



**HAL**  
open science

# On double descent in reinforcement learning with LSTD and random features

David Brellmann, Eloïse Berthier, David Filliat, Goran Frehse

► **To cite this version:**

David Brellmann, Eloïse Berthier, David Filliat, Goran Frehse. On double descent in reinforcement learning with LSTD and random features. ICLR 2024 - The Twelfth International Conference on Learning Representations, May 2024, Vienna, Austria. hal-04628067

**HAL Id: hal-04628067**

**<https://hal.ip-paris.fr/hal-04628067>**

Submitted on 28 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ON DOUBLE DESCENT IN REINFORCEMENT LEARNING WITH LSTD AND RANDOM FEATURES

David Brellmann, Eloïse Berthier, David Filliat & Goran Frehse  
 U2IS, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, FRANCE  
 {first\_name.last\_name}@ensta-paris.fr

## ABSTRACT

Temporal Difference (TD) algorithms are widely used in Deep Reinforcement Learning (RL). Their performance is heavily influenced by the size of the neural network. While in supervised learning, the regime of over-parameterization and its benefits are well understood, the situation in RL is much less clear. In this paper, we present a theoretical analysis of the influence of network size and  $l_2$ -regularization on performance. We identify the ratio between the number of parameters and the number of visited states as a crucial factor and define over-parameterization as the regime when it is larger than one. Furthermore, we observe a double descent phenomenon, i.e., a sudden drop in performance around the parameter/state ratio of one. Leveraging random features and the lazy training regime, we study the regularized Least-Squared Temporal Difference (LSTD) algorithm in an asymptotic regime, as both the number of parameters and states go to infinity, maintaining a constant ratio. We derive deterministic limits of both the empirical and the true Mean-Squared Bellman Error (MSBE) that feature correction terms responsible for the double descent. Correction terms vanish when the  $l_2$ -regularization is increased or the number of unvisited states goes to zero. Numerical experiments with synthetic and small real-world environments closely match the theoretical predictions.

## 1 INTRODUCTION

In recent years, neural networks have seen increased use in Reinforcement Learning (RL) (Mnih et al., 2015; Schulman et al., 2017; Haarnoja et al., 2018). While they can outperform traditional RL algorithms on challenging tasks, their theoretical understanding remains limited. Even for supervised learning, which can be considered a special case of RL with discount factor equal to zero, deep neural networks are still far from being fully understood despite significant research efforts (Arora et al., 2019; Mei et al., 2018; Rotskoff & Vanden-Eijnden, 2018; Lee et al., 2019; Bietti & Mairal, 2019; Cao et al., 2019). The difficulty is further exacerbated in RL by a myriad of new challenges that limit the scope of these works, such as the absence of true targets or the non-i.i.d nature of the collected samples (Kumar et al., 2020; Luo et al., 2020; Lyle et al., 2021; Dong et al., 2020). Temporal-Difference (TD) methods are widely used RL algorithms that frequently use neural networks, are simple, and efficient in practice. We use the regularized Least-squares Temporal Difference (LSTD) algorithm (Bradtke & Barto, 1996), which is easier to analyze since it doesn't use gradient descent, and because it converges to the same solution as other TD algorithms (Bradtke & Barto, 1996; Boyan, 1999; Berthier et al., 2022).

Theoretical studies of TD algorithms often explore asymptotic regimes where the number of samples  $n \rightarrow \infty$  while the number of model parameters  $N$  remains constant (Tsitsiklis & Van Roy, 1996; Sutton, 1988). When TD learning algorithms are applied to neural networks, it is commonly assumed that the number of parameters  $N \rightarrow \infty$  with either a fixed or infinite number of samples without providing details on the relative magnitudes of these parameters (Cai et al., 2019; Agazzi & Lu, 2022; Berthier et al., 2022; Xiao et al., 2021). Inspired by advancements in supervised learning (Louart et al., 2018; Liao et al., 2020), we apply Random Matrix tools and propose a novel double asymptotic regime where the number of parameters  $N$  and the number of distinct visited states  $m$  go to infinity, maintaining a constant ratio, called model complexity. We use a linear model and nonlinear random features (RF) (Rahimi & Recht, 2007) to approximate an overparameterized

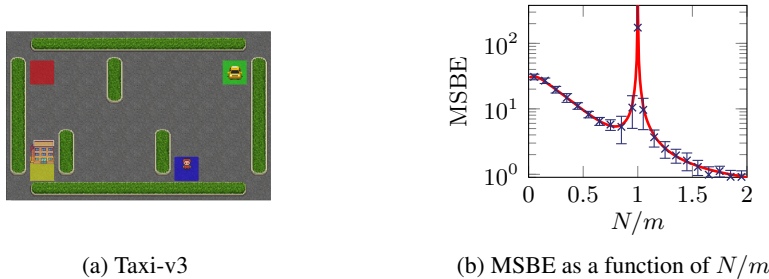


Figure 1: **As the model complexity  $N/m$  (for  $N$  parameters,  $m$  distinct visited states) increases, the MSBE first shows a U-shaped curve, peaking around the interpolation threshold ( $N = m$ ). Double descent refers to the phenomenon for  $N/m > 1$  where the MSBE drops once again.** Continuous lines (red) indicate the theoretical values from Theorem 5.3, the crosses (blue) are numerical results averaged over 30 instances with their standard deviations after the learning with regularized LSTD on Taxi-v3 for  $\gamma = 0.95$ ,  $\lambda = 10^{-9}$ ,  $n = 5000$ ,  $m = 310$ .

single-hidden-layer network in the lazy training regime (Chizat et al., 2019). The results of our theoretical and empirical analyses are outlined below.

**Contributions.** We make the following contributions, taking a step towards a better theoretical understanding of the influence of model complexity  $N/m$  and  $l_2$ -regularization on the performance of Temporal Difference algorithms:

1. We propose a novel double asymptotic regime, where the number of parameters  $N$  and distinct visited states  $m$  go to infinity while maintaining a constant ratio. This leads to a precise assessment of the performance in both over-parameterized ( $N/m > 1$ ) and under-parameterized regimes ( $N/m < 1$ ). This is a nontrivial extension of existing work in supervised learning since several properties essential to proofs, such as the positive definiteness of key matrices, are voided by a discount factor in RL.
2. In the phase transition around  $N/m = 1$ , we observe a peak in the Mean-Squared Bellman Error (MSBE), as illustrated in figure 1, i.e., a double descent phenomenon similar to what has been reported in supervised learning (Mei & Montanari, 2022; Liao et al., 2020).
3. We identify the resolvent of a non-symmetric positive-definite matrix that emerges as a crucial factor in the performance analysis of TD learning algorithms in terms of the MSBE and we provide its deterministic limit form in the double asymptotic regime.
4. We derive analytical equations for both the asymptotic empirical MSBE on the collected transitions and the asymptotic true MSBE. The deterministic forms expose correction terms that we experimentally associate with the double descent phenomenon. We show that the correction terms vanish as the  $l_2$ -regularization is increased or  $N/m$  goes to infinity. We also show that the influence of the  $l_2$ -regularization parameter decreases as  $N/m$  increases.
5. Our theory closely matches empirical results on a range of both toy and small real-world Markov Reward Processes where  $m$  and  $N$  are fixed, but for which the asymptotic regime still gives accurate predictions. Notably, we observe a peak in the true MSBE around  $N/m = 1$  that is not observed in the empirical MSBE. Correction terms, and therefore the difference between true and empirical MSBE, empirically vanish when the number of unvisited states goes to zero.

## 2 RELATED WORK

We review three related approaches to study neural networks in supervised learning or RL. Further technical results from the literature are cited where relevant throughout the paper.

**Neural Tangent Kernel (NTK) regime.** In the NTK regime, one considers that infinitely wide neural networks, with appropriate scaling and initial conditions, behave like the linearization of the

neural network around its initialization (Jacot et al., 2018). However, as highlighted by Chizat et al. (2019), this behavior is not specific to neural networks and is not so much due to over-parameterization than to an implicit choice of scaling. In such a scenario, neural networks can be modeled as a random feature model (Rahimi & Recht, 2007); we adopt this technique in order to abstract from the learning dynamics. The NTK regime was also considered in RL, in both finite and infinite state space, to prove the convergence of infinite-width neural TD learning algorithms towards the global optimum of the MSBE (Cai et al., 2019; Agazzi & Lu, 2022; Liu et al., 2019).

**Mean-Field regime.** Under appropriate initial conditions and scaling, the mean-field analysis models the neural network and its induced feature representation with an empirical distribution, which, at the infinite-width limit, corresponds to a population distribution. The evolution of such a population distribution is characterized by a partial differential equation (PDE) known as the continuity equation and captures Stochastic Gradient Descent (SGD) dynamics as a Wasserstein gradient flow of the objective function (Chizat & Bach, 2018; Rotskoff & Vanden-Eijnden, 2018; Mei et al., 2018). Although more challenging than the NTK regime, the mean-field regime is more realistic since the weights are not restricted to staying in their initial regions (Chizat et al., 2019). The mean-field regime was studied in RL to prove the convergence of infinite-width neural TD learning algorithms towards the global optimum of the MSBE (Zhang et al., 2021; Agazzi & Lu, 2022).

**Double Asymptotic regime.** In the above regimes, the number of data points  $n$  is negligible compared to the number of parameters  $N$  since  $N$  grows to infinity. However, this is rarely the case in practice, which is why the double descent phenomenon can not be explained with the previous regimes (Zhang et al., 2020; Belkin et al., 2018). The double descent phenomenon is characterized by a peak near the interpolation threshold ( $N = n$ ). For this reason, many studies in supervised learning consider a double asymptotic regime (Mei & Montanari, 2022; Louart et al., 2018; Liao et al., 2020; Belkin et al., 2020), where both  $n, N$  go to infinity while maintaining their ratio constant. In the above work, techniques from Random Matrix theory are used to derive a precise description of the phase transition between under- ( $N < n$ ) and over- ( $N > n$ ) parameterization and the double descent phenomenon. Our work extends this approach from supervised learning to RL. Since several key properties from supervised learning are voided by the discount factor, the analysis is substantially more involved in the case of RL. Thomas (2022) investigated off-policy linear TD methods in the limit of large number of states and parameters on a transition matrix of rank 1, and observed a peaking behavior in the MSBE. We consider a more general setting on the on-policy setting and a different ratio in the double asymptotic regime without making assumptions about the rank of the transition matrix.

### 3 PRELIMINARIES

**Notations.** We define  $[n] = \{1, 2, \dots, n\}$ . For a real matrix  $\mathbf{A}$ ,  $[\mathbf{A}]_{ij}$  denotes its  $(i, j)^{th}$  entry. For  $\mathbf{A}$  with real eigenvalues, we denote with  $\nu_{\min}(\mathbf{A})$  the smallest and  $\nu_{\max}(\mathbf{A})$  the largest eigenvalue. The symmetric part of  $\mathbf{A}$  is  $H(\mathbf{A}) = \frac{\mathbf{A} + \mathbf{A}^T}{2}$ . The operator norm of a  $\mathbf{A}$  is written  $\|\mathbf{A}\|$ . The norm induced by  $\mathbf{A}$  on a vector  $\mathbf{v}$  is  $\|\mathbf{v}\|_{\mathbf{A}} = \mathbf{v}^T \mathbf{A} \mathbf{v}$  and  $\|\mathbf{v}\|$  depicts the Euclidean norm of  $\mathbf{v}$ .  $\mathcal{N}(0, 1)$  denotes the standard Gaussian distribution.

**Markov Reward Processes.** We consider a Markov Reward Process (MRP)  $(\mathcal{S}, P, r, \gamma)$ , where  $\mathcal{S} \subseteq \mathbb{R}^d$  is the state space;  $P : \mathcal{S} \times \mathcal{S} \rightarrow [0, 1]$  is the transition kernel (stochastic kernel) and  $P(\mathbf{s}, \mathbf{s}')$  denotes the probability of transitioning to state  $\mathbf{s}'$  from state  $\mathbf{s}$ ;  $r : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$  is the reward function; and  $\gamma \in [0, 1)$  is the discount factor. For notational convenience, the state space  $\mathcal{S}$  is described by the state matrix  $\mathbf{S} \in \mathbb{R}^{d \times |\mathcal{S}|}$ , where each column of  $\mathbf{S}$ , written  $\mathbf{S}_i$ , represents a state in  $\mathcal{S}$ . The transition probability matrix associated with the stochastic kernel  $P$  is  $\mathbf{P} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$ . The goal is to learn the value function  $V : \mathcal{S} \rightarrow \mathbb{R}$ , which maps each state  $\mathbf{s}$  to the expected discounted sum of rewards when starting from  $\mathbf{s}$  and following the dynamics of the MRP defined by  $\mathbf{P}$  as  $V(\mathbf{s}) := \mathbb{E}_{\mathbf{P}} \left[ \sum_{k=1}^{\infty} \gamma^{k-1} r(\mathbf{s}_k, \mathbf{s}_{k+1}) \mid \mathbf{s}_1 = \mathbf{s} \right]$ . It is well-known that the value function is the unique fixed-point of Bellman’s equation

$$\mathbf{V} = \bar{\mathbf{r}} + \gamma \mathbf{P} \mathbf{V}, \quad (1)$$

where  $\mathbf{V} \in \mathbb{R}^{|\mathcal{S}|}$  is a vector whose  $i$ -th element is the value function of the  $i$ -th state  $\mathbf{S}_i$ ; and  $\bar{\mathbf{r}} \in \mathbb{R}^{|\mathcal{S}|}$  is the vector containing the expected rewards, for which  $\bar{r}_i = \mathbb{E}_{\mathbf{P}}[r | \mathbf{S}_i]$  for all  $i \in [|\mathcal{S}|]$ .

**Linear Function Approximation.** In practice, equation 1 cannot be solved since  $P$  is unknown and  $|\mathcal{S}|$  is too large. One common solution is to use Linear Function Approximation (LFA). Using a parameter vector  $\theta \in \mathbb{R}^N$  and a feature matrix  $\Sigma_{\mathcal{S}} \in \mathbb{R}^{N \times |\mathcal{S}|}$ , whose columns are the feature vectors for every state,  $V$  is approximated by  $V \approx \Sigma_{\mathcal{S}}^T \theta$ . In Deep RL, the neural network learns both the feature vectors and the parameter vector. For a given feature matrix, the learning process based on equation 1 amounts to finding a parameter vector  $\theta$  that minimizes the Mean-Squared Bellman error (MSBE)

$$\text{MSBE}(\theta) = \|\bar{r} + \gamma P \Sigma_{\mathcal{S}}^T \theta - \Sigma_{\mathcal{S}}^T \theta\|_{D_{\pi}}^2, \quad (2)$$

where  $\pi \in \mathbb{R}^{|\mathcal{S}|}$  is the stationary distribution induced by the MRP and  $D_{\pi} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$  is its diagonal matrix. Since  $\bar{r} + \gamma P \Sigma_{\mathcal{S}}^T \theta$  may not lie in the span of the bases  $\Sigma_{\mathcal{S}}$ , there may not be a parameter vector  $\theta$  that brings the MSBE to zero.

**Linear Temporal-Difference Methods.** Linear Temporal-Difference (TD) methods are LFA methods that try to minimize the MSBE in equation 2 by replacing the second occurrence of  $\theta$  in equation 2 with an auxiliary vector  $u$ , minimizing on  $u$  and then finding a  $\theta$  close to  $u$  (Dann et al., 2014):

$$u^* = \arg \min_{u \in \mathbb{R}^N} \|\bar{r} + \gamma P \Sigma_{\mathcal{S}}^T \theta - \Sigma_{\mathcal{S}}^T u\|_{D_{\pi}}^2 \quad (\text{projection step}), \quad (3)$$

$$\theta^* = \arg \min_{\theta \in \mathbb{R}^N} \|\Sigma_{\mathcal{S}}^T u^* - \Sigma_{\mathcal{S}}^T \theta\|_{D_{\pi}}^2 \quad (\text{fixed-point step}). \quad (4)$$

The projection step (equation 3) implies that TD methods actually minimize the Mean-Squared Projected Bellman error (MSPBE) rather than the MSBE (Dann et al., 2014). In our asymptotic regime, as the number of features  $N \rightarrow \infty$ , the class of representable value functions becomes richer, and the MSPBE converges to the MSBE (Cai et al., 2019; Agazzi & Lu, 2022).

## 4 SYSTEM MODEL

We now describe the key elements on which we base our asymptotic analysis of the MSBE in TD learning: Random features, the regularized LSTD algorithm, and the double asymptotic regime.

### 4.1 REGULARIZED LSTD WITH RANDOM FEATURES

**Random Features.** We consider value function approximation using the Random Feature (RF) mapping  $\text{RF} : \mathcal{S} \rightarrow \mathbb{R}^N$  defined for all  $s \in \mathcal{S}$  as

$$\text{RF}(s) = \sigma(\mathbf{W}s), \quad (5)$$

where  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  is  $K_{\sigma}$ -Lipschitz continuous and applied component-wise;  $\mathbf{W} = \varphi(\tilde{\mathbf{W}}) \in \mathbb{R}^{N \times d}$  is a random weight matrix fixed throughout training for which  $\tilde{\mathbf{W}} \in \mathbb{R}^{N \times d}$  has independent and identically distributed  $\mathcal{N}(0, 1)$  entries, and  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is  $K_{\varphi}$ -Lipschitz continuous and applied component-wise. From the perspective of neural networks, the  $N$  random features can be seen as  $N$  outputs from a single-hidden-layer neural network. In our asymptotic regime, this simplification becomes even more accurate as the number of features  $N$  of the single layer grows towards infinity and we enter into the lazy training regime, where weights barely deviate from their random initial values (Chizat et al., 2019). In the literature on Deep Learning and double descent, large-width neural networks are often modeled using asymptotic random features (Louart et al., 2018; Liao et al., 2020; Mei & Montanari, 2022), including in RL (Cai et al., 2019; Agazzi & Lu, 2022; Liu et al., 2019). In the following, we denote the random feature matrix of any state matrix  $\mathbf{A} \in \mathbb{R}^{d \times p}$  as  $\Sigma_{\mathbf{A}}$  where RF is applied column-wise, i.e.,  $\Sigma_{\mathbf{A}} = \sigma(\mathbf{W}\mathbf{A})$ .

**Sample Matrices and Empirical MSBE.** We assume that the transition probability matrix  $P$  is unknown during the training phase. Instead, we have a dataset of  $n$  transitions consisting of states, rewards, and next-states drawn from the MRP, i.e.,  $\mathcal{D}_{\text{train}} := \{(s_i, r_i, s'_i)\}_{i=1}^n$  where  $s'_i \sim P(s_i, \cdot)$ . We consider the on-policy setting, where  $\mathcal{D}_{\text{train}}$  is derived from a sample path of the MRP or its stationary distribution  $\pi$ . We collect the states and rewards in the sample matrices

$$\mathbf{X}_n = [s_1, \dots, s_n] \in \mathbb{R}^{d \times n}, \quad \mathbf{r} = [r_1, \dots, r_n]^T \in \mathbb{R}^n, \quad \mathbf{X}'_n = [s'_1, \dots, s'_n] \in \mathbb{R}^{d \times n}. \quad (6)$$

Let  $\hat{\mathcal{S}} \subseteq \mathcal{S}$  be the set of distinct states in  $\mathcal{D}_{\text{train}}$ , which we call visited states, and let  $m = |\hat{\mathcal{S}}|$ . Let  $\hat{\mathbf{S}} \in \mathbb{R}^{d \times m}$  be the state matrix of  $\hat{\mathcal{S}}$ , i.e., each column  $\hat{\mathbf{S}}_i$  of  $\hat{\mathbf{S}}$  describes a state in  $\hat{\mathcal{S}}$ . We denote by  $\Sigma_{\hat{\mathcal{S}}} \in \mathbb{R}^{N \times m}$ ,  $\Sigma_{\mathbf{X}_n} \in \mathbb{R}^{N \times n}$ , and  $\Sigma_{\mathbf{X}'_n} \in \mathbb{R}^{N \times n}$  the random feature matrices of  $\hat{\mathcal{S}}$ ,  $\mathbf{X}_n$ , and  $\mathbf{X}'_n$ , respectively. For the proof of our results, it will be mathematically advantageous to express  $\Sigma_{\mathbf{X}_n}$  and  $\Sigma_{\mathbf{X}'_n}$  as the product of  $\Sigma_{\hat{\mathcal{S}}}$  with auxiliary matrices  $\hat{\mathbf{U}}_n \in \mathbb{R}^{m \times n}$  and  $\hat{\mathbf{V}}_n \in \mathbb{R}^{m \times n}$  as follows:

$$\Sigma_{\mathbf{X}_n} = \sqrt{n} \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \quad \text{and} \quad \Sigma_{\mathbf{X}'_n} = \sqrt{n} \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{V}}_n. \quad (7)$$

Each column  $i$  of  $\sqrt{n} \hat{\mathbf{U}}_n$  is a one-hot vector, where the  $j$ -th element equals 1 if the  $i$ -th state  $\mathbf{s}_i$  of  $\mathbf{X}_n$  is  $\hat{\mathbf{S}}_j$ , and similarly for  $\sqrt{n} \hat{\mathbf{V}}_n$  and  $\mathbf{X}'_n$ . Since  $\mathbf{P}$  is unknown, we aim to find  $\boldsymbol{\theta}$  that minimizes the empirical version of the MSBE (equation 2) obtained with transitions collected in  $\mathcal{D}_{\text{train}}$ :

$$\widehat{\text{MSBE}}(\boldsymbol{\theta}) = \frac{1}{n} \|\mathbf{r} + \gamma \Sigma_{\mathbf{X}'_n}^T \boldsymbol{\theta} - \Sigma_{\mathbf{X}_n}^T \boldsymbol{\theta}\|^2, \quad (8)$$

which uses the Euclidean norm since the distribution is reflected by the samples. Assuming globally stable MRP, a fixed number of features, and all states being visited,  $\widehat{\text{MSBE}}(\boldsymbol{\theta})$  converges to  $\text{MSBE}(\boldsymbol{\theta})$  with probability 1, as the number of collected transitions  $n \rightarrow \infty$ . This follows from the law of large numbers (Stachurski, 2009). In our analysis, we will also consider the case where  $n \rightarrow \infty$  without visiting all states, i.e.,  $m < |\mathcal{S}|$ , such that there can be a significant difference between  $\widehat{\text{MSBE}}(\boldsymbol{\theta})$  and  $\text{MSBE}(\boldsymbol{\theta})$ .

**Regularized Least-Square Temporal-Difference Methods.** Regularized Least-Square Temporal-Difference (LSTD) Methods (Bradtke & Barto, 1996) are linear TD methods that solve an empirical regularized version of equation 3 and 4 with transitions collected in  $\mathcal{D}_{\text{train}}$ :

$$\mathbf{u}^* = \arg \min_{\mathbf{u} \in \mathbb{R}^N} \|\mathbf{r} + \gamma \Sigma_{\mathbf{X}'_n}^T \boldsymbol{\theta} - \Sigma_{\mathbf{X}_n}^T \mathbf{u}\|^2 + \lambda_{m,n} \|\mathbf{u}\|^2, \quad (9)$$

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^N} \|\Sigma_{\mathbf{X}_n}^T \mathbf{u}^* - \Sigma_{\mathbf{X}'_n}^T \boldsymbol{\theta}\|^2, \quad (10)$$

where  $\lambda_{m,n} > 0$  is the effective  $l_2$ -regularization parameter, introduced to mitigate overfitting (Hoffman et al., 2011; Chen et al., 2013). It is well known that for  $\lambda_{m,n} = 0$  and with the number of samples  $n \rightarrow \infty$ , the fixed point of the approximation equation 9 and 10 equals the fixed point of equation 3 and 4 with probability one. Solving the fixed-point of the linear system approximation given by equation 9 and 10 gives

$$\hat{\boldsymbol{\theta}} = \left[ \Sigma_{\mathbf{X}_n} \left[ \Sigma_{\mathbf{X}_n} - \gamma \Sigma_{\mathbf{X}'_n} \right]^T + \lambda_{m,n} \mathbf{I}_N \right]^{-1} \Sigma_{\mathbf{X}_n} \mathbf{r}. \quad (11)$$

Under appropriate learning rates, linear TD methods based on gradient-descent converge towards the same fixed-point  $\hat{\boldsymbol{\theta}}$  (Robbins & Monro, 1951; Dann et al., 2014; Sutton & Barto, 2018). Besides reducing overfitting, an appropriate  $\lambda_{m,n}$  ensures that  $\Sigma_{\mathbf{X}_n} \left[ \Sigma_{\mathbf{X}_n} - \gamma \Sigma_{\mathbf{X}'_n} \right]^T + \lambda_{m,n} \mathbf{I}_N$  is invertible.

## 4.2 DOUBLE ASYMPTOTIC REGIME AND RESOLVENT IN LSTD

We study the regularized LSTD in the following double asymptotic regime:

**Assumption 1** (Double Asymptotic Regime). *As  $N, m, d \rightarrow \infty$ , we have:*

1.  $0 < \lim \min \left\{ \frac{N}{m}, \frac{d}{m}, \frac{m}{|\mathcal{S}|} \right\} < \lim \max \left\{ \frac{N}{m}, \frac{d}{m}, \frac{m}{|\mathcal{S}|} \right\} < \infty$ .

2. *There exists  $K_{\mathcal{S}}, K_r > 0$  such that  $\lim \sup_{|\mathcal{S}|} \|\mathbf{S}\| < K_{\mathcal{S}}$  and  $r(\cdot, \cdot)$  is bounded by  $K_r$ .*

In order to use Random Matrix tools, we rewrite equation 11 as (see proof in Lemma L.9)

$$\hat{\boldsymbol{\theta}} = \frac{1}{mn} \Sigma_{\mathbf{X}_n} \left[ \frac{1}{mn} \left[ \Sigma_{\mathbf{X}_n} - \gamma \Sigma_{\mathbf{X}'_n} \right]^T \Sigma_{\mathbf{X}_n} + \frac{\lambda_{m,n}}{mn} \mathbf{I}_n \right]^{-1} \mathbf{r}. \quad (12)$$

Instead of the effective  $l_2$ -regularization parameter  $\lambda_{m,n}$ , we will use its scaled version  $\lambda = \frac{\lambda_{m,n}}{mn}$  in the remainder of this paper. We observe that  $\hat{\boldsymbol{\theta}} = \frac{1}{mn} \Sigma_{\mathbf{X}_n} \mathbf{Q}_m(\lambda) \mathbf{r}$  depends on the resolvent

$$\mathbf{Q}_m(\lambda) = \left[ \frac{1}{mn} \left[ \Sigma_{\mathbf{X}_n} - \gamma \Sigma_{\mathbf{X}'_n} \right]^T \Sigma_{\mathbf{X}_n} + \lambda \mathbf{I}_n \right]^{-1} = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{\mathcal{S}}}^T \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1} \quad (13)$$

when  $\frac{1}{m}(\hat{U}_n - \gamma\hat{V}_n)^T \Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}} \hat{U}_n + \lambda \mathbf{I}_n$  is invertible, which in general may not be the case. We can guarantee invertibility if the empirical transition model matrix  $\hat{A}_m \in \mathbb{R}^{m \times m}$

$$\hat{A}_m = \hat{U}_n (\hat{U}_n - \gamma \hat{V}_n)^T \quad (14)$$

has a positive-definite symmetric part (see Appendix H for a formal proof). For the remainder of the paper, we therefore make the following assumption on  $\hat{A}_m$ :

**Assumption 2** (Bounded Eigenspectrum). *There exist  $0 < \xi_{\min} < \xi_{\max}$  such that for every  $m$ , all the eigenvalues of  $H(\hat{A}_m)$  are in  $[\xi_{\min}, \xi_{\max}]$ .*

Note that the above assumption is satisfied for regularized pathwise LSTD (Lazaric et al., 2012), and also for sufficiently large  $n$  (see Appendix H).

## 5 ASYMPTOTIC ANALYSIS OF REGULARIZED LSTD

In this section, we present our main theoretical results, which characterize the true and empirical MSBE under Assumptions 1 and 2.

### 5.1 AN EQUIVALENT DETERMINISTIC RESOLVENT

The resolvent  $\mathbf{Q}_m(\lambda)$  (in equation 13) plays a significant role in the performance of regularized LSTD since  $\hat{\theta} = \frac{1}{\sqrt{n}} \frac{1}{m} \Sigma_{\mathcal{S}} \hat{U}_n \mathbf{Q}_m(\lambda) r$ . To assess the asymptotic  $\widehat{\text{MSBE}}(\hat{\theta})$  and true  $\text{MSBE}(\hat{\theta})$ , we first find a deterministic equivalent for the resolvent  $\mathbf{Q}_m(\lambda)$ . A natural deterministic equivalent would be  $\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m(\lambda)]$ , but it involves integration without having a closed form expression (due to the matrix inverse) and is inconvenient for practical computation. Leveraging Random Matrix tools, the following Theorem 5.1 proposes an asymptotic form that is *i.* close to  $\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m(\lambda)]$  under Assumptions 1 and 2, and *ii.* numerically more accessible (for the proof, see Appendix E).

**Theorem 5.1** (Asymptotic Deterministic Resolvent). *Under Assumptions 1 (double asymptotic regime) and 2 (bounded spectrum), let  $\lambda > 0$  and let the deterministic resolvent  $\bar{\mathbf{Q}}_m(\lambda) \in \mathbb{R}^{n \times n}$  be*

$$\bar{\mathbf{Q}}_m(\lambda) = \left[ \frac{N}{m} \frac{1}{1+\delta} (\hat{U}_n - \gamma \hat{V}_n)^T \Phi_{\mathcal{S}} \hat{U}_n + \lambda \mathbf{I}_n \right]^{-1}, \quad (15)$$

where the deterministic Gram feature matrix  $\Phi_{\mathcal{S}} \in \mathbb{R}^{m \times m}$  is

$$\Phi_{\mathcal{S}} = \mathbb{E}_{\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)} \left[ \sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right], \quad (16)$$

and the correction factor  $\delta$  is the unique, positive, solution to

$$\delta = \frac{1}{m} \text{Tr} \left( (\hat{U}_n - \gamma \hat{V}_n)^T \Phi_{\mathcal{S}} \hat{U}_n \left[ \frac{N}{m} \frac{1}{1+\delta} (\hat{U}_n - \gamma \hat{V}_n)^T \Phi_{\mathcal{S}} \hat{U}_n + \lambda \mathbf{I}_n \right]^{-1} \right). \quad (17)$$

Then  $\lim_{m \rightarrow \infty} \|\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m(\lambda)] - \bar{\mathbf{Q}}_m(\lambda)\| = 0$ . The correction factor  $\delta$  diminishes as  $N$  or  $\lambda$  grows (see Lemma J.3 and J.4).

**Remark 1.** Since  $\delta \rightarrow 0$  when  $N/m \rightarrow \infty$ , the correction factor  $\frac{1}{1+\delta}$  arises from our asymptotic regime, which keeps the ratio  $N/m$  asymptotically constant (see Lemma J.1 for existence and uniqueness). Similar correction factors arise in related Random Matrix literature, which, however, mostly deals with positive semi-definite matrices (Couillet & Debbah, 2011; Liu et al., 2019; Liao et al., 2020). Our problem exceeds this frame, so we prove the result, including existence and uniqueness, with a somewhat more involved analysis based on the eigenspectrum of the products of matrices with positive-definite symmetric part and skew-symmetric matrices (see Appendix J).

**Remark 2.** For the case of supervised learning, a comparable proposition is presented by Louart et al. (2018, Theorem 1). It constitutes a special case of Theorem 5.1 with  $\gamma = 0$ , which corresponds to the case where we learn the reward function.

**Remark 3.** Note that the eigenvalues of  $\bar{\mathbf{Q}}_m(\lambda)$  are not necessarily real, which renders many tools from the related Random Matrix literature not applicable, e.g., Stieltjes transforms would provide information on the eigenspectrum density of matrices based on the trace of their resolvents.

## 5.2 ASYMPTOTIC EMPIRICAL MEAN-SQUARED BELLMAN ERROR

TD methods learn by minimizing the empirical MSBE (equation 8) and, under appropriate learning rates, converge towards the empirical MSBE of LSTD, as mentioned in Section 4. It is straightforward to show that this leads to an optimal  $\widehat{\text{MSBE}}(\hat{\theta}) = \frac{\lambda^2}{n} \|\mathbf{Q}_m(\lambda)\mathbf{r}\|^2$  (see Appendix F). Using concentration arguments for Gaussian distributions and Lipschitz applications, as well as Theorem 5.1, we derive the following deterministic form (see proof in Appendix F).

**Theorem 5.2** (Asymptotic Empirical MSBE). *Under the conditions of Theorem 5.1, the deterministic asymptotic empirical MSBE is  $\widehat{\text{MSBE}}(\hat{\theta}) = \frac{\lambda^2}{n} \|\bar{\mathbf{Q}}_m(\lambda)\mathbf{r}\|^2 + \hat{\Delta}$ , with second-order correction factor*

$$\hat{\Delta} = \frac{\lambda^2}{n} \frac{\frac{1}{N} \text{Tr}(\mathbf{Q}_m(\lambda)\Psi_2\mathbf{Q}_m(\lambda)^T)}{1 - \frac{1}{N} \text{Tr}(\Psi_2\mathbf{Q}_m(\lambda)^T\Psi_1\mathbf{Q}_m(\lambda))} \|\bar{\mathbf{Q}}_m(\lambda)\mathbf{r}\|_{\Psi_1}^2, \text{ where} \quad (18)$$

$$\Psi_1 = \frac{N}{m} \frac{1}{1+\delta} \hat{\mathbf{U}}_n^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n, \quad \text{and} \quad \Psi_2 = \frac{N}{m} \frac{1}{1+\delta} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n). \quad (19)$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,  $\widehat{\text{MSBE}}(\hat{\theta}) - \widehat{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0$ .

**Remark 4.** As  $N/m \rightarrow \infty$ , we find  $\widehat{\text{MSBE}}(\hat{\theta}) \rightarrow 0$ .

**Remark 5.** In supervised learning, a comparable proposition is presented by Louart et al. (2018, Theorem 3). It is a special case of Theorem 5.2 with  $\gamma = 0$ , where we learn the reward function  $r$ .

## 5.3 ASYMPTOTIC MEAN-SQUARED BELLMAN ERROR

While the empirical MSBE only takes states from the data set into account, the true MSBE (equation 2) involves all states in  $\mathcal{S}$ . To extend the convergence results from the previous section to this case, we require some further notations. Using a decomposition similar to equation 7, we express  $\Sigma_{\mathbf{X}_n}$  and  $\Sigma_{\mathbf{X}'_n}$  as a product of the random feature matrix of the entire state space  $\Sigma_{\mathcal{S}} \in \mathbb{R}^{N \times |\mathcal{S}|}$  with  $\mathbf{U}_n \in \mathbb{R}^{|\mathcal{S}| \times n}$  and  $\mathbf{V}_n \in \mathbb{R}^{|\mathcal{S}| \times n}$  instead of  $\Sigma_{\hat{S}}, \hat{\mathbf{U}}_n, \hat{\mathbf{V}}_n$ . We obtain a decomposition of the transition model matrix  $\mathbf{A}_n = \mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T$ .  $\mathbf{A}_n$  was used by Boyan (1999) to interpret LSTD as model-based RL. Tsitsiklis & Van Roy (1996) and Nedić & Bertsekas (2003) showed that  $\mathbb{E}[\mathbf{A}_n] \rightarrow \mathbf{D}_{\pi} [\mathbf{I}_{|\mathcal{S}|} - \gamma\mathbf{P}]$  as  $n \rightarrow \infty$ . The bound on the difference  $\|\mathbf{A}_n - \mathbf{D}_{\pi} [\mathbf{I}_{|\mathcal{S}|} - \gamma\mathbf{P}]\|$  as a function of  $n$  was studied by Tagorti & Scherrer (2015). We make the following assumption on this norm:

**Assumption 3.** As  $n, m \rightarrow \infty$ , we have  $\|\mathbf{A}_n - \mathbf{D}_{\pi} [\mathbf{I}_{|\mathcal{S}|} - \gamma\mathbf{P}]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ .

Using an approach similar to that of Theorem 5.2, plus a detailed analysis of operator norms, we obtain the following deterministic form of the asymptotic MSBE (proof in Appendix G):

**Theorem 5.3** (Asymptotic MSBE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSBE is  $\widehat{\text{MSBE}}(\hat{\theta}) = \left\| \bar{\mathbf{r}} + \gamma \frac{1}{\sqrt{n}} \frac{N}{m} \frac{1}{1+\delta} \mathbf{P} \Phi_{\mathcal{S}} \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} - \frac{1}{\sqrt{n}} \frac{N}{m} \frac{1}{1+\delta} \Phi_{\mathcal{S}} \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} \right\|_{\mathbf{D}_{\pi}}^2 + \Delta$ , with second-order correction factor*

$$\Delta = \frac{1}{n} \frac{\text{Tr}(\Lambda_{\mathbf{P}} [\Theta_{\mathcal{S}} \Psi_2 \Theta_{\mathcal{S}}^T - 2\Theta_{\mathcal{S}} (\mathbf{U}_n - \gamma\mathbf{V}_n)^T \Psi_{\mathcal{S}} + \Psi_{\mathcal{S}}])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m(\lambda)^T \Psi_1 \bar{\mathbf{Q}}_m(\lambda))} \|\bar{\mathbf{Q}}_m(\lambda)\mathbf{r}\|_{\Psi_1}^2, \text{ where} \quad (20)$$

$$\Psi_{\mathcal{S}} = \frac{N}{m} \frac{1}{1+\delta} \Phi_{\mathcal{S}}, \quad \Lambda_{\mathbf{P}} = [\mathbf{I}_{|\mathcal{S}|} - \gamma\mathbf{P}]^T \mathbf{D}_{\pi} [\mathbf{I}_{|\mathcal{S}|} - \gamma\mathbf{P}], \quad \text{and} \quad \Theta_{\mathcal{S}} = \Psi_{\mathcal{S}} \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda). \quad (21)$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,  $\text{MSBE}(\hat{\theta}) - \widehat{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0$ .

**Remark 6.** Like the empirical  $\widehat{\text{MSBE}}(\hat{\theta})$  in Theorem 5.2, the true  $\text{MSBE}(\hat{\theta})$  is also influenced by the correction terms  $\delta$  and  $\Delta$ . Note that in asymptotic regimes where  $N/m \rightarrow \infty$  or  $\lambda \rightarrow \infty$ , the correction terms vanish. When  $N/m \rightarrow \infty$ ,  $\widehat{\text{MSBE}}(\hat{\theta})$  is independent of  $\lambda$  as shown in details in Appendix C.

**Remark 7.** When all states have been visited, the common subexpressions in the second-order correction factors  $\hat{\Delta}$  and  $\Delta$  dominate so that  $\hat{\Delta}, \Delta$  become similar (for a proof, see Lemma G.8).



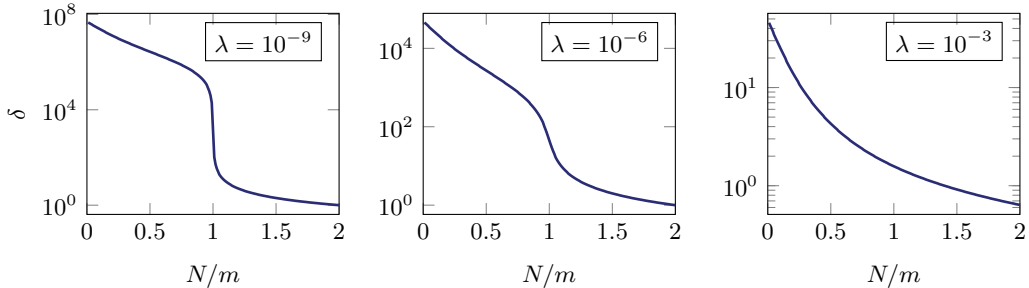


Figure 2: **The correction factor  $\delta$  is a decreasing function of the number of parameters  $N$ . For a small  $l_2$ -regularization parameter  $\lambda$ , we observe a sharp decrease near the interpolation threshold ( $N = m$  for  $m$  distinct visited states). As  $\lambda$  increases, the function becomes smoother and smaller (note the different scales of the y-axis).  $\delta$  is computed with equation 17 on Taxi-v3 with  $\gamma = 0.95, m = 310, n = 5000$ .**

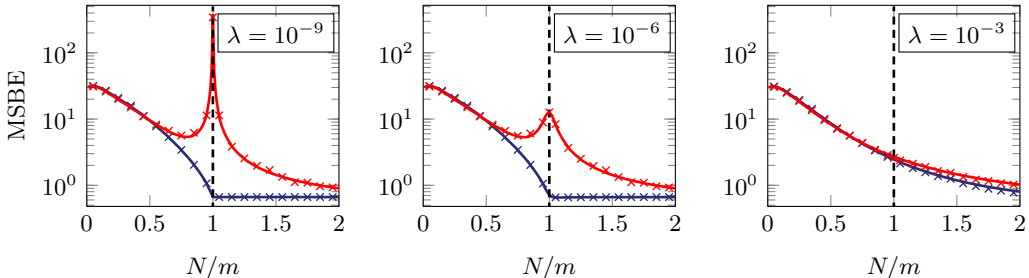


Figure 3: **The double descent phenomenon occurs in the true MSBE (red) of regularized LSTD, peaking around the interpolation threshold ( $N = m$  for  $N$  parameters,  $m$  distinct visited states) when the empirical  $\widehat{\text{MSBE}}$  (blue) vanishes. It diminishes as the  $l_2$ -regularization parameter  $\lambda$  increases. Continuous lines indicate the theoretical values from Theorem 5.2 and Theorem 5.3, the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in Taxi-v3 with  $\gamma = 0.95, m = 310, n = 5000$ .**

## 6 NUMERICAL EXPERIMENTS

In this section, we provide an empirical evaluation, including a discussion of the behavior of the correction factor  $\delta$  from Theorem 5.1, and its impact on the empirical and true MSBE from Theorem 5.2 and Theorem 5.3. Additional experiments can be found in Appendix A.

**Experimental Setup.** We use the recursive regularized LSTD implementation of Dann et al. (2014) on three MRPs: a synthetic ergodic MRP (500 states); a gridworld MRP (400 states) obtained from a random policy in a  $20 \times 20$  gridworld (Ahmed, 2018); and a Taxi-v3 MRP (356 states) obtained from a learned policy in OpenAI gym Taxi-v3 (Towers et al., 2023) (Figure 1a). In all cases, states are described by  $d$ -Gaussian vectors where  $d = 50$ . For the random features,  $\mathbf{W}$  is drawn from a Gaussian distribution and  $\sigma(\cdot) = \max(0, \cdot)$  is the ReLU function. For all experiments,  $\mathcal{D}_{\text{train}} := \{(s_i, r_i, s'_i)\}_{i=1}^n$  is derived from a sample path of  $n$  transitions with the same seed (42). For each instance  $i$ , we sample random features using the seed  $i$ . The following graphs show averages over 30 instances.

**Correction Factor  $\delta$  vs. Model Complexity.** The correction factor  $\delta$  (equation 17) plays a key role in the asymptotic  $\widehat{\text{MSBE}}$  and MSBE. Figure 2 shows  $\delta$  as a function of the model complexity  $N/m$  and for different values of the regularization parameter  $\lambda$ . It confirms that, as stated in Theorem 5.1,  $\delta$  is a decreasing function of  $N/m$ . For a small  $\lambda$ , we observe a sharp decrease at the interpolation threshold ( $N = m$ ). E.g., for  $\lambda = 10^{-9}$ ,  $\delta$  falls from an order of  $10^7$  in under-parameterized models ( $N < m$ ) to an order of  $10^1$  in over-parameterized models ( $N > m$ ). For larger values of  $\lambda$ ,  $\delta$  decreases more smoothly and has smaller values. Further experiments on the behavior of  $\delta$  and experiments for other environments are provided in Appendix A.

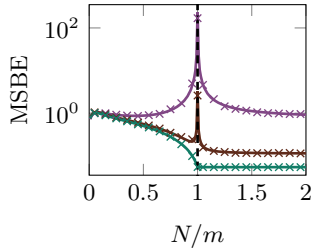


Figure 4: **With more distinct states  $m$  visited, the double descent in the MSBE diminishes, disappearing for  $m = |\mathcal{S}|$ .** MSBE from Theorem 5.3 (lines) and avg. numerical results over 30 instances (crosses) in a synthetic ergodic MRP for  $m = 0.86|\mathcal{S}|$  (purple),  $m = 0.998|\mathcal{S}|$  (maroon), and  $m = |\mathcal{S}|$  (green) with  $\gamma = 0.95$ ,  $s = |\mathcal{S}|$ ,  $n = 3000$ .

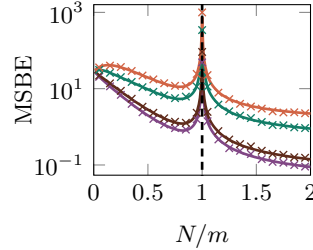


Figure 5: **The discount factor  $\gamma$  has little effect on the double descent in the MSBE.** Results in the Gridworld MRP for  $\gamma = 0$  (purple),  $\gamma = 0.5$  (maroon),  $\gamma = 0.95$  (green), and  $\gamma = 0.99$  (orange) with  $m = 386$ ,  $n = 5000$ . MSBE from Theorem 5.3 (lines) and avg. numerical results over 30 instances (crosses).

**Double Descent Behavior.** As a consequence of the sharp transition of the correction factor  $\delta$  for small  $l_2$ -regularization parameters, as discussed above, Theorem 5.2 and Theorem 5.3 predict a change in behavior of the empirical  $\widehat{\text{MSBE}}$  and true MSBE between the under- and overparameterized regimes. Figure 3 shows both  $\widehat{\text{MSBE}}$  and MSBE as a function of the model complexity  $N/m$  with different  $l_2$ -regularization penalties  $\lambda$  in Taxi-v3. Despite the fact that the equations for  $\widehat{\text{MSBE}}$  and MSBE were derived for the asymptotic case  $N \rightarrow \infty$ , we observe an almost perfect match with the numerically evaluated original definitions in equation 8 and equation 2. For small  $\lambda$ , the true MSBE exhibits a peak around the interpolation threshold  $N = m$ , leading to a double descent phenomenon. In contrast, the empirical  $\widehat{\text{MSBE}}$  is close to its minimum at  $N = m$  and almost constant for  $N \geq m$ , so no double descent is observed. While for the Taxi-v3, the empirical  $\widehat{\text{MSBE}}$  is smaller than the true MSBE, this is not necessarily the case in other environments, where the empirical  $\widehat{\text{MSBE}}$  can be larger overall than the true MSBE (see further experiments in Appendix A). For larger  $\lambda$ , the double descent in the true MSBE disappears and the difference between the true MSBE and the empirical  $\widehat{\text{MSBE}}$  is less pronounced, although it may not vanish. Appendix B shows a similar double descent phenomenon for the Mean-Squared Value Error. All the above observations are in accordance with established results in supervised learning (Liao et al., 2020).

**Impact of the Number of Unvisited States and the Discount Factor  $\gamma$ .** Once all states have been visited, MSBE and  $\widehat{\text{MSBE}}$  have similar behavior, with no peak at the interpolation threshold ( $N = m$ ), see also Remark 7. The experiments in Figure 4 depict this behavior. They also illustrate that the double descent phenomenon diminishes as the number of distinct unvisited states goes to zero. The experiments in Figure 5 illustrate that the discount factor  $\gamma$  has little impact on the double descent phenomenon.

## 7 CONCLUSION

In this work, we have analyzed the performance of regularized LSTD with random features in a novel double asymptotic regime, where the number of parameters  $N$  and distinct visited states  $m$  go to infinity with a constant ratio. We have established deterministic limit forms for both the empirical MSBE and true MSBE that feature correction terms. We have observed that these correction terms are responsible for a double descent phenomenon in the true MSBE, similar to supervised learning, resulting in a sudden drop in performance for  $N = m$ . The correction terms vanish, and so does the double descent phenomenon when the  $l_2$ -regularization is increased or the number of unvisited states goes to zero. Directions for future work include a study of the off-policy setting, extending our results beyond one hidden layer to deep neural networks, and going beyond policy evaluation in order to investigate other RL algorithms, such as Q-Learning.

## 8 ACKNOWLEDGEMENTS

This work was supported by Ecole Polytechnique. We also thank Zhenyu Liao for his helpful discussions on this work.

## REFERENCES

- Andrea Agazzi and Jianfeng Lu. Temporal-difference learning with nonlinear function approximation: lazy training and mean field regimes. In Mathematical and Scientific Machine Learning, pp. 37–74, 2022.
- Zafarali Ahmed. emdp. <https://github.com/zafarali/emdp>, 2018.
- Sanjeev Arora, Simon Du, Wei Hu, Zhiyuan Li, and Ruosong Wang. Fine-grained analysis of optimization and generalization for overparameterized two-layer neural networks. In International Conference on Machine Learning, pp. 322–332, 2019.
- Mikhail Belkin, Daniel Hsu, Siyuan Ma, and Soumik Mandal. Reconciling modern machine learning and the bias-variance trade-off. arXiv preprint arXiv:1812.11118, 321, 2018.
- Mikhail Belkin, Daniel Hsu, and Ji Xu. Two models of double descent for weak features. SIAM Journal on Mathematics of Data Science, 2(4):1167–1180, 2020.
- Eloïse Berthier, Ziad Kobeissi, and Francis Bach. A non-asymptotic analysis of non-parametric temporal-difference learning. Advances in Neural Information Processing Systems, 35:7599–7613, 2022.
- Alberto Bietti and Julien Mairal. On the inductive bias of neural tangent kernels. Advances in Neural Information Processing Systems, 32, 2019.
- Justin A. Boyan. Least-squares temporal difference learning. In International Conference on Machine Learning, pp. 49–56, 1999.
- Steven J. Bradtke and Andrew G. Barto. Linear least-squares algorithms for temporal difference learning. Machine Learning, 22(1):33–57, 1996.
- Qi Cai, Zhuoran Yang, Jason D. Lee, and Zhaoran Wang. Neural temporal-difference learning converges to global optima. Advances in Neural Information Processing Systems, 32, 2019.
- Yuan Cao, Zhiying Fang, Yue Wu, Ding-Xuan Zhou, and Quanquan Gu. Towards understanding the spectral bias of deep learning. arXiv preprint arXiv:1912.01198, 2019.
- Shenglei Chen, Geng Chen, and Ruijun Gu. An efficient L2-norm regularized least-squares temporal difference learning algorithm. Knowledge-Based Systems, 45:94–99, 2013.
- Lénaïc Chizat and Francis Bach. On the global convergence of gradient descent for overparameterized models using optimal transport. Advances in Neural Information Processing Systems, 31, 2018.
- Lénaïc Chizat, Edouard Oyallon, and Francis Bach. On lazy training in differentiable programming. Advances in Neural Information Processing Systems, 32, 2019.
- Romain Couillet and Merouane Debbah. Random Matrix Methods for Wireless Communications. Cambridge University Press, 2011.
- Christoph Dann, Gerhard Neumann, Jan Peters, et al. Policy evaluation with temporal differences: A survey and comparison. Journal of Machine Learning Research, 15:809–883, 2014.
- Kefan Dong, Yuping Luo, Tianhe Yu, Chelsea Finn, and Tengyu Ma. On the expressivity of neural networks for deep reinforcement learning. In International Conference on Machine Learning, pp. 2627–2637, 2020.

- Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, and Sergey Levine. Soft actor-critic algorithms and applications. CoRR, 2018.
- Matthew W. Hoffman, Alessandro Lazaric, Mohammad Ghavamzadeh, and Rémi Munos. Regularized least squares temporal difference learning with nested L2 and L1 penalization. In European Workshop on Reinforcement Learning, pp. 102–114. Springer, 2011.
- Roger A. Horn and Charles R. Johnson. Matrix Analysis. Cambridge University Press, 2012.
- Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. Advances in Neural Information Processing Systems, 31, 2018.
- Aviral Kumar, Rishabh Agarwal, Dibya Ghosh, and Sergey Levine. Implicit under-parameterization inhibits data-efficient deep reinforcement learning. In International Conference on Learning Representations, 2020.
- Alessandro Lazaric, Mohammad Ghavamzadeh, and Rémi Munos. Finite-sample analysis of least-squares policy iteration. Journal of Machine Learning Research, 13:3041–3074, 2012.
- Michel Ledoux. The Concentration of Measure Phenomenon. American Mathematical Soc., 2001.
- Jaehoon Lee, Lechao Xiao, Samuel Schoenholz, Yasaman Bahri, Roman Novak, Jascha Sohl-Dickstein, and Jeffrey Pennington. Wide neural networks of any depth evolve as linear models under gradient descent. Advances in Neural Information Processing Systems, 32, 2019.
- Zhenyu Liao, Romain Couillet, and Michael W. Mahoney. A random matrix analysis of random fourier features: beyond the Gaussian kernel, a precise phase transition, and the corresponding double descent. Advances in Neural Information Processing Systems, 33:13939–13950, 2020.
- Boyi Liu, Qi Cai, Zhuoran Yang, and Zhaoran Wang. Neural trust region/proximal policy optimization attains globally optimal policy. Advances in Neural Information Processing Systems, 32, 2019.
- Cosme Louart, Zhenyu Liao, and Romain Couillet. A random matrix approach to neural networks. The Annals of Applied Probability, 28(2):1190–1248, 2018.
- Xufang Luo, Qi Meng, Di He, Wei Chen, and Yunhong Wang. I4r: Promoting deep reinforcement learning by the indicator for expressive representations. In IJCAI, 2020.
- Clare Lyle, Mark Rowland, and Will Dabney. Understanding and preventing capacity loss in reinforcement learning. In International Conference on Learning Representations, 2021.
- Song Mei and Andrea Montanari. The generalization error of random features regression: Precise asymptotics and the double descent curve. Communications on Pure and Applied Mathematics, 75(4):667–766, 2022.
- Song Mei, Andrea Montanari, and Phan-Minh Nguyen. A mean field view of the landscape of two-layer neural networks. Proceedings of the National Academy of Sciences, 115(33):E7665–E7671, 2018.
- Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Belle-mare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. Nature, 518(7540):529–533, 2015.
- A. Nedić and Dimitri P. Bertsekas. Least squares policy evaluation algorithms with linear function approximation. Discrete Event Dynamic Systems, 13(1-2):79–110, 2003.
- Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. Advances in Neural Information Processing Systems, 20, 2007.
- Herbert Robbins and Sutton Monro. A stochastic approximation method. The Annals of Mathematical Statistics, pp. 400–407, 1951.

- Grant Rotskoff and Eric Vanden-Eijnden. Parameters as interacting particles: long time convergence and asymptotic error scaling of neural networks. Advances in Neural Information Processing Systems, 31, 2018.
- Bernhard Schölkopf and Alexander J Smola. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press, 2002.
- John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.
- John Stachurski. Economic dynamics: theory and computation. MIT Press, 2009.
- 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, 2018.
- Manel Tagorti and Bruno Scherrer. On the rate of convergence and error bounds for LSTD( $\lambda$ ). In International Conference on Machine Learning, pp. 1521–1529. PMLR, 2015.
- Terence Tao. Topics in Random Matrix Theory, volume 132. American Mathematical Soc., 2012.
- Valentin Thomas. On the role of overparameterization in off-policy temporal difference learning with linear function approximation. Advances in Neural Information Processing Systems, 35: 37228–37240, 2022.
- Mark Towers, Jordan K. Terry, Ariel Kwiatkowski, John U. Balis, Gianluca de Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Arjun K.G., Markus Krimmel, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Andrew Tan Jin Shen, and Omar G. Younis. Gymnasium, March 2023. URL <https://zenodo.org/record/8127025>.
- John Tsitsiklis and Benjamin Van Roy. Analysis of temporal-difference learning with function approximation. Advances in Neural Information Processing Systems, 9, 1996.
- Chenjun Xiao, Bo Dai, Jincheng Mei, Oscar A Ramirez, Ramki Gummadi, Chris Harris, and Dale Schuurmans. Understanding and leveraging overparameterization in recursive value estimation. In International Conference on Learning Representations, 2021.
- Roy D. Yates. A framework for uplink power control in cellular radio systems. IEEE Journal on Selected Areas in Communications, 13(7):1341–1347, 1995.
- Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning (still) requires rethinking generalization. Communications of the ACM, 64(3):107–115, 2021.
- Yufeng Zhang, Qi Cai, Zhuoran Yang, Yongxin Chen, and Zhaoran Wang. Can temporal-difference and Q-learning learn representation? A mean-field theory. Advances in Neural Information Processing Systems, 33:19680–19692, 2020.

## A ADDITIONAL EXPERIMENTS

This appendix shows additional empirical results which cannot be put in the main body due to space limitations.

### A.1 $\delta$ IN THE DOUBLE ASYMPTOTIC REGIME

Like Figure 2 in Section 6, Figure 6 depicts the correction factor  $\delta$  (equation 17) as a function of the complexity model  $N/m$  in synthetic ergodic and Gridworld MRPs.  $\delta$  shows a similar behavior than for the one observed in Taxi-v3 in Figure 2.

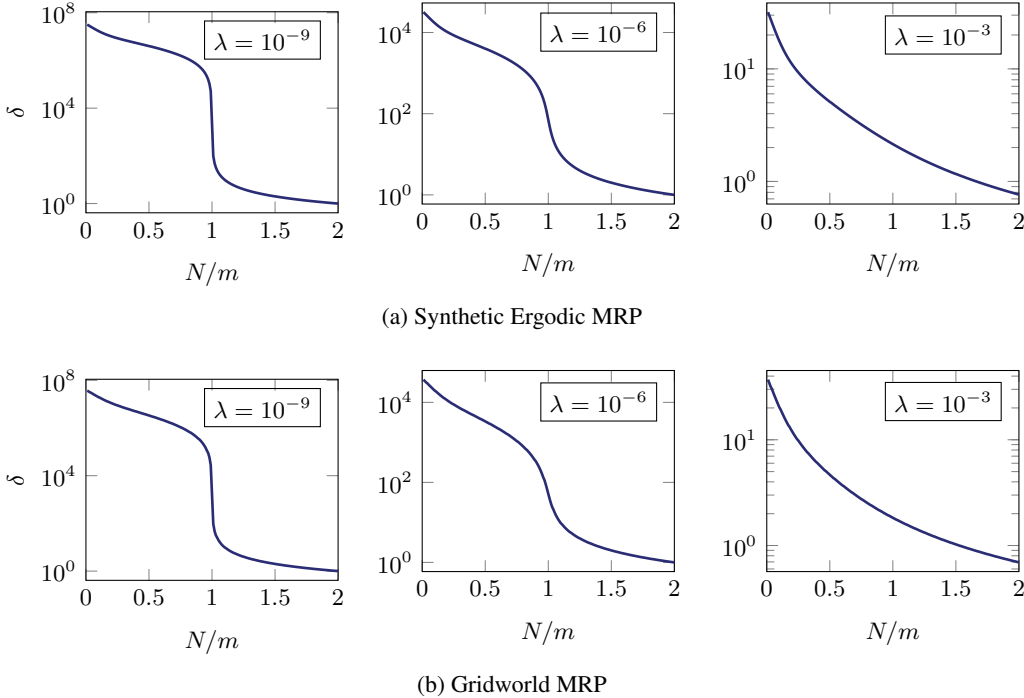


Figure 6: **The correction factor  $\delta$  is a decreasing function of the number of parameters  $N$ . For small  $l_2$ -regularization parameter  $\lambda$ , we observe a sharp decrease near the interpolation threshold ( $N = m$ , for  $m$  distinct visited states). As  $\lambda$  increases, the function becomes smoother and smaller (note the different scales of the y-axis).  $\delta$  is computed with equation 17 in synthetic ergodic and Gridworld MRPs with  $\gamma = 0.95, m = 499, n = 3000$  and  $\gamma = 0.95, m = 386, n = 5000$ , respectively.**

Figure 7 depicts  $\delta$  as a function of the  $l_2$ -regularization parameter for different ratio  $N/m$ . It confirms  $\delta$  decreases monotonically as the  $l_2$ -regularization parameter  $\lambda$  increases, as stated by Lemma J.4. Furthermore, we observe the impact of regularization parameter  $\lambda$  becomes less significant as the model complexity  $N/m$  increases. Indeed, as  $N/m$  increases, we observe a larger initially flat region and smaller values of  $\delta$ .

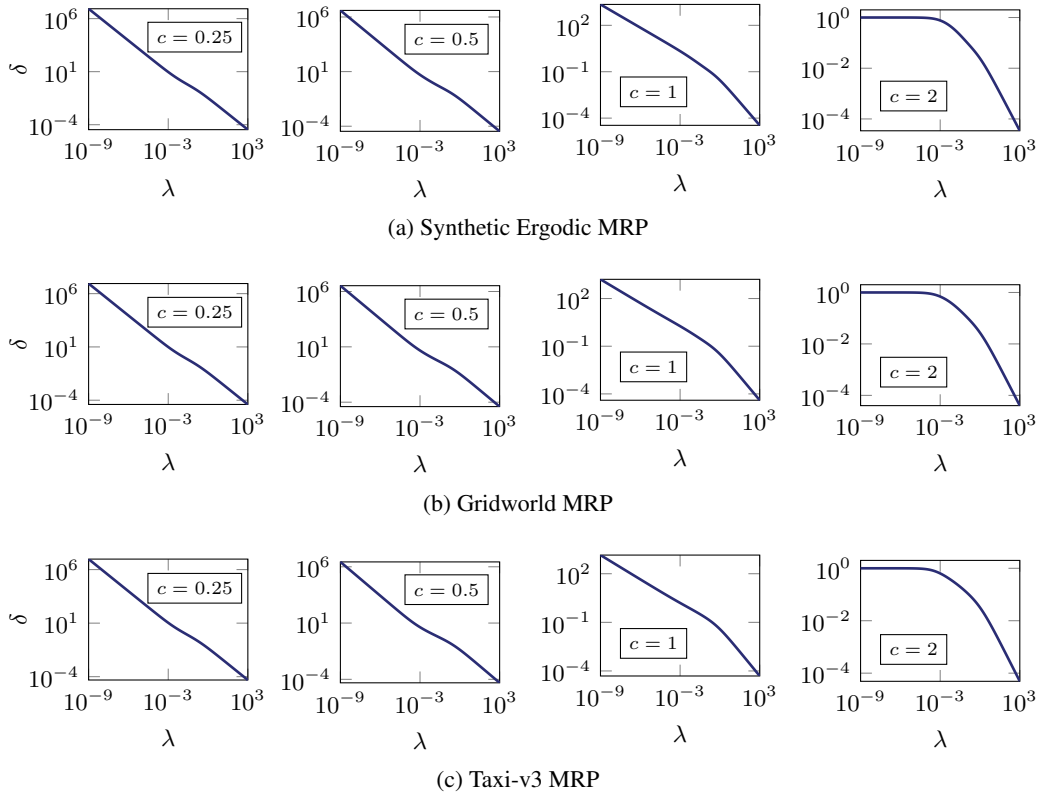


Figure 7: **The correction factor  $\delta$  is a decreasing function of the  $l_2$ -regularization parameter  $\lambda$ . As the model complexity  $c = N/m$  increases, the impact of regularization parameter  $\lambda$  becomes less significant (note the different scales of the y-axis).  $\delta$  is computed with equation 17 in synthetic ergodic, Gridworld and Taxi-v3 MRPs with  $\gamma = 0.95, m = 499, n = 3000$ ,  $\gamma = 0.95, m = 386, n = 5000$  and  $\gamma = 0.95, m = 310, n = 5000$ , respectively.**

## A.2 DOUBLE DESCENT BEHAVIOR

Figure 8 shows both  $\widehat{\text{MSBE}}$  and MSBE as a function of the model complexity  $N/m$  with different  $l_2$ -regularization penalties  $\lambda$ , in synthetic ergodic and Gridworld MRP. Both MSBE and  $\widehat{\text{MSBE}}$  depict a similar double descent behavior for small  $\lambda$  than in Figure 3 in Section 6. We observe the empirical  $\widehat{\text{MSBE}}$  is not necessarily lower for over-parameterized ( $N > m$ ) models.

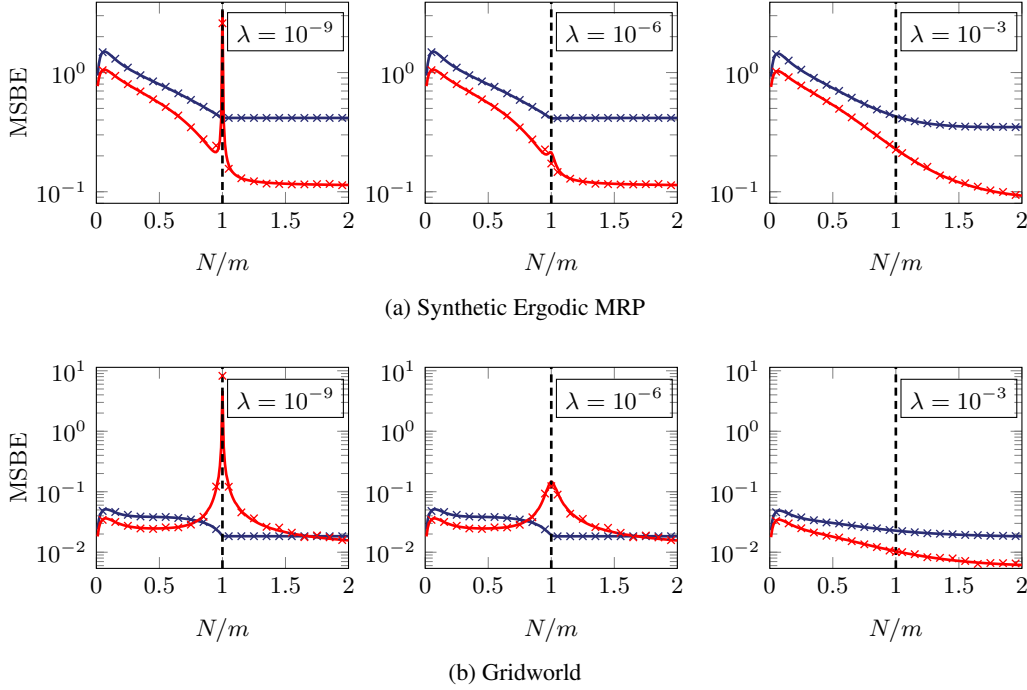
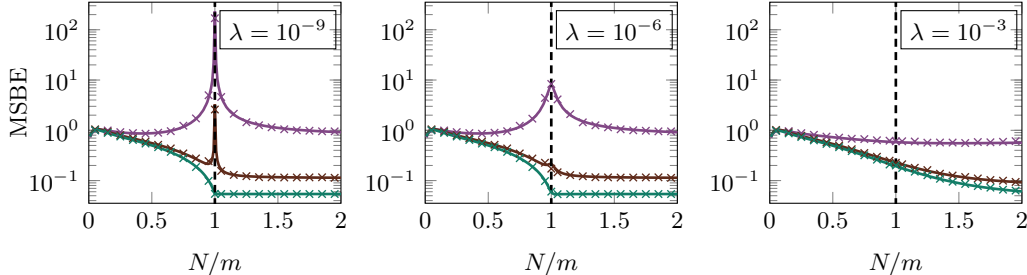


Figure 8: **The double descent phenomenon occurs in the true MSBE (red) of regularized LSTD, peaking around the interpolation threshold ( $N = m$  for  $N$  parameters,  $m$  distinct visited states) when the empirical  $\widehat{\text{MSBE}}$  (blue) vanishes. It diminishes as the  $l_2$ -regularization parameter  $\lambda$  increases.** Continuous lines indicate the theoretical values from Theorem 5.2 and Theorem 5.3, the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic and Gridworld MRP with  $\gamma = 0.95$ ,  $d = 50$ ,  $n = 3000$  and  $n = 5000$ , respectively.

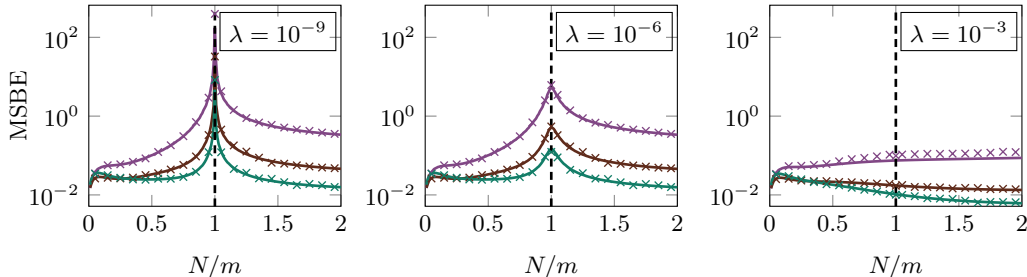


A.3 IMPACT OF THE NUMBER OF UNVISITED STATES AND OF THE DISCOUNT FACTOR  $\gamma$

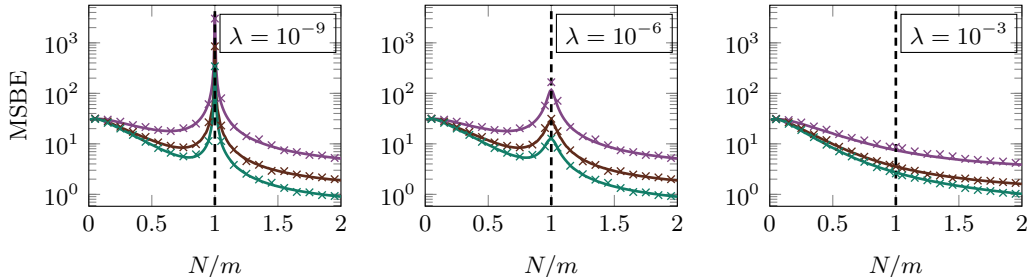
Like Figure 4, Figure 9 depicts the behavior of the true MSBE for different numbers of distinct visited states  $m$  and shows that as the number of distinct unvisited states goes to zero, the double descent phenomenon diminishes.



(a) Synthetic Ergodic MRP for  $m = 0.86|\mathcal{S}|$  (purple),  $m = 0.998|\mathcal{S}|$  (maroon),  $m = |\mathcal{S}|$  (green).



(b) Gridworld MRP for  $m = 0.59|\mathcal{S}|$  (purple),  $m = 0.92|\mathcal{S}|$  (maroon),  $m = 0.97|\mathcal{S}|$  (green).



(c) Taxi-v3 MRP for  $m = 0.57|\mathcal{S}|$  (purple),  $m = 0.79|\mathcal{S}|$  (maroon),  $m = 0.87|\mathcal{S}|$  (green).

Figure 9: **With more distinct states  $m$  visited, the double descent in the MSBE diminishes, disappearing for  $m = |\mathcal{S}|$ .** Continuous lines indicate the theoretical values of MSBE from Theorem 5.3 for different numbers of distinct visited states  $m$ ; the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic, Gridworld and Taxi-v3 MRP with  $\gamma = 0.95, d = 50$ .

Figure 10 describes the impact of the discount factor  $\gamma$  on the double descent phenomenon, and shows it remains true for all  $\gamma$ .

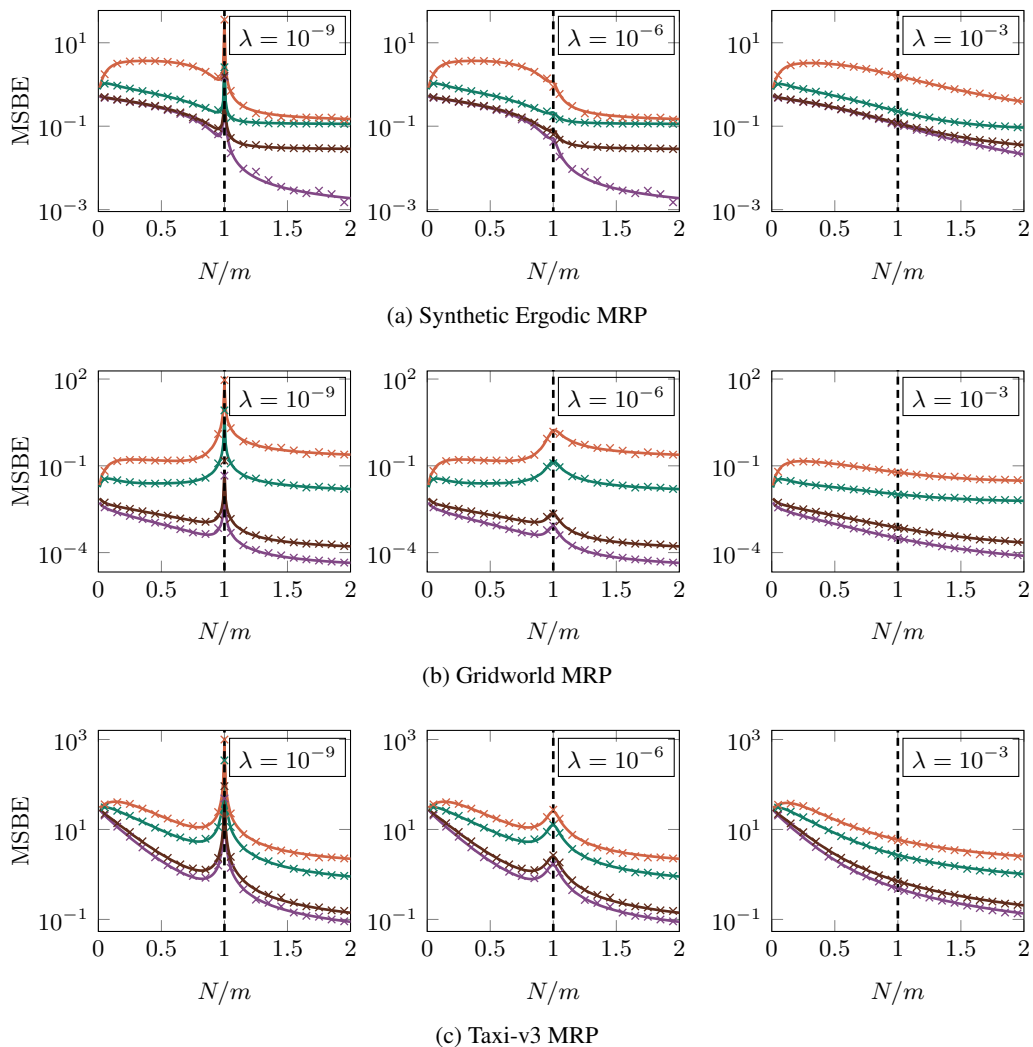


Figure 10: **The discount factor  $\gamma$  has little effect on the double descent in the MSBE.** Continuous lines indicate the theoretical values of MSBE from Theorem 5.3 for  $\gamma = 0$  (purple),  $\gamma = 0.5$  (maroon),  $\gamma = 0.95$  (green), and  $\gamma = 0.99$  (orange); the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic, Gridworld and Taxi-v3 MRPs for  $d = 50, n = 3000$ ,  $d = 50, n = 5000$  and  $d = 50, n = 5000$ , respectively.

A.4 IMPACT OF THE  $l_2$ -REGULARIZATION PARAMETER ON THE MSBE

Figure 11 and Figure 12 depict the empirical  $\widehat{\text{MSBE}}$  and the true MSBE as a function of the  $l_2$ -regularization parameter. In supervised learning, the training error is an increasing function of the  $l_2$ -regularization parameter  $\lambda$  (Liao et al., 2020), whereas the discount factor induces a more intricate behavior in RL as described by Figure 11.

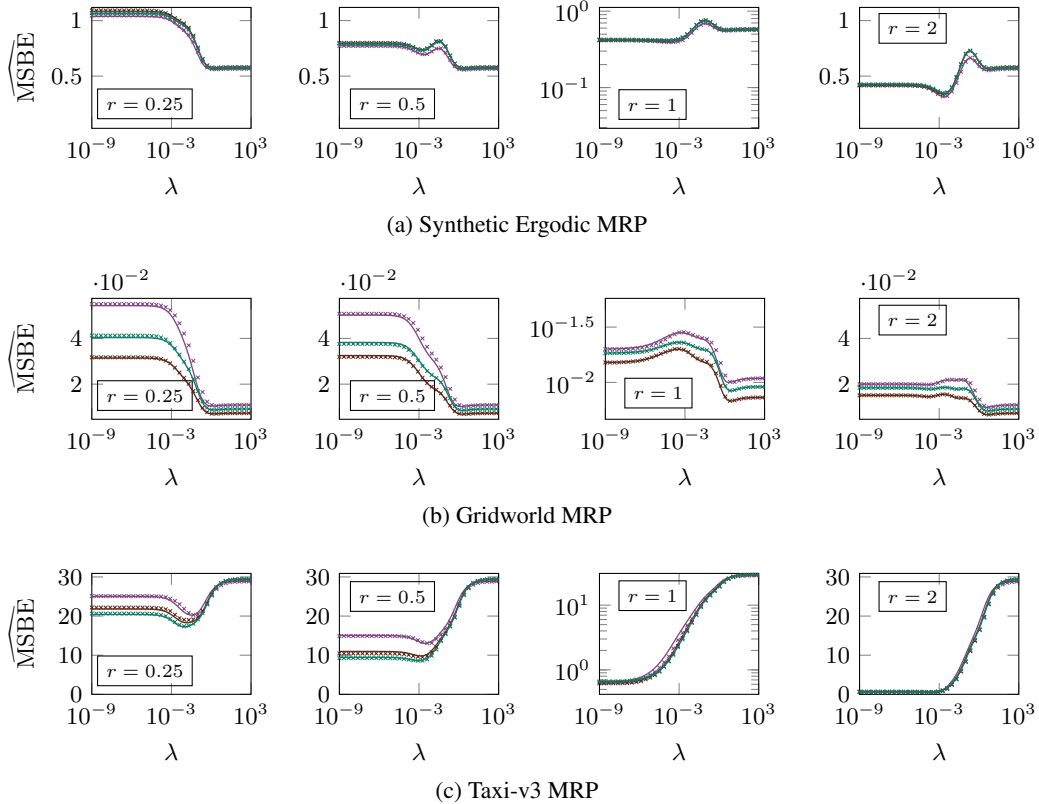


Figure 11:  $\widehat{\text{MSBE}}$  is not an increasing function of the  $l_2$ -regularization parameter  $\lambda$ . Continuous lines indicate the theoretical values of  $\widehat{\text{MSBE}}$  from Theorem 5.2 for  $n = 1000$  (purple),  $n = 3000$  (maroon), and  $n = 5000$  (green); the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic, Gridworld and Taxi-v3 MRPs for  $d = 50, n = 3000, d = 50, n = 5000$  and  $d = 50, n = 5000$ , respectively. Note that the y-axis has a logarithmic scale for  $r = N/m = 1$ .

In Figure 12 at the interpolation threshold ( $N = m$ ), we observe that as  $\lambda$  increases, MSBE decreases and so does the peak observed in the previous experiments. For other ratios, the true MSBE depicts complex behaviors that may differ between under- and over-parameterized models and depends on the environment. Yet, both empirical and true MSBE show similar and opposite trends for the same  $\lambda = \frac{\lambda_{m,n}}{mn}$ , regardless of the number of samples  $n$  and distinct visited states  $m$ . In practice, the effective  $l_2$ -regularization parameter  $\lambda_{m,n}$  is tuned. This suggests that it depends on both the number of transitions collected  $n$  and the number of distinct visited states  $m$ , and not just on the number of samples  $n$  as it is commonly suggested (Hoffman et al., 2011; Chen et al., 2013).

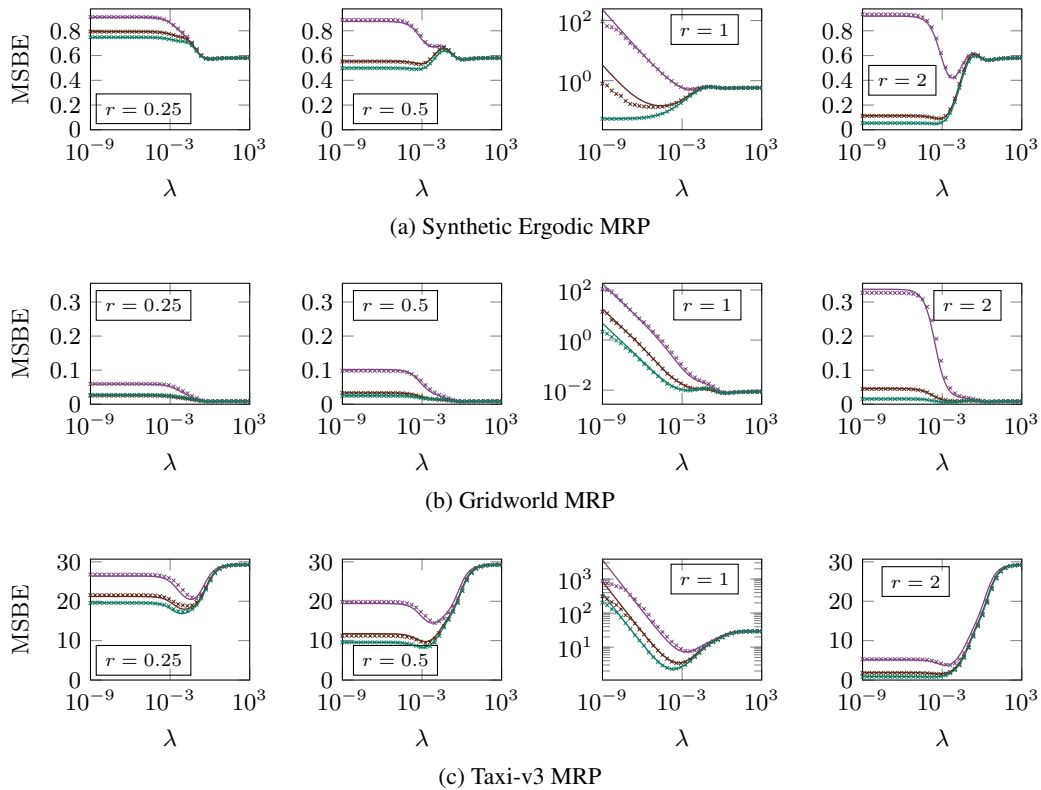


Figure 12: **At the interpolation threshold ( $r = N/m = 1$ ), as  $\lambda$  increases, MSBE decreases and so does the double descent phenomenon.** Continuous lines indicate the theoretical values of MSBE from Theorem 5.3 for  $n = 1000$  (purple),  $n = 3000$  (maroon), and  $n = 5000$  (green); the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic, Gridworld and Taxi-v3 MRPs for  $d = 50, n = 3000, d = 50, n = 5000$  and  $d = 50, n = 5000$ , respectively. Note that the y-axis has a logarithmic scale for  $r = N/m = 1$ .

### A.5 IMPACT OF THE SECOND-ORDER CORRECTION FACTOR $\Delta$ IN THE TRUE MSBE

Figure 13 depicts  $\Delta$  in Theorem 5.3 as a function of the model complexity  $N/m$  for small  $l_2$ -regularization parameter ( $\lambda = 10^{-9}$ ). It shows the double descent phenomenon in the true MSBE is mainly due to the second-order correction term  $\Delta$ .

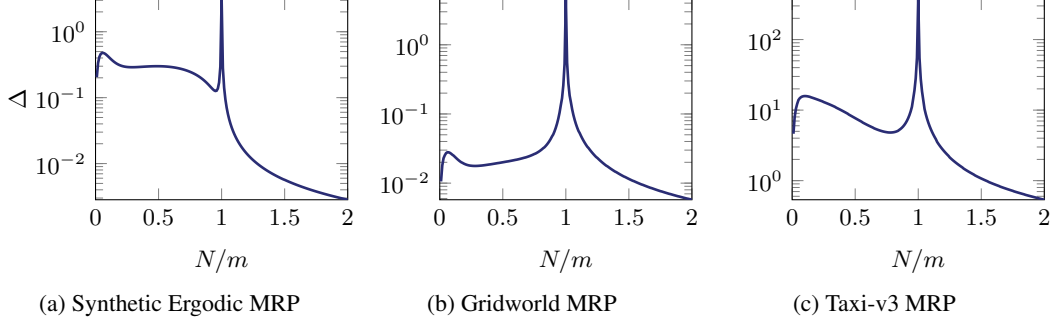


Figure 13: **The double descent phenomenon in the true MSBE is mainly due to the second-order correction term  $\Delta$ .**  $\Delta$  is computed with equation 20 in synthetic ergodic, Gridworld, and Taxi-v3 MRPs with  $\lambda = 10^{-9}, \gamma = 0.95, m = 499, n = 3000$ ,  $\lambda = 10^{-9}, \gamma = 0.95, m = 386, n = 5000$  and  $\lambda = 10^{-9}, \gamma = 0.95, m = 310, n = 5000$ , respectively.

## B MEAN-SQUARED VALUE ERROR

In this section, we study the Mean-Squared Value Error (MSVE) and observe a similar double descent behavior than the one observed for the true MSBE. The MSVE is defined as

$$\text{MSVE}(\hat{\theta}) = \|\mathbf{V} - \Sigma_S^T \hat{\theta}\|_{\mathcal{D}_\pi}^2. \quad (22)$$

Using a similar approach than for Theorem 5.3, we obtain the following deterministic form of the asymptotic MSVE:

**Corollary B.0.1** (Asymptotic MSVE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSVE is  $\overline{\text{MSVE}}(\hat{\theta}) = \left\| \mathbf{V} - \frac{1}{\sqrt{n}} \frac{N}{m} \frac{1}{1+\delta} \Phi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} \right\|_{\mathcal{D}_\pi}^2 + \Delta'$ , with second-order correction factor*

$$\Delta' = \frac{\frac{1}{n} \text{Tr}(\mathcal{D}_\pi [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m(\lambda)^T \Psi_1 \bar{\mathbf{Q}}_m(\lambda))} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|_{\Psi_1}^2. \quad (23)$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,  $\text{MSVE}(\hat{\theta}) - \overline{\text{MSVE}}(\hat{\theta}) \xrightarrow{a.s.} 0$ .

*Proof.* Using  $\mathcal{D}_\pi = [\mathbf{I}_m - \gamma \mathbf{P}]^T \mathcal{D}_\pi \mathcal{D}_\pi^{-1} [\mathbf{I}_m - \gamma \mathbf{P}]^{-1T} \mathcal{D}_\pi [\mathbf{I}_m - \gamma \mathbf{P}]^{-1} \mathcal{D}_\pi^{-1} \mathcal{D}_\pi [\mathbf{I}_m - \gamma \mathbf{P}]$  and a with similar proof than for Theorem 5.3, we find Corollary B.0.1.  $\square$

**Remark 8.** Like  $\overline{\text{MSBE}}(\hat{\theta})$  in Theorem 5.3, the  $\overline{\text{MSVE}}(\hat{\theta})$  is also influenced by the correction terms  $\delta$  and  $\Delta'$ . Note that in asymptotic regimes where  $N/m \rightarrow \infty$  or  $\lambda \rightarrow \infty$ , the correction terms vanish. When  $N/m \rightarrow \infty$ ,  $\overline{\text{MSVE}}(\hat{\theta})$  is independent of  $\lambda$  as shown in details in Appendix C.

Figure 14 shows both the empirical  $\overline{\text{MSVE}}$  and the true MSVE as a function of the model complexity  $N/m$  with different  $l_2$ -regularization penalties  $\lambda$  in the synthetic ergodic, Gridworld and Taxi MRPs. Like the true MSBE, we observe an almost perfect match with the numerically evaluated original definition in equation 22.

For small  $\lambda$ , like for the true MSBE, the true MSVE exhibits a peak around the interpolation threshold  $N = m$ , leading to a double descent phenomenon. In contrast, the empirical  $\overline{\text{MSVE}}$  is close to its minimum at  $N = m$  and almost constant for  $N \geq m$ , so no double descent is observed. Unlike than for the true MSBE, we observe that the empirical  $\overline{\text{MSVE}}$  is always smaller than the true

MSVE. For larger  $\lambda$ , the double descent in the true MSVE disappears and the difference between the true MSVE and the empirical MSVE is less pronounced, although it may not vanish.

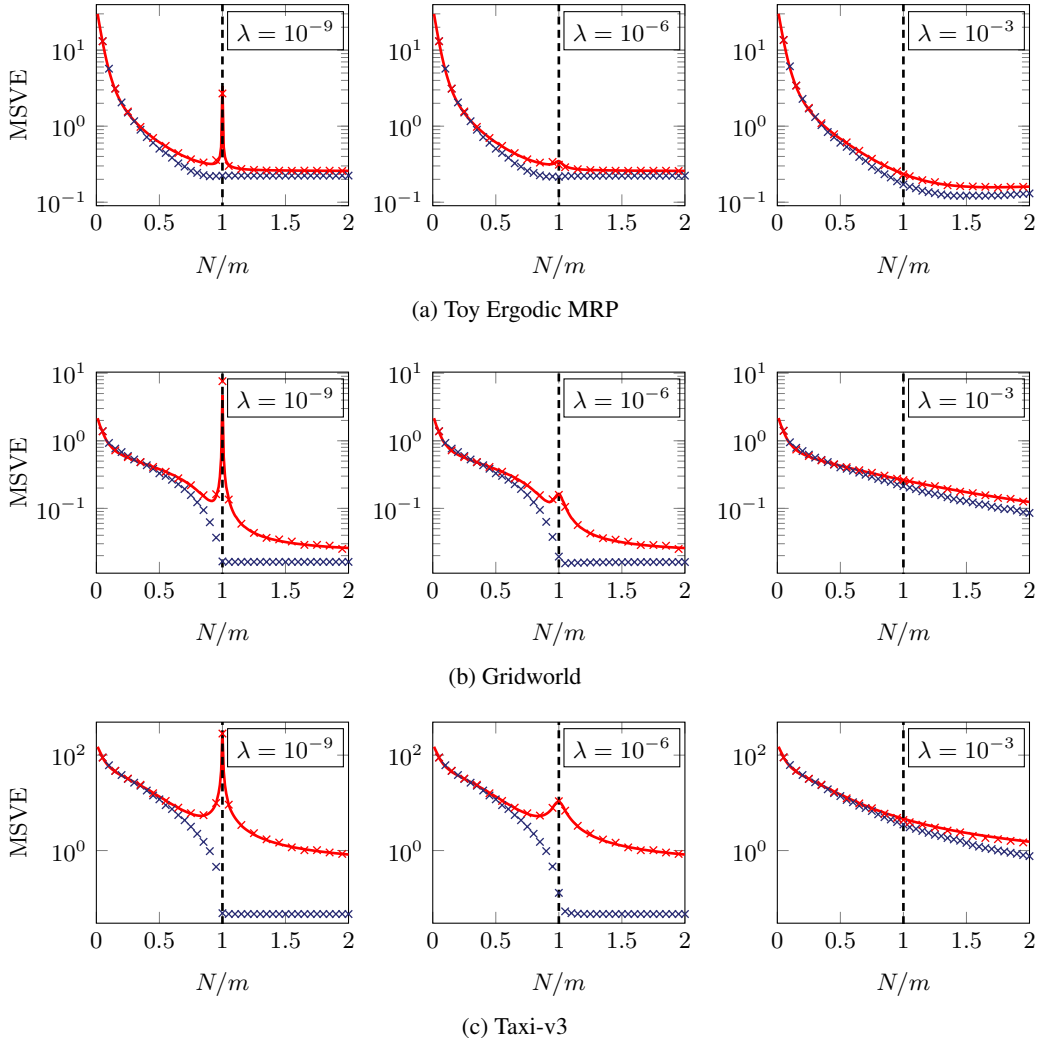


Figure 14: **The double descent phenomenon occurs in the true MSVE (red) of regularized LSTD, peaking around the interpolation threshold ( $N = m$  for  $N$  parameters,  $m$  distinct visited states) when the empirical MSVE (blue) vanishes. It diminishes as the  $l_2$ -regularization parameter  $\lambda$  increases.** Continuous lines indicate the theoretical values from Corollary B.0.1, the crosses are numerical results averaged over 30 instances after the learning with regularized LSTD in synthetic ergodic, Gridworld and Taxi MRP with  $\gamma = 0.95$ ,  $d = 50$ ,  $n = 3000$ ,  $n = 5000$  and  $n = 5000$ , respectively.

## C REFORMULATION OF THE MAIN RESULTS

Let  $\mathcal{C}$  be a compact set such that  $\mathcal{S} \subset \mathcal{C}$ . We extend the stationary distribution  $\pi$  on  $\mathcal{C}$  by setting  $\pi(s) = 0$  for  $s \in \mathcal{S} \setminus \mathcal{C}$ . We denote by  $L^2(\mathcal{C}, \pi)$  the set of squared integrable functions  $f : \mathcal{C} \rightarrow \mathbb{R}$  with respect to the distribution  $\pi$  on  $\mathcal{C}$ , and the norm on  $L^2(\mathcal{C}, \pi)$  defined as  $\|f\|_{L^2(\mathcal{C}, \pi)}^2 = \langle f, f \rangle = \int f(x)^2 \pi(dx)$ . The Frobenius norm of a matrix  $\mathbf{A}$  is denoted as  $\|\mathbf{A}\|_F = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle_F} = \sqrt{\text{Tr}(\mathbf{A}^T \mathbf{A})}$ .

In this section, we aim to reformulate the results from Section 5 in a feature space derived from the eigendecomposition of  $\Phi_S$ . The section is organized into three subsections: in Section C.1 we introduce the new asymptotic feature space, we reformulate the results of Section 5 in the new feature space in Section C.2 and we provide the proofs for these results in Section C.3.

### C.1 ASYMPTOTIC FEATURE SPACE

**Asymptotic Feature Space.** Both Theorem 5.2 and Theorem 5.3 use either the deterministic Gram feature matrix  $\Phi_{\hat{S}}$  (defined in equation 16) or the deterministic Gram feature matrix  $\Phi_S$  that are given by the continuous kernel function  $\Phi : \mathcal{C} \times \mathcal{C} \rightarrow \mathbb{R}$  defined as

$$\Phi(\mathbf{s}, \mathbf{s}') = \mathbb{E}_{\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)} [\sigma(\mathbf{w}^T \mathbf{s})^T \sigma(\mathbf{w}^T \mathbf{s}')].$$

Since  $\Phi$  is continuous and  $\mathcal{C}$  is compact, the Mercer's theorem states (Schölkopf & Smola, 2002) that

$$\Phi(\mathbf{s}, \mathbf{s}') = \sum_{i=1}^M \nu_i \varphi_i(\mathbf{s}) \varphi_i(\mathbf{s}') = \sum_{i=1}^M \omega_i(\mathbf{s}) \omega_i(\mathbf{s}'); \quad (24)$$

where  $\{\nu_i\}_{i=1}^M$  and  $\{\varphi_i\}_{i=1}^M$  are the eigenvalues and eigenfunctions of the Hilbert-Schmidt integral operators  $T_\Phi : L^2(\mathcal{C}, \pi) \rightarrow L^2(\mathcal{C}, \pi), f \mapsto T_\Phi(f)(\mathbf{s}') = \int_{\mathbb{R}^d} \Phi(\mathbf{s}, \mathbf{s}') f(\mathbf{s}) \pi(\mathbf{s}) d\mathbf{s}$  and  $\{\omega_i(\cdot)\}_{i=1}^M = \sqrt{\nu_i} \varphi_i(\cdot)$  are the rescaled eigenfunction.  $\{\omega_i(\cdot)\}_{i=1}^M$  forms an orthogonal basis in  $L^2(\mathcal{C}, \pi)$ . Usually,  $M$  is infinite. In the following of this section, we will find convenient to define a vector representation of functions in the asymptotic feature space defined by the feature map  $\{\omega_i\}_{i=1}^M$ . For any state matrix  $\mathbf{A} \in \mathbb{R}^{d \times p}$ , we denote by  $\Omega_{\mathbf{A}} \in \mathbb{R}^{M \times p}$  the feature matrices of  $\mathbf{A}$  so that  $[\Omega_{\mathbf{A}}]_{ij} = \omega_i(\mathbf{A}_j)$  for  $\mathbf{A}_j$  the  $j^{\text{th}}$  column of  $\mathbf{A}$ . With those new notations, we can decompose  $\Phi_{\hat{S}}$  and  $\Phi_S$  as

$$\Phi_{\hat{S}} = \Omega_{\hat{S}}^T \Omega_{\hat{S}} \quad \text{and} \quad \Phi_S = \Omega_S^T \Omega_S.$$

**Regularized LSTD in the Asymptotic Feature Space.** Let  $\mathbf{v}_{\bar{\theta}} \in \mathbb{R}^M$  the weight vector returned by the regularized LSTD with the rescaled  $l_2$ -regularization parameter  $\lambda \frac{m(1+\delta)}{N}$  for the asymptotic features  $\{\omega_i\}_{i=1}^M$  on transitions collected in  $\mathcal{D}_{\text{train}}$ :

$$\mathbf{v}_{\bar{\theta}} = \frac{1}{\sqrt{n}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \Omega_{\hat{S}}^T \hat{\mathbf{U}}_n \mathbf{r},$$

where

$$\bar{\mathbf{A}} = \Omega_{\hat{S}} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Omega_{\hat{S}}^T. \quad (25)$$

**Second-Order Correction Factor in the Asymptotic Feature Space.** Let  $f_{\Delta}(\cdot)$  be the second-order correction function defined as

$$f_{\Delta}(\mathbf{B}) = \frac{\lambda^2 \frac{m^2(1+\delta)^2}{N^2} \frac{\frac{1}{N} \|\mathbf{v}_{\bar{\theta}}\|^2}{1 - \frac{1}{N} \left\| \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \bar{\mathbf{A}} \right\|_F^2}}{\left\| \mathbf{B} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\|_F^2}.$$

From the definition of  $f_{\Delta}(\mathbf{B})$ , it directly follows that  $f_{\Delta}(\mathbf{B}) \rightarrow 0$  as  $N/m \rightarrow \infty$  or as  $\lambda \rightarrow \infty$ .

### C.2 REFORMULATION OF THE MAIN RESULTS

#### C.2.1 EMPIRICAL MEAN-SQUARED BELLMAN ERROR

**Theorem C.1** (Asymptotic Empirical MSBE). *Under the conditions of Theorem 5.1, the deterministic asymptotic empirical MSBE is*

$$\widehat{\text{MSBE}}(\hat{\boldsymbol{\theta}}) = \frac{1}{n} \left\| \mathbf{r} + \gamma \Omega_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}} - \Omega_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}} \right\|^2 + f_{\Delta}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Omega_{\hat{S}}^T).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\widehat{\text{MSBE}}(\hat{\boldsymbol{\theta}}) - \overline{\widehat{\text{MSBE}}}(\hat{\boldsymbol{\theta}}) \xrightarrow{a.s.} 0.$$

**Remark 9.** The left-hand term  $\frac{1}{n} \|\mathbf{r} + \gamma \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}} - \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}}\|^2$  in  $\overline{\text{MSBE}}(\hat{\boldsymbol{\theta}})$  depicts the empirical MSBE for the regularized LSTD in the asymptotic features  $\{\omega_i\}_{i=1}^M$  with the rescaled  $l_2$ -regularization parameter  $\lambda \frac{m(1+\delta)}{N}$  on transitions collected in  $\mathcal{D}_{\text{train}}$ .

**Remark 10.** As  $N/m$  increases, the influence of the  $l_2$ -regularization parameter  $\lambda$  decreases.

**Remark 11.** As  $N/m \rightarrow \infty$ ,  $\widehat{\overline{\text{MSBE}}}(\hat{\boldsymbol{\theta}}) \rightarrow 0$ .

### C.2.2 MEAN-SQUARED BELLMAN ERROR

**Theorem C.2** (Asymptotic MSBE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSBE is*

$$\overline{\text{MSBE}}(\hat{\boldsymbol{\theta}}) = \|\mathbf{r} + \gamma \mathbf{P} \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}} - \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}}\|_{\mathcal{D}_{\pi}}^2 + f_{\Delta}(\mathbf{D}_{\pi}^{\frac{1}{2}}(\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}) \boldsymbol{\Omega}_{\mathcal{S}}^T).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\text{MSBE}(\hat{\boldsymbol{\theta}}) - \overline{\text{MSBE}}(\hat{\boldsymbol{\theta}}) \xrightarrow{a.s.} 0.$$

**Remark 12.** The left-hand term  $\|\mathbf{r} + \gamma \mathbf{P} \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}} - \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}}\|_{\mathcal{D}_{\pi}}^2$  in  $\overline{\text{MSBE}}(\hat{\boldsymbol{\theta}})$  depicts the true MSBE for the regularized LSTD in the asymptotic features  $\{\omega_i\}_{i=1}^M$  with the rescaled  $l_2$ -regularization parameter  $\lambda \frac{m(1+\delta)}{N}$  on transitions collected in  $\mathcal{D}_{\text{train}}$ .

**Remark 13.** As  $N/m$  increases, the influence of the  $l_2$ -regularization parameter  $\lambda$  decreases.

**Remark 14.** As  $N/m \rightarrow \infty$ ,  $\overline{\text{MSBE}}(\hat{\boldsymbol{\theta}})$  converges to the MSBE of  $\mathbf{v}_{\bar{\theta}}$  without the  $l_2$  regularization parameter.

### C.2.3 MEAN-SQUARED VALUE ERROR

**Corollary C.2.1** (Asymptotic MSVE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSVE is*

$$\overline{\text{MSVE}}(\hat{\boldsymbol{\theta}}) = \|\mathbf{V} - \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}}\|_{\mathcal{D}_{\pi}}^2 + f_{\Delta}(\mathbf{D}_{\pi}^{\frac{1}{2}} \boldsymbol{\Omega}_{\mathcal{S}}^T).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\text{MSVE}(\hat{\boldsymbol{\theta}}) - \overline{\text{MSVE}}(\hat{\boldsymbol{\theta}}) \xrightarrow{a.s.} 0.$$

**Remark 15.** The left-hand term  $\|\mathbf{V} - \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\bar{\theta}}\|_{\mathcal{D}_{\pi}}^2$  in  $\overline{\text{MSVE}}(\hat{\boldsymbol{\theta}})$  depicts the MSVE for the regularized LSTD in the asymptotic features  $\{\omega_i\}_{i=1}^M$  with the rescaled  $l_2$ -regularization parameter  $\lambda \frac{m(1+\delta)}{N}$  on transitions collected in  $\mathcal{D}_{\text{train}}$ .

**Remark 16.** As  $N/m$  increases, the influence of the  $l_2$ -regularization parameter  $\lambda$  decreases.

**Remark 17.** As  $N/m \rightarrow \infty$ ,  $\overline{\text{MSVE}}(\hat{\boldsymbol{\theta}})$  converges to the MSVE of  $\mathbf{v}_{\bar{\theta}}$  without the  $l_2$  regularization parameter.

## C.3 PROOFS OF RESULTS FOUND IN SECTION C.2

**Theorem C.3** (Asymptotic Empirical MSBE). *Under the conditions of Theorem 5.1, the deterministic asymptotic empirical MSBE is*

$$\widehat{\overline{\text{MSBE}}}(\hat{\boldsymbol{\theta}}) = \frac{1}{n} \|\mathbf{r} + \gamma \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}} - \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}}\|^2 + f_{\Delta}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Omega}_{\mathcal{S}}^T).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\widehat{\text{MSBE}}(\hat{\boldsymbol{\theta}}) - \widehat{\overline{\text{MSBE}}}(\hat{\boldsymbol{\theta}}) \xrightarrow{a.s.} 0.$$

*Proof.* We have

$$\begin{aligned} \frac{\lambda^2}{n} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|^2 &= \frac{1}{n} \left\| \mathbf{r} + \gamma \frac{N}{m} \frac{1}{1+\delta} \hat{\mathbf{V}}_n^T \boldsymbol{\Omega}_{\mathcal{S}}^T \boldsymbol{\Omega}_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} - \frac{N}{m} \frac{1}{1+\delta} \hat{\mathbf{U}}_n \boldsymbol{\Omega}_{\mathcal{S}}^T \boldsymbol{\Omega}_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} \right\|^2 \\ &= \frac{1}{n} \|\mathbf{r} + \gamma \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}} - \boldsymbol{\Omega}_{\mathbf{X}}^T \mathbf{v}_{\bar{\theta}}\|^2. \end{aligned}$$



For  $\hat{\Delta}$ , we have

$$\begin{aligned}
& \text{Tr} \left( \bar{\mathbf{Q}}_m(\lambda) \boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m(\lambda)^T \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \text{Tr} \left( \bar{\mathbf{Q}}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_{\mathcal{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m(\lambda)^T \right) \\
&= \frac{m(1+\delta)}{N} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Omega}_{\mathcal{S}}^T \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1T} \boldsymbol{\Omega}_{\mathcal{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \right) \\
&= \frac{m(1+\delta)}{N} \left\| (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Omega}_{\mathcal{S}}^T \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\|_F^2 \\
& \text{Tr} \left( \boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m(\lambda)^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m(\lambda) \right) \\
&= \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Omega}_{\mathcal{S}}^T \boldsymbol{\Omega}_{\mathcal{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m(\lambda)^T \hat{\mathbf{U}}_n^T \boldsymbol{\Omega}_{\mathcal{S}}^T \boldsymbol{\Omega}_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\lambda) \right) \\
&= \text{Tr} \left( \bar{\mathbf{A}}^T \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1T} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \bar{\mathbf{A}} \right) \\
&= \left\| \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \bar{\mathbf{A}} \right\|_F^2,
\end{aligned}$$

and

$$\| \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} \|_{\boldsymbol{\Psi}_1}^2 = \frac{m(1+\delta)}{n} \| \mathbf{v}_{\boldsymbol{\theta}} \|^2$$

□

**Theorem C.4** (Asymptotic MSBE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSBE is*

$$\overline{\text{MSBE}}(\hat{\boldsymbol{\theta}}) = \| \mathbf{r} + \gamma \mathbf{P} \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\boldsymbol{\theta}} - \boldsymbol{\Omega}_{\mathcal{S}}^T \mathbf{v}_{\boldsymbol{\theta}} \|_{\mathbf{D}_{\boldsymbol{\pi}}}^2 + f_{\Delta} \left( \mathbf{D}_{\boldsymbol{\pi}}^{\frac{1}{2}} (\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}) \boldsymbol{\Omega}_{\mathcal{S}}^T \right).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\text{MSBE}(\hat{\boldsymbol{\theta}}) - \overline{\text{MSBE}}(\hat{\boldsymbol{\theta}}) \xrightarrow{a.s.} 0.$$

*Proof.* We can rewrite

$$\begin{aligned}
& \text{Tr} \left( \boldsymbol{\Lambda}_{\mathbf{P}} \boldsymbol{\Theta}_{\mathcal{S}} \boldsymbol{\Psi}_2 \boldsymbol{\Theta}_{\mathcal{S}}^T \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \text{Tr} \left( \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1T} \bar{\mathbf{A}}^T \boldsymbol{\Omega}_{\mathcal{S}} [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}]^T \mathbf{D}_{\boldsymbol{\pi}} [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \left\| [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\|_{F, \mathbf{D}_{\boldsymbol{\pi}}}^2, \\
& \frac{N}{m} \frac{1}{1+\delta} \text{Tr} \left( \boldsymbol{\Lambda}_{\mathbf{P}} \boldsymbol{\Phi}_{\mathcal{S}} \right) = \frac{N}{m} \frac{1}{1+\delta} \left\| [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \right\|_{F, \mathbf{D}_{\boldsymbol{\pi}}}^2,
\end{aligned}$$

and

$$\begin{aligned}
\text{Tr} \left( \boldsymbol{\Lambda}_{\mathbf{P}} \boldsymbol{\Theta}_{\mathcal{S}} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_{\mathcal{S}} \right) &= \frac{N}{m} \frac{1}{1+\delta} \text{Tr} \left( \boldsymbol{\Lambda}_{\mathbf{P}} \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \boldsymbol{\Omega}_{\mathcal{S}} \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \text{Tr} \left( \boldsymbol{\Omega}_{\mathcal{S}} [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}]^T \mathbf{D}_{\boldsymbol{\pi}} [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \left\langle [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T, [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\rangle_{F, \mathbf{D}_{\boldsymbol{\pi}}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& \text{Tr} \left( \boldsymbol{\Lambda}_{\mathbf{P}} \left[ \boldsymbol{\Theta}_{\mathcal{S}} \boldsymbol{\Psi}_2 \boldsymbol{\Theta}_{\mathcal{S}}^T - 2 \boldsymbol{\Theta}_{\mathcal{S}} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_{\mathcal{S}} + \boldsymbol{\Psi}_{\mathcal{S}} \right] \right) \\
&= \frac{N}{m} \frac{1}{1+\delta} \left\| [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T - [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \bar{\mathbf{A}} \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\|_{F, \mathbf{D}_{\boldsymbol{\pi}}}^2 \\
&= \lambda^2 \left\| [\mathbf{I}_{|\mathcal{S}|} - \gamma \mathbf{P}] \boldsymbol{\Omega}_{\mathcal{S}}^T \left[ \bar{\mathbf{A}} + \lambda \frac{m(1+\delta)}{N} \mathbf{I}_M \right]^{-1} \right\|_{F, \mathbf{D}_{\boldsymbol{\pi}}}^2.
\end{aligned}$$

□

## D EVALUATION OF $\Phi_S$ OR $\Phi_{\hat{S}}$

The evaluation of  $\Phi_{\hat{S}} = \mathbb{E}_{\mathbf{w}}[\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \sigma(\mathbf{w}^T \hat{\mathbf{S}})]$  or  $\Phi_S = \mathbb{E}_{\mathbf{w}}[\sigma(\mathbf{w}^T \mathbf{S})^T \sigma(\mathbf{w}^T \mathbf{S})]$  naturally boils down to the evaluation of its individual entries and thus to the calculus, for arbitrary vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$  of

$$\Phi_{ab} = \mathbb{E}[\sigma(\mathbf{w}^T \mathbf{a}) \sigma(\mathbf{w}^T \mathbf{b})] = (2\pi)^{-\frac{p}{2}} \int \sigma(\varphi(\tilde{\mathbf{w}})^T \mathbf{a}) \sigma(\varphi(\tilde{\mathbf{w}})^T \mathbf{b}) e^{-\frac{1}{2} \|\tilde{\mathbf{w}}\|^2} d\tilde{\mathbf{w}}. \quad (26)$$

The evaluation of equation 26 can be obtained through various integration tricks for a wide family of mappings  $\varphi(\cdot)$  and activation functions  $\sigma(\cdot)$ . We provide in Table 1 (found in Louart et al. (2018)) the values of  $\Phi_{ab}$  for  $\mathbf{w} \sim \mathcal{N}(0, \mathbf{I}_d)$  (i.e., for  $\varphi(t) = t$ ) and for a set of activation functions  $\sigma(\cdot)$  not necessarily satisfying the Lipschitz continuity. In experiments in Section 6, we focus only on the ReLU function, i.e.,  $\sigma(t) = \max(t, 0)$ .

Table 1: Values of  $\Phi_{ab}$  for  $\mathbf{w} \sim \mathcal{N}(0, \mathbf{I}_d)$ ,  $\angle(\mathbf{a}, \mathbf{b}) \equiv \frac{\mathbf{a}^T \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|}$  (Louart et al., 2018).

$\sigma(t)$	$\Phi_{ab}$
$t$	$\mathbf{a}^T \mathbf{b}$
$\max(t, 0)$	$\frac{1}{2\pi} \ \mathbf{a}\  \ \mathbf{b}\  \left( \angle(\mathbf{a}, \mathbf{b}) \arccos(-\angle(\mathbf{a}, \mathbf{b})) + \sqrt{1 - \angle(\mathbf{a}, \mathbf{b})^2} \right)$
$ t $	$\frac{2}{\pi} \ \mathbf{a}\  \ \mathbf{b}\  \left( \angle(\mathbf{a}, \mathbf{b}) \arcsin(\angle(\mathbf{a}, \mathbf{b})) + \sqrt{1 - \angle(\mathbf{a}, \mathbf{b})^2} \right)$
$\text{erf}(t)$	$\frac{2}{\pi} \arcsin \left( \frac{2\mathbf{a}^T \mathbf{b}}{\sqrt{(1+2\ \mathbf{a}\ ^2)(1+2\ \mathbf{b}\ ^2)}} \right)$
$1_{\{t>0\}}$	$\frac{1}{2} - \frac{1}{2\pi} \arccos(\angle(\mathbf{a}, \mathbf{b}))$
$\text{sign}(t)$	$\frac{2}{\pi} \arcsin(\angle(\mathbf{a}, \mathbf{b}))$
$\cos(t)$	$\exp(-\frac{1}{2}(\ \mathbf{a}\ ^2 + \ \mathbf{b}\ ^2)) \cosh(\mathbf{a}^T \mathbf{b})$
$\sin(t)$	$\exp(-\frac{1}{2}(\ \mathbf{a}\ ^2 + \ \mathbf{b}\ ^2)) \sinh(\mathbf{a}^T \mathbf{b})$

## E PROOF OF THEOREM 5.1

Under Assumptions 1 and 2, this section is dedicated to prove the asymptotic equivalence between  $\mathbb{E}[\mathbf{Q}_m(\lambda)]$  and

$$\bar{\mathbf{Q}}_m(\lambda) = \left[ \frac{N}{m} \frac{1}{1+\delta} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}$$

defined in Theorem 5.1, when  $N, m \rightarrow \infty$ . In order to prove Theorem 5.1, we shall proceed by introducing an intermediary resolvent  $\tilde{\mathbf{Q}}_m(\lambda)$  (defined in equation 27), and show subsequently under Assumptions 1 and 2 that

$$\|\mathbb{E}[\mathbf{Q}_m(\lambda)] - \tilde{\mathbf{Q}}_m(\lambda)\| \rightarrow 0 \quad \text{and} \quad \|\tilde{\mathbf{Q}}_m(\lambda) - \bar{\mathbf{Q}}_m(\lambda)\| \rightarrow 0,$$

as  $N, m \rightarrow \infty$ .

In order to simplify the notations, we denote by  $\mathbf{Q}_m$  the resolvent  $\mathbf{Q}_m(\lambda)$ . The first half of the proof is dedicated to Lemma E.1, which proposes a first characterization of  $\mathbb{E}[\mathbf{Q}_m]$  by  $\tilde{\mathbf{Q}}_m$  as  $N, m \rightarrow \infty$  under Assumptions 1 and 2. This preliminary step is classical in studying resolvents in Random Matrix literature (Louart et al., 2018; Liao et al., 2020) as the direct comparison of  $\mathbb{E}[\mathbf{Q}_m]$  to  $\bar{\mathbf{Q}}_m$  with the implicit  $\delta$  (equation 17) may be cumbersome.

**Lemma E.1.** *Under Assumptions 1 and 2, let  $\lambda > 0$  and let  $\tilde{\mathbf{Q}}_m(\lambda) \in \mathbb{R}^{n \times n}$  be the resolvent defined as*

$$\tilde{\mathbf{Q}}_m(\lambda) = \left[ \frac{N}{m} \frac{1}{1+\alpha} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}; \quad (27)$$

for the deterministic Gram feature matrix

$$\Phi_{\mathcal{S}} = \mathbb{E}_{\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)} [\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \sigma(\mathbf{w}^T \hat{\mathbf{S}})],$$

and

$$\alpha = \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-(\lambda)]), \quad (28)$$

where

$$\mathbf{Q}_-(\lambda) = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{W}_- \hat{\mathbf{S}})^T \sigma(\mathbf{W}_- \hat{\mathbf{S}}) \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}, \quad (29)$$

for which  $\mathbf{W}_- \in \mathbb{R}^{(N-1) \times d}$  depicts the submatrix of the weight matrix  $\mathbf{W}$  (defined in equation 5) without the first row. Then,

$$\lim_{m \rightarrow \infty} \|\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m(\lambda)] - \tilde{\mathbf{Q}}_m(\lambda)\| = 0.$$

**Remark 18.** *Firstly, we can note that  $\alpha$  is uniformly bounded. Since  $\frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}}) = \mathbb{E} \left[ \frac{1}{m} \|\sigma(\mathbf{w}^T \hat{\mathbf{S}})\|^2 \right]$  and from Lemma K.2, we have*

$$\frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}}) = \int_0^\infty \Pr \left( \frac{1}{m} \|\sigma(\mathbf{w}^T \hat{\mathbf{S}})\|^2 > t \right) dt = \int_0^\infty 2t \Pr \left( \frac{1}{m} \|\sigma(\mathbf{w}^T \hat{\mathbf{S}})\| > t \right) dt = \mathcal{O}(1). \quad (30)$$

We deduce that

$$\alpha = \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-]) \leq \|\hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-]\| (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}}) = \mathcal{O}(1), \quad (31)$$

where we used  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{A}\| \text{Tr}(\mathbf{B})$  for non-negative definite matrix  $\mathbf{B}$  together with Lemma I.1 which asserts the operator norm of the resolvent  $\mathbf{Q}_-$  is uniformly bounded. Furthermore, both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1.

*Proof.* We decompose the matrix  $\Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}}$  as

$$\Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}} = \sum_{i=1}^N \sigma_i \sigma_i^T, \quad (32)$$

where  $\sigma_i = \sigma(\hat{S}^T \mathbf{w}_i) \in \mathbb{R}^m$  for which  $\mathbf{w}_i \in \mathbb{R}^d$  denotes the  $i$ -th row of  $\mathbf{W}$  (defined in equation 5). Using the resolvent identity (Lemma L.2), we write

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_m] - \tilde{\mathbf{Q}}_m &= \mathbb{E} \left[ \mathbf{Q}_m \left[ \tilde{\mathbf{Q}}_m^{-1} - \lambda \mathbf{I}_n - \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{\mathbf{U}}_n \right] \tilde{\mathbf{Q}}_m \right] \\ &= \frac{N}{m} \frac{1}{1+\alpha} \mathbb{E}[\mathbf{Q}_m] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m - \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \right] \tilde{\mathbf{Q}}_m \\ &= \frac{N}{m} \frac{1}{1+\alpha} \mathbb{E}[\mathbf{Q}_m] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m - \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_{-i} \frac{(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n}{1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i} \right] \tilde{\mathbf{Q}}_m, \end{aligned}$$

where the last equality is obtained with the Sherman identity (Lemma L.4) for

$$\mathbf{Q}_{-i} = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{\mathbf{U}}_n - \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1} \quad (33)$$

independent of  $\sigma_i$  and thus  $\mathbf{w}_i$ . Exploiting this independence, we decompose

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_m] - \tilde{\mathbf{Q}}_m &= \frac{N}{m} \frac{1}{1+\alpha} \mathbb{E}[\mathbf{Q}_m] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m - \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \right] \tilde{\mathbf{Q}}_m \\ &\quad + \frac{1}{m} \frac{1}{1+\alpha} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_{-i} \frac{(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n (\frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i - \alpha)}{1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i} \right] \tilde{\mathbf{Q}}_m \\ &= \frac{1}{m} \frac{1}{1+\alpha} \underbrace{\sum_{i=1}^N \mathbb{E}[\mathbf{Q}_m - \mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m}_{=\mathbf{Z}_1} \\ &\quad + \frac{1}{m} \frac{1}{1+\alpha} \underbrace{\sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i - \alpha \right) \right]}_{=\mathbf{Z}_2}. \end{aligned} \quad (35)$$

The last equality is obtained by exploiting the Sherman identity (Lemma L.4) in reverse on the rightmost term, and from the independence of  $\mathbf{Q}_{-i}$  and  $\sigma_i \sigma_i^T$  for the second right-hand term. We want to prove that both  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  have a vanishing spectral norm under Assumptions 1 and 2. With both the resolvent identity (Lemma L.2) and the Sherman identity (Lemma L.4), we rewrite  $\mathbf{Z}_1$  as

$$\begin{aligned} \mathbf{Z}_1 &= \frac{1}{m} \frac{1}{1+\alpha} \sum_{i=1}^N \mathbb{E}[\mathbf{Q}_m - \mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \\ &= -\frac{1}{m^2} \frac{1}{1+\alpha} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \right] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \\ &= -\frac{1}{m^2} \frac{1}{1+\alpha} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \mathbf{D}_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_m \right] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \\ &= -\frac{1}{m^2} \frac{1}{1+\alpha} \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{S}}^T \mathbf{D} \Sigma_{\hat{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m, \end{aligned}$$

where  $\mathbf{D} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = \left( 1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \right). \quad (37)$$

With a similar proof than for Lemma I.1, we can show there exists a  $K_{\tilde{\mathbf{Q}}_m}$  such that, for all  $m$ , we have  $\|\tilde{\mathbf{Q}}_m\| \leq K_{\tilde{\mathbf{Q}}_m}$  and then

$$\left\| \frac{1}{1+\alpha} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \right\| = \left\| \frac{m}{N} (\mathbf{I}_n - \lambda \tilde{\mathbf{Q}}_m) \right\| \leq \frac{m}{N} (1 + \lambda K_{\tilde{\mathbf{Q}}_m}). \quad (38)$$

Furthermore, from Lemma E.4, we have

$$\left\| \frac{1}{m^2} \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \mathbf{D} \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right\| = \mathcal{O} \left( \frac{1}{m} \right). \quad (39)$$

Therefore, by combining both equation 38 and equation 39, we conclude that  $\mathbf{Z}_1$  has a vanishing spectral norm, i.e.,

$$\|\mathbf{Z}_1\| = \left\| \frac{1}{m} \frac{1}{1+\alpha} \sum_{i=1}^N \mathbb{E} [\mathbf{Q}_m - \mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \right\| = \mathcal{O} \left( \frac{1}{m} \right). \quad (40)$$

We want to show now that  $\mathbf{Z}_2$  has also a vanishing operator norm. For  $i \in [N]$ , by setting

$$\mathbf{B}_i = m^{\frac{1}{4}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i - \alpha \right)$$

and

$$\mathbf{C}_i = m^{-\frac{1}{4}} \tilde{\mathbf{Q}}_m^T \hat{\mathbf{U}}_n^T \sigma_i,$$

we decompose  $\mathbf{Z}_2$  with its symmetric and its skew-symmetric part as

$$\begin{aligned} \mathbf{Z}_2 &= \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i - \alpha \right) \right] \\ &= \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} [\mathbf{B}_i \mathbf{C}_i^T] \\ &= \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{B}_i \mathbf{C}_i^T + \mathbf{C}_i \mathbf{B}_i^T}{2} \right] + \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{B}_i \mathbf{C}_i^T - \mathbf{C}_i \mathbf{B}_i^T}{2} \right]. \end{aligned}$$

For the symmetric part, we use the relations  $(\mathbf{B}_i - \mathbf{C}_i)(\mathbf{B}_i - \mathbf{C}_i)^T \succeq 0$  and  $(\mathbf{B}_i + \mathbf{C}_i)(\mathbf{B}_i + \mathbf{C}_i)^T \succeq 0$  to deduce that

$$-\mathbf{B}_i \mathbf{B}_i^T - \mathbf{C}_i \mathbf{C}_i^T \preceq \mathbf{B}_i \mathbf{C}_i^T + \mathbf{C}_i \mathbf{B}_i^T \preceq \mathbf{B}_i \mathbf{B}_i^T + \mathbf{C}_i \mathbf{C}_i^T,$$

where  $\preceq$  is the Loewner order for semi-positive-definite matrices. For the skew-symmetric part, we observe that  $\|\mathbb{E} [\mathbf{B}_i \mathbf{C}_i^T - \mathbf{C}_i \mathbf{B}_i^T]\| = \|i \mathbb{E} [\mathbf{B}_i \mathbf{C}_i^T - \mathbf{C}_i \mathbf{B}_i^T]\|$  for  $i^2 = -1$ . With a similar reasoning than above, using the relations  $(\mathbf{B}_i + i \mathbf{C}_i)(\mathbf{B}_i + i \mathbf{C}_i)^* \succeq 0$  and  $-(\mathbf{B}_i - i \mathbf{C}_i)(\mathbf{B}_i - i \mathbf{C}_i)^* \preceq 0$ , we deduce the relation

$$-\mathbf{B}_i \mathbf{B}_i^T - \mathbf{C}_i \mathbf{C}_i^T \preceq i(\mathbf{B}_i \mathbf{C}_i^T - \mathbf{C}_i \mathbf{B}_i^T) \preceq \mathbf{B}_i \mathbf{B}_i^T + \mathbf{C}_i \mathbf{C}_i^T.$$

From those relations, for both the symmetric and skew-symmetric parts, we have

$$\begin{aligned} \|\mathbf{Z}_2\| &= \left\| \frac{1}{1+\alpha} \frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i - \alpha \right) \right] \right\| \\ &\leq \frac{1}{1+\alpha} \left( \left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{B}_i \mathbf{B}_i^T \right] \right\| + \left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{C}_i \mathbf{C}_i^T \right] \right\| \right). \end{aligned} \quad (41)$$

From Lemma I.4, we know there exists a real  $K'_{\mathbf{Q}_m} > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \right\| \leq K'_{\mathbf{Q}_m}.$$

At this point,

$$\begin{aligned}
& \left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{B}_i \mathbf{B}_i^T \right] \right\| \\
&= \left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{\sqrt{m}} \mathbf{Q}_m \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \mathbf{Q}_m^T \left( \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \boldsymbol{\sigma}_i - \alpha \right)^2 \right] \right\| \\
&= \left\| \sqrt{m} \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D}_2^2 \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \mathbf{Q}_m^T \right] \right\| \\
&\leq \sqrt{m} K_{\mathbf{Q}_m}' \mathbb{E}[\|\mathbf{D}_2^2\|],
\end{aligned}$$

where  $\mathbf{D}_2 \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$[\mathbf{D}_2]_i = \left( \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \left( \hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n \right)^T \boldsymbol{\sigma}_i - \alpha \right).$$

From both Lemma K.5 and the union bound, we have

$$\Pr(\|\mathbf{D}_2\| > t) = \Pr\left(\max_{1 \leq i \leq N} [\mathbf{D}_2]_i > t\right) \leq CN e^{-cm \min(t, t^2)}$$

for some  $c, C > 0$  independent of  $m$  and  $N$ . We have thus

$$\begin{aligned}
\mathbb{E}(\|\mathbf{D}_2\|^2) &= \mathbb{E}\left(\max_{1 \leq i \leq N} [\mathbf{D}_2^2]_i\right) = \int_0^\infty \Pr\left(\max_{1 \leq i \leq N} [\mathbf{D}_2^2]_i > t\right) dt \\
&= \int_0^\infty 2t \Pr\left(\max_{1 \leq i \leq N} [\mathbf{D}_2]_i > t\right) dt \\
&\leq \int_0^\infty 2t CN e^{-cm \min(t, t^2)} dt \\
&= \int_0^1 2t CN e^{-cmt^2} dt + \int_1^\infty 2t CN e^{-cmt} dt \\
&\leq \int_0^\infty 2t CN e^{-cmt^2} dt + \int_0^\infty 2t CN e^{-cmt} dt \\
&= \frac{1}{m} \frac{2C}{c} \int_0^\infty t N e^{-t^2} dt + \frac{1}{m^2} \frac{2C}{c^2} \int_0^\infty t N e^{-t} dt \\
&= \mathcal{O}\left(\frac{1}{m}\right).
\end{aligned}$$

We deduce that

$$\left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{B}_i \mathbf{B}_i^T \right] \right\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

In addition, with a similar proof than for Lemma I.4, we can show there exists a real  $K_{\mathbf{Q}_m}' > 0$  such that, for all  $m$ , we have

$$\left\| \sqrt{\frac{N}{m}} \sqrt{\frac{1}{1+\alpha}} \bar{\mathbf{Z}}^T \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \right\| \leq K_{\mathbf{Q}_m}',$$

where  $\bar{\mathbf{Z}} \bar{\mathbf{Z}}^T$  is the Cholesky decomposition of  $\boldsymbol{\Phi}_{\hat{\mathcal{S}}}$ . Therefore,

$$\begin{aligned}
\left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{C}_i \mathbf{C}_i^T \right] \right\| &= \left\| \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m\sqrt{m}} \tilde{\mathbf{Q}}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \right] \right\| \\
&= \left\| \frac{1}{\sqrt{m}} \frac{N}{m} \tilde{\mathbf{Q}}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Phi}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m \right\| \\
&= \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).
\end{aligned}$$

From equation 41 and above, we deduce that  $\mathbf{Z}_2$  vanishes under the operator norm, i.e.,

$$\|\mathbf{Z}_2\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \quad (42)$$

Using both equation 40 and equation 42 into equation 36, we conclude that

$$\|\mathbb{E}[\mathbf{Q}_m] - \tilde{\mathbf{Q}}_m\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \quad (43)$$

□

To get Theorem 5.1, we start from Lemma E.1 and we show that

$$\|\bar{\mathbf{Q}}_m(\lambda) - \tilde{\mathbf{Q}}_m(\lambda)\| \rightarrow 0,$$

as  $N, m \rightarrow \infty$ .

**Theorem E.2 (Asymptotic Deterministic Resolvent).** *Under Assumptions 1 and 2, let  $\lambda > 0$  and let  $\bar{\mathbf{Q}}_m(\lambda) \in \mathbb{R}^{n \times n}$  be the resolvent defined as*

$$\bar{\mathbf{Q}}_m(\lambda) = \left[ \frac{N}{m} \frac{1}{1 + \delta} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1},$$

where  $\delta$  is the correction factor defined as the unique positive solution to

$$\delta = \frac{1}{m} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \left[ \frac{N}{m} \frac{1}{1 + \delta} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1} \right).$$

Then,

$$\lim_{m \rightarrow \infty} \left\| \mathbb{E}_{\mathbf{W}} [\mathbf{Q}_m(\lambda)] - \bar{\mathbf{Q}}_m(\lambda) \right\| = 0.$$

*Proof.* From Lemma J.1 in Appendix J, we know that  $\delta$  exists and is the unique positive solution of equation 17 under Assumptions 1 and 2. From Lemma E.1 we have a first asymptotic equivalent of  $\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m]$  given by

$$\tilde{\mathbf{Q}}_m = \left[ \frac{N}{m} \frac{1}{1 + \alpha} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1},$$

where

$$\alpha = \frac{1}{m} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-] \right),$$

since

$$\lim_{m \rightarrow \infty} \|\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m] - \tilde{\mathbf{Q}}_m\| = 0.$$

In order to finish the proof of the Theorem, we want to show that

$$\lim_{m \rightarrow \infty} \|\tilde{\mathbf{Q}}_m - \bar{\mathbf{Q}}_m\| = 0. \quad (44)$$

From the resolvent identity (Lemma L.2), we have

$$\|\tilde{\mathbf{Q}}_m - \bar{\mathbf{Q}}_m\| = \frac{N}{m} \frac{|\alpha - \delta|}{(1 + \delta)(1 + \alpha)} \left\| \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m \right\|. \quad (45)$$

Let  $\bar{\mathbf{Z}}\bar{\mathbf{Z}}^T$  be the Cholesky decomposition of  $\Phi_{\hat{\mathcal{S}}}$ . With a similar proof than for Lemma I.4, we can show there exists a real  $K'_{\tilde{\mathbf{Q}}} > 0$  such that, for all  $m$ , we have

$$\left\| \sqrt{\frac{1}{1 + \alpha}} \sqrt{\frac{N}{m}} \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \bar{\mathbf{Z}} \right\| \leq K'_{\tilde{\mathbf{Q}}}.$$

Similarly, we can show there exists a real  $K'_{\bar{\mathbf{Q}}} > 0$  such that, for all  $m$ , we have

$$\left\| \sqrt{\frac{1}{1 + \delta}} \sqrt{\frac{N}{m}} \bar{\mathbf{Z}}^T \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m \right\| \leq K'_{\bar{\mathbf{Q}}}.$$

Therefore,

$$\left\| \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m \right\| \leq \sqrt{(1+\delta)(1+\alpha)} \frac{m}{N} K'_{\tilde{\mathbf{Q}}} K'_{\bar{\mathbf{Q}}}.$$

As a consequence, in order to prove equation 44, it remains to prove that

$$\lim_{m \rightarrow \infty} |\alpha - \delta| = 0.$$

We decompose  $|\alpha - \delta|$  as

$$|\alpha - \delta| = \left| \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n [\mathbb{E}[\mathbf{Q}_-] - \bar{\mathbf{Q}}_m]) \right| \quad (46)$$

$$\leq \underbrace{\left| \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n [\mathbb{E}[\mathbf{Q}_-] - \tilde{\mathbf{Q}}_m]) \right|}_{=Z_1} + \underbrace{\left| \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n [\tilde{\mathbf{Q}}_m - \bar{\mathbf{Q}}_m]) \right|}_{=Z_2}. \quad (47)$$

To show  $Z_1$  vanishes, we write  $\alpha$  as

$$\begin{aligned} \alpha &= \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-]) \\ &= \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_m]) + \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n [\mathbb{E}[\mathbf{Q}_-] - \mathbb{E}[\mathbf{Q}_m]]). \end{aligned}$$

There exists a real  $K > 0$  such that

$$\frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n [\mathbb{E}[\mathbf{Q}_-] - \mathbb{E}[\mathbf{Q}_m]]) \leq K \|\mathbb{E}[\mathbf{Q}_-] - \mathbb{E}[\mathbf{Q}_m]\|;$$

since both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1,  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{A}\| \text{Tr}(\mathbf{B})$  for non-negative definite matrix  $\mathbf{B}$ , and from equation 30 that uniformly bounds  $\frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}})$ . From Lemma E.4, we have

$$\begin{aligned} \|\mathbb{E}[\mathbf{Q}_m - \mathbf{Q}_-]\| &= \left\| \frac{1}{m} \frac{m}{N} \sum_{i=1}^N \mathbb{E}[\mathbf{Q}_m - \mathbf{Q}_{-i}] \right\| \\ &= \left\| \frac{1}{m^2} \frac{m}{N} \mathbb{E} \left[ \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \mathbf{D} \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right\| \\ &= \mathcal{O} \left( \frac{1}{m} \right), \end{aligned}$$

where  $\mathbf{D} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = \left( 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \right).$$

As a consequence, by combining the results above and from Lemma E.1, we conclude for  $Z_1$  that

$$|Z_1| = \left| \alpha - \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m) \right| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

Using the vanishing result of  $Z_1$  into equation 47 and applying the resolvent identity (Lemma L.2) on  $Z_2$ , we get

$$|\alpha - \delta| \leq \frac{N}{m} \frac{|\alpha - \delta|}{(1+\delta)(1+\alpha)} \left| \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m) \right| + \mathcal{O} \left( \frac{1}{\sqrt{m}} \right),$$

which implies that

$$|\alpha - \delta| \left( 1 - \frac{N}{m} \frac{1}{(1+\delta)(1+\alpha)} \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m) \right) = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

It remains to show

$$\limsup_{m \rightarrow \infty} \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)(1+\alpha)} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m) < 1.$$



Let the matrices  $\mathbf{B}_n = (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_\delta \hat{\mathbf{U}}_n$ ,  $\mathbf{B}'_n = \bar{\mathbf{Z}}^T \hat{\mathbf{A}}_m \mathbf{Z}$ ,  $\bar{\mathbf{Q}}'_m = \left[ \frac{N}{m} \frac{1}{1+\delta} \mathbf{B}'_n + \lambda \mathbf{I}_m \right]^{-1}$ , and  $\tilde{\mathbf{Q}}'_m = \left[ \frac{N}{m} \frac{1}{1+\alpha} \mathbf{B}'_n + \lambda \mathbf{I}_m \right]^{-1}$ ; where  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  is the empirical transition model matrix defined in equation 14. Using the Cauchy–Schwarz inequality, we write

$$\begin{aligned} & \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)(1+\alpha)} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m) \\ &= \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)(1+\alpha)} \text{Tr}(\mathbf{B}'_n \tilde{\mathbf{Q}}'_m \mathbf{B}'_n \bar{\mathbf{Q}}'_m) \\ &\leq \sqrt{\underbrace{\frac{N}{m} \frac{1}{m} \frac{1}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m \mathbf{B}'^T_n)}_{Z'_1} \underbrace{\frac{N}{m} \frac{1}{m} \frac{1}{(1+\alpha)^2} \text{Tr}(\mathbf{B}'_n \tilde{\mathbf{Q}}'^T_m \tilde{\mathbf{Q}}'^T_m \mathbf{B}'^T_n)}_{Z'_2}}. \end{aligned}$$

We observe that

$$\begin{aligned} \delta &= \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m) = \frac{1}{m} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m) = \frac{1}{m} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \mathbf{Q}'^T_m \mathbf{Q}_m^{-1T}) \\ &= \frac{1}{m} \frac{N}{m} \frac{1+\delta}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m \mathbf{B}'^T_n) + \frac{\lambda}{m} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m). \end{aligned}$$

Since  $H(\mathbf{B}'_n)$  is at least semi-positive-definite under Assumption 2, we have

$$\text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m) = \text{Tr}(\bar{\mathbf{Q}}'^T_m \mathbf{B}'_n \bar{\mathbf{Q}}'_m) = \text{Tr}(\bar{\mathbf{Q}}'^T_m H(\mathbf{B}'_n) \bar{\mathbf{Q}}'_m) \geq 0.$$

As a consequence we have

$$\frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m \mathbf{B}'^T_n) \leq \frac{\delta - \frac{\lambda}{m} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m)}{1+\delta} \leq \frac{\delta}{1+\delta}.$$

To prove  $\frac{\delta}{1+\delta} < 1$ , it remains to show that  $\delta < \infty$ . With a similar proof than for Lemma I.1, we can show there exists a real  $K_{\bar{\mathbf{Q}}} > 0$  such that, for all  $m$ , we have  $\|\bar{\mathbf{Q}}_m\| \leq K_{\bar{\mathbf{Q}}}$ , and thus

$$\delta = \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m) = \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_\delta \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m) \leq \frac{2}{m} \text{Tr}(\Phi_\delta) \|\bar{\mathbf{Q}}_m(\delta)\| \leq \frac{2}{m} \text{Tr}(\Phi_\delta) K_{\bar{\mathbf{Q}}} < \infty$$

where we used for the first inequality the relation  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{A}\| \text{Tr}(\mathbf{B})$  for non-negative definite matrix  $\mathbf{B}$ . Furthermore, from equation 30,  $\frac{1}{m} \text{Tr}(\Phi_\delta)$  is bounded under Assumptions 1 and 2, and both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1. We thus conclude for  $Z'_1$  that

$$\limsup_m \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}'^T_m \mathbf{B}'^T_n) < 1. \quad (48)$$

With similar arguments, we can show for  $Z'_2$  that

$$\limsup_m \frac{1}{m} \frac{N}{m} \frac{1}{(1+\alpha)^2} \text{Tr}(\mathbf{B}'_n \tilde{\mathbf{Q}}'^T_m \tilde{\mathbf{Q}}'^T_m \mathbf{B}'^T_n) < 1,$$

which concludes the proof that

$$|\alpha - \delta| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \quad (49)$$

Using the result above with equation 45, we get

$$\begin{aligned} \|\tilde{\mathbf{Q}}_m - \bar{\mathbf{Q}}_m\| &= |\alpha - \delta| \left\| \frac{N}{m} \frac{1}{(1+\delta)(1+\alpha)} \tilde{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_\delta \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m \right\| \\ &= \mathcal{O}\left(\frac{1}{\sqrt{m}}\right), \end{aligned}$$

which concludes the proof.  $\square$

**Lemma E.3.** Under Assumptions 1 and 2, let  $\mathbf{D} \in \mathbb{R}^{N \times N}$  be the diagonal matrix defined in equation 37 for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i. \quad (50)$$

Then

$$\mathbb{E} [\|\mathbf{D}\|] = \mathcal{O}(1).$$

*Proof.* Let  $\alpha = \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-}(\lambda)])$  defined in equation 28. From equation 31,  $\alpha$  is uniformly bounded, i.e., there exists a real  $K_\alpha > 0$  such that  $\alpha \leq K_\alpha$ . From both Lemma K.5 and the union bound, we have

$$\Pr(\|\mathbf{D}\| > 1 + \alpha + t) = \Pr\left(\max_{1 \leq i \leq N} \mathbf{D}_i > 1 + \alpha + t\right) \leq CN e^{-cm \min(t, t^2)},$$

for some  $c, C > 0$  independent of  $m$  and  $N$ . Therefore,

$$\begin{aligned} \mathbb{E} [\|\mathbf{D}\|] &= \mathbb{E} \left[ \max_{1 \leq i \leq N} \mathbf{D}_i \right] = \int_0^\infty \Pr\left(\max_{1 \leq i \leq N} \mathbf{D}_i > t\right) dt \\ &= \int_0^{2(1+K_\alpha)} \Pr\left(\max_{1 \leq i \leq N} \mathbf{D}_i > t\right) dt + \int_{2(1+K_\alpha)}^\infty \Pr\left(\max_{1 \leq i \leq N} \mathbf{D}_i > t\right) dt \\ &\leq 2(1+K_\alpha) + \int_{2(1+K_\alpha)}^\infty CN e^{-cm \min((t-(1+K_\alpha))^2, t-(1+K_\alpha))} dt \\ &= 2(1+K_\alpha) + \int_{1+K_\alpha}^\infty CN e^{-cmt} dt \\ &= 2(1+K_\alpha) + \frac{CN}{cm} e^{-Cm(1+K_\alpha)} \\ &= \mathcal{O}(1). \end{aligned}$$

□

**Lemma E.4.** Under Assumptions 1 and 2, let  $\mathbf{D} \in \mathbb{R}^{N \times N}$  be the diagonal matrix defined in equation 37 for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i.$$

Then

$$\left\| \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right\| = \mathcal{O}(1).$$

*Proof.* From Lemma I.4, there exists  $K'_{\mathbf{Q}_m} > 0$  such that, for all  $m$ , we have  $\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \right\| \leq 2K'_{\mathbf{Q}_m}$  and  $\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right\| \leq K'_{\mathbf{Q}_m}$ . Therefore,

$$\left\| \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right\| \leq 2K_{\mathbf{Q}_m}^{\prime 2} \mathbb{E} [\|\mathbf{D}\|].$$

From Lemma E.3, we have

$$\mathbb{E} [\|\mathbf{D}\|] = \mathcal{O}(1).$$

As a consequence, we deduce that

$$\left\| \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right\| = \mathcal{O}(1). \quad (51)$$

□

## F PROOF OF THEOREM 5.2

This section is dedicated to find an asymptotic deterministic limit of the empirical  $\widehat{\text{MSBE}}(\hat{\theta})$  (equation 8) under Assumptions 1 and 2. First, we observe that the empirical MSBE depends on the quadratic form  $\mathbf{r}^T \mathbf{Q}_m(\lambda)^T \mathbf{Q}_m(\lambda) \mathbf{r}$  since

$$\widehat{\text{MSBE}}(\hat{\theta}) = \frac{1}{n} \|\mathbf{r} + \gamma \boldsymbol{\Sigma}_{\mathbf{X}'_n} \hat{\theta} - \boldsymbol{\Sigma}_{\mathbf{X}_n}^T \hat{\theta}\|^2 \quad (52)$$

$$= \frac{1}{n} \left\| \mathbf{r} - \frac{1}{mn} (\boldsymbol{\Sigma}_{\mathbf{X}_n} - \gamma \boldsymbol{\Sigma}_{\mathbf{X}'_n})^T \boldsymbol{\Sigma}_{\mathbf{X}_n} \mathbf{Q}_m(\lambda) \mathbf{r} \right\|^2 \quad (53)$$

$$= \frac{1}{n} \left\| \left[ \frac{1}{mn} (\boldsymbol{\Sigma}_{\mathbf{X}_n} - \gamma \boldsymbol{\Sigma}_{\mathbf{X}'_n})^T \boldsymbol{\Sigma}_{\mathbf{X}_n} + \lambda \mathbf{I}_n - \frac{1}{mn} (\boldsymbol{\Sigma}_{\mathbf{X}_n} - \gamma \boldsymbol{\Sigma}_{\mathbf{X}'_n})^T \boldsymbol{\Sigma}_{\mathbf{X}_n} \right] \mathbf{Q}_m(\lambda) \mathbf{r} \right\|^2 \quad (54)$$

$$= \frac{\lambda^2}{n} \|\mathbf{Q}_m(\lambda) \mathbf{r}\|^2. \quad (55)$$

We determine in Theorem 5.2 a deterministic limit of  $\widehat{\text{MSBE}}(\hat{\theta})$  by combining Theorem 5.1, which provides an asymptotically more tractable approximation of  $\mathbb{E}_{\mathbf{W}}[\mathbf{Q}_m(\lambda)]$  under the form of a fixed-point equation, with concentration arguments. Theorem 5.2 is corollary of Lemma F.2 and of the concentration result of Lemma K.2 in Section K. Both Lemma F.4 and Lemma F.5 are key Lemma used in the proof of Theorem 5.2 and Theorem 5.3.

To simplify the notations, we denote the matrix  $\mathbf{Q}_m$  as the resolvent  $\mathbf{Q}_m(\lambda)$  (defined in equation 13). We define the matrix  $\boldsymbol{\Psi}_{\hat{\delta}} \in \mathbb{R}^{m \times m}$  as

$$\boldsymbol{\Psi}_{\hat{\delta}} = \frac{N}{m} \frac{1}{1 + \delta} \boldsymbol{\Phi}_{\hat{\delta}}.$$

Furthermore, the notation  $\mathbf{A} = \mathbf{B} + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right)$  means that  $\|\mathbf{A} - \mathbf{B}\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right)$ .

**Theorem F.1** (Asymptotic Empirical MSBE). *Under the conditions of Theorem 5.1, the deterministic asymptotic empirical MSBE is*

$$\widehat{\text{MSBE}}(\hat{\theta}) = \frac{\lambda^2}{n} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|^2 + \hat{\Delta},$$

with second-order correction factor

$$\hat{\Delta} = \frac{\lambda^2}{n} \frac{\frac{1}{N} \text{Tr}(\bar{\mathbf{Q}}_m(\lambda) \boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m(\lambda)^T)}{1 - \frac{1}{N} \text{Tr}(\bar{\mathbf{Q}}_m(\lambda) \boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m(\lambda)^T \boldsymbol{\Psi}_1)} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|_{\boldsymbol{\Psi}_1}^2,$$

where

$$\boldsymbol{\Psi}_1 = \hat{\mathbf{U}}_n^T \boldsymbol{\Psi}_{\hat{\delta}} \hat{\mathbf{U}}_n, \quad \text{and} \quad \boldsymbol{\Psi}_2 = (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Psi}_{\hat{\delta}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n).$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,

$$\widehat{\text{MSBE}}(\hat{\theta}) - \overline{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0.$$

*Proof.* From equation 55, we have

$$\widehat{\text{MSBE}}(\hat{\theta}) = \frac{\lambda^2}{n} \|\mathbf{Q}_m \mathbf{r}\|^2 = \frac{\lambda^2}{n} \mathbf{r}^T \mathbf{Q}_m^T \mathbf{Q}_m \mathbf{r}.$$

From Lemma K.6, we have

$$\Pr \left( \left| \frac{\lambda^2}{n} \mathbf{r}^T \mathbf{Q}_m^T \mathbf{Q}_m \mathbf{r} - \frac{\lambda^2}{n} \mathbf{r}^T \mathbb{E}[\mathbf{Q}_m^T \mathbf{Q}_m] \mathbf{r} \right| > t \right) \leq C e^{-cn^2 mt^2},$$

for some  $C, c > 0$  independent of  $m, n$  and  $N$ . Furthermore, from Lemma F.2, we have

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{Q}_m] - \bar{\mathbf{Q}}_m^T \bar{\mathbf{Q}}_m - \frac{\frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \bar{\mathbf{Q}}_m)}{1 - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m \right\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

As a consequence, we have

$$\widehat{\text{MSBE}}(\hat{\theta}) - \overline{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0,$$

as  $m \rightarrow \infty$ . □

**Lemma F.2.** Under Assumptions 1 and 2, let  $\mathbf{Q}_m \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13, let  $\bar{\mathbf{Q}}_m \in \mathbb{R}^{n \times n}$  be the deterministic resolvent defined in equation 15, and let  $\mathbf{M} \in \mathbb{R}^{n \times n}$  be any matrix with a bounded operator norm. Then,

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m - \frac{\frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m \right\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

for  $\Psi_1, \Psi_2 \in \mathbb{R}^{n \times n}$  defined in equation 19.

*Proof.* From Lemma F.6, we have

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] &= \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m + \mathbb{E}\left[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\mathcal{S} \hat{\mathbf{U}}_n \mathbf{Q}_m\right] \\ &\quad - \mathbb{E}\left[\mathbf{Q}_-^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\mathcal{S} \hat{\mathbf{U}}_n \mathbf{Q}_-\right] + \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] \\ &\quad + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

Let

$$\mathbf{M}' = \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\mathcal{S} \hat{\mathbf{U}}_n = \mathbf{M} [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m]$$

With a similar proof than for Lemma I.1, we can show that there exists a real  $K_{\mathbf{Q}}$  such that, for all  $m$ , we have  $\|\bar{\mathbf{Q}}_m\| \leq K_{\mathbf{Q}}$ . We deduce thus that  $\mathbf{M}'$  is a matrix with a bounded operator norm since  $\|\mathbf{M}'\| \leq (1 + \lambda K_{\mathbf{Q}}) \|\mathbf{M}\|$ . From Lemma F.3, we have

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\mathcal{S} \hat{\mathbf{U}}_n \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\mathcal{S} \hat{\mathbf{U}}_n \mathbf{Q}_-] \right\| = \mathcal{O}\left(\frac{1}{m}\right).$$

Therefore,

$$\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] = \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m + \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right).$$

Furthermore, from Lemma F.4, we have

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] \right\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

and from Lemma F.5 we have

$$\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] = \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \frac{1}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right).$$

We conclude thus

$$\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] = \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m + \frac{\frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right).$$

□

**Lemma F.3.** Under Assumptions 1 and 2, let  $\mathbf{M} \in \mathbb{R}^{n \times n}$  be any matrix with a bounded operator norm, let  $\mathbf{Q}_m \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13, and let  $\mathbf{Q}_- \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 29. Then,

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_-] \right\| = \mathcal{O}\left(\frac{1}{m}\right).$$

*Proof.* We observe that

$$\begin{aligned} \left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_-] \right\| &\leq \left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_m] \right\| \\ &\quad + \left\| \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_-] \right\|. \end{aligned}$$

The objective is to show that both terms vanish. By exchangeability arguments, we have

$$\begin{aligned}
& \left\| \mathbb{E} [\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_m] \right\| \\
&= \left\| \frac{1}{N} \mathbb{E} \left[ \sum_{i=1}^N [\mathbf{Q}_m - \mathbf{Q}_{-i}]^T \mathbf{M} \mathbf{Q}_m \right] \right\| \\
&= \left\| \frac{1}{N} \mathbb{E} \left[ \sum_{i=1}^N \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \mathbf{Q}_m \right] \right\| \quad (\text{Lemma L.2}) \\
&= \left\| \frac{1}{N} \mathbb{E} \left[ \sum_{i=1}^N \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m \left( 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \right) \right] \right\| \\
&= \left\| \frac{1}{N} \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_S^T \mathbf{D} \boldsymbol{\Sigma}_S (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m \right] \right\|,
\end{aligned}$$

where  $\mathbf{D} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i.$$

From Lemma E.3, we know

$$\mathbb{E} [\|\mathbf{D}\|] = \mathcal{O}(1).$$

Furthermore, from Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_S \hat{\mathbf{U}}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_S^T \right\| \leq 2K'_Q.$$

We deduce thus

$$\begin{aligned}
\left\| \mathbb{E} [\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_m] \right\| &= \left\| \frac{1}{m} \frac{m}{N} \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_S^T \mathbf{D} \boldsymbol{\Sigma}_S (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m \right] \right\| \\
&= \mathcal{O} \left( \frac{1}{m} \right).
\end{aligned}$$

With a similar reasoning, we can show that

$$\left\| \mathbb{E} [\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_-] \right\| = \mathcal{O} \left( \frac{1}{m} \right),$$

and we conclude thus

$$\left\| \mathbb{E} [\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \mathbf{M} \mathbf{Q}_-] \right\| = \mathcal{O} \left( \frac{1}{m} \right).$$

□

**Lemma F.4.** Under Assumptions 1 and 2, let  $\mathbf{Q}_m \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13, let  $\mathbf{Q}_- \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 29, and let  $\boldsymbol{\Psi}_1 \in \mathbb{R}^{n \times n}$  be the matrix defined in equation 19. Then,

$$\left\| \mathbb{E} [\mathbf{Q}_m^T \boldsymbol{\Psi}_1 \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \right\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

*Proof.* We observe that

$$\begin{aligned}
\left\| \mathbb{E} [\mathbf{Q}_m^T \boldsymbol{\Psi}_1 \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \right\| &\leq \left\| \mathbb{E} [\mathbf{Q}_m^T \boldsymbol{\Psi}_1 \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_m] \right\| \\
&\quad + \left\| \mathbb{E} [\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \right\|.
\end{aligned}$$

The objective is to show that both terms vanish. By exchangeability arguments, we have

$$\begin{aligned}
& \left\| \mathbb{E} [\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] - \mathbb{E} [\mathbf{Q}_{-i}^T \Psi_1 \mathbf{Q}_m] \right\| \\
&= \left\| \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ (\mathbf{Q}_m - \mathbf{Q}_{-i})^T \Psi_1 \mathbf{Q}_m \right] \right\| \\
&= \left\| \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \Psi_1 \mathbf{Q}_m \right] \right\| \quad (\text{Lemma L.2}) \\
&= \left\| \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m \left( 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \right) \right] \right\| \\
&= \left\| \frac{1}{N} \mathbb{E} \left[ \underbrace{\frac{1}{m} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m}_{=Z} \right] \right\|,
\end{aligned}$$

where  $\mathbf{D} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$D_i = 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i.$$

Let the matrices

$$\mathbf{B} = \frac{1}{N} \frac{1}{m^{\frac{1}{4}}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T$$

and

$$\mathbf{C}^T = \frac{1}{m^{\frac{3}{4}}} \Psi_1 \mathbf{Q}_m.$$

We decompose  $\mathbf{Z}$  with its symmetric and its skew-symmetric parts as

$$\mathbf{Z} = \mathbb{E} [\mathbf{B} \mathbf{C}^T] = \mathbb{E} \left[ \frac{\mathbf{B} \mathbf{C}^T + \mathbf{C} \mathbf{B}^T}{2} \right] + \mathbb{E} \left[ \frac{\mathbf{B} \mathbf{C}^T - \mathbf{C} \mathbf{B}^T}{2} \right].$$

With the same reasoning on the symmetric part and the skew-symmetric part than for equation 41, we get for the operator norm

$$\|\mathbf{Z}\| \leq \|\mathbb{E} [\mathbf{B} \mathbf{B}^T]\| + \|\mathbb{E} [\mathbf{C} \mathbf{C}^T]\|.$$

We want to show that both  $\|\mathbb{E} [\mathbf{B} \mathbf{B}^T]\|$  and  $\|\mathbb{E} [\mathbf{C} \mathbf{C}^T]\|$  vanish. We have

$$\mathbb{E} [\mathbf{B} \mathbf{B}^T] = \mathbb{E} \left[ \frac{m^2}{N^2} \frac{1}{m^2 \sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_m^T \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \mathbf{D} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right].$$

From Lemma E.3, we know

$$\mathbb{E} [\|\mathbf{D}\|] = \mathcal{O}(1).$$

Furthermore, from Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \right\| \leq 2K'_Q.$$

We have therefore

$$\|\mathbb{E} [\mathbf{B} \mathbf{B}^T]\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

For  $\mathbb{E} [\mathbf{C} \mathbf{C}^T]$ , we have

$$\mathbb{E} [\mathbf{C} \mathbf{C}^T] = \mathbb{E} \left[ \frac{1}{m \sqrt{m}} \mathbf{Q}_m^T \Psi_1^2 \mathbf{Q}_m \right].$$

Let  $\sigma_{N+1}$  and  $\sigma_{N+2}$  be independent vectors with the same law as  $\sigma_i$ , we have

$$\mathbb{E} \left[ \frac{1}{m\sqrt{m}} \mathbf{Q}_m^T \Psi_1 \Psi_1 \mathbf{Q}_m \right] = \mathbb{E} \left[ \frac{1}{m\sqrt{m}} \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_{N+2} \sigma_{N+2}^T \hat{\mathbf{U}}_n \mathbf{Q}_m \right].$$

Let

$$\mathbf{B}' = \frac{1}{m^{\frac{3}{4}}} \frac{N}{m} \frac{1}{1+\delta} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n$$

and

$$\mathbf{C}'^T = \frac{1}{m^{\frac{3}{4}}} \frac{N}{m} \frac{1}{1+\delta} \hat{\mathbf{U}}_n^T \sigma_{N+2} \sigma_{N+2}^T \hat{\mathbf{U}}_n \mathbf{Q}_m.$$

We decompose  $\mathbb{E} [\mathbf{C}\mathbf{C}^T]$  with its symmetric and its skew-symmetric parts as

$$\mathbb{E} [\mathbf{C}\mathbf{C}^T] = \mathbb{E} [\mathbf{B}'\mathbf{C}'^T] = \mathbb{E} \left[ \frac{\mathbf{B}'\mathbf{C}'^T + \mathbf{C}'\mathbf{B}'^T}{2} \right] + \mathbb{E} \left[ \frac{\mathbf{B}'\mathbf{C}'^T - \mathbf{C}'\mathbf{B}'^T}{2} \right],$$

and we get for the operator norm

$$\|\mathbb{E} [\mathbf{C}\mathbf{C}^T]\| \leq \|\mathbb{E} [\mathbf{B}'\mathbf{B}'^T]\| + \|\mathbb{E} [\mathbf{C}'\mathbf{C}'^T]\|.$$

To prove  $\|\mathbb{E} [\mathbf{C}\mathbf{C}^T]\|$  vanish, we prove both  $\|\mathbb{E} [\mathbf{B}'\mathbf{B}'^T]\|$  and  $\|\mathbb{E} [\mathbf{C}'\mathbf{C}'^T]\|$  vanish. Let  $K = \frac{1}{(1+\delta)^2} \frac{N}{N+1} \frac{N}{m}$ , we write  $\mathbb{E} [\mathbf{B}'\mathbf{B}'^T]$  as

$$\begin{aligned} \mathbb{E} [\mathbf{B}'\mathbf{B}'^T] &= \mathbb{E} \left[ \frac{1}{m\sqrt{m}} \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \\ &= \mathbb{E} \left[ \frac{1}{m\sqrt{m}} \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} \mathbf{Q}_{-N-1}^T \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_{N+1} \sigma_{N+1}^T \hat{\mathbf{U}}_n \mathbf{Q}_{-N-1} \right] \\ &= \mathbb{E} \left[ K \frac{1}{m\sqrt{m}} \sum_{i=1}^{N+1} \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_i \right) \right] \\ &= \mathbb{E} \left[ K \frac{1}{m\sqrt{m}} \sum_{i=1}^{N+1} \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \frac{1}{m} \text{Tr}(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \Phi_{\hat{\mathcal{S}}}) \right] \\ &+ \mathbb{E} \left[ K \frac{1}{m\sqrt{m}} \sum_{i=1}^{N+1} \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \left( \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_i - \frac{1}{m} \text{Tr}(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \Phi_{\hat{\mathcal{S}}}) \right) \right] \\ &= \mathbb{E} \left[ \underbrace{K \frac{1}{m} \text{Tr}(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \Phi_{\hat{\mathcal{S}}}) \frac{1}{m\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}^2 \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m}_{=Z_1} \right] \\ &+ \mathbb{E} \left[ \underbrace{K \frac{1}{m} \text{Tr}(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \Phi_{\hat{\mathcal{S}}}) \frac{1}{m\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}' \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m}_{=Z_2} \right], \end{aligned}$$

where  $\mathbf{D}' \in \mathbb{R}^{N \times N}$  is a diagonal matrices for which, for all  $i \in [N]$ , we have

$$\mathbf{D}'_i = \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \sigma_i - \frac{1}{m} \text{Tr}(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^T \Phi_{\hat{\mathcal{S}}}).$$

From Lemma E.3, from Lemma I.4, and from equation 30, we have

$$\|Z_1\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

From Lemma K.2, we have

$$\mathbb{E} [\|\mathbf{D}'\|] = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

and thus

$$\|Z_2\| = \mathcal{O} \left( \frac{1}{m} \right).$$

We conclude that

$$\|\mathbb{E}[\mathbf{B}'\mathbf{B}'^T]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$$

and

$$\|\mathbb{E}[\mathbf{C}'\mathbf{C}'^T]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

Therefore,

$$\|\mathbb{E}[\mathbf{C}\mathbf{C}^T]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$$

and

$$\left\| \mathbb{E}\left[\mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m\right] - \mathbb{E}\left[\mathbf{Q}_-^T \hat{\mathbf{U}}_n^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_-\right] \right\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

With a similar reasoning, we can show

$$\|\mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

We conclude thus

$$\|\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

□

**Lemma F.5.** *Under Assumptions 1 and 2, let  $\mathbf{Q}_m \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13, let  $\bar{\mathbf{Q}}_m \in \mathbb{R}^{n \times n}$  be the deterministic resolvent defined in equation 15, let  $\Psi_1, \Psi_2 \in \mathbb{R}^{n \times n}$  be the matrices defined in equation 19. Then,*

$$\left\| \mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] - \frac{1}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m \right\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

*Proof.* From Lemma F.6, we know that

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] &= \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathbb{E}\left[\mathbf{Q}_m^T \Psi_1 \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m\right] \\ &\quad - \mathbb{E}\left[\mathbf{Q}_-^T \Psi_1 \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_-\right] + \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] \\ &\quad + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

Exploiting  $\bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_{\mathcal{S}} \hat{\mathbf{U}}_n = \mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m$  in the above equation, and from Lemma F.4, we obtain the simplification

$$\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] = \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right).$$

or equivalently

$$\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] \left(1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-]\right) = \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathcal{O}_{\|\cdot\|}\left(\frac{1}{\sqrt{m}}\right).$$

Let  $\mathbf{B}'_n = \bar{\mathbf{Z}}^T \hat{\mathbf{A}}_m \mathbf{Z}$  and  $\bar{\mathbf{Q}}'_m = \left[\frac{N}{m} \frac{1}{1+\delta} \mathbf{B}'_n + \lambda \mathbf{I}_m\right]^{-1}$ , for which  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  is the empirical transition model matrix (equation 14) and  $\bar{\mathbf{Z}} \bar{\mathbf{Z}}^T = \Phi_{\mathcal{S}}$  is the Cholesky decomposition of  $\Phi_{\mathcal{S}}$ . We have from the cyclic properties of the trace

$$\begin{aligned} &\frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m) \\ &= \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)^2} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \hat{\mathbf{U}}_n^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m) \\ &= \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}_m^T \mathbf{B}'_n{}^T). \end{aligned}$$



From equation 48, we have

$$\limsup_m \frac{1}{m} \frac{N}{m} \frac{1}{(1+\delta)^2} \text{Tr}(\mathbf{B}'_n \bar{\mathbf{Q}}'_m \bar{\mathbf{Q}}_m{}^T \mathbf{B}'_n{}^T) < 1.$$

Therefore,

$$\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] = \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \frac{1}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right).$$

□

**Lemma F.6.** *Under Assumptions 1 and 2, let  $\mathbf{Q}_m \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13, let  $\mathbf{Q}_- \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 29, let  $\bar{\mathbf{Q}}_m \in \mathbb{R}^{n \times n}$  be the deterministic resolvent defined in equation 15, let  $\hat{\mathbf{U}}_n, \hat{\mathbf{V}}_n \in \mathbb{R}^{m \times n}$  be the shift matrices defined in equation 7, and let  $\mathbf{M}$  be either any matrix with a bounded operator norm or  $\mathbf{M} = \Psi_1$ . Then,*

$$\begin{aligned} & \left\| \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] - \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m - \mathbb{E} \left[ \mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\delta \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \right. \\ & \quad \left. + \mathbb{E} \left[ \mathbf{Q}_-^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\delta \hat{\mathbf{U}}_n \mathbf{Q}_- \right] - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] \right\| \\ & = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right), \end{aligned}$$

for  $\Psi_1, \Psi_2 \in \mathbb{R}^{n \times n}$  defined in equation 19.

*Proof.* With the resolvent identity (Lemma L.2), we decompose  $\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m]$  as

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] &= \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m] - \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} (\bar{\mathbf{Q}}_m - \mathbf{Q}_m)] \\ &= \mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m] \\ & \quad - \mathbb{E} \left[ \mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_\delta^T \Sigma_\delta \hat{\mathbf{U}}_n - (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\delta \hat{\mathbf{U}}_n \right] \mathbf{Q}_m \right] \\ &= \underbrace{\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m]}_{=\mathbf{Z}_1} + \mathbb{E} \left[ \mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Psi_\delta \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \\ & \quad - \underbrace{\frac{1}{m} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_m \right]}_{=\mathbf{Z}_2}, \end{aligned} \tag{56}$$

where  $\Sigma_\delta^T \Sigma_\delta = \sum_{i=1}^N \sigma_i \sigma_i^T$  is the same decomposition of  $\Sigma_\delta^T \Sigma_\delta$  than the one used in equation 32. From Theorem 5.1, we have

$$\|\mathbb{E}[\mathbf{Q}_m] - \bar{\mathbf{Q}}_m\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

Therefore, from above and from Lemma F.9 which upper bounds  $\|\mathbf{M} \bar{\mathbf{Q}}_m\|$ , we deduce for  $\mathbf{Z}_1$  that

$$\begin{aligned} \|\mathbf{Z}_1\| - \|\bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m\| &= \|\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m]\| - \|\bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m\| \\ &\leq \|\mathbb{E}[\mathbf{Q}_m] - \bar{\mathbf{Q}}_m\| \|\mathbf{M} \bar{\mathbf{Q}}_m\| \\ &= \mathcal{O} \left( \frac{1}{\sqrt{m}} \right). \end{aligned}$$

We want to find now a deterministic approximation for  $\mathbf{Z}_2$  in equation 57. From the Sherman identity (Lemma L.4) and with the resolvent  $\mathbf{Q}_{-i}$  defined in equation 33 as

$$\mathbf{Q}_{-i} = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_\delta^T \Sigma_\delta \hat{\mathbf{U}}_n - \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i \sigma_i^T \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1},$$



As a consequence, we conclude that

$$\begin{aligned}\mathbb{E}[\mathbf{Q}_m^T \mathbf{M} \mathbf{Q}_m] &= \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m + \mathbb{E} \left[ \mathbf{Q}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Psi}_S \hat{\mathbf{U}}_n \mathbf{Q}_m \right] \\ &\quad - \mathbb{E} \left[ \mathbf{Q}_-^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Psi}_S \hat{\mathbf{U}}_n \mathbf{Q}_- \right] + \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \\ &\quad + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right).\end{aligned}$$

□

**Lemma F.7.** Under Assumptions 1 and 2, let  $\mathbf{Z}_{22} \in \mathbb{R}^{n \times n}$  be the matrix defined in equation 60 as

$$\mathbf{Z}_{22} = \frac{1}{m} \frac{1}{1 + \delta} \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \frac{\frac{1}{m} \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}} \right].$$

Then,

$$\left\| \mathbf{Z}_{22} - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \mathbb{E}[\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \right\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right),$$

where  $\bar{\mathbf{Q}}_m \in \mathbb{R}^{n \times n}$  is the deterministic resolvent defined in equation 15,  $\mathbf{Q}_- \in \mathbb{R}^{n \times n}$  is the resolvent defined in equation 29, and  $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \in \mathbb{R}^{n \times n}$  defined in equation 19.

*Proof.* Let  $\mathbf{D} \in \mathbb{R}^{N \times N}$  be a diagonal matrix for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = 1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i,$$

and  $\mathbf{D}_2 \in \mathbb{R}^{N \times N}$  be another diagonal matrix for which, for all  $i \in [N]$ , we have

$$\begin{aligned}[\mathbf{D}_2]_i &= \frac{\frac{1}{m} \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i} - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \\ &= \frac{\frac{1}{m} \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i} \\ &\quad - \frac{\frac{1}{m} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_S (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m \right)}{1 + \delta} \\ &= \frac{\frac{1}{m} \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i} \\ &\quad - \frac{\frac{1}{m} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_S \right)}{1 + \delta}.\end{aligned}$$

We have

$$\begin{aligned}&\left\| \mathbf{Z}_{22} - \frac{1}{N} \mathbb{E}[\mathbf{Q}_-^T \boldsymbol{\Psi}_1 \mathbf{Q}_-] \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \right\| \\ &= \left\| \mathbf{Z}_{22} - \frac{1}{m} \frac{1}{1 + \delta} \mathbb{E} \left[ \sum_{i=1}^N \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \right] \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \right\| \\ &= \left\| \frac{1}{m} \frac{1}{1 + \delta} \mathbb{E} \left[ \sum_{i=1}^N \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \left( \frac{\frac{1}{m} \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i} \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m) \right) \right] \right\| \\ &= \left\| \frac{1}{m} \frac{1}{1 + \delta} \mathbb{E} \left[ \underbrace{\mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_S^T \mathbf{D}^2 \mathbf{D}_2 \boldsymbol{\Sigma}_S \hat{\mathbf{U}}_n \mathbf{Q}_m}_{=Z_{221}} \right] \right\|.\end{aligned}$$

Let the matrices

$$\mathbf{B} = m^{-\frac{1}{4}} \frac{1}{\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}^2,$$

and

$$\mathbf{C}^T = m^{\frac{1}{4}} \frac{1}{\sqrt{m}} \mathbf{D}_2 \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m.$$

Using the matrices above, we have

$$\mathbf{Z}_{221} = \frac{1}{m} \frac{1}{1+\delta} \mathbb{E}[\mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}^2 \mathbf{D}_2 \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m] = \frac{1}{1+\delta} \mathbb{E}[\mathbf{B} \mathbf{C}^T] = \frac{1}{1+\delta} \mathbb{E}\left[\frac{\mathbf{B} \mathbf{C}^T + \mathbf{C} \mathbf{B}^T}{2}\right],$$

since  $\mathbf{Z}_{221}$  is symmetric. We use the relations  $(\mathbf{B} - \mathbf{C})(\mathbf{B} - \mathbf{C})^T \succeq 0$  and  $(\mathbf{B} + \mathbf{C})(\mathbf{B} + \mathbf{C})^T \succeq 0$  to deduce the following relation

$$-\mathbf{B} \mathbf{B}^T - \mathbf{C} \mathbf{C}^T \preceq \mathbf{B} \mathbf{C}^T + \mathbf{C} \mathbf{B}^T \preceq \mathbf{B} \mathbf{B}^T + \mathbf{C} \mathbf{C}^T.$$

From this relation, we obtain

$$\|\mathbf{Z}_{221}\| \leq \frac{1}{2(1+\delta)} \left( \mathbb{E}[\|\mathbf{B} \mathbf{B}^T\|] + \mathbb{E}[\|\mathbf{C} \mathbf{C}^T\|] \right),$$

where

$$\mathbf{B} \mathbf{B}^T = \frac{1}{m\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}^4 \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m$$

and

$$\mathbf{C} \mathbf{C}^T = \frac{1}{\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \Sigma_{\hat{\mathcal{S}}}^T \mathbf{D}_2^2 \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m.$$

To get the Lemma, we prove that both  $\mathbb{E}[\|\mathbf{B} \mathbf{B}^T\|]$  and  $\mathbb{E}[\|\mathbf{C} \mathbf{C}^T\|]$  vanish. From Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \Sigma_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{\mathcal{S}}}^T \right\| \leq 2K'_Q.$$

Furthermore, from Lemma E.3, we know

$$\mathbb{E}[\|\mathbf{D}^4\|] = \mathcal{O}(1).$$

We conclude that

$$\mathbb{E}[\|\mathbf{B} \mathbf{B}^T\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

For  $\mathbb{E}[\|\mathbf{C} \mathbf{C}^T\|]$ , we remark that

$$\begin{aligned} \Pr([\mathbf{D}_2]_i \geq t) &\leq \Pr\left(\frac{\frac{1}{m} \sigma_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i}{1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i} \right. \\ &\quad \left. - \frac{\frac{1}{m} \text{Tr}(\Phi_{\hat{\mathcal{S}}}(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T)}{1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i} \geq \frac{t}{2}\right) \\ &\quad + \Pr\left(\frac{\frac{1}{m} \text{Tr}(\Phi_{\hat{\mathcal{S}}}(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T)}{1 + \frac{1}{m} \sigma_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma_i} \right. \\ &\quad \left. - \frac{\frac{1}{m} \text{Tr}(\Phi_{\hat{\mathcal{S}}}(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T)}{1 + \delta} \geq \frac{t}{2}\right). \end{aligned}$$

Since  $\|\mathbf{M}\bar{\mathbf{Q}}_m\|$  is bounded from Lemma F.9, with a similar proof than for Lemma K.5, we can prove that

$$\Pr\left(\left|\frac{1}{m}\boldsymbol{\sigma}^T(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)\mathbf{Q}_{-i}^T\mathbf{M}\bar{\mathbf{Q}}_m(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T\boldsymbol{\sigma} - \frac{1}{m}\text{Tr}((\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)\mathbb{E}[\mathbf{Q}_{-i}^T]\mathbf{M}\bar{\mathbf{Q}}_m(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T\boldsymbol{\Phi}_{\mathcal{S}})\right| > t\right) \leq Ce^{-cm \max(t, t^2)},$$

for some  $C, c$  independent of  $N, m$ . Besides, from the proof of Theorem 5.1, we also have

$$\left|\frac{1}{m}\text{Tr}((\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)\mathbb{E}[\mathbf{Q}_{-i}^T]\mathbf{M}\bar{\mathbf{Q}}_m(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T\boldsymbol{\Phi}_{\mathcal{S}}) - \frac{1}{m}\text{Tr}((\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)\bar{\mathbf{Q}}_m^T\mathbf{M}\bar{\mathbf{Q}}_m(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T\boldsymbol{\Phi}_{\mathcal{S}})\right| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

as both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1,  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{A}\| \text{Tr}(\mathbf{B})$  for non-negative definite matrix  $\mathbf{B}$ , and from equation 30 that bounds  $\frac{1}{m}\text{Tr}(\boldsymbol{\Phi}_{\mathcal{S}})$ . From Lemma K.5, we have

$$\Pr\left(\frac{1}{m}\boldsymbol{\sigma}_i^T\hat{\mathbf{U}}_n\mathbf{Q}_{-i}(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T\boldsymbol{\sigma}_i - \alpha > t\right) \leq C'e^{-m c' \max(t, t^2)},$$

for some  $C', c'$  independent of  $N, m$ . From equation 49 in the proof of Theorem 5.1, we have

$$|\alpha - \delta| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

Combining all results above, we deduce that

$$\mathbb{E}(\|\mathbf{D}_2\|^2) = \mathbb{E}\left(\max_{1 \leq i \leq N}[\mathbf{D}_2^2]_i\right) = \int_0^\infty \Pr\left(\max_{1 \leq i \leq N}[\mathbf{D}_2^2]_i > t\right) dt = \mathcal{O}\left(\frac{1}{m}\right),$$

and therefore

$$\mathbb{E}[\|\mathbf{C}\mathbf{C}^T\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

□

**Lemma F.8.** Under Assumptions 1 and 2, let  $\mathbf{Z}_{23} \in \mathbb{R}^{n \times n}$  be the matrix defined in equation 61 as

$$\mathbf{Z}_{23} = \frac{1}{m} \frac{1}{1 + \delta} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \frac{\delta - \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i} \right].$$

Then,

$$\|\mathbf{Z}_{23}\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

*Proof.* Let the matrices

$$\mathbf{B}_i = m^{-\frac{1}{4}} \frac{1}{\sqrt{m}} \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i$$

and

$$\mathbf{C}_i^T = m^{\frac{1}{4}} \frac{1}{\sqrt{m}} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \frac{\delta - \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}{1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i}.$$

We decompose  $\mathbf{Z}_{23}$  with its symmetric and skew-symmetric parts as

$$\begin{aligned} \mathbf{Z}_{23} &= \frac{1}{1 + \delta} \sum_{i=1}^N \mathbb{E}[\mathbf{B}_i \mathbf{C}_i^T] \\ &= \frac{1}{1 + \delta} \sum_{i=1}^N \mathbb{E} \left[ \frac{\mathbf{B}_i \mathbf{C}_i^T + \mathbf{C}_i \mathbf{B}_i^T}{2} \right] + \frac{1}{1 + \delta} \mathbb{E} \left[ \sum_{i=1}^N \frac{\mathbf{B}_i \mathbf{C}_i^T - \mathbf{C}_i \mathbf{B}_i^T}{2} \right]. \end{aligned}$$

With the same reasoning on the symmetric part and on the skew-symmetric part than for equation 41, we get for the operator norm

$$\|\mathbf{Z}_{23}\| \leq \frac{1}{1+\delta} \left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \right] \right\| + \frac{1}{1+\delta} \left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{C}_i \mathbf{C}_i^T \right] \right\|.$$

We want to show that both  $\left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \right] \right\|$  and  $\left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{C}_i \mathbf{C}_i^T \right] \right\|$  vanish. We write  $\mathbb{E} \left[ \sum_{i=1}^N \mathbf{C}_i \mathbf{C}_i^T \right]$  as

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^N \mathbf{C}_i \mathbf{C}_i^T \right] &= \mathbb{E} \left[ \sum_{i=1}^N \frac{1}{\sqrt{m}} \mathbf{Q}_{-i}^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} \frac{(\delta - \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i)^2}{(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i)^2} \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^N \frac{1}{\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_m \left( \delta - \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \right)^2 \right] \\ &= \mathbb{E} \left[ \frac{1}{\sqrt{m}} \mathbf{Q}_m^T \hat{\mathbf{U}}_n^T \boldsymbol{\Sigma}_{\mathcal{S}}^T \mathbf{D}_3^2 \boldsymbol{\Sigma}_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right], \end{aligned}$$

where  $\mathbf{D}_3 \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$[\mathbf{D}_3]_i = \delta - \frac{1}{m} \boldsymbol{\sigma}_i^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i.$$

With a similar proof than for Lemma E.3, and from Lemma E.1 and Theorem 5.1, we find that

$$\mathbb{E} (\|\mathbf{D}_3\|^2) = \mathcal{O} \left( \frac{1}{m} \right).$$

From Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\mathcal{S}}^T \right\| \leq 2K'_Q.$$

We deduce thus

$$\left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{C}_i \mathbf{C}_i^T \right] \right\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

We write  $\mathbb{E} \left[ \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \right]$  as

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \right] &= \mathbb{E} \left[ \sum_{i=1}^N \frac{1}{m\sqrt{m}} \mathbf{Q}_{-i}^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M}^T \mathbf{Q}_{-i} \right] \\ &= \frac{1}{\sqrt{m}} \frac{N}{m} \mathbb{E} \left[ \mathbf{Q}_-^T \mathbf{M} \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_{\mathcal{S}} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) \bar{\mathbf{Q}}_m^T \mathbf{M}^T \mathbf{Q}_- \right], \end{aligned}$$

With a similar proof than for Lemma I.1, we can show there exists a real  $K_{\bar{Q}} > 0$  such that, for all  $m$ , we have

$$\|\bar{\mathbf{Q}}_m\| \leq K_{\bar{Q}}.$$

Let  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  be the empirical transition model matrix defined in equation 14. Under Assumption 2,  $\hat{\mathbf{A}}_m$  is invertible. From Lemma I.3, we have

$$\begin{aligned} \left\| \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_{\mathcal{S}} \right\| &= \left\| \bar{\mathbf{Q}}_m (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Phi}_{\mathcal{S}} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{A}}_m^{-1} \right\| \\ &= \left\| \frac{m}{N} (1 + \delta) [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{A}}_m^{-1} \right\| \\ &\leq 2 \frac{m}{N} \frac{1 + \delta}{\xi_{\min}} (1 + K_{\bar{Q}}). \end{aligned}$$

From above and from Lemma F.9 that upper bounds  $\|M\bar{Q}_m\|$ , we conclude that

$$\left\| \mathbb{E} \left[ \sum_{i=1}^N \mathbf{B}_i \mathbf{B}_i^T \right] \right\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right).$$

□

**Lemma F.9.** *Under Assumptions 1 and 2, let  $\bar{Q}_m \in \mathbb{R}^{n \times n}$  be the deterministic resolvent defined in equation 15, and let  $M$  be either any matrix with a bounded operator norm or  $M = \hat{U}_n^T \Psi_{\hat{S}} \hat{U}_n$ . Then there exists a real  $K > 0$  such that, for all  $m$ , we have*

$$\|M\bar{Q}_m\| \leq K.$$

*Proof.* With a similar proof than for Lemma I.1, we can show there exists a real  $K_{\bar{Q}} > 0$  such that, for all  $m$ , we have

$$\|\bar{Q}_m\| \leq K_{\bar{Q}}.$$

In the case where  $M$  is a matrix with a bounded operator norm, i.e.,  $\|M\| \leq K_M$  we have

$$\|M\bar{Q}_m\| \leq K_M K_{\bar{Q}}.$$

Otherwise, when  $M = \hat{U}_n^T \Psi_{\hat{S}} \hat{U}_n$ , we consider  $\hat{A}_m = \hat{U}_n (\hat{U}_n - \gamma \hat{V}_n)^T$  the empirical transition model matrix defined in equation 14. Under Assumption 2,  $\hat{A}_m$  is invertible. From Lemma I.3, we have

$$\begin{aligned} \|M\bar{Q}_m\| &= \left\| \frac{N}{m} \frac{1}{1+\delta} \hat{U}_n^T \Phi_{\hat{S}} \hat{U}_n \bar{Q}_m \right\| \\ &= \left\| \frac{N}{m} \frac{1}{1+\delta} \hat{U}_n^T \hat{A}_m^{-1} \hat{U}_n (\hat{U}_n - \gamma \hat{V}_n)^T \Phi_{\hat{S}} \hat{U}_n \bar{Q}_m \right\| \\ &= \left\| \hat{U}_n^T \hat{A}_m^{-1} \hat{U}_n [\mathbf{I}_n - \lambda \bar{Q}_m] \right\| \\ &\leq \frac{1}{\xi_{\min}} (1 + \lambda K_{\bar{Q}}). \end{aligned}$$

□

## G PROOF OF THEOREM 5.3

To simplify the notations, we denote the matrix  $Q_m$  as the resolvent  $Q_m(\lambda)$  (defined in equation 13), and we set  $p = |\mathcal{S}|$ . We define the matrices  $\Psi_{\hat{S}} \in \mathbb{R}^{m \times m}$  and  $\Psi_{\mathcal{S}} \in \mathbb{R}^{p \times p}$  as

$$\Psi_{\hat{S}} = \frac{N}{m} \frac{1}{1+\delta} \Phi_{\hat{S}} \quad \text{and} \quad \Psi_{\mathcal{S}} = \frac{N}{m} \frac{1}{1+\delta} \Phi_{\mathcal{S}}.$$

We also add the notation  $\mathbf{A} = \mathbf{B} + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right)$  which means that  $\|\mathbf{A} - \mathbf{B}\| = \mathcal{O} \left( \frac{1}{\sqrt{m}} \right)$ .

Under Assumptions 1, 2 and 3, this section is dedicated to find an asymptotic deterministic version of the true MSBE( $\hat{\theta}$ ) defined in equation 2 with a similar approach than the one used in Appendix F for  $\widehat{\text{MSBE}}(\hat{\theta})$ . In proofs of both Theorem 5.1 and Theorem 5.2, we constantly use the fact that  $\|Q_m\|$  is uniformly bounded by a constant  $K_Q > 0$ . We can also easily bound the operator norm of  $\frac{1}{m} \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{U}_n Q_m$  since the empirical transition model matrix  $\hat{A}_m = \hat{U}_n (\hat{U}_n - \gamma \hat{V}_n)^T$  (equation 14) is invertible under Assumption 2. Indeed,

$$\begin{aligned} \left\| \frac{1}{m} \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{U}_n Q_m \right\| &= \left\| \frac{1}{m} \hat{A}_m^{-1} \hat{U}_n (\hat{U}_n - \gamma \hat{V}_n)^T \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{U}_n Q_m \right\| = \left\| \hat{A}_m^{-1} \hat{U}_n [\mathbf{I}_n - \lambda Q_m] \right\| \\ &\leq \frac{1}{\xi_{\min}} (1 + \lambda K_Q) \end{aligned}$$

However, we do not have such simple control for  $\frac{1}{m}\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m$  since  $\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T$  is not invertible until all states are visited. Furthermore, from Corollary K.1.1, only a  $\mathcal{O}(\sqrt{m})$  upper bound can be derived for  $\left\|\frac{1}{\sqrt{m}}\Sigma_S\right\|$  or  $\left\|\frac{1}{\sqrt{p}}\Sigma_S\right\|$ . Nonetheless, with the additional Assumption 3, we can bound  $\frac{1}{m}\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P})\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m$  as stated by Lemma G.1. Fortunately for us, the proof of Theorem 5.3 indicates that controlling the operator norm of  $\frac{1}{m}\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P})\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m$  is sufficient.

**Lemma G.1.** *Under Assumptions 1, 2 and 3, there exists  $K > 0$  such that, for all  $m$ , we have*

$$\left\|\frac{1}{m}\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P})\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\| \leq K.$$

*Proof.* The main idea of the proof is to use a similar reasoning than for  $\left\|\frac{1}{m}\hat{\mathbf{A}}_m\Sigma_S^T\Sigma_S\mathbf{Q}_m\right\|$ . To this end, we use the triangular inequality, the decomposition  $\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P}) = \mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T + \mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P}) - \mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T$ , and Assumption 3 as follows:

$$\begin{aligned} \left\|\frac{1}{m}\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P})\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\| &\leq \left\|\frac{1}{m}\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\| \\ &\quad + \left\|\frac{1}{m}[\mathbf{D}_\pi(\mathbf{I}_p - \gamma\mathbf{P}) - \mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T]\Sigma_S^T\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\| \\ &\leq \underbrace{\left\|\mathbf{U}_n[\mathbf{I}_n - \lambda\mathbf{Q}_m]\right\|}_{=Z_1} \\ &\quad + \underbrace{\left\|\frac{1}{\sqrt{m}}\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T\Sigma_S^T - \frac{1}{\sqrt{m}}\mathbf{D}_\pi[\mathbf{I}_{|S|} - \gamma\mathbf{P}]\Sigma_S^T\right\|}_{=Z_2} \left\|\frac{1}{\sqrt{m}}\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\|. \end{aligned}$$

From Lemma I.1, we know there exists  $K_Q > 0$  such that, for all  $m$ , we have  $\|\mathbf{Q}_m\| \leq K_Q$ . Therefore, for the left-hand part  $Z_1$ , we have

$$\|Z_1\| = \|\mathbf{U}_n[\mathbf{I}_n - \lambda\mathbf{Q}_m]\| \leq 1 + \lambda K_Q.$$

From Assumption 3, for the right-hand part  $Z_2$ , we have

$$\begin{aligned} &\left\|\frac{1}{\sqrt{m}}\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T\Sigma_S^T - \frac{1}{\sqrt{m}}\mathbf{D}_\pi[\mathbf{I}_{|S|} - \gamma\mathbf{P}]\Sigma_S^T\right\| \\ &\leq \left\|\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T - \mathbf{D}_\pi[\mathbf{I}_{|S|} - \gamma\mathbf{P}]\right\| \left\|\frac{1}{\sqrt{m}}\Sigma_S\right\| \\ &= \mathcal{O}(1), \end{aligned}$$

since  $\|\Sigma_S\| = \mathcal{O}(m)$  from Corollary K.1.1. Furthermore, from Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\|\frac{1}{\sqrt{m}}\Sigma_S\mathbf{U}_n\mathbf{Q}_m\right\| = \left\|\frac{1}{\sqrt{m}}\Sigma_S\hat{\mathbf{U}}_n\mathbf{Q}_m\right\| \leq K'_Q.$$

We have  $\|Z_2\| = \mathcal{O}(1)$  which concludes the proof.  $\square$

Assumption 3 may hold for sufficiently large  $n$  since  $\mathbb{E}[\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T] \rightarrow \mathbf{D}_\pi[\mathbf{I}_p - \gamma\mathbf{P}]$  as  $n \rightarrow \infty$  (Tsitsiklis & Van Roy, 1996; Nedić & Bertsekas, 2003). Indeed, Tagorti & Scherrer (2015) has established that we can control  $\|\mathbf{U}_n(\mathbf{U}_n - \gamma\mathbf{V}_n)^T - \mathbf{D}_\pi[\mathbf{I}_{|S|} - \gamma\mathbf{P}]\|$  with a sufficiently large  $n$ .

With the additional Assumption 3, we can now present the following result on the asymptotic Mean-Squared Bellman error.



**Theorem G.2** (Asymptotic MSBE). *Under Assumptions 1, 2, and 3, the deterministic asymptotic MSBE is*

$$\overline{\text{MSBE}}(\hat{\theta}) = \left\| \bar{\mathbf{r}} + \gamma \frac{1}{\sqrt{n}} \mathbf{P} \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} - \frac{1}{\sqrt{n}} \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \mathbf{r} \right\|_{\mathbf{D}_\pi}^2 + \Delta,$$

with second-order correction factor

$$\Delta = \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m(\lambda)^T \Psi_1 \bar{\mathbf{Q}}_m(\lambda))} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|_{\Psi_1}^2,$$

where

$$\begin{aligned} \Lambda_P &= [\mathbf{I}_{|S|} - \gamma \mathbf{P}]^T \mathbf{D}_\pi [\mathbf{I}_{|S|} - \gamma \mathbf{P}], \\ \Theta_S &= \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda). \end{aligned}$$

As  $N, m, d \rightarrow \infty$  with asymptotic constant ratio  $N/m$ ,  $\text{MSBE}(\hat{\theta}) - \overline{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0$ .

*Proof.* We decompose  $\text{MSBE}(\hat{\theta})$  as

$$\text{MSBE}(\hat{\theta}) = \|\bar{\mathbf{r}} + \gamma \mathbf{P} \Sigma_S^T \hat{\theta} - \Sigma_S^T \hat{\theta}\|_{\mathbf{D}_\pi}^2 = \|\bar{\mathbf{r}} + [\gamma \mathbf{P} - \mathbf{I}_p] \Sigma_S^T \hat{\theta}\|_{\mathbf{D}_\pi}^2 \quad (64)$$

$$= \left\| \bar{\mathbf{r}} - \frac{1}{m\sqrt{n}} [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right\|_{\mathbf{D}_\pi}^2 \quad (65)$$

$$= \|\bar{\mathbf{r}}\|_{\mathbf{D}_\pi}^2 \quad (66)$$

$$- \underbrace{\frac{2}{m\sqrt{n}} \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r}}_{=Z_2} \quad (67)$$

$$+ \underbrace{\left\| \frac{1}{m\sqrt{n}} [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right\|_{\mathbf{D}_\pi}^2}_{=Z_3}. \quad (68)$$

We want to find an asymptotic equivalent for both  $Z_2$  and  $Z_3$ . From Lemma G.3, we have

$$\mathbb{E}[Z_2] = \frac{2}{\sqrt{n}} \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

For  $Z_3$ , we have

$$Z_3 = \left\| \frac{1}{m\sqrt{n}} [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right\|_{\mathbf{D}_\pi}^2 = \frac{1}{nm^2} \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r}.$$

From Lemma G.4, we have

$$\begin{aligned} \mathbb{E}[Z_3] &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \Psi_S \Lambda_P \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \\ &+ \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 \\ &+ \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

With a similar proof than for Lemma K.6 we can deduce that

$$\text{MSBE}(\hat{\theta}) - \overline{\text{MSBE}}(\hat{\theta}) \xrightarrow{a.s.} 0,$$

as  $m \rightarrow \infty$ . □

**Lemma G.3.** *Under Assumptions 1, 2 and 3, let  $Z_2 \in \mathbb{R}$  defined in equation 67 as*

$$Z_2 = \frac{1}{m\sqrt{n}} \mathbb{E}[\bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r}].$$

Then

$$\left| Z_2 - \frac{1}{\sqrt{n}} \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \right| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

for  $\bar{\mathbf{Q}}_m$  the deterministic resolvent defined in equation 15, and  $\Psi_S \in \mathbb{R}^{p \times p}$  defined in equation 21.

*Proof.* As in equation 32, we decompose the matrix  $\Sigma_S^T \Sigma_S$  as

$$\Sigma_S^T \Sigma_S = \sum_{i=1}^N \sigma_i \sigma_i^T,$$

where  $\sigma_i = \sigma(\mathbf{S}^T \mathbf{w}_i) \in \mathbb{R}^m$  for which  $\mathbf{w}_i \in \mathbb{R}^d$  denotes the  $i$ -th row of  $\mathbf{W}$  defined in equation 5. Let  $\mathbf{Q}_{-i} \in \mathbb{R}^{n \times n}$  be the following resolvent

$$\mathbf{Q}_{-i} = \left[ \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Sigma_S^T \Sigma_S \mathbf{U}_n - \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i \sigma_i^T \mathbf{U}_n + \lambda \mathbf{I}_n \right]^{-1},$$

independent of  $\sigma_i$  and thus  $\mathbf{w}_i$ . From the Sherman identity (Lemma L.4), we have

$$\begin{aligned} Z_2 &= \frac{1}{m\sqrt{n}} \mathbb{E} \left[ \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right] \\ &= \frac{1}{m\sqrt{n}} \mathbb{E} \left[ \sum_{i=1}^N \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right] \\ &= \frac{1}{m\sqrt{n}} \mathbb{E} \left[ \sum_{i=1}^N \frac{\bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} \mathbf{r}}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i} \right]. \end{aligned}$$

Let  $\mathbf{D} \in \mathbb{R}^{N \times N}$  be a diagonal matrix for which, for all  $i \in [N]$ , we have

$$\mathbf{D}_i = \delta - \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i.$$

We replace  $1 + \frac{1}{m} \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j$  by  $1 + \delta$  as following

$$\begin{aligned} Z_2 &= \frac{1}{m\sqrt{n}} \frac{1}{1 + \delta} \mathbb{E} \left[ \underbrace{\sum_{i=1}^N \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} \mathbf{r}}_{Z_{21}} \right] \\ &\quad + \frac{1}{m\sqrt{n}} \frac{1}{1 + \delta} \mathbb{E} \left[ \underbrace{\sum_{i=1}^N \frac{\bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} \mathbf{D}_i \mathbf{r}}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i}}_{Z_{22}} \right]. \end{aligned}$$

We have  $Z_{22}$  vanishing since  $\mathbb{E}[\|\mathbf{D}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$  and from Lemma G.1. From Theorem 5.1, we have thus

$$\begin{aligned} Z_2 &= \frac{1}{m\sqrt{n}} \frac{1}{1 + \delta} \mathbb{E} \left[ \sum_{i=1}^N \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} \mathbf{r} \right] + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\ &= \frac{1}{\sqrt{n}} \frac{N}{m} \frac{1}{1 + \delta} \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Phi_S \mathbf{U}_n \mathbb{E}[\mathbf{Q}_{-}] \mathbf{r} + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\ &= \frac{1}{\sqrt{n}} \bar{\mathbf{r}}^T \mathbf{D}_\pi [\mathbf{I}_p - \gamma \mathbf{P}] \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

□

**Lemma G.4.** Under Assumptions 1, 2 and 3, let  $\Lambda_P \in \mathbb{R}^{p \times p}$  be the matrix defined in equation 21, and let  $Z_3 \in \mathbb{R}$  be defined in equation 68 as

$$Z_3 = \mathbb{E} \left[ \frac{1}{nm^2} \bar{\mathbf{r}}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \Sigma_S^T \Sigma_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \right].$$

Then

$$\left| Z_3 - \frac{1}{n} \bar{\mathbf{r}}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \Psi_S \Lambda_P \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} - \frac{1}{n} \frac{\frac{1}{N} \text{Tr} \left( \Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S] \right)}{1 - \frac{1}{N} \text{Tr} (\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 \right| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

where  $\bar{\mathbf{Q}}_m$  is the deterministic resolvent defined in equation 15,  $\Psi_1, \Psi_2 \in \mathbb{R}^{n \times n}$  are defined in equation 19,  $\Psi_S \in \mathbb{R}^{p \times p}$  and  $\Theta_S \in \mathbb{R}^{p \times n}$  are defined in equation 21.

*Proof.* As in equation 32, we decompose the matrix  $\Sigma_S^T \Sigma_S$  as

$$\Sigma_S^T \Sigma_S = \sum_{i=1}^N \sigma_i \sigma_i^T,$$

where  $\sigma_i = \sigma(\mathbf{S}^T \mathbf{w}_i) \in \mathbb{R}^m$  for which  $\mathbf{w}_i \in \mathbb{R}^d$  denotes the  $i$ -th row of  $\mathbf{W}$  defined in equation 5. Let  $\mathbf{Q}_{-i} \in \mathbb{R}^{n \times n}$  be the following resolvent

$$\mathbf{Q}_{-i} = \left[ \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Sigma_S^T \Sigma_S \mathbf{U}_n - \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i \sigma_i^T \mathbf{U}_n + \lambda \mathbf{I}_n \right]^{-1}$$

independent of  $\sigma_i$  and thus  $\mathbf{w}_i$ . From the Sherman identity (Lemma L.4), we decompose  $Z_3$  as

$$Z_3 = \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \Sigma_S^T \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right] \quad (69)$$

$$= \sum_{i,j=1}^N \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right] \quad (70)$$

$$= \sum_{i,j=1}^N \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \sigma_i \sigma_i^T}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i} \Lambda_P \frac{\sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{1 + \frac{1}{m} \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j} \mathbf{r} \right] \quad (71)$$

$$= \underbrace{\sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \sigma_i \sigma_i^T}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i} \Lambda_P \frac{\sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{1 + \frac{1}{m} \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j} \mathbf{r} \right]}_{=Z_{31}} + \underbrace{\sum_{i=1}^N \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i}}{\left(1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i\right)^2} \mathbf{r} \right]}_{=Z_{32}}. \quad (72)$$

From Lemma G.5, we have

$$\begin{aligned} Z_{31} &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \Psi_S \Lambda_P \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \\ &\quad + \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 \\ &\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

For the second term  $Z_{32}$ , we have from Lemma G.6,

$$Z_{32} = \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P \Psi_S)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

We conclude that

$$\begin{aligned} Z_3 &= Z_{31} + Z_{32} \\ &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \Psi_S \Lambda_P \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \\ &\quad + \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S])}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 \\ &\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

□

**Lemma G.5.** Under Assumptions 1, 2 and 3, let  $Z_{31} \in \mathbb{R}$  defined in equation 72 as

$$Z_{31} = \frac{1}{nm^2} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \frac{\mathbf{r}^T \mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{r}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right) \left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)} \right],$$

Then

$$\left| Z_{31} - \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \boldsymbol{\Psi}_S \boldsymbol{\Lambda}_P \boldsymbol{\Psi}_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} - \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\boldsymbol{\Lambda}_P [\boldsymbol{\Theta}_S \boldsymbol{\Psi}_2 \boldsymbol{\Theta}_S^T - 2\boldsymbol{\Theta}_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_S])}{1 - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\boldsymbol{\Psi}_1}^2} \right| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

where  $\bar{\mathbf{Q}}_m$  is the deterministic resolvent defined in equation 15,  $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2 \in \mathbb{R}^{n \times n}$  are defined in equation 19,  $\boldsymbol{\Psi}_S \in \mathbb{R}^{p \times p}$  and  $\boldsymbol{\Theta}_S \in \mathbb{R}^{p \times n}$  are defined in equation 21.

*Proof.* Using the Sherman identity (Lemma L.4), we decompose  $Z_{31}$  as

$$\begin{aligned} Z_{31} &= \frac{1}{nm^2} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \frac{\mathbf{r}^T \mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{r}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right) \left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)} \right] \\ &= \frac{1}{nm^2} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j} \mathbf{r} \right] \\ &= \frac{1}{nm^2} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j} \mathbf{r}}_{=Z_{311}} \right] \\ &\quad - \frac{1}{nm^3} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)^2} \mathbf{r}}_{=Z_{312}} \right]. \end{aligned}$$

We want to find an asymptotic equivalent for both  $Z_{311}$  and  $Z_{312}$ . For  $Z_{312}$ , we have

$$\begin{aligned} Z_{312} &= \frac{1}{nm^3} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)^2} \mathbf{r} \right] \\ &= \frac{1}{nm} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \left(\frac{1}{m^2} \boldsymbol{\sigma}_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j\right)}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)^2} \mathbf{r} \right], \end{aligned}$$

where  $\boldsymbol{\Sigma}_S^{-j} = \sigma(\mathbf{W}^{-j} \mathbf{S}) \in \mathbb{R}^{(N-1) \times n}$  for which  $\mathbf{W}^{-j} \in \mathbb{R}^{(N-1) \times d}$  depicts the same matrix than the weight matrix  $\mathbf{W}$  defined in equation 5 without the  $j^{\text{th}}$  row. Let  $\mathbf{D}_{312} \in \mathbb{R}^{N \times N}$  be a diagonal matrix for which, for all  $j \in [N]$ , we have

$$[\mathbf{D}_{312}]_j = \frac{1}{m^2} \boldsymbol{\sigma}_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_j - \frac{1}{m^2} \text{Tr} \left( (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S \right).$$

From Lemma G.1, we know there exists a real  $K_1 > 0$  such that, for all  $m$ , we have  $\|\mathbf{D}_\pi[\mathbf{I}_p - \gamma \mathbf{P}] \boldsymbol{\Sigma}_S^T \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m\| \leq K_1$ . Therefore, we deduce that

$$\left\| \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \right\| = \mathcal{O}(1).$$

From Lemma K.2, we deduce that  $\mathbb{E}[\|D_{312}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ . Therefore, we get

$$\begin{aligned}
Z_{312} &= \frac{1}{nm} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S)}{(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j)^2} \mathbf{r} \right] \\
&+ \frac{1}{nm} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} [D_{312}]_j}{(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j)^2} \mathbf{r} \right] \\
&= \frac{1}{nm} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S)}{(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j)^2} \mathbf{r} \right] \\
&+ \frac{1}{nm} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^T D_{312} \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right] \\
&= \frac{1}{nm} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S)}{(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j)^2} \mathbf{r} \right] \\
&+ \mathcal{O}\left(\frac{1}{\sqrt{m}}\right), \tag{73}
\end{aligned}$$

where the last equality is obtained since  $\mathbb{E}[\|D_{312}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , and since we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Sigma}_S^T \right\| \leq 2K'_Q.$$

from Lemma I.4. We replace  $1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j$  by  $1 + \delta$  in  $Z_{312}$  as following

$$\begin{aligned}
Z_{312} &= \frac{1}{nm^3} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S) \mathbf{r} \right] \\
&+ \frac{1}{nm^3} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\sigma}_j \mathbf{D}'_j \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-jT} \boldsymbol{\Sigma}_S^{-j} \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S)}{(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j)^2} \mathbf{r} \right] \\
&+ \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\
&= \underbrace{\frac{N}{nm^3} \frac{1}{(1 + \delta)^2} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_-^T \mathbf{U}_n^T \boldsymbol{\Phi}_S \mathbf{U}_n \mathbf{Q}_- \mathbf{r} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-T} \boldsymbol{\Sigma}_S^- \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S) \right]}_{=Z_{3121}} \\
&+ \underbrace{\frac{1}{nm^3} \frac{1}{(1 + \delta)^2} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^T \mathbf{D}' \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^{-T} \boldsymbol{\Sigma}_S^- \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S) \right]}_{=Z_{3122}} \\
&+ \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),
\end{aligned}$$

where  $\mathbf{D}' \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $j \in [N]$ , we have

$$\mathbf{D}'_j = (1 + \delta)^2 - \left(1 + \frac{1}{m} \boldsymbol{\sigma}_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_j\right)^2.$$

With a similar proof than for equation 30, we can show  $\frac{1}{m} \text{Tr}(\Phi_S) = \frac{p}{m} \frac{1}{p} \text{Tr}(\Phi_S)$  is uniformly bounded under Assumption 3. Combining  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{A}\| \text{Tr}(\mathbf{B})$  for non-negative definite matrix  $\mathbf{B}$  and Lemma G.1, we have  $\frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S) = \mathcal{O}(1)$ . From all these upper bounds, and since it can be shown that  $\mathbb{E}[\|\mathbf{D}'\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , we deduce the second term,  $Z_{3122}$ , vanishes and thus

$$Z_{312} = \frac{1}{nm^2} \frac{1}{1+\delta} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_-^T \Psi_1 \mathbf{Q}_- \mathbf{r} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S) \right] + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

Let  $\mathbf{Q}_{-ij} \in \mathbb{R}^{n \times n}$  be the resolvent defined as

$$\mathbf{Q}_{-ij} = \left[ \frac{1}{m} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Sigma_S^{-ijT} \Sigma_S^{-ij} \mathbf{U}_n + \lambda \mathbf{I}_n \right]^{-1}, \quad (74)$$

where  $\Sigma_S^{-ij} = \sigma(\mathbf{W}^{-ij} \mathbf{S}) \in \mathbb{R}^{(N-2) \times n}$  for which  $\mathbf{W}^{-ij} \in \mathbb{R}^{(N-2) \times d}$  depicts the same matrix than the weight matrix  $\mathbf{W}$  defined in equation 5 without the  $i^{\text{th}}$  and  $j^{\text{th}}$  row. Using the Sherman identity (Lemma L.4), the term  $\frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S)$  in  $Z_{312}$  can be rewritten as

$$\begin{aligned} & \frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S) \\ &= \frac{1}{m^2} \text{Tr} \left( \sum_{i \neq j} (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S \right) \\ &= \frac{1}{m^2} \text{Tr} \left( \sum_{i \neq j} \frac{(\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i} \right) \\ &= \frac{1}{m^2} \frac{1}{1+\delta} \text{Tr} \left( \sum_{i \neq j} (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S \right) \\ &+ \frac{1}{m^2} \frac{1}{1+\delta} \text{Tr} \left( \sum_{i \neq j} \frac{(\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S (\delta - \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i)}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i} \right) \\ &= \frac{N}{m^2} \frac{1}{1+\delta} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{--}^T \mathbf{U}_n^T \Phi_S \Lambda_P \Phi_S) \\ &+ \frac{1}{m^2} \frac{1}{1+\delta} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{-j}^T \mathbf{U}_n^T \Sigma_S^{-jT} \mathbf{D} \Sigma_S^{-j} \Lambda_P \Phi_S), \end{aligned}$$

where  $\mathbf{D} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$D_i = \delta - \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i.$$

From the uniform boundness of  $\frac{1}{m} \text{Tr}(\Phi_S) = \frac{1}{K_r} \frac{1}{p} \text{Tr}(\Phi_S)$ , from Lemma G.1, we have  $\frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S) = \mathcal{O}(1)$ . Since the operator norm of  $\mathbb{E}[\|\mathbf{D}\|]$  is of order  $\mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , we deduce the second term vanishes, and thus

$$\begin{aligned} & \frac{1}{m^2} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_-^T \mathbf{U}_n^T \Sigma_S^{-T} \Sigma_S^- \Lambda_P \Phi_S) \\ &= \frac{N}{m^2} \frac{1}{1+\delta} \text{Tr}((\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_{--}^T \mathbf{U}_n^T \Phi_S \Lambda_P \Phi_S) + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

Applying Lemma F.4 and Lemma F.5, we deduce for  $Z_{312}$  that

$$\begin{aligned} Z_{312} &= \frac{1}{n} \frac{\frac{1}{N} \text{Tr}((U_n - \gamma V_n) \bar{Q}_m^T U_n^T \Psi_S \Lambda_P \Psi_S)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{Q}_m^T \Psi_1 \bar{Q}_m)} \|\bar{Q}_m \mathbf{r}\|_{\Psi_1}^2 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\ &= \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Psi_S (U_n - \gamma V_n) \Theta_S^T \Lambda_P)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{Q}_m^T \Psi_1 \bar{Q}_m)} \|\bar{Q}_m \mathbf{r}\|_{\Psi_1}^2 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right). \end{aligned}$$

Now, we want to find an asymptotic equivalent for  $Z_{311}$ . We replace  $1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j$  by  $1 + \delta$  in  $Z_{311}$  as following

$$\begin{aligned} Z_{311} &= \frac{1}{nm^2} \sum_{i=1}^N \sum_{j \neq i} \mathbb{E} \left[ \mathbf{r}^T \frac{Q_{-j}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \sigma_j \sigma_j^T U_n Q_{-j}}{1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j} \mathbf{r} \right] \\ &= \frac{1}{nm^2} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{Q_{-j}^T U_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T U_n Q_{-j}}{1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j} \mathbf{r} \right] \\ &= \frac{1}{nm^2} \frac{1}{1 + \delta} \sum_{j=1}^N \mathbb{E} \left[ \underbrace{\mathbf{r}^T Q_{-j}^T U_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T U_n Q_{-j} \mathbf{r}}_{=Z_{3111}} \right] \\ &\quad + \frac{1}{nm^2} \frac{1}{1 + \delta} \sum_{j=1}^N \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{Q_{-j}^T U_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T U_n Q_{-j} D_j}{1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j} \mathbf{r}}_{=Z_{3112}} \right], \end{aligned}$$

where  $D \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $j \in [N]$ , we have

$$D_j = \delta - \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j.$$

We observe that

$$\begin{aligned} Q_{-j}^T U_n^T \frac{\Sigma_S^{-jT} \Sigma_S^{-j}}{m} &= Q_{-j}^T U_n^T \frac{\Sigma_S^T \Sigma_S}{m} - Q_{-j}^T U_n^T \frac{\sigma_j \sigma_j^T}{m} \\ &= Q_m^T U_n^T \frac{\Sigma_S^T \Sigma_S}{m} + \frac{\frac{1}{m} Q_{-j}^T U_n^T \sigma_j \sigma_j^T (U_n - \gamma V_n) Q_{-j}^T U_n^T \Sigma_S^T \Sigma_S}{1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j} - Q_{-j}^T U_n^T \frac{\sigma_j \sigma_j^T}{m} \\ &= Q_m^T U_n^T \frac{\Sigma_S^T \Sigma_S}{m} \\ &\quad + Q_m^T U_n^T \sigma_j \sigma_j^T (U_n - \gamma V_n) Q_m^T \left( 1 + \frac{1}{m} \sigma_j^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_j \right) U_n^T \frac{\Sigma_S^T \Sigma_S}{m} \\ &\quad - Q_{-j}^T U_n^T \frac{\sigma_j \sigma_j^T}{m} \\ &= Q_m^T U_n^T \frac{\Sigma_S^T \Sigma_S}{m} + \frac{Q_m^T U_n^T \sigma_j \sigma_j^T (U_n - \gamma V_n) Q_m^T U_n^T \Sigma_S^T \Sigma_S}{1 - \frac{1}{m} \sigma_j^T U_n Q_m (U_n - \gamma V_n)^T \sigma_j} - Q_{-j}^T U_n^T \frac{\sigma_j \sigma_j^T}{m}. \end{aligned}$$

From above, we expand  $Z_{3112}$  as

$$\begin{aligned}
Z_{3112} &= \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T \mathbf{U}_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{D}_j}{1 + \frac{1}{m} \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j} \mathbf{r} \right] \\
&= \frac{1}{nm^2} \frac{1}{1+\delta} \mathbb{E} \left[ \underbrace{\mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \Sigma_S^T \mathbf{D} \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r}}_{=Z_{31121}} \right] \\
&+ \sum_{j=1}^N \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{\mathbf{Q}_m^T \mathbf{U}_n^T \sigma_j \sigma_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_m \mathbf{D}_j}{nm^2(1+\delta) \left(1 - \frac{1}{m} \sigma_j^T \mathbf{U}_n^T \mathbf{Q}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j\right)} \mathbf{r}}_{=Z_{31122}} \right] \\
&- \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \mathbb{E} \left[ \underbrace{\mathbf{r}^T \mathbf{Q}_{-j}^T \mathbf{U}_n^T \sigma_j \sigma_j^T \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_m \mathbf{D}_j \left(1 + \frac{1}{m} \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j\right)}_{=Z_{31123}} \mathbf{r} \right].
\end{aligned}$$

We have  $Z_{31121} = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$  since  $\mathbb{E}[\|\mathbf{D}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$  and from Lemma G.1. Subsequently, we rewrite  $Z_{31122}$  as

$$Z_{31122} = \frac{1}{nm} \frac{1}{1+\delta} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \mathbf{D}_{31122} \Sigma_S \mathbf{U}_n \mathbf{Q}_m \mathbf{r} \right],$$

with  $\mathbf{D}_{31122} \in \mathbb{R}^{N \times N}$  a diagonal matrix for which, for all  $j \in [N]$ , we have

$$[\mathbf{D}_{31122}]_j = \frac{1}{m} \frac{\mathbf{D}_j \sigma_j^T (\mathbf{U}_n - \gamma \mathbf{V}_n) \mathbf{Q}_m^T \mathbf{U}_n^T \Sigma_S^T \Sigma_S \Lambda_P \sigma_j}{1 - \frac{1}{m} \sigma_j^T \mathbf{U}_n^T \mathbf{Q}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_j}.$$

It can be shown that  $\mathbb{E}[\|\mathbf{D}_{31122}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , and we can deduce  $Z_{31122} = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ . Similarly, we have  $Z_{31123} = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ .  $Z_{3112}$  vanishes, and thus

$$\begin{aligned}
Z_{311} &= Z_{3111} + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\
&= \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T \mathbf{U}_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{r} \right] + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).
\end{aligned}$$

It remains to handle  $Z_{3111}$  for which we have from the Sherman identity (Lemma L.4),

$$\begin{aligned}
Z_{3111} &= \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T \mathbf{U}_n^T \Sigma_S^{-jT} \Sigma_S^{-j} \Lambda_P \sigma_j \sigma_j^T \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{r} \right] \\
&= \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S \mathbf{U}_n \mathbf{Q}_{-j} \mathbf{r} \right] \\
&= \frac{1}{nm^2} \frac{1}{1+\delta} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{\mathbf{Q}_{-ij}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S \mathbf{U}_n \mathbf{Q}_{-ij}}{1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i}}_{=Z_{31111}} \mathbf{r} \right] \\
&- \frac{1}{nm^3} \frac{1}{1+\delta} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \underbrace{\mathbf{r}^T \frac{\mathbf{Q}_{-ij}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij}}{\left(1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i\right)^2}}_{=Z_{31112}} \mathbf{r} \right].
\end{aligned}$$



Again, we replace  $1 + \frac{1}{m} \sigma_i^T U_n Q_{-ij} (U_n - \gamma V_n)^T \sigma_i$  by  $1 + \delta$  in  $Z_{31111}$  as following

$$\begin{aligned}
Z_{31111} &= \frac{1}{nm^2} \frac{1}{1 + \delta} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-ij}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S U_n Q_{-ij}}{1 + \frac{1}{m} \sigma_i^T U_n Q_{-ij} (U_n - \gamma V_n)^T \sigma_i} \mathbf{r} \right] \\
&= \frac{1}{nm^2} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-ij}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S U_n Q_{-ij} \mathbf{r} \right] \\
&\quad + \frac{1}{nm^2} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-ij}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S U_n Q_{-ij} \mathbf{D}_i}{1 + \frac{1}{m} \sigma_i^T U_n Q_{-ij} (U_n - \gamma V_n)^T \sigma_i} \mathbf{r} \right] \\
&= \frac{1}{n} \frac{N^2}{m^2} \frac{1}{(1 + \delta)^2} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{--}^T U_n^T \Phi_S \Lambda_P \Phi_S U_n \mathbf{Q}_{--} \mathbf{r} \right] \\
&\quad + \frac{1}{nm^2} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S U_n Q_{-j} \mathbf{D}_i \mathbf{r} \right] \\
&\quad + \frac{1}{nm^3} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-j}^T U_n^T \sigma_i \sigma_i^T \Lambda_P \Phi_S U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_i \sigma_i^T U_n Q_{-j}}{1 - \frac{1}{m} \sigma_i^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_i} \mathbf{D}_i \mathbf{r} \right] \\
&= \frac{1}{n} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{--}^T U_n^T \Psi_S \Lambda_P \Psi_S U_n \mathbf{Q}_{--} \mathbf{r} \right] \\
&\quad + \frac{1}{nm^2} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T U_n^T \Sigma_S^{-jT} \mathbf{D} \Sigma_S^{-j} \Lambda_P \Phi_S U_n Q_{-j} \mathbf{r} \right] \\
&\quad + \frac{1}{nm} \frac{1}{(1 + \delta)^2} \sum_{j=1}^N \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-j}^T U_n^T \Sigma_S^{-jT} \mathbf{D}_{31111} \Sigma_S^j U_n Q_{-j} \mathbf{r} \right],
\end{aligned}$$

where  $\mathbf{D}_{31111} \in \mathbb{R}^{N \times N}$  is a diagonal matrix for which, for all  $i \in [N]$ , we have

$$[\mathbf{D}_{31111}]_i = \frac{1}{m} \frac{\sigma_i^T \Lambda_P \Phi_S U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_i (\delta - \frac{1}{m} \sigma_i^T U_n Q_{-ij} (U_n - \gamma V_n)^T \sigma_i)}{1 - \frac{1}{m} \sigma_i^T U_n Q_{-j} (U_n - \gamma V_n)^T \sigma_i},$$

and  $\mathbf{Q}_{--}$  is a resolvent with the same law than  $\mathbf{Q}_{-ij}$ . With similar arguments that before, we can show that  $\mathbb{E}[\|\mathbf{D}\|]$  and  $\mathbb{E}[\|\mathbf{D}_{31111}\|]$  are of order  $\mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , and therefore

$$Z_{31111} = \frac{1}{n} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{--}^T U_n^T \Psi_S \Lambda_P \Psi_S U_n \mathbf{Q}_{--} \mathbf{r} \right] + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

By extending Lemma F.2 to the matrix  $\Lambda'_P = U_n^T \Psi_S \Lambda_P \Psi_S U_n$ , and from Lemma F.4 we obtain

$$\begin{aligned}
Z_{31111} &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T U_n^T \Psi_S \Lambda_P \Psi_S U_n \bar{\mathbf{Q}}_m \mathbf{r} \\
&\quad + \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P \Theta_S \Psi_2 \Theta_S^T)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).
\end{aligned}$$

Following the same reasoning for  $Z_{31112}$ , and from Lemma K.2, we have

$$\begin{aligned}
Z_{31112} &= \frac{1}{nm^3} \frac{1}{1+\delta} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-ij}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-ij}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \mathbf{r} \right] \\
&= \frac{1}{nm^3} \frac{1}{(1+\delta)^3} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-ij} \mathbf{r} \right] \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\
&= \frac{1}{nm^2} \frac{1}{(1+\delta)^3} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-ij} \mathbf{r} \left( \frac{1}{m} \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i \right) \right] \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\
&= \frac{1}{nm^2} \frac{1}{(1+\delta)^3} \sum_{j=1}^N \sum_{i \neq j} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{-ij}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-ij} \mathbf{r} \frac{1}{m} \text{Tr}(\boldsymbol{\Lambda}_P \boldsymbol{\Phi}_S \mathbf{U}_n \mathbf{Q}_{-ij} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Phi}_S) \right] \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right) \\
&= \frac{1}{n} \frac{1}{N} \mathbb{E} \left[ \mathbf{r}^T \mathbf{Q}_{--}^T \boldsymbol{\Psi}_1 \mathbf{Q}_{--} \mathbf{r} \text{Tr}(\boldsymbol{\Lambda}_P \boldsymbol{\Psi}_S \mathbf{U}_n \mathbf{Q}_{--} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_S) \right] \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),
\end{aligned}$$

where the last equality is obtained with a similar reasoning than for equation 73. From Lemma F.4 and Lemma F.5, we have

$$Z_{31112} = \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\boldsymbol{\Lambda}_P \boldsymbol{\Theta}_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_S)}{1 - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\boldsymbol{\Psi}_1}^2 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

We conclude for  $Z_{311}$  that

$$\begin{aligned}
Z_{311} &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \boldsymbol{\Psi}_S \boldsymbol{\Lambda}_P \boldsymbol{\Psi}_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \\
&\quad + \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\boldsymbol{\Lambda}_P [\boldsymbol{\Theta}_S \boldsymbol{\Psi}_2 \boldsymbol{\Theta}_S^T - \boldsymbol{\Theta}_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_S])}{1 - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\boldsymbol{\Psi}_1}^2 \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),
\end{aligned}$$

and for  $Z_{31}$  that

$$\begin{aligned}
Z_{31} &= \frac{1}{n} \mathbf{r}^T \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \boldsymbol{\Psi}_S \boldsymbol{\Lambda}_P \boldsymbol{\Psi}_S \mathbf{U}_n \bar{\mathbf{Q}}_m \mathbf{r} \\
&\quad + \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\boldsymbol{\Lambda}_P [\boldsymbol{\Theta}_S \boldsymbol{\Psi}_2 \boldsymbol{\Theta}_S^T - 2\boldsymbol{\Theta}_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Psi}_S])}{1 - \frac{1}{N} \text{Tr}(\boldsymbol{\Psi}_2 \bar{\mathbf{Q}}_m^T \boldsymbol{\Psi}_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\boldsymbol{\Psi}_1}^2 \\
&\quad + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).
\end{aligned}$$

□

**Lemma G.6.** Under Assumptions 1, 2 and 3, let  $Z_{32} \in \mathbb{R}$  defined in equation 72 as

$$Z_{32} = \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{nm^2} \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \boldsymbol{\Lambda}_P \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \mathbf{r} \right].$$

Then

$$\left| Z_{32} - \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P \Psi_S)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{Q}_m^T \Psi_1 \bar{Q}_m)} \|\bar{Q}_m \mathbf{r}\|_{\Psi_1}^2 \right| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

where  $\bar{Q}_m$  is the deterministic resolvent defined in equation 15,  $\Psi_1, \Psi_2 \in \mathbb{R}^{n \times n}$  are defined in equation 19, and  $\Psi_S \in \mathbb{R}^{p \times p}$  is defined in equation 21.

*Proof.* We decompose  $Z_{32}$  as

$$\begin{aligned} Z_{32} &= \frac{1}{nm^2} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \Lambda_P \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \mathbf{r} \right] \\ &= \frac{1}{nm} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} \frac{1}{m} \text{Tr}(\Phi_S \Lambda_P)}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \mathbf{r} \right] \\ &\quad + \frac{1}{nm} \sum_{i=1}^N \mathbb{E} \left[ \mathbf{r}^T \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} \frac{1}{m} (\boldsymbol{\sigma}_i^T \Lambda_P \boldsymbol{\sigma}_i - \text{Tr}(\Phi_S \Lambda_P))}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \mathbf{r} \right] \\ &= \frac{1}{n} \frac{\text{Tr}(\Phi_S \Lambda_P)}{m} \mathbf{r}^T \underbrace{\sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i}}{\left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2} \right]}_{=Z_{321}} \mathbf{r} \\ &\quad + \frac{1}{n} \mathbf{r}^T \underbrace{\sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_m \frac{1}{m} (\boldsymbol{\sigma}_i^T \Lambda_P \boldsymbol{\sigma}_i - \text{Tr}(\Phi_S \Lambda_P)) \right]}_{Z_{322}} \mathbf{r}. \end{aligned}$$

We want to show  $Z_{322}$  vanishes and find an asymptotic equivalent for  $Z_{321}$ . Let  $\mathbf{D}_{322} \in \mathbb{R}^{N \times N}$  be a diagonal matrix for which, for all  $i \in [N]$ , we have

$$[\mathbf{D}_{322}]_i = \frac{1}{m} \boldsymbol{\sigma}_i^T \Lambda_P \boldsymbol{\sigma}_i - \frac{1}{m} \text{Tr}(\Phi_S \Lambda_P).$$

We rewrite  $Z_{322}$  as

$$\begin{aligned} Z_{322} &= \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_m \frac{1}{m} (\boldsymbol{\sigma}_i^T \Lambda_P \boldsymbol{\sigma}_i - \text{Tr}(\Phi_S \Lambda_P)) \right] \\ &= \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \mathbf{U}_n^T \boldsymbol{\Sigma}_S^T \mathbf{D}_{322} \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m \right] \end{aligned}$$

From Lemma I.4, we know there exists a real  $K'_Q > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_S \mathbf{U}_n \mathbf{Q}_m \right\| \leq K'_Q$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\Sigma}_S^T \right\| \leq 2K'_Q.$$

Using Lemma K.2 we show that  $\mathbb{E}[\|\mathbf{D}_{322}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ , and we deduce that

$$\|\mathbf{Z}_{322}\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

We want to find an asymptotic equivalent for  $Z_{321}$ . Let  $\mathbf{D}_{321} \in \mathbb{R}^{N \times N}$  be a diagonal matrix for which, for all  $i \in [N]$ , we have

$$[\mathbf{D}_{321}]_i = (1 + \delta)^2 - \left(1 + \frac{1}{m} \boldsymbol{\sigma}_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \boldsymbol{\sigma}_i\right)^2.$$

We replace  $1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i$  by  $1 + \delta$  as following

$$\begin{aligned}
\mathbf{Z}_{321} &= \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \frac{\mathbf{Q}_{-i}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i}}{\left(1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i\right)^2} \right] \\
&= \frac{1}{(1 + \delta)^2} \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_{-i}^T \mathbf{U}_n^T \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} \right] \\
&\quad + \frac{1}{(1 + \delta)^2} \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m^T \mathbf{U}_n^T \sigma_i \sigma_i^T \mathbf{U}_n \mathbf{Q}_m \left( (1 + \delta)^2 - \left(1 + \frac{1}{m} \sigma_i^T \mathbf{U}_n \mathbf{Q}_{-i} (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \sigma_i\right)^2 \right) \right] \\
&= \frac{N}{m} \frac{1}{(1 + \delta)^2} \mathbb{E} [\mathbf{Q}_-^T \mathbf{U}_n^T \Phi_S \mathbf{U}_n \mathbf{Q}_-] + \frac{1}{(1 + \delta)^2} \mathbb{E} \left[ \frac{1}{m} \mathbf{Q}_m \mathbf{U}_n^T \Sigma_S^T \mathbf{D}_{321} \Sigma_S \mathbf{U}_n \mathbf{Q}_m \right] \\
&= \frac{1}{1 + \delta} \mathbb{E} [\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-] + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right).
\end{aligned}$$

The last equality is obtained since we can show that  $\mathbb{E}[\|\mathbf{D}_{321}\|] = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ . We have from Lemma F.4

$$\|\mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] - \mathbb{E}[\mathbf{Q}_-^T \Psi_1 \mathbf{Q}_-]\| = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right),$$

and from Lemma F.5

$$\mathbf{Z}_{321} = \frac{1}{1 + \delta} \mathbb{E}[\mathbf{Q}_m^T \Psi_1 \mathbf{Q}_m] = \frac{1}{1 + \delta} \frac{1}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m + \mathcal{O}_{\|\cdot\|} \left( \frac{1}{\sqrt{m}} \right).$$

We conclude that

$$\mathbf{Z}_{32} = \frac{1}{n} \frac{\frac{1}{N} \text{Tr}(\Lambda_P \Psi_S)}{1 - \frac{1}{N} \text{Tr}(\Psi_2 \bar{\mathbf{Q}}_m^T \Psi_1 \bar{\mathbf{Q}}_m)} \|\bar{\mathbf{Q}}_m \mathbf{r}\|_{\Psi_1}^2 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

□

**Lemma G.7.** *When all states have been visited, the empirical transition model matrix  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) = \mathbf{U}_n (\mathbf{U}_n - \gamma \mathbf{V}_n)$  defined in equation 14 is invertible.*

*Proof.* Let  $c : \mathcal{S} \rightarrow \mathbb{N}$  and  $c' : \mathcal{S} \rightarrow \mathbb{N}$  be defined such that, for all  $i \in [p]$ ,  $c(\mathbf{S}_i)$  and  $c'(\mathbf{S}_i)$  represent the number of times  $\mathbf{S}_i$  occurs in  $\mathbf{X}_n$  and  $\mathbf{X}'_n$ , respectively. If all states have been visited ( $m = p$ ), for all  $i$  in  $[p]$ , we have thus  $c(\mathbf{S}_i) > 0$ . The structure of  $\sqrt{n} \mathbf{U}_n \in \mathbb{R}^{m \times n}$  indicates each column  $i$  of  $\mathbf{U}_n$  is a one-hot vector, where its  $j$ -th element is 1 if the  $i$ -th state  $\mathbf{s}_i$  of  $\mathbf{X}_n$  is  $\mathbf{S}_j$ . Conversely, each row  $i$  of  $\sqrt{n} \mathbf{U}_n$  has a  $j$ -th element is one if the  $j$ -th state  $\mathbf{s}_j$  of  $\mathbf{X}_n$  is  $\mathbf{S}_i$ . A similar correspondence holds for  $\sqrt{n} \mathbf{V}_n$  and  $\mathbf{X}'_n$ . From interpretations of  $\mathbf{U}_n$  and  $\mathbf{V}_n$ , we deduce  $n \mathbf{U}_n \mathbf{U}_n^T \in \mathbb{R}^{m \times m}$  and  $n \mathbf{V}_n \mathbf{V}_n^T \in \mathbb{R}^{m \times m}$  are diagonal matrices where the  $i$ -th element of its diagonal are  $c(\mathbf{S}_i)$  and  $c'(\mathbf{S}_i)$ , respectively. In the same way,  $n \mathbf{U}_n \mathbf{V}_n^T \in \mathbb{R}^{m \times m}$  is matrix for which  $[n \mathbf{U}_n \mathbf{V}_n^T]_{ij}$  is  $c(\mathbf{S}_i \rightarrow \mathbf{S}_j)$  which represents the number of times the state  $\mathbf{S}_i$  follows  $\mathbf{S}_j$  in  $\mathcal{D}_{\text{train}}$ . We are going to prove  $\hat{\mathbf{A}}_m$  is invertible by using the Gershgorin circle theorem to show  $\hat{\mathbf{A}}_m$  is strictly diagonally dominant, i.e.,  $|[\hat{\mathbf{A}}_m]_{ii}| > \sum_{i \neq j} |[\hat{\mathbf{A}}_m]_{ij}|$ . From the interpretations of  $\mathbf{U}_n \mathbf{U}_n^T$  and  $\mathbf{U}_n \mathbf{V}_n^T$ , we have

$$[\hat{\mathbf{A}}_m]_{ii} = [\mathbf{U}_n \mathbf{U}_n^T]_{ii} - \gamma [\mathbf{U}_n \mathbf{V}_n^T]_{ii} = \frac{c(\mathbf{S}_i) - \gamma c(\mathbf{S}_i \rightarrow \mathbf{S}_i)}{n} > 0, \quad \forall i \in [n],$$

and

$$[\hat{\mathbf{A}}_m]_{ij} = -\gamma [\mathbf{U}_n \mathbf{V}_n^T]_{ij} = \frac{-\gamma c(\mathbf{S}_i \rightarrow \mathbf{S}_j)}{n} < 0, \quad \forall i \neq j.$$

To prove  $\hat{\mathbf{A}}_m$  is invertible it remains to show  $\sum_j [\hat{\mathbf{A}}_m]_{ij} = \sum_j [\mathbf{U}_n (\mathbf{U}_n - \gamma \mathbf{V}_n)^T]_{ij} > 0$  for all  $i \in [m]$ . Let  $i \in [m]$ , we have

$$\sum_j [\mathbf{U}_n (\mathbf{U}_n - \gamma \mathbf{V}_n)^T]_{ij} = \frac{c(\mathbf{S}_i)}{n} - \gamma \sum_j \frac{c(\mathbf{S}_i \rightarrow \mathbf{S}_j)}{n} = (1 - \gamma) \frac{c(\mathbf{S}_i)}{n} > 0,$$

which concludes the proof. □

**Lemma G.8.** Let  $\Delta$  be the second-order correction factor of  $\overline{\text{MSBE}}(\hat{\theta})$  defined in equation 20. If all states have been visited then

$$\Delta = \frac{\lambda^2 \frac{1}{N} \text{Tr} \left( \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m \Psi_2 \bar{\mathbf{Q}}_m^T \right)}{n \left( 1 - \frac{1}{N} \text{Tr} \left( \Psi_2 \bar{\mathbf{Q}}_m(\lambda)^T \Psi_1 \bar{\mathbf{Q}}_m(\lambda) \right) \right)} \|\bar{\mathbf{Q}}_m(\lambda) \mathbf{r}\|_{\Psi_1}^2,$$

where  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) = \mathbf{U}_n(\mathbf{U}_n - \gamma \mathbf{V}_n)$  is the empirical transition model matrix defined in equation 14.

*Proof.* When all states have been visited, we have  $\mathbf{U}_n = \hat{\mathbf{U}}_n$ ,  $\mathbf{V}_n = \hat{\mathbf{V}}_n$  and  $\Sigma_S = \Sigma_{\hat{S}}$ . Furthermore, from Lemma G.7,  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n) = \mathbf{U}_n(\mathbf{U}_n - \gamma \mathbf{V}_n)$  is invertible. We write

$$\begin{aligned} \Theta_S &= \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) = \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S \mathbf{U}_n \bar{\mathbf{Q}}_m(\lambda) \\ &= \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m] \end{aligned}$$

Using the equality above and the cyclic properties of the trace we conclude that

$$\begin{aligned} &\text{Tr} \left( \Lambda_P [\Theta_S \Psi_2 \Theta_S^T - 2\Theta_S (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S] \right) \\ &= \text{Tr} \left( \Lambda_P [\hat{\mathbf{A}}_m^{-1} \mathbf{U}_n [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m] \Psi_2 [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m]^T \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \right. \\ &\quad \left. - 2\hat{\mathbf{A}}_m^{-1} \mathbf{U}_n [\mathbf{I}_n - \lambda \bar{\mathbf{Q}}_m] (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S + \Psi_S] \right) \\ &= \lambda^2 \text{Tr} \left( \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m \Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \right) - \lambda \text{Tr} \left( \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S \right) \\ &\quad - \lambda \text{Tr} \left( \Lambda_P \Psi_S (\mathbf{U}_n - \gamma \mathbf{V}_n) \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \right) + 2\lambda \text{Tr} \left( \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m (\mathbf{U}_n - \gamma \mathbf{V}_n)^T \Psi_S \right) \\ &= \lambda^2 \text{Tr} \left( \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m \Psi_2 \bar{\mathbf{Q}}_m^T \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \right) \\ &= \lambda^2 \text{Tr} \left( \mathbf{U}_n^T \hat{\mathbf{A}}_m^{-1T} \Lambda_P \hat{\mathbf{A}}_m^{-1} \mathbf{U}_n \bar{\mathbf{Q}}_m \Psi_2 \bar{\mathbf{Q}}_m^T \right). \end{aligned}$$

□

## H EXISTENCE OF THE RESOLVENT $\mathbf{Q}_m(\lambda)$

In this section, we show that Assumption 2 guarantees the existence of the resolvent  $\mathbf{Q}_m(\lambda)$  (Lemma H.1), but also that Assumption 2 may be true in practice under certain conditions (Lemma H.2).

**Lemma H.1.** Under Assumption 2, for any  $\lambda > 0$ , the resolvent  $\mathbf{Q}_m(\lambda)$  defined in equation 13 exists.

*Proof.* From Assumption 2, we know that  $\nu_{\min}(H(\hat{\mathbf{A}}_m)) > \xi_{\min} > 0$ , and thus  $H(\Sigma_{\hat{S}} \hat{\mathbf{A}}_m \Sigma_{\hat{S}}^T)$  is at least semi-positive-definite. From the Min-Max theorem, we deduce that the eigenvalues of  $\Sigma_{\hat{S}} \hat{\mathbf{A}}_m \Sigma_{\hat{S}}^T$  have nonnegative real-parts. Consequently, the eigenvalues of  $\frac{1}{m}(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{\mathbf{U}}_n$  have nonnegative real parts since both  $\frac{1}{m}(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{S}}^T \Sigma_{\hat{S}} \hat{\mathbf{U}}_n$  and  $\Sigma_{\hat{S}} \hat{\mathbf{A}}_m \Sigma_{\hat{S}}^T$  share the same nonzero eigenvalues from the Weinstein–Aronszajn identity (Lemma L.6). □

**Lemma H.2.** Let  $c : \hat{S} \rightarrow \mathbb{N}$  and  $c' : \hat{S} \rightarrow \mathbb{N}$  be defined such that, for all  $i \in [m]$ ,  $c(\hat{S}_i)$  and  $c'(\hat{S}_i)$  represent the number of times  $\hat{S}_i$  occurs in  $\mathbf{X}_n$  and  $\mathbf{X}'_n$ , respectively. If for all  $i \in [m]$ ,  $c(\hat{S}_i) \geq \gamma c'(\hat{S}_i)$  then the symmetric part of the empirical transition model matrix  $\hat{\mathbf{A}}_m$  (defined in equation 14) is positive-definite.

*Proof.* The structure of  $\sqrt{n} \hat{\mathbf{U}}_n \in \mathbb{R}^{m \times n}$  indicates each column  $i$  of  $\hat{\mathbf{U}}_n$  is a one-hot vector, where its  $j$ -th element is 1 if the  $i$ -th state  $\mathbf{s}_i$  of  $\mathbf{X}_n$  is  $\hat{S}_j$ . Conversely, each row  $i$  of  $\sqrt{n} \hat{\mathbf{U}}_n$  has a  $j$ -th element equal to one if the  $j$ -th state  $\mathbf{s}_j$  of  $\mathbf{X}_n$  is  $\hat{S}_i$ . A similar correspondence holds for  $\sqrt{n} \hat{\mathbf{V}}_n$  and  $\mathbf{X}'_n$ .

From interpretations of  $\hat{U}_n$  and  $\hat{V}_n$ , we deduce that  $n\hat{U}_n\hat{U}_n^T \in \mathbb{R}^{m \times m}$  and  $n\hat{V}_n\hat{V}_n^T \in \mathbb{R}^{m \times m}$  are diagonal matrices where the  $i$ -th element of its diagonal is equal to  $c(\hat{S}_i)$  and  $c'(\hat{S}_i)$ , respectively. In the same way,  $n\hat{U}_n\hat{V}_n^T \in \mathbb{R}^{m \times m}$  is matrix for which  $[n\hat{U}_n\hat{V}_n^T]_{ij}$  is  $c(\hat{S}_i \rightarrow \hat{S}_j)$ , i.e., the number of times the state  $\hat{S}_i$  follows  $\hat{S}_j$  in  $\mathcal{D}_{\text{train}}$ . We want to prove  $H(\hat{A}_m) = \frac{\hat{A}_m + \hat{A}_m^T}{2}$  is positive-definite by using the Gershgorin circle theorem and by showing  $H(\hat{A}_m)$  is strictly diagonally dominant, i.e.,  $|[H(\hat{A}_m)]_{ii}| > \sum_{i \neq j} |[H(\hat{A}_m)]_{ij}|$ . From interpretations of  $\hat{U}_n\hat{U}_n^T$  and  $\hat{U}_n\hat{V}_n^T$ , we have for all  $i \in [n]$

$$[H(\hat{A}_m)]_{ii} = [\hat{U}_n\hat{U}_n^T]_{ii} - \gamma[\hat{U}_n\hat{V}_n^T]_{ii} = \frac{c(\hat{S}_i) - \gamma c(\hat{S}_i \rightarrow \hat{S}_i)}{n} > 0,$$

and for all  $i \neq j$

$$[H(\hat{A}_m)]_{ij} = \frac{-\gamma[\hat{U}_n\hat{V}_n^T]_{ij} - \gamma[\hat{U}_n\hat{V}_n^T]_{ji}}{2} = \frac{-\gamma c(\hat{S}_i \rightarrow \hat{S}_j) - \gamma c(\hat{S}_j \rightarrow \hat{S}_i)}{2n} < 0.$$

To prove that  $H(\hat{A}_m)$  is positive-definite it remains to show that

$$\begin{aligned} \sum_j [H(\hat{A}_m)]_{ij} &= \sum_{j \neq i} [H(\hat{A}_m)]_{ij} + [H(\hat{A}_m)]_{ii} \\ &= \sum_j \left[ \frac{\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T}{2} \right]_{ij} + \sum_j \left[ \frac{\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T}{2} \right]_{ji} > 0 \end{aligned}$$

for all  $i \in [m]$ . Let  $i \in [m]$ , we have

$$\sum_j [\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T]_{ij} = \frac{c(\hat{S}_i)}{n} - \gamma \sum_j \frac{c(\hat{S}_i \rightarrow \hat{S}_j)}{n} = (1 - \gamma) \frac{c(\hat{S}_i)}{n} > 0$$

and

$$\sum_j [\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T]_{ji} = \frac{c(\hat{S}_i)}{n} - \gamma \sum_j \frac{c(\hat{S}_j \rightarrow \hat{S}_i)}{n} = \frac{c(\hat{S}_i) - \gamma c'(\hat{S}_i)}{n} > 0,$$

since  $c(\hat{S}_i) \geq \gamma c'(\hat{S}_i)$  for all  $i \in [m]$ . We deduce for all  $i \in [m]$  that

$$\sum_j [H(\hat{A}_m)]_{ij} = \sum_j \left[ \frac{\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T}{2} \right]_{ij} + \sum_j \left[ \frac{\hat{U}_n(\hat{U}_n - \gamma\hat{V}_n)^T}{2} \right]_{ji} > 0,$$

and thus  $H(\hat{A}_m)$  is strictly diagonally dominant and positive-definite.  $\square$

**Remark 19.** Conditions of Lemma H.2 may hold in practice. If  $\mathcal{D}_{\text{train}}$  is derived from a sample path of the MRP, where  $s'_{i+1} = s_i$  for all  $i \in [n-1]$ , and if  $\hat{S}_l$  depicts the distinct visited state corresponding to the last next state visited  $s'_n$  in  $\mathcal{D}_{\text{train}}$ , then we have  $c(\hat{S}_i) = c'(\hat{S}_i)$  for all  $i \neq l$  and  $c(\hat{S}_l) = c'(\hat{S}_l) - 1$ . For sufficiently large  $n$ , we may have  $c(\hat{S}_l) \geq \frac{\gamma}{1-\gamma}$  which satisfies conditions of Lemma H.2. Similarly, conditions of Lemma H.2 are satisfied for the pathwise LSTD algorithm, where  $\mathcal{D}_{\text{train}}$  is perturbed slightly by setting the feature of the next state of the last transition to zero (Lazaric et al., 2012) to get  $c(\hat{S}_l) \geq c'(\hat{S}_l)$ .

## I TECHNICAL DETAILS ON THE RESOLVENT $\mathbf{Q}_m(\lambda)$

The objective of this section is to prove that the operator norm of  $\mathbf{Q}_m(\lambda)$  is uniformly upper bounded under Assumption 2. Indeed, controlling the operator norm of  $\mathbf{Q}_m(\lambda)$  is crucial for proving the theorems in Section 5. When  $\gamma = 0$ , which corresponds to the supervised learning case on the reward function, the result is straightforward with Lemma L.7 since  $\frac{1}{m} \Sigma_{\mathbf{X}_n}^T \Sigma_{\mathbf{X}_n}$  is positive-definite (Louart et al., 2018; Liao et al., 2020). In the RL setting, the conclusion is less straightforward as the resolvent is no longer that of a symmetric positive-definite matrix. This issue is further exacerbated by the lack of results in the literature concerning the upper bounds for operator norm of resolvents of non-positive-definite matrices. Lemma I.1 aims to propose a solution for the RL setting under Assumptions 1 and 2. Proof of the widely used Lemma I.4 is also presented at the end of this section.

**Lemma I.1.** Under Assumptions 1 and 2, let  $\lambda > 0$  and let  $\mathbf{Q}_m(\lambda) \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13 as

$$\mathbf{Q}_m(\lambda) = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}.$$

Then there exists a real  $K > 0$  such that, for all  $m$ , we have

$$\|\mathbf{Q}_m(\lambda)\| \leq K.$$

*Proof.* Under Assumption 2, the empirical transition model matrix  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  (equation 14) is invertible since the symmetric part of  $\hat{\mathbf{A}}_m$  is positive-definite. Let

$$0 < \epsilon < \lambda \min \left\{ \frac{1}{\xi_{\max}}, \frac{\xi_{\min}}{4} \right\},$$

for  $\xi_{\min}, \xi_{\max} > 0$  defined in Assumption 2. We rewrite equation 13 as

$$\begin{aligned} \mathbf{Q}_m(\lambda) &= \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1} \\ &= \left[ (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right] \hat{\mathbf{U}}_n + \underbrace{\lambda \mathbf{I}_n - \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n}_{=\mathbf{B}_n} \right]^{-1}. \end{aligned}$$

To apply the Woodbury identity (Lemma L.3) on  $\mathbf{Q}_m(\lambda)$ , we check that both

$$\mathbf{B}_n = \lambda \mathbf{I}_n - \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n,$$

and

$$\begin{aligned} \mathbf{M}_m &= \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1} + \hat{\mathbf{U}}_n \mathbf{B}_n^{-1} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \\ &= \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1} + \left[ \lambda \mathbf{I}_n - \epsilon \hat{\mathbf{A}}_m \right]^{-1} \hat{\mathbf{A}}_m \\ &= \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1} + \left[ \lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m \right]^{-1} \end{aligned}$$

are non-singular, since  $\frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m$  is non-singular. Given that  $H(\hat{\mathbf{A}}_m)$  is positive-definite,  $\hat{\mathbf{A}}_m$  has eigenvalues with positive real parts. Consequently, by the Weinstein–Aronszajn identity (Lemma L.6),  $(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n$  has non-zero eigenvalues with positive real parts. As  $\epsilon < \frac{\lambda}{\xi_{\max}} \leq \frac{\lambda}{\nu_{\max}(H(\hat{\mathbf{A}}_m))} \leq \frac{\lambda}{\text{Re}(\nu_{\max}(\hat{\mathbf{A}}_m))}$ , we deduce that the matrix  $\mathbf{B}_n = \lambda \mathbf{I}_n - \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n$  has eigenvalues with positive real parts and is non-singular. In order to prove that the matrix  $\mathbf{M}_m$  is non-singular, we propose to show  $\mathbf{x}^T \mathbf{M}_m \mathbf{x} > 0$  for all non-zero  $\mathbf{x} \in \mathbb{R}^m$ . Since  $\left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1}$  is at least positive-semi-definite, the statement  $\mathbf{x}^T \mathbf{M}_m \mathbf{x} > 0$  for all non-zero  $\mathbf{x} \in \mathbb{R}^m$  may be restated as

$$\begin{aligned} &\text{for all non-zero } \mathbf{x} \in \mathbb{R}^m, \quad \mathbf{x}^T \left[ \lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m \right]^{-1} \mathbf{x} > 0 \\ \text{iff} &\text{ for all non-zero } \mathbf{x} \in \mathbb{R}^m, \quad \mathbf{x}^T \left[ \lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m \right] \mathbf{x} > 0 \\ \text{iff} &\text{ for all non-zero } \mathbf{x} \in \mathbb{R}^m, \quad \mathbf{x}^T H(\hat{\mathbf{A}}_m^{-1}) \mathbf{x} - \frac{\epsilon}{\lambda} \mathbf{x}^T \mathbf{x} > 0 \\ \text{iff} &\nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) > \frac{\epsilon}{\lambda}. \end{aligned}$$

By construction of  $\hat{\mathbf{U}}_n$  and  $\hat{\mathbf{V}}_n$ , we have both  $\|\hat{\mathbf{U}}_n\| \leq 1$  and  $\|\hat{\mathbf{V}}_n\| \leq 1$ . We deduce thus

$$\|\hat{\mathbf{A}}_m\| = \|\hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T\| < 2.$$

Since  $H(\hat{\mathbf{A}}_m^{-1}) = [\hat{\mathbf{A}}_m^{-1}]^T H(\hat{\mathbf{A}}_m) \hat{\mathbf{A}}_m^{-1}$ , we deduce from Ostrowski's Theorem (Lemma L.5) that

$$\nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) \geq \frac{\nu_{\min}(H(\hat{\mathbf{A}}_m))}{\|\hat{\mathbf{A}}_m\|^2} \geq \frac{\xi_{\min}}{4}.$$

Since  $\epsilon < \frac{\lambda \xi_{\min}}{4}$ , we have  $\mathbf{x}^T \mathbf{M}_m \mathbf{x} > 0$  for all non-zero  $\mathbf{x} \in \mathbb{R}^m$ , and thus  $\mathbf{M}_m$  is non-singular. As a consequence, we apply the Woodbury identity (Lemma L.3) on the resolvent  $\mathbf{Q}_m(\lambda)$  to get

$$\begin{aligned} \mathbf{Q}_m(\lambda) &= \left[ (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right] \hat{\mathbf{U}}_n + \mathbf{B}_n \right]^{-1} \\ &= \mathbf{B}_n^{-1} - \mathbf{B}_n^{-1} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbf{M}_m^{-1} \hat{\mathbf{U}}_n \mathbf{B}_n^{-1}. \end{aligned}$$

Multiplying the equation above by  $\mathbf{B}_n = \lambda \mathbf{I}_n - \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n$  on both sides, and after manipulating terms to isolate  $\mathbf{Q}_m$  on the left-hand side gives

$$\begin{aligned} \mathbf{Q}_m(\lambda) &= \frac{1}{\lambda^2} \left[ \mathbf{B}_n - (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbf{M}_m^{-1} \hat{\mathbf{U}}_n \right. \\ &\quad \left. + \lambda \epsilon \left[ (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) + \mathbf{Q}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n \right] \right. \\ &\quad \left. - \epsilon^2 (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n \right] \\ &= \frac{1}{\lambda^2} \left[ \mathbf{B}_n - (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbf{M}_m^{-1} \hat{\mathbf{U}}_n \right. \\ &\quad \left. + \lambda \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \hat{\mathbf{U}}_n \right. \\ &\quad \left. + \lambda \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{A}}_m + \lambda \mathbf{I}_m \right]^{-1} \hat{\mathbf{U}}_n \right. \\ &\quad \left. - \epsilon^2 (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n \right]. \end{aligned}$$

Applying the operator norm on the equality above, we find

$$\begin{aligned} \|\mathbf{Q}_m(\lambda)\| &\leq \frac{1}{\lambda^2} \left[ \lambda + 2\epsilon + 2\|\mathbf{M}_m^{-1}\| \right. \\ &\quad \left. + 2\lambda\epsilon \left\| \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \right\| + 2\lambda\epsilon \left\| \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{A}}_m + \lambda \mathbf{I}_m \right]^{-1} \right\| \right. \\ &\quad \left. + 4\epsilon^2 \left\| \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \right\| \right], \end{aligned} \tag{75}$$

since

$$\|\mathbf{B}_n\| = \|\lambda \mathbf{I}_n - \epsilon (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \hat{\mathbf{U}}_n\| \leq \lambda + 2\epsilon.$$

From Lemma L.2, we have

$$\left\| \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \right\| \leq \frac{1}{\lambda} \frac{4}{\xi_{\min}^2},$$

and

$$\left\| \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{A}}_m + \lambda \mathbf{I}_m \right]^{-1} \right\| \leq \frac{1}{\lambda} \frac{4}{\xi_{\min}^2}.$$



To finish the proof, we find an upper bound for  $\|\mathbf{M}_m^{-1}\|$ . By denoting by  $\mathbf{Z}^T \mathbf{Z}$  the Cholesky decomposition of the positive-semi-definite matrix  $\left[\frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m\right]^{-1}$ , we reuse the Woodbury identity (Lemma L.3) to rewrite  $\mathbf{M}_m^{-1}$  as

$$\begin{aligned} \mathbf{M}_m^{-1} &= \left[ \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1} + [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m]^{-1} \right]^{-1} \\ &= \left[ \mathbf{Z}^T \mathbf{Z} + [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m]^{-1} \right]^{-1} \\ &= [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m] - [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m] \mathbf{Z}^T \left[ \mathbf{Z} [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m] \mathbf{Z}^T + \mathbf{I}_m \right]^{-1} \mathbf{Z} [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m]. \end{aligned}$$

From Lemma L.7,

$$\left\| \left[ \mathbf{Z} [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m] \mathbf{Z}^T + \mathbf{I}_m \right]^{-1} \right\| \leq 1,$$

since  $H(\mathbf{Z} [\lambda \hat{\mathbf{A}}_m^{-1} - \epsilon \mathbf{I}_m] \mathbf{Z}^T)$  is positive-semi-definite, and from Lemma L.3 we have

$$\|\hat{\mathbf{A}}_m^{-1}\| \leq \frac{1}{\xi_{\min}}.$$

Besides,

$$\|\mathbf{Z}\|^2 = \nu_{\max} \left( \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \epsilon \mathbf{I}_m \right]^{-1} \right) \leq \frac{1}{\epsilon}.$$

We deduce for the operator norm of  $\mathbf{M}_m^{-1}$  that

$$\|\mathbf{M}_m^{-1}\| \leq \left( \frac{\lambda}{\xi_{\min}} + \epsilon \right) + \frac{1}{\epsilon} \left( \frac{\lambda}{\xi_{\min}} + \epsilon \right)^2.$$

Setting  $\epsilon = \frac{\lambda}{2\epsilon'} < \lambda \min \left\{ \frac{1}{\xi_{\max}}, \frac{\xi_{\min}}{4} \right\}$  for  $\epsilon' > \frac{1}{2} \min \left\{ \frac{1}{\xi_{\max}}, \frac{\xi_{\min}}{4} \right\}$  and putting upper bounds of  $\|\mathbf{M}_m^{-1}\|$ ,  $\left\| \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \right\|$ ,  $\left\| \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{A}}_m + \lambda \mathbf{I}_m \right]^{-1} \right\|$  into equation 75 give

$$\begin{aligned} \|\mathbf{Q}_m(\lambda)\| &\leq \frac{1}{\lambda^2} \left[ \lambda + \frac{\lambda}{\epsilon'} + \lambda \left( \frac{2}{\xi_{\min}} + \frac{1}{\epsilon'} \right) + \lambda \epsilon' \left( \frac{2}{\xi_{\min}} + \frac{1}{\epsilon'} \right)^2 + \lambda \frac{8}{\xi_{\min}^2 \epsilon'} + \lambda \frac{4}{\xi_{\min}^2 \epsilon'^2} \right] \\ &= \frac{1}{\lambda} \left[ 1 + \frac{1}{\epsilon'} + \left( \frac{2}{\xi_{\min}} + \frac{1}{\epsilon'} \right) + \epsilon' \left( \frac{2}{\xi_{\min}} + \frac{1}{\epsilon'} \right)^2 + \frac{8}{\xi_{\min}^2 \epsilon'} + \frac{4}{\xi_{\min}^2 \epsilon'^2} \right]. \end{aligned}$$

□

**Remark 20.** From the proof of Lemma I.1, eigenspectrum constraints on the empirical transition model matrix  $\hat{\mathbf{A}}_m$  in Assumption 2 ensure the resolvent  $\mathbf{Q}_m(\lambda)$  is uniformly bounded.

**Lemma I.2.** Under Assumptions 1 and 2, let  $\lambda > 0$  and let  $\mathbf{Q}'_m(\lambda), \mathbf{Q}''_m(\lambda) \in \mathbb{R}^{m \times m}$  be the following resolvents

$$\mathbf{Q}'_m(\lambda) = \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1}$$

and

$$\mathbf{Q}''_m(\lambda) = \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{A}}_m + \lambda \mathbf{I}_m \right]^{-1},$$

where  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \in \mathbb{R}^{m \times m}$  is the empirical transition model matrix (equation 14). Then, for all  $m$ , we have

$$\|\mathbf{Q}'_m(\lambda)\| \leq \frac{1}{\lambda} \frac{4}{\xi_{\min}^2} \quad \text{and} \quad \|\mathbf{Q}''_m(\lambda)\| \leq \frac{1}{\lambda} \frac{4}{\xi_{\min}^2}.$$

*Proof.* Since the symmetric part of the empirical transition model matrix  $\hat{\mathbf{A}}_m$  is positive-definite under Assumption 2, the matrix  $\hat{\mathbf{A}}_m$  is non-singular. We write thus

$$\begin{aligned} \|\mathbf{Q}'_m(\lambda)\| &= \left\| \left[ \frac{1}{m} \hat{\mathbf{A}}_m \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \mathbf{I}_m \right]^{-1} \right\| \\ &= \left\| \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \hat{\mathbf{A}}_m^{-1} \right]^{-1} \hat{\mathbf{A}}_m^{-1} \right\| \\ &\leq \left\| \left[ \frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \hat{\mathbf{A}}_m^{-1} - \lambda \nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) \mathbf{I}_m + \lambda \nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) \mathbf{I}_m \right]^{-1} \right\| \|\hat{\mathbf{A}}_m^{-1}\| \\ &= \frac{1}{\lambda} \frac{1}{\nu_{\min}(H(\hat{\mathbf{A}}_m^{-1}))} \|\hat{\mathbf{A}}_m^{-1}\|. \end{aligned}$$

The last inequality is obtained with Lemma L.7 since  $H\left(\frac{1}{m} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} + \lambda \hat{\mathbf{A}}_m^{-1} - \lambda \nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) \mathbf{I}_m\right)$  is positive-semi-definite. By construction of both  $\hat{\mathbf{U}}_n$  and  $\hat{\mathbf{V}}_n$ , we have  $\|\hat{\mathbf{U}}_n\| \leq 1$  and  $\|\hat{\mathbf{V}}_n\| \leq 1$ . We deduce that

$$\|\hat{\mathbf{A}}_m\| = \|\hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T\| < 2.$$

Since  $H(\hat{\mathbf{A}}_m^{-1}) = [\hat{\mathbf{A}}_m^{-1}]^T H(\hat{\mathbf{A}}_m) \hat{\mathbf{A}}_m^{-1}$ , we deduce from Ostrowski's theorem (Lemma L.5) that

$$\nu_{\min}(H(\hat{\mathbf{A}}_m^{-1})) \geq \frac{\nu_{\min}(H(\hat{\mathbf{A}}_m))}{\|\hat{\mathbf{A}}_m\|^2} \geq \frac{\xi_{\min}}{4}.$$

Furthermore, from Lemma I.3, we have  $\|\hat{\mathbf{A}}_m^{-1}\| \leq \frac{1}{\xi_{\min}}$ . We conclude that

$$\|\mathbf{Q}'_m(\lambda)\| \leq \frac{1}{\lambda} \frac{4}{\xi_{\min}^2}.$$

With a similar reasoning, we can find the same upper bound for  $\|\mathbf{Q}''_m(\lambda)\|$ .  $\square$

**Lemma I.3.** Let  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  be the empirical transition model matrix defined in equation 14. Under Assumption 2, for all  $m$ , we have

$$\|\hat{\mathbf{A}}_m^{-1}\| \leq \frac{1}{\xi_{\min}}.$$

*Proof.* We rewrite  $\hat{\mathbf{A}}_m$  as

$$\hat{\mathbf{A}}_m^{-1} = \left[ [\hat{\mathbf{A}}_m - \nu_{\min}(H(\hat{\mathbf{A}}_m)) \mathbf{I}_m] + \nu_{\min}(H(\hat{\mathbf{A}}_m)) \mathbf{I}_m \right]^{-1}.$$

Since the matrix  $H\left([\hat{\mathbf{A}}_m - \nu_{\min}(H(\hat{\mathbf{A}}_m)) \mathbf{I}_m]\right)$  is positive-semi-definite, we apply Lemma L.7 on  $\hat{\mathbf{A}}_m^{-1}$  to get

$$\|\hat{\mathbf{A}}_m^{-1}\| \leq \frac{1}{\nu_{\min}(H(\hat{\mathbf{A}}_m))} \leq \frac{1}{\xi_{\min}}. \quad \square$$

**Lemma I.4.** Under Assumption 1 and 2, let  $\lambda > 0$  and let  $\mathbf{Q}_m(\lambda) \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 13 as

$$\mathbf{Q}_m(\lambda) = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}.$$

Then there exists a real  $K > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}} \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) \right\| \leq K$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}^T \right\| \leq 2K.$$

*Proof.* From Lemma I.1, we know there exists a real  $K > 0$  such that, for all  $m$ , we have  $\|\mathbf{Q}_m(\lambda)\| \leq K$ . Since the symmetric part of the empirical transition model matrix  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T$  (equation 14) is positive-definite under Assumption 2, the matrix  $\hat{\mathbf{A}}_m$  is non-singular. Furthermore, from Lemma I.3 we have  $\|\hat{\mathbf{A}}_m^{-1}\| \leq \frac{1}{\xi_{\min}}$ , and both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1. We deduce that

$$\begin{aligned} \left\| \frac{1}{\sqrt{m}} \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) \right\| &= \left\| \frac{1}{m} \mathbf{Q}_m(\lambda)^T \hat{\mathbf{U}}_n^T \Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) \right\|^{\frac{1}{2}} \\ &= \left\| \frac{1}{m} \mathbf{Q}_m(\lambda)^T \hat{\mathbf{U}}_n^T \hat{\mathbf{A}}_m^{-1} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n \mathbf{Q}_m(\lambda) \right\|^{\frac{1}{2}} \\ &= \left\| \mathbf{Q}_m(\lambda)^T \hat{\mathbf{U}}_n^T \hat{\mathbf{A}}_m^{-1} \hat{\mathbf{U}}_n [\mathbf{I}_n - \lambda \mathbf{Q}_m(\lambda)] \right\|^{\frac{1}{2}} \\ &\leq \sqrt{\frac{K(1+K)}{\xi_{\min}}}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} &\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \right\| \\ &= \left\| \frac{1}{m} \mathbf{Q}_m(\lambda) (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Sigma_{\mathcal{S}}^T \Sigma_{\mathcal{S}} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \hat{\mathbf{A}}_m^{-1} (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n) \mathbf{Q}_m(\lambda)^T \right\|^{\frac{1}{2}} \\ &= \left\| [\mathbf{I}_n - \lambda \mathbf{Q}_m(\lambda)] (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \hat{\mathbf{A}}_m^{-1} (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n) \mathbf{Q}_m(\lambda)^T \right\|^{\frac{1}{2}} \\ &\leq 2\sqrt{\frac{K(1+K)}{\xi_{\min}}}. \end{aligned}$$

□

## J ABOUT THE EXISTENCE, POSITIVENESS, AND UNIQUENESS OF $\delta$

This section is dedicated to prove that the fixed-point solution  $\delta$  of equation 17 is unique and positive under Assumptions 1 and 2. This result is proven in the following Lemma.

**Lemma J.1.** *Under Assumptions 1 and 2, for all  $m$ , let  $\delta$  be the solution to the fixed-point equation 17 defined as*

$$\delta = \frac{1}{m} \text{Tr} \left( (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \left[ \frac{N}{m} \frac{1}{1+\delta} (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1} \right).$$

*Then  $\delta$  exists, is positive, and is unique.*

*Proof.* For ease of notations, we define the matrix  $\mathbf{B}_n = (\hat{\mathbf{U}}_n - \gamma\hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n$ . The proof is based on the use of Lemma L.8 on the mapping  $f : \delta \mapsto \frac{1}{m} \text{Tr}(\mathbf{B}_n \mathbf{Q}_m(\delta))$ . To apply Lemma L.8, we need to show *i.*  $f$  is positive on  $[0, \infty)$ , *ii.*  $f$  is monotonically increasing, *iii.*  $f$  is scalable, and *iv.*  $f$  admits  $x_0 \in [0, \infty)$  such that  $x_0 \geq f(x_0)$ . Following this plan, we are going to show first *i.*, i.e.,  $f(\delta) > 0$  for all  $\delta > 0$ . By denoting  $\nu_j(\mathbf{B}_n \mathbf{Q}_m(\delta))$  the  $j$ -th eigenvalues of the matrix  $\mathbf{B}_n \mathbf{Q}_m(\delta)$ ,

we have

$$\begin{aligned}
\nu_j(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta)) &= \nu_j \left( \mathbf{B}_n \left[ \frac{N}{m} \frac{1}{1+\delta} \mathbf{B}_n + \lambda \mathbf{I}_n \right]^{-1} \right) \\
&= \nu_j(\mathbf{B}_n) \nu_j \left( \left[ \frac{N}{m} \frac{1}{1+\delta} \mathbf{B}_n + \lambda \mathbf{I}_n \right]^{-1} \right) \quad (\text{from the Schur decomposition of } \mathbf{B}_n) \\
&= \frac{\nu_j(\mathbf{B}_n)}{\frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda} \\
&= \frac{1}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^2} \left( \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 + \lambda \nu_j(\mathbf{B}_n) \right).
\end{aligned} \tag{76}$$

Let  $\hat{\mathbf{A}}_m = \hat{\mathbf{U}}_n(\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$  be the transition model matrix defined in equation 14, and  $\bar{\mathbf{Z}} \bar{\mathbf{Z}}^T$  be the Cholesky decomposition of  $\Phi_{\mathcal{S}}$ . From the Weinstein–Aronszajn identity (Lemma L.6), the matrices  $\mathbf{B}_n = (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n$  and  $\bar{\mathbf{Z}}^T \hat{\mathbf{A}}_m \bar{\mathbf{Z}}$  share the same non-zero eigenvalues. Under Assumption 2, the matrix  $H(\bar{\mathbf{Z}}^T \hat{\mathbf{A}}_m \bar{\mathbf{Z}})$  is at least semi-positive-definite, which implies that non-zero real parts of eigenvalues of  $\bar{\mathbf{Z}}^T \hat{\mathbf{A}}_m \bar{\mathbf{Z}}$  are positive. We deduce that  $\text{Re}(\nu_j(\mathbf{B}_n)) \geq 0$ , for all  $j \in [m]$ . As a consequence,

$$\begin{aligned}
f(\delta) &= \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta)) \\
&= \frac{1}{m} \sum_{j=1}^n \frac{1}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^2} \left( \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 + \lambda \nu_j(\mathbf{B}_n) \right) \\
&= \frac{1}{m} \sum_{j=1}^n \frac{1}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^2} \left( \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 + \lambda \text{Re}(\nu_j(\mathbf{B}_n)) \right) \\
&> 0.
\end{aligned} \tag{77}$$

To prove *ii.*, i.e.,  $f$  is monotonically increasing on  $[0, \infty)$ , we show the derivative  $f'$  of  $f$  is positive on  $[0, \infty)$ . Let  $\delta > 0$ ,

$$\begin{aligned}
f'(\delta) &= \frac{1}{m} \left( \sum_{j=1}^n \frac{\nu_j(\mathbf{B}_n)}{\frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda} \right)' \\
&= \frac{1}{m} \sum_{j=1}^n \frac{\frac{N}{m} \frac{1}{(1+\delta)^2}}{\left( \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right)^2} \nu_j(\mathbf{B}_n)^2 \\
&= \frac{1}{m} \sum_{j=1}^n \frac{\frac{N}{m} \frac{1}{(1+\delta)^2}}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^4} \left( \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 + \lambda \nu_j(\mathbf{B}_n) \right)^2 \\
&= \frac{1}{m} \sum_{j=1}^n \frac{\frac{N}{m} \frac{1}{(1+\delta)^2}}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^4} \left( \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} |\nu_j(\mathbf{B}_n)|^4 + 2\lambda \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 \nu_j(\mathbf{B}_n) + \lambda^2 \nu_j(\mathbf{B}_n)^2 \right) \\
&= \frac{1}{m} \sum_{j=1}^n \frac{\frac{N}{m} \frac{1}{(1+\delta)^2}}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^4} \underbrace{\left( \frac{N^2}{m^2} \frac{1}{(1+\delta)^2} |\nu_j(\mathbf{B}_n)|^4 + 2\lambda \frac{N}{m} \frac{1}{1+\delta} |\nu_j(\mathbf{B}_n)|^2 \text{Re}(\nu_j(\mathbf{B}_n)) \right)}_{(1)} \\
&\quad + \underbrace{\sum_{j=1}^n \lambda^2 \frac{\frac{N}{m} \frac{1}{(1+\delta)^2}}{\left| \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right|^4} \text{Re}(\nu_j(\mathbf{B}_n)^2))}_{(2)}
\end{aligned}$$

Since real parts of eigenvalues of  $\mathbf{B}_n$  are positive, (1) is clearly positive. Since  $\text{Tr}(\mathbf{B}_n^2) > 0$  (Lemma J.2) and thus (2) is positive, we can conclude *ii.* We can use a similar proof for the scalability in *iii.*, i.e.,  $\alpha f(\delta) > f(\alpha\delta)$ ,  $\forall \alpha > 1$ . Let  $\alpha > 1$  and  $\delta > 0$ ,

$$\alpha f(\delta) - f(\alpha\delta) = \alpha \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta)) - \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\alpha\delta)) \quad (78)$$

$$= \frac{1}{m} \text{Tr}(\mathbf{B}_n [\alpha \bar{\mathbf{Q}}_m(\delta) - \bar{\mathbf{Q}}_m(\alpha\delta)]) \quad (79)$$

$$= \frac{1}{m} \text{Tr} \left( \alpha \mathbf{B}_n \bar{\mathbf{Q}}_m(\delta) \left[ \frac{N}{m} \left( \frac{1}{1+\alpha\delta} - \frac{1}{\alpha(1+\delta)} \right) \mathbf{B}_n + \left( \lambda - \frac{\lambda}{\alpha} \right) \mathbf{I}_n \right] \bar{\mathbf{Q}}_m(\alpha\delta) \right) \quad (80)$$

$$= \underbrace{\alpha \frac{1}{m} \frac{N}{m} \left( \frac{1}{1+\alpha\delta} - \frac{1}{\alpha(1+\delta)} \right)}_{>0} \underbrace{\text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta) \mathbf{B}_n \bar{\mathbf{Q}}_m(\alpha\delta))}_{(1)} + \underbrace{\alpha \frac{1}{m} \left( \lambda - \frac{\lambda}{\alpha} \right)}_{>0} \underbrace{\text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta) \bar{\mathbf{Q}}_m(\alpha\delta))}_{(2)}. \quad (81)$$

To prove *iii.*, we can show that both (1) and (2) in equation 81 are positive. We prove in *ii.* that  $\text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta') \mathbf{B}_n \bar{\mathbf{Q}}_m(\delta)) > 0$  for any  $\delta' > \delta$ . Since  $\alpha\delta > \delta$ , we also deduce (1) is positive. For (2), we can write

$$\begin{aligned} & \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta) \bar{\mathbf{Q}}_m(\alpha\delta)) \\ &= \sum_{j=1}^n \nu_j(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta) \bar{\mathbf{Q}}_m(\alpha\delta)) \\ &= \sum_{j=1}^n \frac{\nu_j(\mathbf{B}_n)}{\left( \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right) \left( \frac{N}{m} \frac{1}{1+\alpha\delta} \nu_j(\mathbf{B}_n) + \lambda \right)} \\ &= \sum_{j=1}^n c_j \left( \left( \frac{N^2}{m^2} \frac{|\nu_j(\mathbf{B}_n)|^2}{(1+\delta)(1+\alpha\delta)} + \lambda^2 \right) \text{Re}(\nu_j(\mathbf{B}_n)) + \frac{N}{m} \left( \frac{\lambda}{1+\delta} + \frac{\lambda}{1+\alpha\delta} \right) |\nu_j(\mathbf{B}_n)|^2 \right) \\ &> 0, \end{aligned}$$

where

$$c_j = \frac{1}{\left| \left( \frac{N}{m} \frac{1}{1+\delta} \nu_j(\mathbf{B}_n) + \lambda \right) \left( \frac{N}{m} \frac{1}{1+\alpha\delta} \nu_j(\mathbf{B}_n) + \lambda \right) \right|^2}.$$

In order to apply Lemma L.8, we still need to demonstrate *iv.*, i.e.,  $f$  admits  $x_0 \in [0, \infty)$  such that  $x_0 \geq f(x_0)$ . To prove *iv.*, it is sufficient to notice that if  $f$  is bounded, i.e.,  $\forall \delta, f(\delta) \leq C$ . Let  $\delta > 0$ , we have

$$\begin{aligned} f(\delta) &= \frac{1}{m} \text{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m(\delta)) = \frac{1}{m} \text{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\delta)) \\ &= \frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}} \hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\delta) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T) \\ &\leq \frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}}) \|\hat{\mathbf{U}}_n \bar{\mathbf{Q}}_m(\delta) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T\| \\ &\leq \frac{2}{n} \text{Tr}(\Phi_{\mathcal{S}}) \|\bar{\mathbf{Q}}_m(\delta)\| \\ &= \mathcal{O}(1), \end{aligned}$$

where we used for the first inequality  $|\text{Tr}(\mathbf{AB})| \leq \|\mathbf{B}\| \text{Tr}(\mathbf{A})$  for non-negative definite matrix  $\mathbf{A}$ . The last inequality is obtained since  $\frac{1}{m} \text{Tr}(\Phi_{\mathcal{S}})$  is uniformly bounded under Assumptions 1 and 2 (see equation 30). Furthermore, both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1 and, with a similar proof than for Lemma L.1, we can show there exists a real  $K_{\bar{\mathbf{Q}}} > 0$  such that, for all  $m$  and for all  $\delta \in [0, \infty)$ , we have  $\|\bar{\mathbf{Q}}_m(\delta)\| \leq K_{\bar{\mathbf{Q}}}$ . Since all hypotheses required on  $f$  to apply Lemma L.8 are satisfied, we can apply this Lemma which concludes the proof.  $\square$

**Lemma J.2.** *We have*

$$\mathrm{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n) > 0.$$

*Proof.* Let  $\mathbf{A} = \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T$ . We denote by  $S(\mathbf{A}) = \frac{\mathbf{A} - \mathbf{A}^T}{2}$  the skew-symmetric part of  $\mathbf{A}$ . We have

$$\begin{aligned} & \mathrm{Tr}((\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n) \\ &= \mathrm{Tr}(\Phi_{\delta} \mathbf{A} \Phi_{\delta} \mathbf{A}) \\ &= \mathrm{Tr}(\Phi_{\delta} \mathbf{A} \Phi_{\delta} H(\mathbf{A})) + \mathrm{Tr}(\Phi_{\delta} \mathbf{A} \Phi_{\delta} S(\mathbf{A})) \\ &= \mathrm{Tr}(\Phi_{\delta} H(\mathbf{A}) \Phi_{\delta} H(\mathbf{A})) + \mathrm{Tr}(\Phi_{\delta} S(\mathbf{A}) \Phi_{\delta} H(\mathbf{A})) + \mathrm{Tr}(\Phi_{\delta} H(\mathbf{A}) \Phi_{\delta} S(\mathbf{A})) + \mathrm{Tr}(\Phi_{\delta} S(\mathbf{A}) \Phi_{\delta} S(\mathbf{A})) \\ &= \mathrm{Tr}(\Phi_{\delta} H(\mathbf{A}) \Phi_{\delta} H(\mathbf{A})) + \mathrm{Tr}(\Phi_{\delta} S(\mathbf{A}) \Phi_{\delta} S(\mathbf{A})) > 0. \end{aligned}$$

□

**Lemma J.3.** *Under Assumptions 1 and 2, let  $\delta$  be the correction factor defined in equation 17.  $\delta$  is a decreasing function with respect to  $N$ .*

*Proof.* For ease of notations, we define the matrix  $\mathbf{B}_n = (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n$  and we denote by  $\mathbf{Q}_m$  the resolvent  $\mathbf{Q}_m(\lambda)$ . The derivative of  $\delta$  as function of  $N$  is denoted as  $\delta'(N)$  and defined as

$$\delta'(N) = -\frac{1}{m} \frac{\frac{\frac{1}{m} \mathrm{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m)}{(1+\delta)}}{1 - \frac{N}{m} \frac{\frac{1}{m} \mathrm{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m)}{(1+\delta)^2}}$$

For all  $N$ , we have  $\delta'(N) \leq 0$  since  $\frac{\frac{1}{m} \mathrm{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m)}{(1+\delta)} > 0$  and  $\frac{N}{m} \frac{\frac{1}{m} \mathrm{Tr}(\mathbf{B}_n \bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m)}{(1+\delta)^2} < 1$  using a similar reasoning than for equation 48. □

**Lemma J.4.** *Under Assumptions 1 and 2, let  $\delta$  be the correction factor defined in equation 17.  $\delta$  is a decreasing function with respect to  $\lambda$ .*

*Proof.* For ease of notations, we define the matrix  $\mathbf{B}_n = (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Phi_{\delta} \hat{\mathbf{U}}_n$  and we denote by  $\bar{\mathbf{Q}}_m$  the resolvent  $\bar{\mathbf{Q}}_m(\lambda)$ . The derivative of  $\delta$  as function of  $\lambda$  is denoted as  $\delta'(\lambda)$  and defined as

$$\delta'(\lambda) = -\frac{1}{m} \mathrm{Tr}(\bar{\mathbf{Q}}_m \mathbf{B}_n \bar{\mathbf{Q}}_m)$$

For all  $\lambda$ , we have  $\delta'(\lambda) \leq 0$  using a similar reasoning than for *iii.* in Lemma J.1. □

## K CONCENTRATION RESULTS

The following section is dedicated to a set of concentration results used for the proofs of Theorems. Preliminary results yield a concentration of measure properties for the random feature matrix  $\Sigma_{\delta} \in \mathbb{R}^{N \times m}$ , which stem from the concentration inequality of Lemma L.1 for Lipschitz applications of a Gaussian vector. Essentially, the guideline of the proofs involves the following steps; given  $\mathbf{W}_{ij} = \varphi(\tilde{\mathbf{W}}_{ij})$ , for which  $\tilde{\mathbf{W}}_{ij} \sim \mathcal{N}(0, 1)$  and  $\varphi$  a Lipschitz function, the normal concentration of  $\tilde{\mathbf{W}}$  is transferred to  $\mathbf{W}$ . This process induces a normal concentration of the random vector  $\sigma(\mathbf{w}^T \hat{\mathbf{S}})$ , for  $\mathbf{w} = \varphi(\tilde{\mathbf{w}})$  and  $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ , and of the matrix  $\Sigma_{\delta}$ . This implies that Lipschitz functionals of  $\sigma(\mathbf{w}^T \hat{\mathbf{S}})$  or  $\Sigma_{\delta}$  also concentrate. As highlighted earlier, these concentration results have multiple consequences on convergence of random variables, and are traditionally employed in Random Matrix theory and in Theorem 5.1. We start by revisiting Lemma K.1 and Lemma K.2, which are derived from Lemma L.1 and that were previously introduced in Louart et al. (2018). Subsequently, we provide intermediary Lemma K.3 and Lemma K.4 to reach the principal results of this section articulated by Lemma K.5 and Lemma K.6, which are employed in proof of Theorems. In the remainder of this section, we denote by  $\|\cdot\|_F$  the Frobenius norm of a matrix.

**Lemma K.1.** Let  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  be a  $K_\sigma$ -Lipschitz continuous function, let  $\mathbf{X} \in \mathbb{R}^{d \times m}$  be a matrix, and let  $\mathbf{w} = \varphi(\tilde{\mathbf{w}})$  be a vector for which  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is a  $K_\varphi$ -Lipschitz continuous function and  $\tilde{\mathbf{w}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ . Let

$$t_0 = |\sigma(0)| + K_\sigma K_\varphi \|\mathbf{X}\| \sqrt{\frac{d}{m}}.$$

Then, for all  $t \geq 4t_0$ , we have

$$\Pr \left( \left\| \frac{1}{\sqrt{m}} \sigma(\mathbf{w}^T \mathbf{X}) \right\| \geq t \right) \leq C e^{-\frac{cm t^2}{2K_\sigma^2 K_\varphi^2 \|\mathbf{X}\|^2}},$$

for some  $C, c > 0$  are independent of all other parameters.

*Proof.* The proof of this Lemma can be found in the first half of proof of Louart et al. (2018, Lemma 2), and is based on Lemma L.1.  $\square$

**Corollary K.1.1.** (Louart et al., 2018, Remark 2) Let  $\mathbf{X} \in \mathbb{R}^{d \times m}$  and let  $\Sigma_{\mathbf{X}} = \sigma(\mathbf{W} \mathbf{X}) \in \mathbb{R}^{N \times m}$  be its random features matrix defined as in equation 5. For all  $t \geq 4t_0$ , we have

$$\Pr \left( \left\| \frac{1}{m} \Sigma_{\mathbf{X}} \right\| \geq t \right) \leq C N e^{-\frac{cm t^2}{2N \|\mathbf{X}\|^2}},$$

where  $t_0 = |\sigma(0)| + \|\mathbf{X}\| \sqrt{\frac{d}{m}}$ .

From the previous Lemma, we deduce the following key concentration result.

**Lemma K.2.** (Louart et al., 2018, Lemma 2) Let  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  be a  $K_\sigma$ -Lipschitz continuous function, let  $\mathbf{X} \in \mathbb{R}^{d \times m}$  be a matrix, and let  $\mathbf{w} = \varphi(\tilde{\mathbf{w}})$  be a vector for which  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is a  $K_\varphi$ -Lipschitz continuous function and  $\tilde{\mathbf{w}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ . Let  $\mathbf{A} \in \mathbb{R}^{m \times m}$  be a matrix independent of  $\mathbf{w}$  such that  $\|\mathbf{A}\| \leq K_A$ . Then, we have

$$\begin{aligned} & \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}^T \mathbf{X})^T \mathbf{A} \sigma(\mathbf{w}^T \mathbf{X}) - \frac{1}{m} \text{Tr}(\mathbf{A} \mathbb{E}[\sigma(\mathbf{w}^T \mathbf{X}) \sigma(\mathbf{w}^T \mathbf{X})^T]) \right| > t \right) \\ & \leq C e^{-\frac{cm}{2K_\sigma^2 K_\varphi^2 \|\mathbf{X}\|^2} \min\left(\frac{t^2}{26t_0^2 K_A^2}, \frac{t}{K_A}\right)}, \end{aligned}$$

for  $t_0 = |\sigma(0)| + \sqrt{\frac{d}{m}} K_\sigma K_\varphi \|\mathbf{X}\|$ , and  $c, C \in \mathbb{R}$  independent of all other parameters.

**Lemma K.3.** Let  $f : \mathbb{R}^{N \times d} \rightarrow \mathbb{R}$ ,  $\mathbf{W} \mapsto f(\mathbf{W})$  be a  $K_f$ -Lipschitz function with respect to the Frobenius norm for which  $\mathbf{W} = \varphi(\tilde{\mathbf{W}})$  is the matrix defined in equation 5. Then, we have

$$\Pr(|f(\mathbf{W}) - \mathbb{E}[f(\mathbf{W})]| > t) \leq C e^{-\frac{ct^2}{K_f^2 K_\varphi^2}},$$

for some  $C, c > 0$ .

*Proof.* The vectorization of  $\tilde{\mathbf{W}}$ ,  $\text{vec}(\tilde{\mathbf{W}}) = [\tilde{\mathbf{W}}_{11}, \dots, \tilde{\mathbf{W}}_{nd}] \in \mathbb{R}^{N \times d}$  is a Gaussian vector. A  $K_f$ -Lipschitz function  $f$  of  $\mathbf{W}$  with respect to the Frobenius norm is also a  $K_f$ -Lipschitz function of  $\text{vec}(\mathbf{W})$  with respect to the Euclidean norm. Applying Lemma L.1 gives

$$\Pr(|f(\mathbf{W}) - \mathbb{E}[f(\mathbf{W})]| > t) = \Pr\left(|f(\varphi(\tilde{\mathbf{W}})) - \mathbb{E}[f(\varphi(\tilde{\mathbf{W}}))]| > t\right) \leq C e^{-\frac{ct^2}{K_\varphi^2 K_f^2}},$$

for some  $C, c > 0$ .  $\square$

**Lemma K.4.** Under Assumptions 1 and 2, let  $\lambda > 0$ , let  $\mathbf{W} \in \mathbb{R}^{N \times d}$ , and let the resolvent

$$\mathbf{Q}_m(\mathbf{W}) = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{\mathcal{S}}}(\mathbf{W})^T \Sigma_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}$$

defined as in equation 13. Let  $\boldsymbol{\sigma} \in \mathbb{R}^m$  independent of  $\mathbf{W}$  such that  $\frac{1}{\sqrt{m}} \|\boldsymbol{\sigma}\| \leq \sqrt{K_v}$  for  $K_v > 0$ . Then

$$\Pr \left( \left| \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} - \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_m(\mathbf{W})] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| > t \right) \leq C e^{-cmt^2},$$

for some  $C, c > 0$  independent of  $m$  and  $N$ .

*Proof.* Let the function  $f : \mathbf{W} \mapsto \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma}$ . We want to show  $f$  is Lipschitz in order to apply Lemma K.3. From Lemma I.1, we know there exists a real  $K > 0$  such that, for all  $m$  and for any  $\mathbf{W}$ , we have

$$\|\mathbf{Q}_m(\mathbf{W})\| \leq K.$$

Furthermore, both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1. Let  $\mathbf{H} \in \mathbb{R}^{N \times d}$ , we have

$$\begin{aligned} & |f(\mathbf{W} + \mathbf{H}) - f(\mathbf{W})| \\ &= \left| \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \left[ \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W}) \right] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| \\ &= \left| \frac{1}{m^2} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\ &\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W})^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| \\ &= \left| \frac{1}{m^2} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right. \right. \\ &\quad \left. \left. + \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right]^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| \\ &\leq \left| \frac{1}{m^2} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\ &\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| \\ &\quad + \left| \frac{1}{m^2} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\ &\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right]^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| \\ &\leq 2K_v K \left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \right\| \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \\ &\quad + 4K_v K \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \right\|. \end{aligned}$$

From Lemma I.4, we know there exists a real  $K' > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \right\| \leq K'$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \right\| \leq 2K'.$$

From those results, we conclude the Lipschitz continuity of  $f$  since

$$\begin{aligned} |f(\mathbf{W} + \mathbf{H}) - f(\mathbf{W})| &\leq 8K_v K K' \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \\ &\leq 8K_v K K' \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\|_F \\ &\leq \frac{8K_v K K' K_\sigma}{\sqrt{m}} \|\mathbf{H}\mathbf{S}\|_F \\ &= \frac{8K_v K K' K_\sigma}{\sqrt{m}} \sqrt{\text{Tr}(\mathbf{H}\mathbf{S}\mathbf{S}^T\mathbf{H}^T)} \\ &\leq \frac{8K_v K K' K_\sigma}{\sqrt{m}} \|\mathbf{S}\| \|\mathbf{H}\|_F. \end{aligned}$$



The last inequality is obtained because  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{B}\| \text{Tr}(\mathbf{A})$  for some semi-positive-definite matrix  $\mathbf{A}$ . We prove that  $f$  is Lipschitz with parameter  $\frac{8K_v K K' K_\sigma}{\sqrt{m}} \|\mathbf{S}\|$ , and applying Lemma L.1 gives

$$\begin{aligned} & \Pr \left( \left| \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} - \frac{1}{m} \boldsymbol{\sigma}^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_m(\mathbf{W})] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\sigma} \right| > t \right) \\ & \leq C e^{-\frac{cm t^2}{2^6 \kappa_\sigma^2 \kappa^2 \kappa'^2 \kappa_\sigma^2 \kappa_\varphi^2 \|\mathbf{S}\|^2}}, \end{aligned}$$

for some  $C, c > 0$  independent of other parameters.  $\square$

**Lemma K.5.** *Under Assumptions 1 and 2, let  $\mathbf{Q}_- \in \mathbb{R}^{n \times n}$  be the resolvent defined in equation 29, let  $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$  be a Gaussian vector independent of  $\mathbf{Q}_-$ , and let  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  be a real 1-Lipschitz function. Then*

$$\begin{aligned} & \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}_i^T \hat{\mathbf{S}}) \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \text{Tr} \left( \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbb{E}[\sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \sigma(\mathbf{w}_i^T \hat{\mathbf{S}})] \right) \right| > t \right) \\ & \leq C e^{-cm \max(t^2, t)}, \end{aligned}$$

for some  $C, c > 0$  independent of  $N, m$ .

*Proof.* We can observe that

$$\begin{aligned} & \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}_i^T \hat{\mathbf{S}}) \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \text{Tr} \left( \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbb{E}[\sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \sigma(\mathbf{w}_i^T \hat{\mathbf{S}})] \right) \right| > t \right) \\ & \leq \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right| > \frac{t}{2} \right) \\ & + \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \text{Tr} \left( \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbb{E}[\sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \sigma(\mathbf{w}_i^T \hat{\mathbf{S}})] \right) \right| > \frac{t}{2} \right). \end{aligned} \tag{82}$$

From Lemma I.1, there exists a real  $K > 0$  such that, for all  $m$ , we have

$$\|\mathbf{Q}_{-i}\| \leq K.$$

Besides, both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1. We thus bound the probability of the right-hand part with Lemma K.2 as

$$\begin{aligned} & \Pr \left( \left| \frac{1}{m} \sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \text{Tr} \left( \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbb{E}[\sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \sigma(\mathbf{w}_i^T \hat{\mathbf{S}})] \right) \right| > t \right) \\ & \leq C e^{-\frac{cm}{2\kappa_\sigma^2 \kappa_\varphi^2 \|\mathbf{S}\|^2} \min\left(\frac{t^2}{2^8 t_0^2 \kappa^2}, \frac{t}{2\kappa}\right)}, \end{aligned} \tag{83}$$

for  $t_0 = |\sigma(0)| + \sqrt{\frac{d}{m} K_\sigma K_\varphi \|\hat{\mathbf{S}}\|}$ , and  $c, C \in \mathbb{R}$  independent of all other parameters. Let define the real  $K' > 0$  and let  $\mathcal{A}_{K'}$  be the probability space defined as

$$\mathcal{A}_{K'} = \{\mathbf{w} \in \mathbb{R}^m, \|\sigma(\mathbf{w}^T \hat{\mathbf{S}})\| \leq K' \sqrt{m}\}.$$

From Lemma K.1, we bound the second term  $\Pr(\mathcal{A}_{K'}^c)$  as

$$\Pr(\mathcal{A}_{K'}^c) = \Pr(\{\|\sigma(\mathbf{w}^T \hat{\mathbf{S}})\| > K' \sqrt{m}\}) \leq C' e^{-\frac{c' m K'^2}{2\kappa_\sigma^2 \kappa_\phi^2 \|\mathbf{x}\|^2}},$$

for some  $c', C' > 0$  independent of other parameters. Conditioning the random variable of interest with respect to  $\mathcal{A}_{K'}$  and its complementary  $\mathcal{A}_{K'}^c$  gives with Lemma K.4

$$\begin{aligned} & \Pr\left(\left|\frac{1}{m}\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbf{Q}_- (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) - \frac{1}{m}\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}})\right| > t\right) \\ & \leq \Pr\left(\left|\frac{1}{m}\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbf{Q}_- (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}}) \right. \right. \\ & \quad \left. \left. - \frac{1}{m}\sigma(\mathbf{w}^T \hat{\mathbf{S}})^T \hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_-] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\mathbf{w}^T \hat{\mathbf{S}})\right| > t \cap \mathcal{A}_{K'}\right) + \Pr(\mathcal{A}_{K'}^c) \\ & \leq C'' e^{-c' m t^2} + C' e^{-\frac{c' m K'^2}{2\kappa_\sigma^2 \kappa_\phi^2 \|\hat{\mathbf{S}}\|^2}}, \end{aligned} \tag{84}$$

where  $c'', C'' > 0$ . Combing both equation 83 and equation 84 with equation 82 gives

$$\begin{aligned} & \Pr\left(\left|\frac{1}{m}\sigma(\mathbf{w}_i^T \hat{\mathbf{S}}) \hat{\mathbf{U}}_n \mathbf{Q}_{-i} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \right. \right. \\ & \quad \left. \left. - \frac{1}{m} \text{Tr}\left(\hat{\mathbf{U}}_n \mathbb{E}[\mathbf{Q}_{-i}] (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \mathbb{E}[\sigma(\hat{\mathbf{S}}^T \mathbf{w}_i) \sigma(\mathbf{w}_i^T \hat{\mathbf{S}})]\right)\right| > t\right) \\ & \leq C e^{-\frac{cm}{2\kappa_\sigma^2 \kappa_\phi^2 \|\hat{\mathbf{S}}\|^2} \min\left(\frac{t^2}{2\sigma_0^2 \kappa^2}, \frac{t}{4\kappa}\right)} + C'' e^{-\frac{c' m t^2}{4}} + C' e^{-\frac{c' m K'^2}{2\kappa_\sigma^2 \kappa_\phi^2 \|\mathbf{x}\|^2}}. \end{aligned} \tag{85}$$

□

**Lemma K.6.** Under Assumptions 1 and 2, let  $\lambda > 0$ , let  $\mathbf{W} \in \mathbb{R}^{N \times d}$ , and let the resolvent

$$\mathbf{Q}_m(\mathbf{W}) = \left[ \frac{1}{m} (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \Sigma_{\hat{\mathbf{S}}}(\mathbf{W})^T \Sigma_{\hat{\mathbf{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n + \lambda \mathbf{I}_n \right]^{-1}$$

defined as in equation 13. Let  $\mathbf{u} \in \mathbb{R}^n$  such that  $\|\mathbf{u}\| \leq K_u$  for  $K_u > 0$ . Then

$$\Pr\left(\left|\frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W}) \mathbf{u} - \frac{\lambda^2}{n} \mathbf{u}^T \mathbb{E}[\mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W})] \mathbf{u}\right| > t\right) \leq C e^{-cn^2 m t^2},$$

for some  $C, c > 0$  independent of  $m, n$  and  $N$ .

*Proof.* Let the function  $f : \mathbf{W} \mapsto \frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W}) \mathbf{u}$ . We want to show  $f$  is Lipschitz in order to apply Lemma K.3. From Lemma I.1, we know there exists a real  $K > 0$  such that, for all  $m$  and  $\mathbf{W}$ , we have

$$\|\mathbf{Q}_m(\mathbf{W})\| \leq K.$$

Furthermore, both  $\|\hat{\mathbf{U}}_n\|$  and  $\|\hat{\mathbf{V}}_n\|$  are upper bounded by 1. Let  $\mathbf{H} \in \mathbb{R}^{N \times d}$ , we have

$$\begin{aligned} |f(\mathbf{W} + \mathbf{H}) - f(\mathbf{W})| &= \left| \frac{\lambda^2}{n} \mathbf{u}^T [\mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W})] \mathbf{u} \right| \\ &\leq \underbrace{\left| \frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T [\mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W})] \mathbf{u}^T \right|}_{(1)} \\ &\quad + \underbrace{\left| \frac{\lambda^2}{n} \mathbf{u}^T [\mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W})]^T \mathbf{Q}_m(\mathbf{W}) \mathbf{u} \right|}_{(2)} \end{aligned}$$

For (1), we have

$$\begin{aligned}
& \left| \frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \left[ \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W}) \right] \mathbf{u}^T \right| \\
&= \left| \frac{\lambda^2}{n} \frac{1}{m} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\
&\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W})^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \mathbf{u}^T \right| \\
&= \left| \frac{\lambda^2}{n} \frac{1}{m} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right. \right. \\
&\quad \left. \left. + \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right]^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \mathbf{u}^T \right| \\
&\leq \left| \frac{\lambda^2}{n} \frac{1}{m} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\
&\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \mathbf{u}^T \right| \\
&\quad + \left| \frac{\lambda^2}{n} \frac{1}{m} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) \right. \right. \\
&\quad \left. \left. - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right]^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \mathbf{u}^T \right| \\
&\leq \frac{\lambda^2}{n} K^2 K_u^2 \left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \right\| \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \\
&\quad + \frac{2\lambda^2}{n} K^2 K_u^2 \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \right\|
\end{aligned}$$

From Lemma L.4, we know there exists a real  $K' > 0$  such that, for all  $m$ , we have

$$\left\| \frac{1}{\sqrt{m}} \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \hat{\mathbf{U}}_n \mathbf{Q}_m(\mathbf{W}) \right\| \leq K'$$

and

$$\left\| \frac{1}{\sqrt{m}} \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) (\hat{\mathbf{U}}_n - \gamma \hat{\mathbf{V}}_n)^T \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H})^T \right\| \leq 2K'.$$

From those results, we conclude for (1) that

$$\begin{aligned}
& \left| \frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W} + \mathbf{H})^T \left[ \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W}) \right] \mathbf{u}^T \right| \\
&\leq \frac{4\lambda^2 K^2 K_u^2 K'}{n} \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\| \\
&\leq \frac{4\lambda^2 K^2 K_u^2 K'}{n} \left\| \frac{1}{\sqrt{m}} \left[ \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W} + \mathbf{H}) - \boldsymbol{\Sigma}_{\hat{\mathcal{S}}}(\mathbf{W}) \right] \right\|_F \\
&\leq \frac{4\lambda^2 K^2 K_u^2 K' K_\sigma}{n\sqrt{m}} \|\mathbf{H}\mathbf{S}\|_F \\
&= \frac{4\lambda^2 K^2 K_u^2 K' K_\sigma}{n\sqrt{m}} \sqrt{\text{Tr}(\mathbf{H}\mathbf{S}\mathbf{S}^T\mathbf{H}^T)} \\
&\leq \frac{4\lambda^2 K^2 K_u^2 K' K_\sigma}{n\sqrt{m}} \|\mathbf{S}\| \|\mathbf{H}\|_F.
\end{aligned}$$

The last inequality is obtained because  $|\text{Tr}(\mathbf{A}\mathbf{B})| \leq \|\mathbf{B}\| \text{Tr}(\mathbf{A})$  for some semi-positive-definite matrix  $\mathbf{A}$ . With a similar reasoning, we can prove for (2) that

$$\left| \frac{\lambda^2}{n} \mathbf{u}^T \left[ \mathbf{Q}_m(\mathbf{W} + \mathbf{H}) - \mathbf{Q}_m(\mathbf{W}) \right]^T \mathbf{Q}_m(\mathbf{W}) \mathbf{u} \right| \leq \frac{4\lambda^2 K^2 K_u^2 K' K_\sigma}{n\sqrt{m}} \|\mathbf{S}\| \|\mathbf{H}\|_F.$$

We thus prove that  $f$  is Lipschitz with parameter  $\frac{8\lambda^2 K^2 K_u^2 K' K_\sigma}{n\sqrt{m}} \|S\|$ , and applying Lemma L.1 gives

$$\Pr \left( \left| \frac{\lambda^2}{n} \mathbf{u}^T \mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W}) \mathbf{u} - \frac{\lambda^2}{n} \mathbf{u}^T \mathbb{E}[\mathbf{Q}_m(\mathbf{W})^T \mathbf{Q}_m(\mathbf{W})] \mathbf{u} \right| > t \right) \leq C e^{-\frac{cn^2 m t^2}{2^6 \lambda^2 K^4 K_u^4 K'^2 K_\sigma^2 \|S\|^2}},$$

for some  $C, c > 0$  independent of other parameters.  $\square$

## L INTERMEDIARY LEMMAS

**Lemma L.1** (Normal Concentration). (*Ledoux, 2001, Corollary 2.6, Propositions 1.3, 1.8*) or (*Tao, 2012, Theorem 2.1.12*) For  $d \in \mathbb{N}$ , consider  $\mu$  the canonical Gaussian probability on  $\mathbb{R}^d$  defined through its density  $d\mu(\mathbf{w}) = (2\pi)^{-\frac{d}{2}} e^{-\frac{1}{2}\|\mathbf{w}\|^2}$  and  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  a  $L_f$ -Lipschitz function. Then

$$\mu \left( \left\{ \left| f - \int f d\mu \right| \geq t \right\} \right) \leq C e^{-c \frac{t^2}{L_f^2}}, \quad (86)$$

where  $C, c > 0$  are independent of  $d$  and  $L_f$ .

**Lemma L.2** (Resolvent Identity). For invertible matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ ,

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{B} - \mathbf{A})\mathbf{B}^{-1}$$

**Lemma L.3** (Sherman–Morrison–Woodbury Matrix Identity). (*Horn & Johnson, 2012, Theorem 0.7.4*) Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  be a non-singular matrix with a known inverse  $\mathbf{A}^{-1}$ ; let  $\mathbf{M} = \mathbf{A} + \mathbf{UCV}$ , in which  $\mathbf{U} \in \mathbb{R}^{k \times n}$ ,  $\mathbf{V} \in \mathbb{R}^{n \times k}$ , and  $\mathbf{C}^{k \times k}$  is non-singular. If  $\mathbf{M}$  and  $\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}$  are non-singular then

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}, \quad (87)$$

In particular  $(\mathbf{A} + \mathbf{UV})^{-1}\mathbf{U} = \mathbf{A}^{-1}\mathbf{U}(\mathbf{I}_n + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}$  and  $\mathbf{V}(\mathbf{A} + \mathbf{UV})^{-1} = (\mathbf{I}_n + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$ .

**Lemma L.4** (Sherman–Morrison Formula). Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  be a non-singular matrix with a known inverse  $\mathbf{A}^{-1}$ ; let  $\mathbf{M} = \mathbf{A} + \mathbf{u}\mathbf{v}^T$ , in which  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ . If  $\mathbf{M}$  is non-singular and  $1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u} \neq 0$  then

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}}. \quad (88)$$

In particular,  $(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} \mathbf{u} = \frac{\mathbf{A}^{-1} \mathbf{u}}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}}$  and  $\mathbf{v}^T (\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \frac{\mathbf{v}^T \mathbf{A}^{-1}}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}}$ . This Lemma is an extension of Lemma L.3.

**Lemma L.5** (Ostrowski’s Theorem). (*Horn & Johnson, 2012, Theorem 4.5.9*) Let  $\mathbf{A}, \mathbf{S} \in \mathbb{R}^{n \times n}$  with  $\mathbf{A}$  Hermitian and  $\mathbf{S}$  nonsingular. Let the eigenvalues of  $\mathbf{A}$ ,  $\mathbf{S}\mathbf{A}\mathbf{S}^T$ , and  $\mathbf{S}\mathbf{S}^T$  be arranged in nondecreasing order. Let  $\sigma_1 \geq \dots \geq \sigma_n > 0$  be the singular values of  $\mathbf{S}$ . For each  $k \in [n]$  there is a positive real number  $\theta_k \in [\sigma_n^2, \sigma_1^2]$  such that

$$\nu_k(\mathbf{S}\mathbf{A}\mathbf{S}^T) = \theta_k \nu_k(\mathbf{A})$$

**Lemma L.6** (Weinstein–Aronszajn Identity). For  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$  and  $\lambda \in \mathbb{R} \setminus \{0\}$ ,

$$\det(\mathbf{A}\mathbf{B} - \lambda \mathbf{I}_m) = (-\lambda)^{m-n} \det(\mathbf{B}\mathbf{A} - \lambda \mathbf{I}_n).$$

It follows that the non-zero eigenvalues of  $\mathbf{A}\mathbf{B}$  and  $\mathbf{B}\mathbf{A}$  are the same.

**Lemma L.7.** Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  and  $\lambda > 0$ .

$$\|(\mathbf{A} + \lambda \mathbf{I}_n)^{-1}\| \leq \frac{1}{\lambda}$$

if and only if  $\mathbf{A}\mathbf{A}^T + \lambda(\mathbf{A} + \mathbf{A}^T)$  is positive definite. In particular, for matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  whose the Hermitian part  $H(\mathbf{A}) = \frac{\mathbf{A} + \mathbf{A}^T}{2}$  is semi-positive-definite we have

$$\|(\mathbf{A} + \lambda \mathbf{I}_n)^{-1}\| \leq \frac{1}{\lambda}$$

*Proof.*

$$\begin{aligned}
\|(\mathbf{A} + \lambda \mathbf{I}_n)^{-1}\|^2 &= \nu_{max} \left( (\mathbf{A} + \lambda \mathbf{I}_n)^{-1T} (\mathbf{A} + \lambda \mathbf{I}_n)^{-1} \right) \\
&= \nu_{max} \left( [(\mathbf{A} + \lambda \mathbf{I}_n) (\mathbf{A}^T + \lambda \mathbf{I}_n)]^{-1} \right) \\
&= \nu_{max} \left( (\mathbf{A} \mathbf{A}^T + \lambda(\mathbf{A} + \mathbf{A}^T) + \lambda^2 \mathbf{I}_n)^{-1} \right) \\
&= \nu_{min} \left( (\mathbf{A} \mathbf{A}^T + \lambda(\mathbf{A} + \mathbf{A}^T) + \lambda^2 \mathbf{I}_n) \right)^{-1}
\end{aligned} \tag{89}$$

where  $\nu_{max}(\mathbf{B})$  and  $\nu_{min}(\mathbf{B})$  denotes the maximum eigenvalue and minimum eigenvalues of a matrix  $\mathbf{B}$ . Since  $\mathbf{A}$  is positive-definite the matrix  $\mathbf{A} \mathbf{A}^T + \lambda(\mathbf{A} + \mathbf{A}^T)$  is semi-positive-definite and has positive nonzeros eigenvalues. Therefore,  $\nu_{min} \left( (\mathbf{A} \mathbf{A}^T + \lambda(\mathbf{A} + \mathbf{A}^T) + \lambda^2 \mathbf{I}_n) \right) > \lambda^2$  and  $\|(\mathbf{A} + \lambda \mathbf{I}_n)^{-1}\| \leq \frac{1}{\lambda}$

□

**Lemma L.8.** (*Yates, 1995, Theorem 2*) If a mapping  $f : [0, \infty) \rightarrow [0, \infty)$

- is monotonically increasing, i.e  $x \geq x' \implies f(x) \geq f(x')$ ,
- is scalable, i.e  $\forall \alpha > 1, \alpha f(x) > f(\alpha x)$ ,
- admits  $x_0 \in [0, \infty)$  such that  $x_0 \geq f(x_0)$ ,

then  $f$  has a unique fixed-point.

**Lemma L.9.** Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\mathbf{B} \in \mathbb{R}^{n \times m}$ . If  $\mathbf{A} \mathbf{B} + \lambda \mathbf{I}_m$  is invertible, then

$$[\mathbf{A} \mathbf{B} + \lambda \mathbf{I}_m]^{-1} \mathbf{A} = \mathbf{A} [\mathbf{B} \mathbf{A} + \lambda \mathbf{I}_n]^{-1}.$$

*Proof.* We have

$$\mathbf{A} [\mathbf{B} \mathbf{A} + \lambda \mathbf{I}_n] = [\mathbf{A} \mathbf{B} + \lambda \mathbf{I}_m] \mathbf{A}$$

Since both  $\mathbf{A} \mathbf{B}$  and  $\mathbf{B} \mathbf{A}$  share the same non-zero eigenvalues from Lemma L.6, we deduce  $\mathbf{B} \mathbf{A} + \lambda \mathbf{I}_n$  is also invertible. By multiplying the equation above with both the inverse of  $[\mathbf{B} \mathbf{A} + \lambda \mathbf{I}_n]$  and  $[\mathbf{A} \mathbf{B} + \lambda \mathbf{I}_m]$ , we get

$$[\mathbf{A} \mathbf{B} + \lambda \mathbf{I}_m]^{-1} \mathbf{A} = \mathbf{A} [\mathbf{B} \mathbf{A} + \lambda \mathbf{I}_n]^{-1}$$

□