)^{\frac{p}{2}}\right].\end{aligned}$$

By Hölder's inequality, we obtain:

$$\begin{aligned}\left(\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\xi_{t+1}\right\|_2^2\right)^{\frac{p}{2}} &\leq \left(\sum_{t=0}^{T-1}e^{\frac{p}{p-2}\eta\lambda t}\right)^{\frac{p}{2}-1}\left(\sum_{t=0}^{T-1}\left\|\xi_{t+1}\right\|_2^p\right),\\ \left(\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\Xi_{t+1}\theta^*\right\|_2^2\right)^{\frac{p}{2}} &\leq \left(\sum_{t=0}^{T-1}e^{\frac{p}{p-2}\eta\lambda t}\right)^{\frac{p}{2}-1}\left(\sum_{t=0}^{T-1}\left\|\Xi_{t+1}\theta^*\right\|_2^p\right),\\ \left(\sum_{t=0}^{T-1}e^{\eta\lambda t}\left\|\Xi_{t+1}r_t\right\|_2^2\right)^{\frac{p}{2}} &\leq \left(\sum_{t=0}^{T-1}e^{\frac{p}{p-2}\eta\lambda t}\right)^{\frac{p}{2}-1}\left(\sum_{t=0}^{T-1}\left\|\Xi_{t+1}r_t\right\|_2^p\right).\end{aligned}$$

For the geometric series, it is easy to see that $\left(\sum_{t=0}^{T-1} e^{\frac{p}{p-2}\eta\lambda t}\right)^{\frac{p}{2}-1} \leq \frac{1}{(\eta\lambda)^{\frac{p}{2}-1}} e^{\eta\lambda pT/2}$.

This yields:

$$\mathbb{E}(S_2(T))^{\frac{p}{2}} \leq \|U^{-1}\|_{\text{op}}^{p}\frac{3^{\frac{p}{2}}}{(\eta\lambda)^{\frac{p}{2}-1}}e^{\eta\lambda pT}\left(\sum_{t=0}^{T-1}\mathbb{E}\left\|\xi_{t+1}\right\|_2^p+\sum_{t=0}^{T-1}\mathbb{E}\left\|\Xi_{t+1}r_t\right\|_2^p+\sum_{t=0}^{T-1}\mathbb{E}\left\|\Xi_{t+1}\theta^*\right\|_2^p\right).$$

By Assumption 3′, we have:

$$\mathbb{E}\left\|\xi_{t+1}\right\|_2^p \leq p^{p\beta}(\sigma_b\sqrt{d})^p, \quad \mathbb{E}\left\|\Xi_{t+1}v\right\|_2^p \leq p^{p\alpha}\sigma_A^p\left\|v\right\|_2^p.$$

Putting the pieces together, we have:

$$\mathbb{E}(S_2(T))^p \leq \frac{e^{\eta\lambda pT/2}}{(\eta\lambda)^{\frac{p}{2}}}\|U^{-1}\|_{\text{op}}^p\left(Tp^{p\beta}(\sigma_b\sqrt{d})^p + Tp^{p\alpha}(\sigma_A\left\|\theta^*\right\|_2)^p + p^{p\alpha}\sigma_A^p\|U\|_{\text{op}}^p\sum_{t=0}^{T-1}\mathbb{E}\left\|U^{-1}r_t\right\|_2^p\right).$$

Defining $H_T := e^{-\frac{\lambda\eta T}{2}}\left(\mathbb{E}\sup_{0\le t\le T}\left(e^{\lambda\eta t}\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{p}{2}}\right)^{\frac{2}{p}}$, clearly we have the upper bound $(\mathbb{E}\left\|U^{-1}r_T\right\|_2^p)^{\frac{p}{2}} \le H_T$. By the decomposition of the Lyapunov function, we get:

$$H_T \le (\mathbb{E}\left\|U^{-1}r_0\right\|_2^p)^{\frac{2}{p}} + 6\eta e^{-\frac{\eta\eta T}{2}}(\mathbb{E}\sup_{1\le t\le T}|S_1(t)|^{\frac{p}{2}})^{\frac{2}{p}} + 6\eta^2 e^{-\frac{\eta\eta T}{2}}(\mathbb{E}S_2(T)^{\frac{p}{2}})^{\frac{2}{p}}.$$

Based on the upper bounds for S_1 and S_2, we have

$$\eta^2 e^{-\frac{\eta\eta T}{2}}(\mathbb{E}S_2(T)^{\frac{p}{2}})^{\frac{2}{p}} \le C\frac{\eta}{\lambda}\left(\|U^{-1}\|_{\text{op}}^2 T^{\frac{2}{p}}(p^{2\beta}\sigma_b^2 d + p^{2\alpha}\sigma_A^2\left\|\theta^*\right\|_2^2) + p^{2\alpha}\kappa^2(U)\sigma_A^2\left(\sum_{t=0}^{T-1}H_t^{\frac{p}{2}}\right)^{\frac{2}{p}}\right),$$

$$\eta e^{-\frac{\eta\eta T}{2}}(\mathbb{E}\sup_{1\le t\le T}|S_1(t)|^{\frac{p}{2}})^{\frac{2}{p}} \le C\sqrt{\frac{p\eta}{\lambda}}\left(\sum_{t=0}^{T-1}((p^{\beta}\sigma_b\sqrt{d} + p^{\alpha}\sigma_A\left\|\theta^*\right\|_2)\|U^{-1}\|_{\text{op}}H_t)^{\frac{p}{4}} + (p^{\alpha}\sigma_A\kappa(U)H_t)^{\frac{p}{2}}\right)^{\frac{2}{p}}.$$

Letting $R_T := \sup_{0\le t\le T}H_t$, and noting that the upper bounds above are non-decreasing in T, we have:

$$\begin{aligned} R_T \le H_0 + C\frac{\eta}{\lambda}T^{\frac{2}{p}}\left(\|U^{-1}\|_{\text{op}}^2(p^{2\beta}\sigma_b^2 d + p^{2\alpha}\left\|\theta^*\right\|_2^2) + p^{2\alpha}\kappa^2(U)\sigma_A^2 R_T\right) \\ + C\sqrt{\frac{p\eta}{\lambda}}T^{\frac{2}{p}}\left(\|U^{-1}\|_{\text{op}}(p^{\beta}\sigma_b\sqrt{d} + p^{\alpha}\sigma_A\left\|\theta^*\right\|_2)\sqrt{R_T} + p^{\alpha}\sigma_A\kappa(U)R_T\right).\end{aligned}$$

Take $p \ge 2\log T$ and $\eta \le \frac{\lambda}{18C^2e^2p^{2\alpha+1}\kappa^2(U)\sigma_A^2}$, we obtain that:

$$R_T \le H_0 + Ce\frac{\eta}{\lambda}\|U^{-1}\|_{\text{op}}^2(p^{\beta}\sigma_b\sqrt{d} + p^{\alpha}\sigma_A\left\|\theta^*\right\|_2)^2 + Ce\sqrt{\frac{p\eta}{\lambda}}\|U^{-1}\|_{\text{op}}(p^{\beta}\sigma_b\sqrt{d} + p^{\alpha}\sigma_A\left\|\theta^*\right\|_2) + \frac{1}{2}R_T,$$

and therefore:

$$\max_{0\le t\le T}(\mathbb{E}\left\|r_t\right\|_2^p)^{\frac{2}{p}} \le \|U^{-1}\|_{\text{op}}^2 R_T \lesssim \kappa^2(U)\left((\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^p)^{\frac{2}{p}} + \frac{\eta}{\lambda}(p^{2\beta+1}\sigma_b^2 d + p^{2\alpha+1}\sigma_A^2\left\|\theta^*\right\|_2^2)\right).$$

Thus, we have completed the proof of Lemma 14.

F.2. Proof of Lemma 15

The remainder of our effort is devoted to proving the bounds on the terms $\{I_1, I_2, I_3\}$ claimed in Lemma 15.

F.2.1. UPPER BOUNDS ON I_1

We begin by observing that

$$\mathbb{E}\sum_{t=0}^{T-1}(v^\top\Xi_{t+1}\theta_t)^2 = \mathbb{E}\sum_{t=0}^{T-1}v^\top\mathbb{E}(\Xi_{t+1}\otimes\Xi_{t+1}^\top|\mathcal{F}_t)[\theta_t\theta_t^\top, v] = \langle\mathbb{E}[\Xi_A vv^\top\Xi_A^\top],\,\mathbb{E}\left(\sum_{t=0}^{T-1}\theta_t\theta_t^\top\right)\rangle.$$

In order to deal with the concentration behavior of this term, we define the two processes:

$$\Psi_T := \sum_{t=0}^{T-1} \mathbb{E}\left((\theta_t^\top \Xi_{t+1} v)^2 | \mathcal{F}_t \right), \quad \text{and} \quad \Upsilon_T := \sum_{t=0}^{T-1} (\theta_t^\top \Xi_{t+1} v)^2 - \Psi_T.$$

By definition, it is easy to see that Υ is a martingale. Applying the BDG inequality and Hölder's inequality, we have:

$$\begin{aligned}
\mathbb{E} \sup_{0 \le t \le T-1} |\Upsilon_t|^{\frac{p}{2}} &\le (Cp)^{\frac{p}{4}} \mathbb{E} \langle \Upsilon \rangle_T^{\frac{p}{4}} \\
&= (Cp)^{\frac{p}{4}} \mathbb{E} \left(\sum_{t=0}^{T-1} \left((\theta_t^\top \Xi_{t+1} v)^2 - \mathbb{E}((\theta_t^\top \Xi_{t+1} v)^2 \mid \mathcal{F}_t)^2 \right) \right)^{\frac{p}{4}} \\
&\le (Cp)^{\frac{p}{4}} \mathbb{E} \left(\sum_{t=0}^{T-1} (\theta_t^\top \Xi_{t+1} v)^4 \right)^{\frac{p}{4}} \\
&\le (Cp)^{\frac{p}{4}} T^{\frac{p}{4}-1} \sum_{t=0}^{T-1} \mathbb{E} |\theta_t^\top \Xi_{t+1} v|^p \\
&\le (Cp)^{\frac{p}{4}} T^{\frac{p}{4}} \sigma_A^p p^{\alpha p} \max_{0 \le t \le T-1} \mathbb{E} \|\theta_t\|_2^p .
\end{aligned}$$

As for the process $\{\Psi_T\}_{T \ge 1}$, a straightforward calculation yields:

$$\Psi_T = \sum_{t=0}^{T-1} \mathbb{E}\left((\theta_t^\top \Xi_{t+1} v)^2 \mid \mathcal{F}_t \right) = \langle \mathbb{E}(\Xi_A v)(\Xi_A v)^\top, \sum_{t=0}^{T-1} \theta_t \theta_t^\top \rangle.$$

The summation $\sum_{t=0}^{T-1} \theta_t \theta_t^\top$ involves terms that are functions of an ergodic Markov chain. Thus, metric ergodicity concentration inequalities based on Ricci curvature techniques can show its concentration around its expectation. We first study the expectation of this process. Let $(\widetilde{\theta}_t)_{t \ge 0}$ be a stationary chain which starts from π_η, couple the processes $(\theta_t)_{t \ge 0}$ and $(\widetilde{\theta}_t)_{t \ge 0}$ in the manner defined by Lemma 12. By definition, there is $\mathbb{E} \widetilde{\theta}_t \widetilde{\theta}_t^\top = \mathbb{E}_{\pi_\eta} \theta \theta^\top$. For any matrix L, we have

$$\begin{aligned}
&\left| \frac{1}{T} \mathbb{E} \left(\sum_{t=0}^{T-1} \langle \theta_t \theta_t^\top, L \rangle \right) - \mathbb{E}_{\pi_\eta} \langle \theta \theta^\top, L \rangle \right| \le \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} \left| \theta_t^\top L \theta_t - \widetilde{\theta}_t^\top L \widetilde{\theta}_t \right| \\
&\le \frac{1}{T} \sum_{t=0}^{T-1} \left(\mathbb{E} \left| (\theta_t - \widetilde{\theta})^\top L (\theta_t - \widetilde{\theta}) \right| + 2\mathbb{E} \left| (\theta_t - \widetilde{\theta}_t)^\top L \widetilde{\theta}_t \right| \right) \\
&\le \frac{1}{T} \sum_{t=0}^{T-1} \left(\|L\|_{\mathrm{op}} \mathbb{E} \left\| \theta_t - \widetilde{\theta}_t \right\|_2^2 + 2 \|L\|_{\mathrm{op}} \sqrt{\mathbb{E} \left\| \theta_t - \widetilde{\theta}_t \right\|_2^2} \cdot \sqrt{\mathbb{E} \left\| \widetilde{\theta}_t \right\|_2^2} \right).
\end{aligned}$$

By Lemma 12, for this coupling, we have:

$$\mathbb{E} \left\| \theta_t - \widetilde{\theta}_t \right\|_2^2 \le \kappa^2(U) e^{-\lambda \eta t} \mathcal{W}_2^2(\mathcal{L}(\theta_0), \pi_\eta).$$

By definition, we have $\mathbb{E}\left\|\widetilde{\theta}_t\right\|_2^2 = \mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2$, and it is easy to see that $\mathcal{W}_2^2(\mathcal{L}(\theta_0), \pi_\eta) \leq \mathbb{E}\|\theta_0 - \theta^*\|_2^2 + \mathbb{E}_{\pi_\eta}\|\theta - \theta^*\|_2^2 \leq \|\theta_0 - \theta^*\|_2^2 + \mathrm{trace}(\Lambda_\eta^*)$. Plugging into the above inequality, we obtain:

$$\left|\frac{1}{T}\mathbb{E}\left(\sum_{t=0}^{T-1}\langle\theta_t\theta_t^\top,\, L\rangle\right) - \mathbb{E}_{\pi_\eta}\langle\theta\theta^\top,\, L\rangle\right| \leq \frac{2\|L\|_{\mathrm{op}}\kappa^2(U)}{T}\left(\mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2 + \|\theta_0 - \theta^*\|_2^2\right)\sum_{t=0}^{T-1} e^{-\frac{\lambda\eta t}{2}}$$
$$\leq \frac{4\|L\|_{\mathrm{op}}\kappa^2(U)}{\lambda\eta T}\left(\mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2 + \|\theta_0 - \theta^*\|_2^2\right).$$

In particular, for the matrix $L = \mathbb{E}((\Xi_A v)(\Xi_A v)^\top)$, we have:

$$\left|\frac{1}{T}\mathbb{E}\Psi_T - v^\top\mathbb{E}(\Xi_A(\Lambda_\eta^* + \theta^*\theta^{*\top})\Xi_A^\top)v\right| \leq \frac{4v_A^2\kappa^2(U)}{\lambda\eta T}\left(\mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2 + \|\theta_0 - \theta^*\|_2^2\right).$$

To obtain a high-probability upper bound for the deviation $\Psi_T - \mathbb{E}\Psi_T$, we use the following ergodic concentration inequality:

Lemma 16 *Under Assumption 1, Assumption 2 and Assumption 3′, for a given initial point θ_0, for a matrix L and given $\delta > 0, T > \log\delta^{-1}$, if the step size η satisfies Eq (11a), with probability $1-\delta$, we have:*

$$\left|\frac{1}{T}\sum_{t=1}^{T}(\theta_t^\top L\theta_t - \mathbb{E}\theta_t^\top L\theta_t)\right| \leq C\|L\|_{op}\frac{\kappa^2(U)}{\lambda}B\left(\sigma_A(B + \|\theta^*\|_2)\log^\alpha\frac{T}{\delta} + \sigma_b\sqrt{d}\log^\beta\frac{T}{\delta}\right)\sqrt{\frac{\log\delta^{-1}}{T}},$$

where $B := \|\theta_0 - \theta^*\|_2 + \frac{\eta}{\lambda}(\sigma_b\sqrt{d}\log^{\beta+1/2}\frac{T}{\delta} + \sigma_A\|\theta^*\|_2\log^{\alpha+1/2}\frac{T}{\delta})$.

The proof of this lemma is postponed to Appendix J.1.

By Lemma 16, for any $\delta > 0$, for $B = \|\theta_0 - \theta^*\|_2 + \frac{\eta}{\lambda}(\sigma_b\sqrt{d}\log^{\beta+1/2}\frac{T}{\delta} + \sigma_A\|\theta^*\|_2\log^{\alpha+1/2}\frac{T}{\delta})$, for $\eta < \frac{\lambda}{\rho^2(\bar{A}) + C\kappa^2(U)\sigma_A^2\log^{2\alpha+1}T/\delta}$, with probability $1-\delta$, we have:

$$|\Psi_T - \mathbb{E}\Psi_T| \leq C\|\mathbb{E}((\Xi_A v)(\Xi_A v)^\top)\|_{\mathrm{op}}\frac{\kappa^2(U)}{\lambda}B\left(\sigma_A(B + \|\theta^*\|_2)\log^\alpha\frac{T}{\delta} + \sigma_b\sqrt{d}\log^\beta\frac{T}{\delta}\right)\sqrt{T\log\delta^{-1}} := Q_\delta.$$

Note that this bound holds true only for a fixed failure probability δ. In order to obtain the moment bounds on Ψ, we also use a coarse estimate: $|\Psi_T - \mathbb{E}\Psi_T| \leq T\|\mathbb{E}((\Xi_A v)(\Xi_A v)^\top)\|_{\mathrm{op}}\max_{0\leq t\leq T-1}.\|\theta_t\|_2^2$. Putting them together, we have:

$$\mathbb{E}|\Psi_T - \mathbb{E}\Psi_T|^{\frac{p}{2}} \leq Q_\delta^{\frac{p}{2}} + \mathbb{E}\left(|\Psi_T - \mathbb{E}\Psi_T|^{\frac{p}{2}}\mathbf{1}_{|\Psi_T - \mathbb{E}\Psi_T| > Q_\delta}\right)$$
$$\leq Q_\delta^{\frac{p}{2}} + \sqrt{\delta T^p\|\mathbb{E}((\Xi_A v)(\Xi_A v)^\top)\|_{\mathrm{op}}^p\mathbb{E}\max_{0\leq t\leq T-1}.\|\theta_t\|_2^{2p}}.$$

By Lemma 14, we have $(\mathbb{E}\max_{0\leq t\leq T-1}.\|\theta_t\|_2^p)^{\frac{1}{p}} \leq C\kappa^2(U)(\|\theta_0 - \theta^*\|_2 + \|\theta^*\|_2 + \frac{T\eta}{\lambda}(\sigma_b\sqrt{d}p^{\beta+1/2} + \sigma_A\|\theta^*\|_2 p^{\alpha+1/2})$. Choosing some $\delta \in \left(0, (CT)^{-p}\right)$, we obtain that:

$$\left(\mathbb{E}|\Psi_T - \mathbb{E}\Psi_T|^{\frac{p}{2}}\right)^{\frac{2}{p}}$$
$$\leq C\|\mathbb{E}((\Xi_A v)(\Xi_A v)^\top)\|_{\mathrm{op}}\frac{\kappa^2(U)}{\lambda}B_p\left(\sigma_A(B_p + \|\theta^*\|_2)(p\log T)^\alpha + \sigma_b\sqrt{d}(p\log T)^\beta\right)\sqrt{pT\log T},$$

where $B_p = \|\theta_0 - \theta^*\|_2 + \frac{\eta}{\lambda}(\sigma_b \sqrt{d}(p \log T)^{\beta+1/2} + \sigma_A \|\theta^*\|_2 (p \log T)^{\alpha+1/2})$ is defined in Eq (34).

Recall that we can decompose I_1 into three parts:

$$I_1 \leq \mathbb{E}\left(\Upsilon_T + \Psi_T\right)^{\frac{p}{2}} \leq 3^{\frac{p}{2}} \left(\mathbb{E}|\Upsilon_T|^{\frac{p}{2}} + (\mathbb{E}\Psi_T)^{\frac{p}{2}} + \mathbb{E}|\Psi_T - \mathbb{E}\Psi_T|^{\frac{p}{2}}\right).$$

Using the bounds for three terms derived above, we obtain:

$$(I_1)^{\frac{2}{p}} \leq 3Tv^\top \mathbb{E}(\Xi_A(\Lambda_\eta^* + \theta^*\theta^{*\top})\Xi_A^\top)v + \frac{12 v_A^2 \kappa^2(U)}{\lambda\eta}\left(\mathrm{trace}(\Lambda_\eta^*) + \|\theta^*\|_2^2 + \|\theta_0 - \theta^*\|_2^2\right)$$
$$+ C\|\mathbb{E}((\Xi_A v)(\Xi_A v)^\top)\|_{\mathrm{op}} \frac{\kappa^2(U)}{\lambda} B_p \left(\sigma_A (B_p + \|\theta^*\|_2)(p \log T)^\alpha + \sigma_b \sqrt{d}(p \log T)^\beta\right) \sqrt{pT \log T}$$
$$+ \sqrt{CpT}\sigma_A^2 p^{2\alpha} \kappa^2(U) B_p^2.$$

F.2.2. UPPER BOUNDS ON I_2:

Define $\xi_T := \sum_{t=0}^{T-1}(v^\top \xi_{t+1})^2$, we have $\mathbb{E}\xi_T = v^\top \Sigma_\xi v T$. It is easy to see that $\xi_t - \mathbb{E}\xi_t$ is a martingale difference sequence, and thus by standard sub-exponential martingale concentration inequalities and Assumption 3′, for $p \geq 2$, we have:

$$\mathbb{E}\left((v^\top \xi_t)^2 - \mathbb{E}(v^\top \xi_t)^2\right)^p \leq \mathbb{E}(v^\top \xi_t)^{2p} \leq p^{2\beta p} \sigma_b^{2p}.$$

By the martingale concentration inequality in Lemma 20, for any $\delta > 0$, we have:

$$\mathbb{P}\left(\frac{1}{T}|\xi_T - \mathbb{E}\xi_T| > C_\beta \sigma_b^2 \left(\sqrt{\frac{\log \delta^{-1}}{T}} + \frac{\log^{1+2\beta} T/\delta}{T}\right)\right) < \delta.$$

Integrating the expression, we obtain the upper bound:

$$I_2 \leq (2v^\top \Sigma_\xi v T)^{\frac{p}{2}} + 2\int_0^{+\infty} \mathbb{P}\left(|\xi_T - \mathbb{E}\xi_T| \geq \varepsilon\right) \varepsilon^{\frac{p}{2}-1} d\varepsilon$$
$$\leq (2v^\top \Sigma_\xi v T)^{\frac{p}{2}} + C_\beta^p \sigma_b^p \left((pT)^{\frac{p}{4}} + (p \log T)^{\frac{p}{2}(1+2\beta)}\right).$$

Appendix G. Proof of Theorem 5

We prove a stronger version of the theorem that involves the quantity $Q(v; \delta)$ defined in Eq (14). It is easy to see that $\sigma_{max}^2 \log \frac{d}{\delta}$ is an upper bound on $Q((\mathbf{e}_j^\top \Gamma^*(\eta)\mathbf{e}_j)_{j=1}^d; \delta)$. So the version stated in Theorem 5 is implied by the stronger version.

In order to prove the theorem, we require an auxiliary lemma that provides an almost-sure bound for the ℓ_∞ norm of the process.

Let $(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_d)$ denote the standard orthonormal basis of $\mathbb{R}^d$. We consider the projection of error terms onto the set of vectors $v_i := (A^{-1})^\top \mathbf{e}_i$ for $i = 1, 2, \cdots, d$. We first note that by Assumption 5, we have:

$$\|v_i\|_1 - 1 \leq \|v_i - \mathbf{e}_i\|_1 \leq \left\|v_i - A^\top v_i\right\|_1 = \sup_{\|u\|_\infty \leq 1} v_i^\top (I_d - A)u \leq (1 - \bar{\lambda}) \|v_i\|_1,$$

and consequently, $\|v_i\|_1 \leq \bar{\lambda}^{-1}$.

We consider the martingales $M_t^{(v_i)}$ for each $i = 1, 2, \cdots, d$. Similar to the proof of Theorem 3, we use the BDG inequality and decompose the deviation into three terms:

$$\mathbb{E} \sup_{0 \leq t \leq T} \left| M_t^{(v_i)} \right|^p \leq (Cp)^{\frac{p}{2}} \mathbb{E} \langle M_t^{(v_i)} \rangle_T^{\frac{p}{2}} \leq (6Cp)^{\frac{p}{2}} \left(I_1 + I_2 + I_3 \right),$$

where $I_1 := \mathbb{E} \left(\sum_{t=0}^{T-1} (v_i^\top \Xi_{t+1} \theta_t)^2 \right)^{\frac{p}{2}}$, along with

$$I_2 := \mathbb{E} \left(\sum_{t=0}^{T-1} (\xi_{t+1}^\top v_i)^2 \right)^{\frac{p}{2}}, \quad \text{and} \quad I_3 := \mathbb{E} \left| \sum_{s=0}^{T-1} (v_i^\top \Xi_{s+1} \theta_s^\top)(v_i^\top \xi_{s+1}) \right|^{\frac{p}{2}}.$$

Similar to the proof of Theorem 3, by Cauchy-Schwartz, we know that $I_3 \leq \sqrt{I_1 I_2} \leq (I_1 + I_2)/2$. We now give upper bounds on the terms I_1 and I_2, respectively.

Upper bound for I_2: For the term I_2, note that the terms $(\xi_t^\top v_i)$ are i.i.d. random variables. And by Assumption 4, $\left| \xi_t^\top v \right| \leq \|\xi_t\|_\infty \cdot \|v_i\|_1 \leq \bar{\lambda}^{-1}$. A simple application of Hoeffding's inequality leads to:

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\left| \frac{1}{T} \sum_{t=0}^{T-1} (\xi_t^\top v_i)^2 - \mathbb{E} (\xi_b^\top v_i)^2 \right| > \varepsilon \right) \leq 2 \exp \left(-T \varepsilon^2 \bar{\lambda}^4 \right),$$

which can be easily converted into a moment bound:

$$I_2^{\frac{2}{p}} \leq C \left(T \cdot \mathbb{E} (\xi_b^\top v_i)^2 + p \sqrt{T} \bar{\lambda}^{-2} \right).$$

Upper bound for I_1: As in the proof of Lemma 15, we decompose the sequence into a martingale term and a predictable sequence. Let $\Psi_T := \sum_{t=1}^T \mathbb{E} \left((v_i^\top \Xi_{t+1} \theta_t)^2 | \mathcal{F}_t \right)$, and let $\Upsilon_T := \sum_{t=1}^T (v_i^\top \Xi_{t+1} \theta_t)^2 - \Psi_T$. By definition, it is easy to see that Υ is a martingale. Note that for each term in Υ, by Lemma 13 and Assumption 4, we have:

$$\left| (v_i^\top \Xi_{t+1} \theta_t)^2 - \mathbb{E}((v_i^\top \Xi_{t+1} \theta_t)^2 | \mathcal{F}_t) \right| \leq 2 \left| (v_i^\top \Xi_{t+1} \theta_t)^2 \right| \leq 2 \|v_i\|_1^2 \cdot \|\Xi_{t+1} \theta_t\|_\infty^2 \leq 2 \|v_i\|_1^2 \cdot \|\theta_t\|_\infty^2 \leq 2 \bar{\lambda}^{-4}.$$

By the Azuma-Hoeffding inequality, we obtain:

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\frac{1}{T} |\Upsilon_T| \geq \varepsilon \right) \leq 2 \exp(-T \varepsilon^2 \bar{\lambda}^{-8} / 4),$$

which can easily be converted to a moment bound:

$$\left(\mathbb{E} |\Upsilon_T|^{\frac{p}{2}} \right)^{\frac{2}{p}} \leq Cp \sqrt{T} \bar{\lambda}^{-4}.$$

Now we turn to an upper bound for the term Ψ_T. Define $\psi(\theta) := \mathbb{E}(v_i^\top \Xi_A \theta)^2$. Note that Ψ_T is the partial sum of function ψ applied to the Markov process $(\theta_t)_{t \geq 0}$. We seek to use the ergodic concentration inequalities based on Ricci curvature techniques (Joulin and Ollivier, 2010).

First, we note that for $\theta_1, \theta_2 \in [-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$, we have:

$$\begin{aligned}\psi(\theta_1) - \psi(\theta_2) &= \mathbb{E}(v_i^\top \Xi_A \theta_1)^2 - \mathbb{E}(v_i^\top \Xi_A \theta_2)^2 \\ &= \mathbb{E}\left((v_i^\top \Xi_A \theta_1)(v_i^\top \Xi_A(\theta_1 - \theta_2))\right) + \mathbb{E}\left((v_i^\top \Xi_A \theta_2)(v_i^\top \Xi_A(\theta_1 - \theta_2))\right) \\ &\leq \|v_i\|_1^2\, \mathbb{E}(\|\Xi_A \theta_1\|_\infty \cdot \|\Xi_A(\theta_1 - \theta_2)\|_\infty) + \|v_i\|_1^2\, \mathbb{E}(\|\Xi_A \theta_2\|_\infty \cdot \|\Xi_A(\theta_1 - \theta_2)\|_\infty) \\ &\leq \bar{\lambda}^{-3} \|\theta_1 - \theta_2\|_\infty .\end{aligned}$$

So ψ is $\bar{\lambda}^{-3}$-Lipschitz under the $\|\cdot\|_\infty$ norm, within the region $[-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$.

Denote by $\mathcal{T}$ the transition kernel of the Markov chain $(\theta_t)_{t\geq 0}$. By Assumption 5, when we take the synchronous coupling by using the same oracle for the process starting at two different points, there is:

$$\mathcal{W}_{\|\cdot\|_\infty, 1}(\mathcal{T}\delta_{\theta_1}, \mathcal{T}\delta_{\theta_2}) \leq \mathbb{E}\,\|(I - \eta A_t)(\theta_1 - \theta_2)\|_\infty \leq (1 - \eta\bar{\lambda})\,\|\theta_1 - \theta_2\|_\infty .$$

So the Markov chain $(\theta_t)_{t\geq 0}$ is a $\mathcal{W}_1$ contraction with parameter $(1 - \eta\bar{\lambda})$ under ℓ_∞ norm. Finally, by Assumption 4, we note that:

$$\mathrm{diam}_{\|\cdot\|_\infty}\left(\mathrm{supp}(\mathcal{T}\delta_\theta)\right) \leq \eta\left(1 + \|\theta\|_\infty\right).$$

So the support size of the one-step transition kernel within the region $[-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$ is uniformly bounded by $2\eta\bar{\lambda}^{-1}$.

We apply the ergodic concentration inequality from Theorem 4 in Joulin and Ollivier (2010) (which is restated in Proposition 19 for completeness), and obtain the following concentration inequality:

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{1}{T}\sum_{t=0}^{T}(\psi(\theta_t) - \mathbb{E}\psi(\theta_t))\right| > \bar{\lambda}^{-3}\varepsilon\right) \leq \begin{cases} 2\exp\left(-\frac{\varepsilon^2 T \bar{\lambda}^2}{128\eta^2}\right) & \varepsilon < \frac{8}{3}\bar{\lambda}^{-1}, \\ 2\exp\left(-\frac{\varepsilon T \bar{\lambda}}{24\eta}\right), & \varepsilon > \frac{8}{3}\bar{\lambda}^{-1}. \end{cases}$$

This tail probability bound can be easily translated into a moment bound:

$$\left(\mathbb{E}\,|\Psi_T|^{\frac{p}{2}}\right)^{\frac{2}{p}} \leq 2\mathbb{E}\Psi_T + C\bar{\lambda}^{-4}\eta\left(\sqrt{Tp} + p\right),$$

for a universal constant $C > 0$.

For the term $\mathbb{E}\Psi_T$, the $\mathcal{W}_1$ contraction implies that:

$$\left|\mathbb{E}\psi(\theta_t) - \mathbb{E}_{\pi_\eta}\psi(\theta)\right| \leq \bar{\lambda}^{-3}(1 - \eta\bar{\lambda})^t \mathbb{E}\,\|\theta_0 - \theta\|_\infty \leq \bar{\lambda}^{-5}(1 - \eta\bar{\lambda})^t.$$

So we obtain $\mathbb{E}\Psi_T \leq T\mathbb{E}_{\pi_\eta}\psi(\theta) + \sum_{t=0}^{T} \bar{\lambda}^{-4}(1 - \eta\bar{\lambda})^t \leq T((\mathbf{e}_i^\top \theta^*)^2 + \mathbf{e}_i^\top \Lambda_\eta^* \mathbf{e}_i) + \frac{1}{\eta\bar{\lambda}^5}$.

Putting these results together, we have:

$$I_1^{\frac{2}{p}} \leq CT((\mathbf{e}_i^\top \theta^*)^2 + \mathbf{e}_i^\top \Lambda_\eta^* \mathbf{e}_i) + C\bar{\lambda}^{-4} p\eta\sqrt{T} + C\bar{\lambda}^{-5}\eta^{-1}.$$

Obtaining the final bound: Combining the upper bounds for I_1 and I_2, we obtain:

$$\left(\mathbb{E} \sup_{0 \le t \le T} \left| M_t^{(v_i)} \right|^p \right)^{\frac{2}{p}} \le CpT \mathbf{e}_i^\top \Gamma^* \mathbf{e}_i + C(\bar{\lambda}^{-4}\eta + \lambda^{-2}) p\sqrt{T} + C\bar{\lambda}^{-5}\eta^{-1}.$$

For the term $\frac{A^{-1}(\theta_0 - \theta^*)}{\eta T}$, we note that by Lemma 13, we have $\|\theta_0 - \theta_T\|_\infty \le 2\bar{\lambda}^{-1}$, and furthermore, we note that for any $v \in \mathbb{R}^d$, we have:

$$\left\|A^{-1}v\right\|_\infty = \left\|(I - A)A^{-1}v\right\|_\infty + \|v\|_\infty \le \|v\|_\infty + (1 - \bar{\lambda}) \left\|A^{-1}v\right\|_\infty,$$

which leads to the fact that $\left\|A^{-1}v\right\|_\infty \le \bar{\lambda}^{-1} \|v\|_\infty$ for any v, and consequently, we have the bound $\left\|A^{-1}(\theta_0 - \theta_T)\right\|_\infty \le \frac{2}{\bar{\lambda}^2}$ almost surely.

Putting these results together, we obtain:

$$\left(\mathbb{E} \left| \sqrt{T} \mathbf{e}_i^\top (\bar{\theta}_T - \theta^*) \right|^p \right)^{\frac{1}{p}} \le C\sqrt{p \mathbf{e}_i^\top \Gamma^*(\eta) \mathbf{e}_i} + C(\bar{\lambda}^{-2}\eta + \bar{\lambda}^{-1})\sqrt{p} T^{-\frac{1}{4}} + C\bar{\lambda}^{-\frac{5}{2}}\eta^{-1}.$$

Converting this bound into a high-probability bound and taking a union bound over the d coordinates, for any $Q > 0$, we obtain:

$$\mathbb{P}\left(\sqrt{T} \left\|\bar{\theta}_T - \theta^*\right\|_\infty \ge C\sqrt{Q} + C\frac{\bar{\lambda}^{-2}\eta + \bar{\lambda}^{-1}}{T^{\frac{1}{4}}} \sqrt{\log \frac{d}{\delta}} + \frac{C\bar{\lambda}^{-\frac{5}{2}}}{\eta\sqrt{T}} \right) \le \frac{\delta}{2} + \sum_{i=1}^d \exp\left(- \frac{Q}{\mathbf{e}_i^\top \Gamma^*(\eta) \mathbf{e}_i} \right).$$

Take $Q = Q\left((\mathbf{e}_i^\top \Gamma^*(\eta) \mathbf{e}_i)_{i=1}^d; \delta/2 \right)$ to obtain the result.

Appendix H. Proof of Theorem 6

The proof is also based on the telescope identity (25). The key ingredient in the proof is an upper bound on the second moment of $\|\theta_t - \theta^*\|_2$, as stated in the following:

Lemma 17 *Under Assumptions 2′, 3 and 1, given a step size* $\eta \le \frac{1}{(\rho(\bar{A}) + 3\kappa(U) v_A)\sqrt{T}}$, *for any integer* $t \in [0, T]$, *we have*

$$\mathbb{E} \|\theta_t - \theta^*\|_2^2 \le e\kappa^2(U) \left(\mathbb{E} \|\theta_0 - \theta^*\|_2^2 + \eta^2 t (v_b^2 d + v_A^2 \|\theta^*\|_2^2) \right),$$

where the matrix U *has columns composed of the eigenvectors of* $\bar{A}$.

See Appendix H.1 for the proof of this claim.

Taking Lemma 17 as given, we now prove Theorem 6. By equation (25), we have:

$$\mathbb{E} \left\|\bar{A}(\bar{\theta}_T - \theta^*)\right\|_2^2 \le \frac{4}{\eta^2 T^2} \left(\mathbb{E} \|\theta_0 - \theta^*\|_2^2 + \mathbb{E} \|\theta_T - \theta^*\|_2^2 \right) + \frac{2}{T^2} \mathbb{E} \|M_T\|_2^2.$$

By Lemma 17, we have:

$$\mathbb{E} \|\theta_T - \theta^*\|_2^2 \le e\kappa^2(U) \left(\mathbb{E} \|\theta_0 - \theta^*\|_2^2 + 3\eta^2 T (v_b^2 d + v_A^2 \|\theta^*\|_2^2) \right).$$

For the martingale term, note that:

$$
\begin{aligned}
\mathbb{E}\left\|M_T\right\|_2^2 &= \mathbb{E}\sum_{t=0}^{T-1}\left\|e_{t+1}(\theta_t)\right\|_2^2 \\
&\leq 3\mathbb{E}\sum_{t=0}^{T-1}\left(\mathbb{E}(\|b_{t+1}-b\|_2^2 \mid \mathcal{F}_t) + \mathbb{E}(\left\|(A_{t+1}-\bar{A})(\theta_t-\theta^*)\right\|_2^2 \mid \mathcal{F}_t) + \mathbb{E}(\left\|(A_{t+1}-\bar{A})\theta^*\right\|_2^2 \mid \mathcal{F}_t)\right) \\
&\leq 3\mathbb{E}\sum_{t=0}^{T-1}\left(v_b^2 d + v_A^2\left\|\theta_t-\theta^*\right\|_2^2 + v_A^2\left\|\theta^*\right\|_2^2\right) \\
&\leq 3Tv_b^2 d + 3Tv_A^2\left\|\theta^*\right\|_2^2 + 3Tv_A^2 e\kappa^2(U)\left(\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^2 + \eta^2 T(v_b^2 d + v_A^2\left\|\theta^*\right\|_2^2)\right).
\end{aligned}
$$

Since $\eta \in \left(0, \frac{1}{\sqrt{T}(\rho(\bar{A})+3\kappa(U)v_A)}\right)$, we have:

$$
\mathbb{E}\left\|M_T\right\|_2^2 \leq 3Tv_A^2 e\kappa^2(U)\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^2 + (3+e)T(v_b^2 d + v_A^2\left\|\theta^*\right\|_2^2).
$$

Putting together the pieces yields

$$
\mathbb{E}\left\|\bar{A}(\bar{\theta}_T-\theta^*)\right\|_2^2 \leq C\left(\frac{\kappa^2(U)}{\eta^2 T^2}\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^2 + \frac{v_b^2 d + v_A^2\left\|\theta^*\right\|_2^2}{T} + \frac{v_A^2\kappa^2(U)}{T}\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^2\right).
$$

Setting the step size as $\eta = \frac{1}{(\rho(\bar{A})+3\kappa(U)v_A)\sqrt{T}}$ yields the claim.

H.1. Proof of Lemma 17

By Assumption 2′, the matrix $\bar{A}$ is diagonalizable. Accordingly, we can write $\bar{A} = UDU^{-1}$, and the remaining part of Assumption 2′ implies that $D + D^{\mathsf{H}} \succeq 0$.

We use the function $f(\theta) = \left\|U^{-1}(\theta-\theta^*)\right\|_2^2$ as a Lyapunov function. From the process dynamics (2), we can write

$$
U^{-1}(\theta_{t+1}-\theta^*) = U^{-1}(I_d-\eta\bar{A})(\theta_t-\theta^*) + \eta U^{-1}\Xi_{t+1}(\theta_t-\theta^*) + \eta U^{-1}\xi_{t+1} - \eta U^{-1}\Xi_{t+1}\theta^*.
$$

Using this decomposition, we can write

$$
\mathbb{E}[\left\|U^{-1}(\theta_{t+1}-\theta^*)\right\|_2^2] = T_1 + \eta^2 T_2 + 2\eta T_3,
$$

where

$$
\begin{aligned}
T_1 &:= \mathbb{E}\left\|U^{-1}(I-\eta\bar{A})(\theta_t-\theta^*)\right\|_2^2 &\text{(36a)}\\
T_2 &:= \mathbb{E}\left\|U^{-1}(\Xi_{t+1}(\theta_t-\theta^*) + \xi_{t+1} - \Xi_{t+1}\theta^*)\right\|_2^2 &\text{(36b)}\\
T_3 &:= \mathbb{E}\left(\langle U^{-1}(I-\eta\bar{A})(\theta_t-\theta^*),\, U^{-1}(\Xi_{t+1}(\theta_t-\theta^*) + \xi_{t+1} - \Xi_{t+1}\theta^*)\rangle\right). &\text{(36c)}
\end{aligned}
$$

We upper bound each these three terms in succession.

Bounding T_1: Using Assumption 2′, we have:

$$
\begin{aligned}
T_1 &= \mathbb{E}(U^{-1}(\theta_t-\theta^*))^{\mathsf{H}}\left(I_d - 2\eta\left(U^{-1}\bar{A}U + (U^{-1}\bar{A}U)^{\mathsf{H}}\right) + \eta^2(U^{-1}\bar{A}U)^{\mathsf{H}}(U^{-1}\bar{A}U)\right)U^{-1}(\theta_t-\theta^*) \\
&\leq \mathbb{E}\left\|U^{-1}(\theta_t-\theta^*)\right\|_2^2 + \eta^2\rho^2(\bar{A})\mathbb{E}\left\|U^{-1}(\theta_t-\theta^*)\right\|_2^2.
\end{aligned}
$$

Bounding T_2: By Young's inequality and Assumption 3, we find that

$$\begin{aligned} T_2 &= \mathbb{E}\left\|U^{-1}(\Xi_{t+1}(\theta_t-\theta^*)+\xi_{t+1}-\Xi_{t+1}\theta^*)\right\|_2^2 \\ &\le 3\|U^{-1}\|_{\text{op}}^2\mathbb{E}\left(\|\Xi_{t+1}(\theta_t-\theta^*)\|_2^2+\|\xi_{t+1}\|_2^2+\|\Xi_{t+1}\theta^*\|_2^2\right) \\ &\le 3\|U^{-1}\|_{\text{op}}^2\left(\|U\|_{\text{op}}^2 v_A^2\mathbb{E}\left\|U^{-1}(\theta_t-\theta^*)\right\|_2^2+v_b^2 d+v_A^2\|\theta^*\|_2^2\right). \end{aligned}$$

Bounding T_3: In this case, we have

$$T_3=\mathbb{E}\left(\langle U^{-1}(I_d-\eta\bar{A})(\theta_t-\theta^*),\, U^{-1}\mathbb{E}\left(\Xi_{t+1}(\theta_t-\theta^*)+\xi_{t+1}-\Xi_{t+1}\theta^*\mid\mathcal{F}_t\right)\rangle\right)=0.$$

This yields:

$$\mathbb{E}\left\|U^{-1}(\theta_{t+1}-\theta^*)\right\|_2^2\le(1+\eta^2\rho^2(\bar{A})+3\eta^2\kappa^2(U)v_A^2)\mathbb{E}\left\|U^{-1}(\theta_t-\theta^*)\right\|_2^2+3\|U^{-1}\|_{\text{op}}^2(v_b^2 d+v_A^2\|\theta^*\|_2^2).$$

Solving the recursion, for $\eta\le\frac{1}{(\rho(\bar{A})+3\kappa(U)v_A)\sqrt{T}}$, we obtain:

$$\begin{aligned} &\mathbb{E}\left\|U^{-1}(\theta_T-\theta^*)\right\|_2^2 \\ &\le\exp\left(\eta^2T(\rho^2(\bar{A})+3\kappa^2(U)v_A^2)\right)\mathbb{E}\left\|U^{-1}(\theta_0-\theta^*)\right\|_2^2 \\ &\qquad+3\eta^2\|U^{-1}\|_{\text{op}}^2(v_b^2 d+v_A^2\|\theta^*\|_2^2)\sum_{t=0}^{T-1}\exp\left(\eta^2 t(\rho^2(\bar{A})+3\kappa^2(U)v_A^2)\right) \\ &\le e\left(\mathbb{E}\left\|U^{-1}(\theta_0-\theta^*)\right\|_2^2+3\eta^2T\|U^{-1}\|_{\text{op}}^2(v_b^2 d+v_A^2\|\theta^*\|_2^2)\right). \end{aligned}$$

Noting that $\|\theta_T-\theta^*\|_2\le\|U\|_{\text{op}}\cdot\left\|U^{-1}(\theta_T-\theta^*)\right\|_2$, we obtain the final result.

Appendix I. Properties of the process $\{\theta_t\}_{t\ge0}$

In this appendix, we prove a number of claims about the basic properties of the process $\{\theta_t\}_{t\ge0}$.

I.1. Proof of Lemma 10

Recall that we use $r_t=\theta_t-\theta^*$ to denote the error in the process at time t. We make use of the function $f(r)=\mathbb{E}\left\|U^{-1}r\right\|_2^2$ for a Lyapunov-type analysis. Observe that the error satisfies the recursion

$$r_{t+1}=r_t-\eta(A_{t+1}\theta_t-b_{t+1})=(I_d-\eta\bar{A})r_t-\eta\Xi_{t+1}\theta_t+\eta\xi_{t+1}.$$

Turning to the squared Euclidean norm, we have

$$\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^2=\mathbb{E}\left\|U^{-1}(I_d-\eta\bar{A})r_t\right\|_2^2+\eta^2\mathbb{E}\left\|U^{-1}(\Xi_{t+1}\theta_t+\xi_{t+1})\right\|_2^2,$$

where we have expanded the quadratic term and used the i.i.d. condition (Assumption 1). Examining the first term, we have

$$\begin{aligned} \left\|U^{-1}(I_d-\eta\bar{A})r_t\right\|_2^2 &= \left\|(I_d-\eta U^{-1}AU)U^{-1}r_t\right\|_2^2 \\ &= \left\|U^{-1}r_t\right\|_2^2-\eta(U^{-1}r_t)^{\mathsf{H}}(D+D^{\mathsf{H}})U^{-1}r_t+\|D^{\mathsf{H}}D\|_{\text{op}}\left\|U^{-1}r_t\right\|_2^2 \\ &\le\left\{1-2\eta\lambda+\eta^2\rho^2(\bar{A})\right\}\left\|U^{-1}r_t\right\|_2^2. \end{aligned}$$

For the second term, by Assumption 3 and Assumption 1, we have:

$$\begin{aligned}\mathbb{E}\left\|U^{-1}(\Xi_{t+1}\theta_t+\xi_{t+1})\right\|_2^2 &\le \|U^{-1}\|_{\mathrm{op}}^2\mathbb{E}\left\|\Xi_{t+1}(\theta^*+r_t)+\xi_{t+1}\right\|_2^2\\ &= \|U^{-1}\|_{\mathrm{op}}^2\left(\mathbb{E}\left\|\Xi_{t+1}(\theta^*+r_t)\right\|_2^2+\mathbb{E}\left\|\xi_{t+1}\right\|_2^2\right) \le \|U^{-1}\|_{\mathrm{op}}^2\left(v_A^2(\|\theta^*\|_2+\mathbb{E}\left\|r_t\right\|_2^2)+v_b^2 d\right).\end{aligned}$$

Putting the pieces together and using the fact that $\eta \in \left(0, \frac{\lambda}{\rho^2(\bar{A})+\kappa^2(U)v_A^2}\right)$, we find that

$$\begin{aligned}\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^2 &\le (1-2\eta\lambda+\eta^2(\rho^2(\bar{A})+\kappa^2(U)v_A^2))\mathbb{E}\left\|U^{-1}r_t\right\|_2^2+\eta^2\|U^{-1}\|_{\mathrm{op}}^2(v_A^2\left\|\theta^*\right\|_2^2+v_b^2 d)\\ &\le (1-\eta\lambda)\mathbb{E}\left\|U^{-1}r_t\right\|_2^2+\eta^2\|U^{-1}\|_{\mathrm{op}}^2(v_A^2\left\|\theta^*\right\|_2^2+v_b^2 d).\end{aligned}$$

By induction, it is easy to show that for any $t \ge 0$,

$$\mathbb{E}\left\|U^{-1}r_t\right\|_2^2 \le \mathbb{E}\left\|U^{-1}(\theta_0-\theta^*)\right\|_2^2+\frac{\eta}{\lambda}\|U^{-1}\|_{\mathrm{op}}^2(v_A^2\left\|\theta^*\right\|_2^2+v_b^2 d),$$

and consequently, we have the bound

$$\mathbb{E}\left\|r_t\right\|_2^2 \le \kappa^2(U)\left(\mathbb{E}\left\|\theta_0-\theta^*\right\|_2^2+\frac{\eta}{\lambda}(v_A^2\left\|\theta^*\right\|_2^2+v_b^2 d)\right).$$

Proof of the bound (26b): In establishing this bound, we use the fact that for scalars $A > 0$, $z \in (-A, +\infty)$ and $\alpha \in (0,1)$, we have

$$(A+z)^{1+\alpha} \le A^{1+\alpha}+(1+\alpha)A^\alpha z+|z|^{1+\alpha}.$$

The proof of this inequality is straightforward: by homogeneity, we only need to prove for the case of $A = 1$. Let $f(z) := 1+(1+\alpha)z+|z|^{1+\alpha}-(1+z)^{1+\alpha}$ for $z \in (-1, +\infty)$. It is easy to see that $f'(z) > 0$ for $z > 0$ and $f'(z) < 0$ for $z < 0$.

By Assumption 2, we have

$$\left\|U^{-1}r_{t+1}\right\|_2^2 \le (1-2\eta\lambda)\left\|U^{-1}r_t\right\|_2^2+2\eta\mathrm{Re}(\langle U^{-1}(1-\eta\bar{A})r_t,\, U^{-1}e_{t+1}(\theta_t)\rangle)+\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2.$$

Taking the $(1+\alpha/2)$-order moment, by the scalar inequality, we obtain:

$$\begin{aligned}\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^{2+\alpha} &\le (1-2\eta\lambda)\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}+\mathbb{E}\left|2\eta\mathrm{Re}(\langle U^{-1}(1-\eta\bar{A})r_t,\, U^{-1}e_{t+1}(\theta_t)\rangle)+\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2\right|^{1+\alpha}\\ &\quad+\mathbb{E}\left[\left((1-2\eta\lambda)\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{\alpha}{2}}\left(2\eta\mathrm{Re}(\langle U^{-1}(1-\eta\bar{A})r_t,\, U^{-1}e_{t+1}(\theta_t)\rangle)+\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2\right)\right].\end{aligned}$$

Note that $\mathbb{E}(e_{t+1}(\theta_t)|\mathcal{F}_t)=0$. The last term equals $\mathbb{E}\left[\left((1-2\eta\lambda)\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{\alpha}{2}}\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2\right]$.

By the existence of $(2+\alpha)$-order moment, there exists constant $M_1, M_2 > 0$ such that:

$$\begin{aligned}\mathbb{E}\left|2\eta\mathrm{Re}(\langle U^{-1}(1-\eta\bar{A})r_t,\, U^{-1}e_{t+1}(\theta_t)\rangle)+\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2\right|^{1+\alpha} &\le \eta^{1+\alpha}\left(M_1+M_2\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}\right)\\ \mathbb{E}\left[\left((1-2\eta\lambda)\left\|U^{-1}r_t\right\|_2^2\right)^{\frac{\alpha}{2}}\eta^2\left\|U^{-1}e_{t+1}\right\|_2^2\right] &\le \eta^2\left(M_1+M_2\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}\right).\end{aligned}$$

Thus we obtain:

$$\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^{2+\alpha} \le (1-2\eta\lambda)\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}+(\eta^{1+\alpha}+\eta^2)\left(M_1+M_2\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}\right).$$

For $\eta < \eta_0 = \frac{1}{2}(\lambda/M_2)^{\frac{1}{\alpha}}$, we have: $\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^{2+\alpha} \le (1-\eta\lambda)\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha}+\eta^{1+\alpha}M_1$. An induction proof argument leads to $\mathbb{E}\left\|U^{-1}r_t\right\|_2^{2+\alpha} \le \mathbb{E}\left\|U^{-1}r_0\right\|_2^{2+\alpha}+\frac{\eta^\alpha}{\lambda}M_1$ for any $t \ge 0$.

I.2. Proof of Lemma 11

In proving this lemma, we make use of Lemma 12; for $z_t := U^{-1} r_t$, there exists a pathwise coupling such that for any starting points $z_0^{(1)}, z_0^{(2)}$, we have $\mathbb{E}\left\|z_{t+1}^{(1)} - z_{t+1}^{(2)}\right\|_2^2 \leq e^{-\lambda\eta}\mathbb{E}\left\|z_t^{(1)} - z_t^{(2)}\right\|_2^2$. (Note that the proof of Lemma 12 does not use any results from this proof.)

We first show the existence and uniqueness of the stationary distribution, as well as the existence of the second moment. Then we calculate the first and second moment under the stationary distribution.

I.2.1. PROOF OF EXISTENCE

Since $\mathbb{R}^d$ is separable and complete, the Wasserstein space $\mathcal{W}^2$ is complete (Villani, 2008). Therefore, it suffices to show that $\{\mathcal{L}(\theta_t)\}_{t=0}^{+\infty}$ is a Cauchy sequence in this space.

Given $\mu \in \mathcal{W}^2$ and taking $\theta_0 \sim \mu$, take any positive integer $N > 0$, for any $k \geq N$ and $m \geq 0$, and we seek to upper bound $\mathcal{W}_2(\mathcal{L}(\theta_k), \mathcal{L}(\theta_{k+m}))$. Consider the process with two different initial points $\theta_0^{(1)} \sim \mu$ and $\theta_0^{(2)} \sim \mathcal{L}(\theta_m)$, coupled in an arbitrary way. By Lemma 12, we have:

$$\mathcal{W}_2\left(\mathcal{L}(\theta_k^{(1)}), \mathcal{L}(\theta_k^{(2)})\right) \leq e^{-\frac{\lambda\eta k}{2}}\kappa(U)\sqrt{\mathbb{E}\left\|\theta_0^{(1)} - \theta_0^{(2)}\right\|_2^2} \leq e^{-\frac{\lambda\eta N}{2}}\kappa(U)\sqrt{2\sup_{t\geq 0}\mathbb{E}\left\|\theta_t - \theta^*\right\|_2^2}.$$

Moreover, by Lemma 10, we have $\sup_{t\geq 0}\mathbb{E}\left\|\theta_t - \theta^*\right\|_2^2 \leq \kappa^2(U)\left(\mathbb{E}\left\|\theta_0 - \theta^*\right\|_2^2 + \frac{\eta}{\lambda}(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d)\right)$ is a finite constant independent of N. Therefore, $(\mathcal{L}(\theta_t))_{t\geq 0}$ is a Cauchy sequence in the space $\mathcal{W}^2$. The limit exists in $\mathcal{W}^2$.

I.2.2. PROOF OF UNIQUENESS

Suppose that there were two stationary measures $\pi^{(1)}$ and $\pi^{(2)}$, let $\theta_t^{(i)} \sim \pi^{(i)}$ for $i = 1, 2$, with an optimal coupling such that:

$$\mathbb{E}\left\|\theta_t^{(1)} - \theta_t^{(2)}\right\|_2^2 = \mathcal{W}_2^2(\pi^{(1)}, \pi^{(2)}).$$

By stationarity, we have $\theta_{t+1}^{(i)} \sim \pi^{(i)}$, and consequently:

$$\mathcal{W}_2^2(\pi^{(1)}, \pi^{(2)}) \leq \mathbb{E}\left\|\theta_{t+1}^{(1)} - \theta_{t+1}^{(2)}\right\|_2^2 \leq e^{-\eta\lambda}\mathbb{E}\left\|\theta_t^{(1)} - \theta_t^{(2)}\right\|_2^2 = e^{-\eta\lambda}\mathcal{W}_2^2(\pi^{(1)}, \pi^{(2)}),$$

which implies $\mathcal{W}_2(\pi^{(1)}, \pi^{(2)}) = 0$ and therefore $\pi^{(1)} = \pi^{(2)}$.

I.2.3. FIRST MOMENT UNDER THE STATIONARY DISTRIBUTION

Let $\theta_t \sim \pi_\eta$. Consider a stationary chain $(\theta_t)_{t\geq 0}$ starting at θ_0. By stationarity, we have $\mathcal{L}(\theta_{t+1}) = \mathcal{L}(\theta_t) = \pi_\eta$. Note that $\theta_{t+1} = \theta - \eta(A_{t+1}\theta_t - b_{t+1})$, taking expectations, we have:

$$\mathbb{E}(\theta_t) = \mathbb{E}(\theta_{t+1}) = \mathbb{E}\left(\theta_t - \eta(A_{t+1}\theta_t - b_t)\right) = \mathbb{E}\left(\theta_t - \eta\mathbb{E}(A_{t+1}\theta_t - b_{t+1}|\mathcal{F}_t)\right) = \mathbb{E}\left(\theta_t - \eta(A\theta_t - b)\right).$$

Therefore, we have $\bar{A}\mathbb{E}_{\pi_\eta}(\theta) - b = 0$, which implies $\theta = \theta^*$ since $\bar{A}$ is non-degenerate.

I.2.4. SECOND MOMENT UNDER THE STATIONARY DISTRIBUTION

Let $\theta_t \sim \pi_\eta$. Consider a stationary chain $(\theta_t)_{t\geq 0}$ starting at θ_0. By stationarity, we have $\mathcal{L}(\theta_{t+1}) = \mathcal{L}(\theta_t) = \pi_\eta$. Note that $\theta_{t+1} = \theta - \eta(A_{t+1}\theta_t - b_{t+1})$, and consequently, we have:

$$(\theta_{t+1} - \theta^*) = (I - \eta\bar{A})(\theta_t - \theta^*) - \eta\Xi_{t+1}(\theta_t - \theta^*) + \eta\xi_{t+1} - \eta\Xi_{t+1}\theta^*.$$

As we have shown, $\mathbb{E}_{\pi_\eta}\theta = \theta^*$. Let $r_t := \theta_t - \theta^*$, taking conditional second moments of both sides of the equation, we obtain:

$$\begin{aligned}\mathbb{E}\left(r_{t+1}r_{t+1}^\top \mid \mathcal{F}_t\right) &= (I_d - \eta\bar{A})r_t r_t^\top (I_d - \eta\bar{A})^\top + \eta^2\mathbb{E}(\Xi_{t+1}r_t r_t^\top \Xi_{t+1}^\top|\mathcal{F}_t)\\ &\quad + \eta^2\mathbb{E}\left(\Xi_{t+1}r_t(\xi_{t+1} + \Xi_{t+1}\theta^*)^\top + (\xi_{t+1} + \Xi_{t+1}\theta^*)r_t^\top\Xi_{t+1}^\top \mid \mathcal{F}_t\right)\\ &\quad + \eta^2\mathbb{E}((\xi_{t+1} + \Xi_{t+1}\theta^*)(\xi_{t+1} + \Xi_{t+1}\theta^*)^\top \mid \mathcal{F}_t).\end{aligned}$$

Let $\Lambda := \mathbb{E}_{\pi_\eta}\left(r_t r_t^\top\right)$. Taking the expectation of both sides, note that by Assumption 1:

$$\mathbb{E}\left(\Xi_{t+1}r_t\xi_{t+1}^\top \mid \mathcal{F}_t\right) = 0, \quad \mathbb{E}((\xi_{t+1} + \Xi_{t+1}\theta^*)(\xi_{t+1} + \Xi_{t+1}\theta^*)^\top \mid \mathcal{F}_t) = \Sigma_\xi + \mathbb{E}(\Xi_A\theta^*\theta^{*\top}\Xi_A^\top),$$
$$\mathbb{E}\left(\Xi_{t+1}r_t(\Xi_{t+1}\theta^*)^\top\right) = \mathbb{E}\left(\Xi_A \otimes \Xi_A\right) \cdot \mathrm{vec}(\mathbb{E}(r_t)\theta^{*\top}) = \mathbb{E}\left(\Xi_A \otimes \Xi_A\right) \cdot \mathrm{vec}(0 \cdot \theta^{*\top}) = 0.$$

Simplifying this equation yields

$$\Lambda = (I_d - \eta\bar{A})\Lambda(I_d - \eta\bar{A})^\top + \eta^2\mathbb{E}(\Xi_A\Lambda\Xi_A^\top) + \eta^2\Sigma_\xi + \eta^2\mathbb{E}(\Xi_A\theta^*\theta^{*\top}\Xi_A^\top),$$

which means:

$$\bar{A}\Lambda + \Lambda\bar{A}^\top = \eta\bar{A}\Lambda\bar{A}^\top + \eta\mathbb{E}(\Xi_A\Lambda\Xi_A^\top) + \eta\Sigma^*.$$

By flattening the tensors, we can write the equation in a matrix-vector form:

$$\left(I_d \otimes \bar{A} + \bar{A}^\top \otimes I_d - \eta\bar{A} \otimes \bar{A} - \eta\mathbb{E}(\Xi_A \otimes \Xi_A)\right)\mathrm{vec}(\Lambda) = \eta\mathrm{vec}(\Sigma^*),$$

where $\oplus$ denotes the Kronecker sum and $\otimes$ denotes the Kronecker product.

To provide an upper bound on the trace of the solution to this matrix equation, which is the covariance under the stationary distribution, we note that in the proof of Lemma 10, we use a contraction inequality:

$$\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^2 \leq (1 - \lambda\eta)\mathbb{E}\left\|U^{-1}r_t\right\|_2^2 + \eta^2\|U^{-1}\|_{\mathrm{op}}^2(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d).$$

If $\theta_t \sim \pi_\eta$, we have $\theta_{t+1} \sim \pi_\eta$, and hence

$$\mathbb{E}_{\pi_\eta}\left\|U^{-1}(\theta - \theta^*)\right\|_2^2 \leq (1 - \lambda\eta)\mathbb{E}_{\pi_\eta}\left\|U^{-1}(\theta - \theta^*)\right\|_2^2 + \eta^2\|U^{-1}\|_{\mathrm{op}}^2(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d),$$

which implies the claimed bound:

$$\mathbb{E}_{\pi_\eta}\left\|\theta - \theta^*\right\|_2^2 \leq \frac{\eta}{\lambda}\kappa^2(U)(v_A^2\left\|\theta^*\right\|_2^2 + v_b^2 d).$$

I.3. Proof of Lemma 12

Given two different starting points $x^{(i)} \in \mathbb{R}^d$ for $i = 1, 2$, let $\{\theta_t^{(i)}\}_{t\geq 0}$ be the process starting at $x^{(i)}$, and let the two processes to be driven by the same sequences of noise variables ξ_b and Ξ_A, so that $A_t^{(1)} = A_t^{(2)}$ and $b_t^{(1)} = b_t^{(2)}$ almost surely.

By Lemma 1, we can write $\bar{A} = UD^\top U^{-1}$, such that $D + D^{\mathsf{H}} \succeq \lambda I_d$. Introducing the shorthand $r_t := \theta_t^{(1)} - \theta_t^{(2)}$, some algebra leads to the recursive relation

$$r_{t+1} = \theta_{t+1}^{(1)} - \theta_{t+1}^{(2)} = \theta_t^{(1)} - \eta\left(\bar{A}\theta_t^{(1)} - b + \Xi_{t+1}\theta_t^{(1)} - \xi_{t+1}\right) - \theta_t^{(2)} + \eta\left(\bar{A}\theta_t^{(2)} - b + \Xi_{t+1}\theta_t^{(2)} - \xi_{t+1}\right)$$
$$= (I_d - \eta\bar{A} - \eta\Xi_{t+1})r_t.$$

Define the Lyapunov function $f(r) = \mathbb{E}\left\|U^{-1}r\right\|_2^2$. By Assumptions 2 and 3, note that $\rho(\bar{A}) = \sqrt{\|D^{\mathsf{H}}D\|_{\text{op}}}$ and $\kappa(U) = \|U\|_{\text{op}}\|U^{-1}\|_{\text{op}}$, we have:

$$\begin{aligned}
&\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^2 \\
&= \mathbb{E}\left(r_t^{\mathsf{H}}(I_d - \eta\bar{A} - \eta\Xi_{t+1})^\top (U^{-1})^{\mathsf{H}}U^{-1}(I_d - \eta\bar{A} - \eta\Xi_t)r_t\right) \\
&= \mathbb{E}\left((U^{-1}r_t)^{\mathsf{H}}(I_d - \eta D - \eta U^{-1}\Xi_{t+1}U)^{\mathsf{H}}(I_d - \eta D - \eta U^{-1}\Xi_{t+1}U)(U^{-1}r_t)\right) \\
&= \mathbb{E}\left\|(I_d - \eta D)U^{-1}r_t\right\|_2^2 + \eta^2\mathbb{E}\left\|U\Xi_{t+1}r_t\right\|_2^2 \\
&\leq \mathbb{E}\left\|U^{-1}r_t\right\|_2^2 - \eta\mathbb{E}(U^{-1}r_t)^{\mathsf{H}}(D + D^{\mathsf{H}})(U^{-1}r_t) + \eta^2\|D^{\mathsf{H}}D\|_{\text{op}}\mathbb{E}\left\|U^{-1}r_t\right\|_2^2 + \eta^2\|U\|_{\text{op}}^2\mathbb{E}\left\|\Xi_{t+1}r_t\right\|_2^2 \\
&\leq \mathbb{E}\left\|U^{-1}r_t\right\|_2^2 - 2\eta\lambda\mathbb{E}\left\|U^{-1}r_t\right\|_2^2 + \eta^2\rho(\bar{A})^2\mathbb{E}\left\|U^{-1}r_t\right\|_2^2 + \kappa^2(U)v_A^2\mathbb{E}\left\|U^{-1}r_t\right\|_2^2.
\end{aligned}$$

For $\eta \in \left(0, \frac{\lambda}{\rho(\bar{A})^2 + \kappa(U)^2 v_A^2}\right)$, we have $\mathbb{E}\left\|U^{-1}r_{t+1}\right\|_2^2 \leq (1 - \eta\lambda)\mathbb{E}\left\|U^{-1}r_t\right\|_2^2$ for any $t \geq 0$. Consequently, we have the coupling estimate:

$$\mathbb{E}\left\|r_T\right\|_2^2 \leq \|U\|_{\text{op}}^2\left\|U^{-1}r_T\right\|_2^2 \leq \|U\|_{\text{op}}^2 e^{-\eta\lambda T}\left\|U^{-1}r_0\right\|_2^2 \leq e^{-\eta\lambda T}\kappa^2(U)\mathbb{E}\left\|r_0\right\|_2^2,$$

which completes the proof of the lemma.

I.4. Proof of Lemma 13

We first prove the almost-sure upper bounds on the iterates. Note that for $\theta_t \in [-\bar{\lambda}^{-1}, \bar{\lambda}^{-1}]^d$, we have the following sequence of inequalities almost surely:

$$\|\theta_{t+1}\|_\infty = \|\theta_t - \eta(A_{t+1}\theta_t - b_{t+1})\|_\infty \leq \|(1-\eta)\theta_t\|_\infty + \eta\|(I_d - A_{t+1})\theta_t\|_\infty + \eta\|b_{t+1}\|_\infty$$
$$\leq (1-\eta)\|\theta_t\|_\infty + \eta(1-\bar{\lambda})\|\theta_t\|_\infty + \eta \leq (1-\eta\bar{\lambda})\bar{\lambda}^{-1} + \eta = \bar{\lambda}^{-1}.$$

The result then follows by induction.

We then prove the ℓ_∞ contraction bound. We take a synchronous coupling where the two processes use the same sequence of stochastic oracles. We have:

$$\left\|\theta_{t+1}^{(1)} - \theta_{t+1}^{(2)}\right\|_\infty = \left\|(I - \eta A_{t+1})(\theta_t^{(1)} - \theta_t^{(2)})\right\|_\infty$$
$$\leq (1-\eta)\left\|\theta_t^{(1)} - \theta_t^{(2)}\right\|_\infty + \eta\left\|(I - A)(\theta_t^{(1)} - \theta_t^{(2)})\right\|_\infty \leq (1-\eta\bar{\lambda})\left\|\theta_t^{(1)} - \theta_t^{(2)}\right\|_\infty,$$

which proves the coupling bound.

Appendix J. Proof of Concentration Inequalities

In this section, we present the concentration inequalities used in the proof of our main theorems. We first state and prove a concentration inequality for time averages of functions of a Markov chain, following the general results from Joulin and Ollivier (2010). Then, we state and prove a concentration inequality for heavy-tailed martingales.

J.1. Concentration inequalities involving metric ergodocity

In this section, we prove Lemma 16, the metric ergodic concentration inequality for the LSA process, which plays an important role in our analysis. To prove it, we need the following general result, which asserts the concentration inequalities under uniform upper bounds on the tail of the iterates and stochastic oracles.

Lemma 18 *Under Assumption 1, Assumption 2 and Assumption 3, for given $T > 0$, if for any $\delta > 0$, there exists $R(\delta), r(\delta) > 0$ such that:*

- $\mathbb{P}\left(\max_{0\le t\le T} \left\|U^{-1}\theta_t\right\|_2 > R(\delta)\right) < \delta$.
- $\mathbb{P}\left(\max_{0\le t\le T} \left\|U^{-1}(\Xi_{t+1}\theta_t - \xi_{t+1})\right\|_2 > r(\delta)\right) < \delta$,

then, for any matrix $L \in \mathbb{R}^{d\times d}$ and any $\delta \in \left(0, (T^2 \|L\|_{op}^2 \max_{t\le T} \mathbb{E}\,\|\theta_t\|_2^4)^{-1}\right)$, we have:

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T}(\theta_t^\top L\theta_t - \mathbb{E}\theta_t^\top L\theta_t)\right| > C\|L\|_{op}\|U\|_{op}^2 \frac{R(\delta)r(\delta)}{\lambda}\left(\sqrt{\frac{\log\delta^{-1}}{T}} + \frac{\log\delta^{-1}}{T}\right)\right) \le 3\delta.$$

Lemma 16 is actually an instantiation of Lemma 18, which provides concrete upper bounds on the quantities $R(\delta)$ and $r(\delta)$ based on the tail assumption $3'$. In the following, we first prove Lemma 18, and then prove Lemma 16 by verifying the conditions in the general lemma.

J.1.1. PROOF OF LEMMA 18

In order to prove this lemma, we make use of the following known result due to Joulin and Ollivier Joulin and Ollivier (2010):

Proposition 19 (Theorem 4 Joulin and Ollivier (2010), special case) *Let $(X_t)_{t\ge1}$ be a discrete-time Markov chain with transition kernel P, defined on a space $\mathcal{X}$ equipped with the metric $d(\cdot,\cdot)$. Assume that $\forall x, y \in \mathcal{X}$, $\mathcal{W}_{1,d}(P_x, P_y) \le (1-\kappa)d(x,y)$ for some $\kappa > 0$. Assume furthermore that $\sigma_\infty := \sup_{x\in\mathcal{X}} \operatorname{diam}(\operatorname{supp}(P_x))$. For any function f that is 1-Lipschitz on $\mathcal{X}$ with respect to $d(\cdot,\cdot)$, given a trajectory $(X_t)_{1\le t\le T}$ of the Markov chain, we have:*

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T}(f(X_t) - \mathbb{E}f(X_t))\right| > r\right) \le \begin{cases} 2\exp\left(-\frac{r^2T}{32}\cdot\frac{\kappa^2}{\sigma_\infty^2}\right) & r < \frac{4\sigma_\infty}{3\kappa} \\ 2\exp\left(-\frac{r\kappa T}{12\sigma_\infty}\right) & r \ge \frac{4\sigma_\infty}{3\kappa} \end{cases}.$$

Proposition 19 requires bounded noise and global Lipschitzness, neither of which is satisfied by the process θ_t with a quadratic function f. In order to circumvent this limitation, we use a standard truncation argument.

Under the assumptions of Lemma 18, for any $\delta > 0$, define a stopping time

$$\tau(\delta) := \inf\left\{t \geq 1 : \left\|U^{-1}\theta_t\right\|_2 > R(\delta) \text{ or } \left\|U^{-1}(\Xi_t\theta_t - \xi_t)\right\|_2 > r(\delta)\right\}.$$

Let $A = UDU^{-1}$ be its eigendecomposition. By the proof of Lemma 12, when $\eta < \frac{\lambda}{2(\rho^2(\bar{A})+\kappa^2(U)v_A^2)}$, the Markov process $(U^{-1}\theta_t)_{t\geq 0}$ satisfies:

$$\mathcal{W}_1(P_x, P_y) \leq \mathcal{W}_2(P_x, P_y) \leq (1 - \eta\lambda/2)\left\|x - y\right\|_2, \quad \forall x, y \in \mathbb{R}^d.$$

We define a killed Markov process $\vartheta_t := U^{-1}\theta_t$ for $t < \tau(\delta)$, which gets killed at time $\tau(\delta)$. The one-step transition of the process ϑ_t is defined as $\vartheta_t \mapsto \vartheta_t - \eta U^{-1}(AU\vartheta_t - b) - U^{-1}(\Xi_t U\vartheta_t - \xi_t)$, whose support has a diameter bounded by $2\eta r(\delta)$ before being killed. Note that the Wasserstein contraction property remains true for the killed process. The assumptions in Lemma 18 guarantee that $\mathbb{P}(\tau(\delta) \leq T) < 2\delta$. By definition, we have $\|\vartheta_t\|_2 \leq R(\delta)$. Finally, for the function $f : \mathbb{B}(0, R(\delta)) \to \mathbb{R}$ with $f(\vartheta) := \vartheta^\top U^\top LU\vartheta$, we have:

$$\left\|\nabla f(\vartheta)\right\|_2 \leq 2\|L\|_{\mathrm{op}}\|U\|_{\mathrm{op}}^2 \left\|\vartheta\right\|_2 \leq 2\|U\|_{\mathrm{op}}^2\|L\|_{\mathrm{op}}R(\delta).$$

Applying Proposition 19, for any $\varepsilon > 0$, we obtain:

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T}(\vartheta_t^\top U^\top LU\vartheta_t \mathbf{1}_{t<\tau(\delta)} - \mathbb{E}\vartheta_t^\top U^\top LU\vartheta_t)\right| > 2\varepsilon\|L\|_{\mathrm{op}} \cdot \|U\|_{\mathrm{op}}^2 R(\delta)\right)$$
$$\leq \begin{cases} 2\exp\left(-\frac{\varepsilon^2 T}{32} \cdot \frac{(\lambda)^2}{16(r(\delta))^2}\right), & \varepsilon < \frac{16r(\delta)}{3\lambda} \\ 2\exp\left(-\frac{\varepsilon\lambda T}{48r(\delta)}\right), & \varepsilon \geq \frac{16r(\delta)}{3\lambda}. \end{cases}$$

On the event $\{T < \tau(\delta)\}$, we have $\vartheta_t = U^{-1}\theta_t$ for $t = 1, 2, \cdots, T$. It remains to bound the difference between $\mathbb{E}\vartheta_t^\top U^\top LU\vartheta_t$ and $\mathbb{E}\theta_t^\top L\theta_t$. Note that:

$$\begin{aligned}|\mathbb{E}\vartheta_t^\top U^\top LU\vartheta_t - \mathbb{E}\theta_t^\top L\theta_t| &= |\mathbb{E}(\theta_t^\top L\theta_t\mathbf{1}_{t<\tau}) - \mathbb{E}\theta_t^\top L\theta_t| \leq \|L\|_{\mathrm{op}}\mathbb{E}(\|\theta_t\|_2^2\,\mathbf{1}_{\tau<t}) \\ &\leq \|L\|_{\mathrm{op}}\sqrt{\mathbb{E}(\|\theta_t\|_2^4)\mathbb{E}(\mathbf{1}_{\tau<t}^2)} \leq \|L\|_{\mathrm{op}}\sqrt{\delta\mathbb{E}\left\|\theta_t\right\|_2^4}.\end{aligned}$$

Putting together the pieces yields the claimed result.

J.1.2. PROOF OF LEMMA 16

The proof involves verifying the assumptions in Lemma 18. For the high-probability bound on $\max_{0\leq t\leq T}\left\|U^{-1}\theta_t\right\|_2$, we note that by the proof of Lemma 14, for $p \geq 2\log T$ we have:

$$\begin{aligned}\mathbb{E}\max_{0\leq t\leq T}\left\|U^{-1}\theta_t\right\|_2^p &\leq \sum_{t=1}^{T}\mathbb{E}\left\|U^{-1}\theta_t\right\|_2^p \\ &\leq T\|U^{-1}\|_{\mathrm{op}}^p\left(\|\theta_0 - \theta^*\|_2 + \frac{\eta}{\lambda}(\sigma_b\sqrt{d}p^{\beta+1/2} + \sigma_A\left\|\theta^*\right\|_2 p^{\alpha+1/2})\right)^p.\end{aligned}$$

Taking $p = C\log\frac{T}{\delta}$ for a universal constant $C > 0$ and applying Markov inequality, we have:

$$\mathbb{P}\left(\max_{0\leq t\leq T}\left\|U^{-1}\theta_t\right\|_2 > B\right) < \delta.$$

In order to verify the second condition, we note that by Assumption 3′, conditionally on $\mathcal{F}_t$, the Markov inequality yields:

$$\mathbb{P}\left(\|\Xi_{t+1}\theta_t\|_2 > \sigma_A \|\theta_t\|_2 \log^\alpha \delta^{-1} | \mathcal{F}_t\right) < \delta, \quad \mathbb{P}\left(\|\xi_{t+1}\|_2 > \sigma_b \sqrt{d} \log^\beta \delta^{-1} | \mathcal{F}_t\right) < \delta.$$

Combined with high probability bounds on θ_t and take union bound over $t \in \{1, 2, \cdots, T\}$, we obtain the final result.

J.2. A concentration inequality for heavy-tailed martingales

In this appendix, we state and prove a useful concentration inequality for heavy-tailed martingales.

Lemma 20 *For a (scalar) martingale difference sequence $(X_t : t \geq 1)$ adapted to filtration $(\mathcal{F}_t)_{t\geq 0}$, if we have $\forall p \geq 2$, $\mathbb{E}(|X_t|^p | \mathcal{F}_{t-1})^{\frac{1}{p}} \leq p^\gamma \sigma$ almost surely for some $\gamma, \sigma > 0$, for any $\delta > 0$, we have*

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^T X_t\right| > C_\gamma \sigma \left(\sqrt{\frac{\log \delta^{-1}}{T}} + \frac{\log^{1+\gamma} T/\delta}{T}\right)\right) < \delta.$$

Proof For a constant $M > 0$ which will be determined later, define $\tilde{X}_t := X_t \mathbf{1}_{|X_t| \leq M}$ be the truncated version of the process. By the Bernstein inequality for martingales (Freedman, 1975), for any $K > 0$, we have:

$$\forall \varepsilon > 0, \ \mathbb{P}\left(\left|\sum_{t=1}^T \tilde{X}_t - \mathbb{E}(\tilde{X}_t \mid \mathcal{F}_{t-1})\right| > \varepsilon, \ \sum_{t=1}^T \mathrm{var}\left(\tilde{X}_t | \mathcal{F}_{t-1}\right) < K\right) \leq 2 \exp\left(-\frac{\varepsilon^2}{2K + 2M\varepsilon/3}\right).$$

On the other hand, note that for $z > (2e)^\gamma \sigma$, we have

$$\mathbb{P}\left(|X_t| > z\right) \leq \inf_{p \geq 2} \frac{p^{p\gamma}\sigma^p}{z^p} = \exp\left(-\frac{\gamma}{e}\left(\frac{z}{\sigma}\right)^{\frac{1}{\gamma}}\right).$$

Consequently, we have $\mathbb{P}\left(X_t \neq \tilde{X}_t | \mathcal{F}_{t-1}\right) \leq \exp\left(-\frac{\gamma}{e}\left(\frac{z}{\sigma}\right)^{\frac{1}{\gamma}}\right)$.

Furthermore, we note that

$$\left|\mathbb{E}\left(\tilde{X}_t | \mathcal{F}_{t-1}\right)\right| \leq \mathbb{E}\left(|X_t - \tilde{X}_t| \big| \mathcal{F}_{t-1}\right) \leq 2 \int_M^{+\infty} \exp\left(-\frac{\gamma}{e}\left(\frac{z}{\sigma}\right)^{\frac{1}{\gamma}}\right) dz \leq C_\gamma \left(\frac{M}{\sigma}\right)^{1-\frac{1}{\gamma}} \exp\left(-\frac{\gamma}{e}\left(\frac{M}{\sigma}\right)^{\frac{1}{\gamma}}\right).$$

For the conditional second moment, we have:

$$\mathrm{var}(\tilde{X}_t | \mathcal{F}_{t-1}) \leq \mathbb{E}(\tilde{X}_t^2 | \mathcal{F}_{t-1}) \leq \mathbb{E}(X_t^2 | \mathcal{F}_{t-1}) \leq 2^{2\gamma}\sigma^2, \quad \text{a.s.}$$

Choosing $K = 2^{2\gamma}\sigma^2 T$, we have:

$$\forall \varepsilon > 0, \ \mathbb{P}\left(\left|\sum_{t=1}^T \tilde{X}_t - \mathbb{E}(\tilde{X}_t | \mathcal{F}_{t-1})\right| > \varepsilon\right) \leq 2 \exp\left(-\frac{\varepsilon^2}{C_\gamma \sigma^2 T + 2M\varepsilon/3}\right)$$

Putting together the pieces, we find that

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^T X_t\right| > C_\gamma \sigma \sqrt{\frac{\log \delta^{-1}}{T}} + \frac{M \log \delta^{-1}}{T} + C_\gamma \left(\frac{M}{\sigma}\right)^{1-\frac{1}{\gamma}} e^{-\frac{\gamma}{e}(\frac{M}{\sigma})^{1/\gamma}}\right) \leq \delta + T \exp\left(-\frac{\gamma}{e}\left(\frac{M}{\sigma}\right)^{\frac{1}{\gamma}}\right).$$

Setting $M = C_\gamma \sigma \log^\gamma(\frac{T}{\delta})$ yields the claim. ■

Appendix K. Proof of Deterministic Properties of Matrices

In this section, we prove some auxiliary deterministic facts about square matrices. We first prove Lemma 1, which guarantees the existence of a good similarity transformation for Huriwtz matrices. Then, we state and prove Proposition 21, which asserts that such nice property does not hold in the critical case without the diagonalizability condition. In particular, the Polyak-Ruppert procedure fails for certain non-diagonalizable matrices with pure imaginary eigenvalues. Finally, we compute the eigen-values for the asymmetric stochastic approximation matrix used in momentum SGD, as discussed in Section C.1.

K.1. Proof of Lemma 1

In this appendix, we prove Lemma 1. This lemma is a standard fact in linear algebra; for instance, see Section 1.8 in Perko (2013). We include the proof for completeness and so as to extract the behavior of λ.

When the matrix $\bar{A}$ is diagonalizable, we can write $\bar{A} = UDU^{-1}$, which implies the stronger lower bound $D + D^{\mathsf{H}} \succeq 2 \min_{i \in [d]} \mathrm{Re}(\lambda_i(\bar{A}))$. For a non-diagonalizable matrix $\bar{A}$, we instead write $\bar{A} = UJU^{-1}$, where the matrix $J = \mathrm{diag}(\lambda_i I_{d_i} + J_{d_i})_{i=1}^k$ contains the Jordan decomposition. For each Jordan block, we note that for $Q_i := \mathrm{diag}(1, \mathrm{Re}(\lambda_i/2), \cdots, \mathrm{Re}(\lambda_i/2)^{d_i - 1})$, we have

$$Q_i^{-1}(\lambda_i I_{d_i} + J_{d_i}) Q_i = \lambda_i I_{d_i} + \mathrm{Re}(\lambda_i/2) J_{d_i} := B_i.$$

We note that A is similar to $\mathrm{diag}(B_1, B_2, \cdots, B_k)$. We only need to study the eigenvalues of $B_i + B_i^{\mathsf{H}}$. A straightforward calculation yields:

$$B_i + B_i^{\mathsf{H}} = \frac{1}{2}\mathrm{Re}(\lambda_i) \begin{bmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \cdots & 0 \\ & & \cdots & & \\ 0 & \cdots & 1 & 4 & 1 \\ 0 & \cdots & 0 & 1 & 4 \end{bmatrix} := \mathrm{Re}(\lambda_i) T_{d_i}.$$

Note that the matrix T_{d_i} is a symmetric tridiagonal Toeplitz matrix, whose eigenvalues are given by the formula $\lambda_j(T_{d_i}) = 4 + 2\cos\left(\frac{j\pi}{(d_i+1)}\right) \geq 2$. Therefore, we have $B_i + B_i^{\mathsf{H}} \succeq \mathrm{Re}(\lambda_i)$, which completes the proof.

K.2. Necessity of diagonalizable $\bar{A}$ in the critical case

In this appendix, we demonstrate that the diagonalizability condition in Assumption 2′ cannot be removed. More precisely, we show that even in the case of deterministic observations (i.e., $A_t = \bar{A}$ and $b_t = b$ for all iterations t), there is a choice of matrix $\bar{A}$ and initial vector θ_0 for which the Polyak-Ruppert iterates behave badly.

Proposition 21 *For any dimension $d \geq 2$ and given initial vector $\theta_0 = [0, 0, \cdots, 0, 1]^\top$, there exists a matrix $\bar{A} \in \mathbb{C}^{d \times d}$ with $\min_{i \in [d]} \mathrm{Re}(\lambda_i(\bar{A})) \geq 0$ and $\min_i |\lambda_i(\bar{A})| \geq 1$ such that for any positive step size η and any iteration $T \geq 4$, the Polyak-Ruppert averaged iterate satisfies the lower bound*

$$\left\|\bar{\theta}_T - \theta^*\right\|_2 \geq \frac{1}{2}. \tag{37}$$

The proof is based on an explicit construction. Consider the d-dimensional matrix

$$J_d := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Define the matrix $\bar{A} = -iI_d - J_d$. In this deterministic setting, we have:

$$\theta_T - \theta^* = (I_d - \eta\bar{A})^T(\theta_0 - \theta^*) = ((1+\eta i)I_d + \eta J_d)^T(\theta_0 - \theta^*) = \sum_{\ell=0}^{\min(d,T)} \eta^\ell (1+\eta i)^{T-\ell} \binom{T}{\ell} J_d^\ell (\theta_0 - \theta^*).$$

Take $\theta^* = 0$. Given our initialization $\theta_0 = [0, 0, \cdots, 0, 1]^\top$. for all $T \geq d-1$, we have $\theta_T = \sum_{\ell=0}^{d-1} \eta^\ell (1+\eta i)^{T-\ell} \binom{T}{\ell} e_{d-\ell}$, and consequently, we have:

$$-(\bar{\theta}_T - \theta^*) = \frac{1}{T}\sum_{t=1}^{T}\sum_{\ell=0}^{d-1} \eta^\ell \binom{t}{\ell} e_{d-\ell} = \sum_{\ell=0}^{d-2} e_{d-\ell}\eta^\ell \frac{1}{T}\sum_{t=\ell}^{T}(1+\eta i)^{t-\ell}\binom{t}{\ell}.$$

Consider the coefficient in the $(d-1)$-th coordinate, which corresponds to the case with $\ell = 1$, we have:

$$-e_{d-1}^{\mathsf{H}}(\bar{\theta}_T - \theta^*) = \frac{\eta}{T}\sum_{t=1}^{T}(1+\eta i)^{t-1} t = \left(-i + \frac{1}{T}\right)(1+\eta i)^T + \frac{i-1}{T}$$

Therefore, for $T \geq 4$, we have:

$$\left\|\bar{\theta}_T - \theta^*\right\|_2 \geq |e_{d-1}^{\mathsf{H}}(\bar{\theta}_T - \theta^*)_{d-1})| \geq \left|\left(i + \frac{1}{T}\right)(1+\eta i)^T\right| - \frac{\sqrt{2}}{T} \geq (1+\eta^2)^{\frac{T}{2}} - \frac{\sqrt{2}}{T} \geq \frac{1}{2},$$

which completes the proof.

K.3. Eigenvalue computation for momentum SGD

Since $\bar{A}$ is real symmetric and positive definite, it is guaranteed to have a spectral decomposition of the form $\bar{A} = UDU^{-1}$, where U is a orthonormal matrix and $D = \text{diag}\{\lambda_i(\bar{A})\}_{i=1}^d$. Using this fact, we can write

$$\begin{aligned} \tilde{A} &= \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} 0 & I_d \\ -D & \alpha I_d + \eta D \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}^{-1} \\ &= \left(\begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} P_0\right) \text{diag}\left(\begin{bmatrix} 0 & 1 \\ -\lambda_i & \alpha + \eta\lambda_i \end{bmatrix}\right)_{i=1}^d \left(\begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} P_0\right)^{-1}, \end{aligned}$$

where P_0 is a permutation matrix which turns the order $(1, 2, \cdots, 2d)$ into $(1, d{+}1, 2, d{+}2, \cdots, d, 2d)$. It can be seen that P_0 is orthonormal.

For $\alpha \in \mathbb{R}_+ \setminus \{2\sqrt{\lambda_i} - \eta\lambda_i\}_{i=1}^d$, each 2×2 block has distinct eigenvalues, which makes it diagonalizable. In particular, we have:

$$\begin{bmatrix} 0 & 1 \\ -\lambda_i & \alpha + \eta\lambda_i \end{bmatrix} = \begin{bmatrix} \lambda_i & -\nu_i^+ \\ \lambda_i & -\nu_i^- \end{bmatrix} \cdot \begin{bmatrix} \nu_i^+ & 0 \\ 0 & \nu_i^- \end{bmatrix} \cdot \begin{bmatrix} \lambda_i & -\nu_i^+ \\ \lambda_i & -\nu_i^- \end{bmatrix}^{-1},$$

where $\nu_i^\pm = \frac{(\alpha+\eta\lambda_i)\pm\sqrt{(\alpha+\eta\lambda_i)^2-4\lambda_i}}{2}$.