A Proofs and Derivations

In this appendix, we report the proofs and derivations of the results presented in the main paper.

A.1 Proofs of Section 4

Theorem 4.1. *If $\boldsymbol{\Sigma}$ is positive definite, the optimization problem* (5) *can be restated as:*

$$\min_{\substack{\boldsymbol{\omega}\in\mathbb{R}^q_+ \\ \|\boldsymbol{\omega}\|_1=1}} \left\| \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \boldsymbol{\omega} \right\|^2_{\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}}, \tag{6}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{I}_d$ is the identity matrix of order d. Furthermore, the approximating Jacobian $\mathbf{M}(\boldsymbol{\omega})$ is given by:

$$\begin{aligned} \operatorname{vec}(\mathbf{M}(\boldsymbol{\omega})) =& \Big\{ \mathbf{I}_{dq} - \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d) \left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1} \\ & \times (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \Big\} \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right). \end{aligned}$$

Proof. The proof is analogous of that of Theorem 1 of (Manton et al., 2003). We report it using our notation for completeness. Let $\boldsymbol{\omega} \in \mathbb{R}^q$ be a fixed vector, we are to solve the following optimization problem:

$$\min_{\substack{\mathbf{M}\in\mathbb{R}^{d\times q} \\ \mathbf{M}\boldsymbol{\omega}=\mathbf{0}}} \left\| \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \mathbf{M}\right) \right\|^2_{\boldsymbol{\Sigma}^{-1}}$$

We employ Lagrange multipliers, leading to the Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{M}, \boldsymbol{\lambda}) &= \left\| \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \mathbf{M}\right) \right\|^2_{\boldsymbol{\Sigma}^{-1}} + \boldsymbol{\lambda}^T \mathbf{M}\boldsymbol{\omega} \\ &= \left\| \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) - \operatorname{vec}(\mathbf{M}) \right\|^2_{\boldsymbol{\Sigma}^{-1}} + \boldsymbol{\lambda}^T (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \operatorname{vec}(\mathbf{M}), \end{aligned}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^d$ is the Lagrange multiplier and we exploited the properties of the vectorization operator and the Kronecker product to derive the second equation. We notice that the Lagrangian function $\mathcal{L}$ is convex w.r.t. to $\operatorname{vec}(\mathbf{M})$. Thus, we make the gradient vanish:

$$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial \operatorname{vec}(\mathbf{M})} = 2\boldsymbol{\Sigma}^{-1} \left(\operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) - \operatorname{vec}(\mathbf{M}) \right) + (\boldsymbol{\omega}\otimes\mathbf{I}_d)\boldsymbol{\lambda} = \mathbf{0} \\ &\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} = (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \operatorname{vec}(\mathbf{M}) = \mathbf{0}. \end{aligned}$$

From the first equation, we obtain an expression for $\operatorname{vec}(\mathbf{M})$ as a function of $\boldsymbol{\lambda}$:

$$\operatorname{vec}(\mathbf{M}) = \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) + \frac{1}{2}\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\boldsymbol{\lambda}.$$

By substituting into the second equation, we get the value of $\boldsymbol{\lambda}$:

$$\begin{aligned} &(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) + \frac{1}{2}(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\boldsymbol{\lambda} = \mathbf{0} \\ &\implies \boldsymbol{\lambda} = -2\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right). \end{aligned}$$

Finally, we get the expression for $\operatorname{vec}(\mathbf{M})$:

$$\begin{aligned} \operatorname{vec}(\mathbf{M}) &= \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) - \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) \\ &= \left\{ \mathbf{I}_{dq} - \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T \right\} \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right). \end{aligned}$$

We can now substitute the value of $\operatorname{vec}(\mathbf{M})$ into the loss function:

$$\begin{aligned}
\left\|\operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}) - \mathbf{M}\right)\right\|^2_{\boldsymbol{\Sigma}^{-1}} &= \left\|\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\right\|^2_{\boldsymbol{\Sigma}^{-1}} \\
&= \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)^T(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} \\
&\quad\times\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right) \\
&= \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)^T(\boldsymbol{\omega}\otimes\mathbf{I}_d)\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right) \\
&= \left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right)^T\left[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)\right]^{-1}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right) \\
&= \left\|\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_{[(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d)]^{-1}}
\end{aligned}$$

where we employed the properties of the vectorization operator and the Kronecker product in the last but one line and the definition of norm in the last line. □

Corollary 4.1. *Let* $\mathbf{Q} \in \mathbb{R}^{d\times d}$ *be a positive definite matrix and let* $\mathbb{1}_q$ *denote the q-dimensional vector of all ones. If* $\boldsymbol{\Sigma} = \mathbb{1}_q\mathbb{1}_q^T \otimes \mathbf{Q}$*, then objective function* (6) *is convex. Furthermore, if* $\mathbf{Q} = \mathbf{I}_d$*, then the objective function* (6) *is equivalent to* (2) *with* $p = 2$*.*

Proof. When $\boldsymbol{\Sigma} = \mathbb{1}_q\mathbb{1}_q^T \otimes \mathbf{Q}$, we can provide the following derivation exploiting the properties of the Kronecker product and recalling that $\boldsymbol{\omega}^T\mathbb{1}_q = 1$ because of the enforced constraints:

$$\begin{aligned}
(\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\boldsymbol{\Sigma}(\boldsymbol{\omega}\otimes\mathbf{I}_d) &= (\boldsymbol{\omega}\otimes\mathbf{I}_d)^T\left(\mathbb{1}_q\mathbb{1}_q^T\otimes\mathbf{Q}\right)(\boldsymbol{\omega}\otimes\mathbf{I}_d) \\
&= (\boldsymbol{\omega}^T\mathbb{1}_q\mathbb{1}_q^T\otimes\mathbf{I}_d\mathbf{Q})(\boldsymbol{\omega}\otimes\mathbf{I}_d) \\
&= (\mathbb{1}_q^T\otimes\mathbf{Q})(\boldsymbol{\omega}\otimes\mathbf{I}_d) \\
&= \mathbb{1}_q^T\boldsymbol{\omega}\otimes\mathbf{Q}\mathbf{I}_d \\
&= \mathbf{Q}.
\end{aligned}$$

Therefore, the objective function becomes $\left\|\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_{\mathbf{Q}^{-1}}$ which is clearly convex in $\boldsymbol{\omega}$, as $\mathbf{Q}$ is positive definite. Moreover, if we take $\mathbf{Q} = \mathbf{I}_d$, then we have $\left\|\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_{\mathbf{I}_d} = \left\|\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_2$, that is the objective function (2) optimized by GIRL when $p = 2$. □

A.2 Proofs of Section 4.1

Lemma A.1. *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ *any pair of vectors, then it holds that:*

$$\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|_2} - \frac{\mathbf{y}}{\|\mathbf{y}\|_2}\right\|_2 \leq \frac{2\,\|\mathbf{x}-\mathbf{y}\|_2}{\max\{\|\mathbf{x}\|_2, \|\mathbf{y}\|_2\}}.$$

Proof. The result follows from the following sequence of algebraic manipulations:

$$\begin{aligned}
\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|_2} - \frac{\mathbf{y}}{\|\mathbf{y}\|_2}\right\|_2 &= \left\|\frac{\mathbf{x}}{\|\mathbf{x}\|_2} - \frac{\mathbf{y}}{\|\mathbf{y}\|_2} \pm \frac{\mathbf{y}}{\|\mathbf{x}\|_2}\right\|_2 \\
&\leq \frac{\|\mathbf{x}-\mathbf{y}\|_2}{\|\mathbf{x}\|_2} + \frac{|\|\mathbf{x}\|_2 - \|\mathbf{y}\|_2|}{\|\mathbf{x}\|_2} \\
&\leq 2\frac{\|\mathbf{x}-\mathbf{y}\|_2}{\|\mathbf{x}\|_2},
\end{aligned}$$

where we applied the triangular inequality in the second line and the reverse triangular inequality in the last one, i.e., $|\|\mathbf{x}\|_2 - \|\mathbf{y}\|_2| \leq \|\mathbf{x}-\mathbf{y}\|_2$. By observing that, for symmetry reasons, the same derivation can be performed getting $\|\mathbf{y}\|_2$ at the denominator, we get the result. □

Lemma A.2. *Let* $\mathbf{A} = (\mathbf{a}_1|\dots|\mathbf{a}_q)$, $\mathbf{B} = (\mathbf{b}_1|\dots|\mathbf{b}_q) \in \mathbb{R}^{d\times q}$ *be two matrices of rank* $q-1$ *such that* $s_{q-1}(\mathbf{B}) > 0$, *where* s_{q-1} *denotes the* $(q-1)$*-th singular value. Let* $\mathcal{A} = \operatorname{span}(\{\mathbf{a}_1,\dots,\mathbf{a}_q\})$ *and* $\mathcal{B} = \operatorname{span}(\{\mathbf{b}_1,\dots,\mathbf{b}_q\})$ *be the vector spaces generated by the columns of* $\mathbf{A}$ *and* $\mathbf{B}$ *respectively. Then, the cosine of the (principal) angle* α *between the corresponding orthogonal complements* $\mathcal{A}^\perp$ *and* $\mathcal{B}^\perp$ *is lower bounded by:*

$$\cos\alpha = \cos\sphericalangle\left(\mathcal{A}^\perp,\mathcal{B}^\perp\right) \ge 1 - \frac{2}{s_{q-1}(\mathbf{A})^2} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \|\mathbf{A}-\mathbf{B}\mathbf{\Pi}\|_F^2,$$

where Perm_q *is the set of all permutation matrices of order* q *and* $\|\cdot\|_F$ *denotes the Frobenius norm.*

Proof. Since both matrices $\mathbf{A}$ and $\mathbf{B}$ have rank $q-1$, the orthogonal complements $\mathcal{A}^\perp$ and $\mathcal{B}^\perp$ have dimension 1. Since the principal angles (which in this case is just one) of the orthogonal complements are essentially the same as those of the correponding spaces (Knyazev et al., 2010), we reduce the problem to the computation of $\sphericalangle(\mathcal{A},\mathcal{B})$. In particular, we are interested in the maximum (and only non-zero) principal angle α, whose cosine can be conveniently defined as (Taslaman, 2014):

$$\cos\alpha = \min_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1}} \max_{\substack{\mathbf{y}\in\mathbb{R}^q\\ \|\mathbf{B}\mathbf{y}\|_2=1}} (\mathbf{A}\mathbf{x})^T\mathbf{B}\mathbf{y} = 1 - \frac{1}{2}\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1}} \min_{\substack{\mathbf{y}\in\mathbb{R}^q\\ \|\mathbf{B}\mathbf{y}\|_2=1}} \|\mathbf{A}\mathbf{x}-\mathbf{B}\mathbf{y}\|_2^2,$$

where the identity follows from recalling that $\|\mathbf{a}-\mathbf{b}\|_2^2 = \|\mathbf{a}\|_2^2 + \|\mathbf{b}\|_2^2 - 2\mathbf{a}^T\mathbf{b}$. Consider now the set $\mathcal{X} = \{\mathbf{x}\in\mathbb{R}^q : \|\mathbf{A}\mathbf{x}\|_2 = 1\}$. Since $\mathbf{A}$ is not full rank, the set $\mathcal{X}$ will contain vectors with non-zero projection onto the null-space of $\mathbf{A}$. Thus, for any $\mathbf{x}\in\mathcal{X}$ we can write $\mathbf{x} = \mathbf{x}^\perp + \mathbf{x}^\parallel$, where $\mathbf{x}^\perp \in \mathrm{null}(\mathbf{A})$ and $\mathbf{x}^\parallel \perp \mathrm{null}(\mathbf{A})$. Furthermore, we have that $\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x}^\perp + \mathbf{x}^\parallel) = \mathbf{A}\mathbf{x}^\parallel$, by definition of null space. Therefore, for the computation of the min, we can limit our search of $\mathbf{x}$ to the set $\{\mathbf{x}\in\mathbb{R}^q : \|\mathbf{A}\mathbf{x}\|_2 = 1 \wedge \mathbf{x}\perp\mathrm{null}(\mathbf{A})\}$. Let $\mathbf{\Pi}$ be a permutation matrix, we now consider the following sequence of inequalities:

$$\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\substack{\mathbf{y}\in\mathbb{R}^q\\ \|\mathbf{B}\mathbf{y}\|_2=1}} \|\mathbf{A}\mathbf{x}-\mathbf{B}\mathbf{y}\|_2 \le \max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \left\|\mathbf{A}\mathbf{x} - \frac{\mathbf{B}\mathbf{\Pi}\mathbf{x}}{\|\mathbf{B}\mathbf{\Pi}\mathbf{x}\|_2}\right\|_2^2 \tag{8}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \frac{\|\mathbf{A}\mathbf{x}-\mathbf{B}\mathbf{\Pi}\mathbf{x}\|_2}{\max\{1,\|\mathbf{B}\mathbf{\Pi}\mathbf{x}\|_2\}} \tag{9}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \|\mathbf{A}\mathbf{x}-\mathbf{B}\mathbf{\Pi}\mathbf{x}\|_2 \tag{10}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \left\|\sum_{i=1}^q x_i\left(\mathbf{a}_i - \mathbf{b}_{\pi(i)}\right)\right\|_2 \tag{11}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \sum_{i=1}^q |x_i|\left\|\mathbf{a}_i - \mathbf{b}_{\pi(i)}\right\|_2 \tag{12}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \|\mathbf{x}\|_2 \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \sqrt{\sum_{i=1}^q \left\|\mathbf{a}_i - \mathbf{b}_{\pi(i)}\right\|_2^2} \tag{13}$$

$$\le 2\max_{\substack{\mathbf{x}\in\mathbb{R}^q\\ \|\mathbf{A}\mathbf{x}\|_2=1\\ \mathbf{x}\perp\mathrm{null}(\mathbf{A})}} \|\mathbf{x}\|_2 \min_{\mathbf{\Pi}\in\mathrm{Perm}_q} \|\mathbf{A}-\mathbf{B}\mathbf{\Pi}\|_F, \tag{14}$$

where line (8) follows from bounding the min over $\mathbf{y}$ with a specific choice of $\mathbf{y} = \mathbf{\Pi}\mathbf{x}$. Line (9) is obtained from Lemma A.1 and line (10) derives from bounding the maximum at the denominator with its first argument. Line (11) follows from the definition of permutation matrix, having denoted with $\pi : \{1,\dots,q\} \to \{1,\dots,q\}$ the permutation realized by $\mathbf{\Pi}$. Line (12) follows from expanding the expression at the previous line, while line (13) is an application of Cauchy-Swartz inequality. Finally, line (14) is obtained from the definition of Frobenius norm.

To conclude, we bound the norm $\|\mathbf{x}\|_2$ under the constraints $\|\mathbf{Ax}\|_2 = 1$ and $\mathbf{x} \perp \mathrm{null}(\mathbf{A})$. For this purpose, we consider the singular value decomposition of $\mathbf{A} = \mathbf{USV}^T$, where $\mathbf{S} = \mathrm{diag}(s_1, \ldots, s_{q-1}, 0)$ and $s_{q-1} > 0$ for the hypothesis. Moreover, let $\mathbf{V} = (\mathbf{v}_1 | \ldots | \mathbf{v}_q)$, we know that $\mathrm{null}(\mathbf{A}) = \mathrm{span}(\{\mathbf{v}_q\})$. Therefore, our chosen $\mathbf{x}$ is orthogonal to $\mathbf{x}^T \mathbf{v}_q = \mathbf{0}$. We now consider the matrix-vector product norm:

$$\|\mathbf{Ax}\|_2^2 = \left\|\mathbf{USV}^T\mathbf{x}\right\|_2^2 = \mathbf{x}^T\mathbf{VSU}^T\mathbf{USV}^T\mathbf{x} = \mathbf{x}^T\mathbf{VS}^2\mathbf{V}^T\mathbf{x} = \sum_{i=1}^{q-1} s_i^2 (\mathbf{x}^T\mathbf{v}_i)^2$$
$$\geq s_{q-1}^2 \sum_{i=1}^{q-1} (\mathbf{x}^T\mathbf{v}_i)^2 = s_{q-1}^2 \|\mathbf{x}\|_2^2,$$

where we exploited the fact that $\mathbf{U}$ is a unitary matrix and the fact that $\sum_{i=1}^{q-1}(\mathbf{x}^T\mathbf{v}_i)^2 = \|\mathbf{x}\|_2^2$, being the vectors of $\mathbf{V}$ an orthonormal basis. Using this result, and recalling that $\|\mathbf{Ax}\|_2 = 1$, we can upper bound the value of $\|\mathbf{x}\|_2$, to get the result. □

Lemma A.3. *Let $\mathbf{M}(\widehat{\boldsymbol{\omega}})$ be the approximate Jacobian recovered by Σ-GIRL run with the covariance matrix $\boldsymbol{\Sigma}$, starting from the sample Jacobian $\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$. Let $\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ be the true Jacobian. Then, it holds that:*

$$\|\mathrm{vec}\left(\mathbf{M}(\widehat{\boldsymbol{\omega}}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\|_2^2 \leq 4\|\boldsymbol{\Sigma}\|_2 \left\|\mathrm{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\right\|_{\boldsymbol{\Sigma}^{-1}}^2. \tag{15}$$

Proof. Given a vector $\mathbf{x}$, we upper bound the norm $\|\mathbf{x}\|_2$ with $\|\mathbf{x}\|_{\boldsymbol{\Sigma}^{-1}}$:

$$\|\mathbf{x}\|_{\boldsymbol{\Sigma}^{-1}}^2 = \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} \geq s_{\min}\left(\boldsymbol{\Sigma}^{-1}\right)\mathbf{x}^T\mathbf{x} = s_{\min}\left(\boldsymbol{\Sigma}^{-1}\right)\|\mathbf{x}\|_2^2,$$

where $s_{\min}(\cdot)$ is the minimum singular value of a matrix. Now, since $s_{\min}\left(\boldsymbol{\Sigma}^{-1}\right) = \frac{1}{s_{\max}(\boldsymbol{\Sigma})} = \frac{1}{\|\boldsymbol{\Sigma}\|_2}$, we have that $\|\mathbf{x}\|_2^2 \leq \|\boldsymbol{\Sigma}\|_2 \|\mathbf{x}\|_{\boldsymbol{\Sigma}^{-1}}^2$. Additionally, if $\boldsymbol{\Sigma}$ is the covariance matrix that is used for recovering $\mathbf{M}(\widehat{\boldsymbol{\omega}})$ it follows that the distance between $\mathbf{M}(\widehat{\boldsymbol{\omega}})$ and $\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ cannot be larger than twice the distance between $\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ and $\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$:

$$\|\mathrm{vec}\left(\mathbf{M}(\widehat{\boldsymbol{\omega}}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\|_{\boldsymbol{\Sigma}^{-1}} \leq 2\left\|\mathrm{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\right\|_{\boldsymbol{\Sigma}^{-1}}. \tag{16}$$

Putting these two inequalities together, we get the result. □

Theorem 4.2. *Let $\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ be an unbiased estimate of the Jacobian $\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ obtained with the trajectories $D = \{\tau_1, \ldots, \tau_n\}$. Let $\frac{1}{n}\boldsymbol{\Sigma} = \mathbb{C}\mathrm{ov}[\mathrm{vec}(\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}))]$ be the true covariance matrix of the estimated Jacobian. Let $\widehat{\boldsymbol{\omega}}$ be the weight vector recovered by Σ-GIRL run with covariance matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{\omega}^E$ be the expert's weight vector. If $\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ and $\mathbf{M}(\widehat{\boldsymbol{\omega}})$ have rank $q-1$ and $s_{q-1}(\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})) = s > 0$, where $s_{q-1}(\cdot)$ denotes the $(q-1)$-th singular value, then it holds that:*

$$\mathbb{E}\left[\left\|\widehat{\boldsymbol{\omega}} - \boldsymbol{\omega}^E\right\|_2\right] \leq \sqrt{\frac{16dq\,\|\boldsymbol{\Sigma}\|_2}{s^2 n}},$$

where the expectation is taken w.r.t. the randomness of the trajectories in D used to compute $\widehat{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$.

Proof. From the proof of Theorem 13.2 of (Pirotta, 2016), we know that:

$$\left\|\widehat{\boldsymbol{\omega}} - \boldsymbol{\omega}^E\right\|_2 \leq \sqrt{2(1 - \cos\alpha)}, \tag{17}$$

where α is the angle between the two vectors $\widehat{\boldsymbol{\omega}}$ and $\boldsymbol{\omega}^E$. We now provide a bound for $\cos\alpha$. Since $\widehat{\boldsymbol{\omega}}$ and $\boldsymbol{\omega}^E$ belong to the orthogonal complements of the column spaces generated by $\mathbf{M}(\widehat{\boldsymbol{\omega}})$ and $\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})$ respectively. From Lemma A.2, we have that:

$$\cos\alpha \geq 1 - \frac{2}{s_{q-1}(\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}))^2} \min_{\boldsymbol{\Pi}\in\mathrm{Perm}_q} \|\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}) - \mathbf{M}(\widehat{\boldsymbol{\omega}})\boldsymbol{\Pi}\|_F^2.$$

We now consider the following sequence of derivations:

$$\min_{\boldsymbol{\Pi}\in\mathrm{Perm}_q} \|\nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta}) - \mathbf{M}(\widehat{\boldsymbol{\omega}})\boldsymbol{\Pi}\|_F^2 \leq \|\mathbf{M}(\widehat{\boldsymbol{\omega}}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\|_F^2 \tag{18}$$
$$\leq \|\mathrm{vec}\left(\mathbf{M}(\boldsymbol{\omega}) - \nabla_{\boldsymbol{\theta}}\boldsymbol{\psi}(\boldsymbol{\theta})\right)\|_2^2 \tag{19}$$

$$\leq 4 \left\| \boldsymbol{\Sigma} \right\|_2 \left\| \mathrm{vec} \left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right\|_{\boldsymbol{\Sigma}^{-1}}^2, \tag{20}$$

where line (18) is obtained from selecting $\boldsymbol{\Pi} = \mathbf{I}_q$. Line (19) derives from observing that the Frobenius norm of a matrix equals the L^2-norm of the corresponding vectorization. Finally, line (20) follows from Lemma A.3. Putting this latter result into Equation (17), we have:

$$\left\| \widehat{\boldsymbol{\omega}} - \boldsymbol{\omega}^E \right\|_2 \leq \sqrt{\frac{16 \left\| \boldsymbol{\Sigma} \right\|_2}{s_{q-1}(\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}))^2} \left\| \mathrm{vec} \left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right\|_{\boldsymbol{\Sigma}^{-1}}^2}.$$

Now we compute the expectation of the norm of the difference:

$$\begin{aligned}
\mathbb{E}\left[\left\| \widehat{\boldsymbol{\omega}} - \boldsymbol{\omega}^E \right\|_2 \right] &\leq \mathbb{E}\left[\sqrt{\frac{16 \left\| \boldsymbol{\Sigma} \right\|_2}{s_{q-1}(\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}))^2} \left\| \mathrm{vec} \left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right\|_{\boldsymbol{\Sigma}^{-1}}^2} \right] \\
&\leq \sqrt{\frac{16 \left\| \boldsymbol{\Sigma} \right\|_2}{s_{q-1}(\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}))^2} \mathbb{E}\left[\left\| \mathrm{vec} \left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right\|_{\boldsymbol{\Sigma}^{-1}}^2 \right]},
\end{aligned}$$

where the last passage follows from Jensen inequality. To conclude, we compute the expectation inside the square root by observing that it is the expectation of a zero-mean random vector under the norm induced by its true covariance matrix. Thus, by renaming $\mathbf{x} = \mathrm{vec} \left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \in \mathbb{R}^{dq}$ and recalling that $\mathbb{E}[\mathbf{x}\mathbf{x}^T] = \mathbb{C}\mathrm{ov}[\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})] = \frac{\boldsymbol{\Sigma}}{n}$ we have:

$$\begin{aligned}
\mathbb{E}\left[\left\| \mathbf{x} \right\|_{\boldsymbol{\Sigma}^{-1}}^2 \right] &= \mathbb{E}\left[\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \right] = \mathbb{E}\left[\mathrm{tr}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}) \right] \\
&= \mathbb{E}\left[\mathrm{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{x}\mathbf{x}^T) \right] = \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1} \mathbb{E}\left[\mathbf{x}\mathbf{x}^T \right] \right) = \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1} \frac{\boldsymbol{\Sigma}}{n} \right) = \frac{dq}{n}.
\end{aligned}$$

□

Bound Discussion The bound on the error of the recovered weights depends on the size of the Jacobian matrix dq, on the L^2-norm of the true covariance matrix $\boldsymbol{\Sigma}$, as more uncertain gradients make the estimation of the true Jacobian harder, and on s_{q-1} (the last but one singular value of the true Jacobian matrix). The dependence on s_{q-1} is related to the reward feature space. This quantity replaces the quantity ρ of Theorem 13.2 of Pirotta (2016). The difference is ρ is a property of the estimated Jacobian, whereas s depends is a property of the true Jacobian matrix.

A.3 Proofs of Section 5

We obtain the expression of function $Q(\Omega, \Omega^{\mathrm{old}})$ following a derivation analogous to the one presented in Bilmes et al. (1998). We denote with $\mathbf{y} = (y_1, \dots, y_m)$ the realization of the random vector $\mathbf{Y}$.

$$\begin{aligned}
Q(\Omega, \Omega^{\mathrm{old}}) &= \mathop{\mathbb{E}}_{\mathbf{Y} \sim p(\cdot|\mathbf{X}; \Omega^{\mathrm{old}})} \left[\log \mathcal{L}(\Omega | \mathbf{D}, \mathbf{Y}) \right] \\
&= \sum_{\mathbf{y}} \log(\mathcal{L}(\Omega | \mathbf{D}, \mathbf{y})) p(\mathbf{y} | \mathbf{D}, \Omega^{\mathrm{old}}) \\
&= \sum_{\mathbf{y}} \sum_{i=1}^{m} \log(\alpha_{y_i} p_{y_i}(D_i | \boldsymbol{\omega}_{y_i})) \prod_{i'=1}^{m} p(y_{i'} | D_{i'}, \Omega^{\mathrm{old}}) \\
&= \sum_{y_1=1}^{k} \cdots \sum_{y_m=1}^{k} \sum_{i=1}^{m} \log(\alpha_{y_i} p_{y_i}(D_i | \boldsymbol{\omega}_{y_i})) \prod_{i'=1}^{m} p(y_{i'} | D_{i'}, \Omega^{\mathrm{old}}) = \\
&= \sum_{j=1}^{k} \sum_{i=1}^{m} \log(\alpha_j p(D_i | \boldsymbol{\omega}_j)) \sum_{y_1=1}^{k} \cdots \sum_{y_m=1}^{k} \mathbb{1}_{\{j = y_i\}} \prod_{i'=1}^{m} p(y_{i'} | D_{i'}, \Omega)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{k} \sum_{i=1}^{m} \log(\alpha_j p(D_i|\boldsymbol{\omega}_j)) z_{ij} \\
&= \sum_{j=1}^{k} \sum_{i=1}^{m} z_{ij} \left(\log \alpha_j + \log p(D_i|\boldsymbol{\omega}_j) \right).
\end{aligned}
$$

B Details on Optimization Problem (6)

B.1 Approximation of Σ as in Corollary 4.1

In this appendix, we provide a way to represent a generic $\boldsymbol{\Sigma}$ as a matrix of the form $\mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}$ as in Corollary 4.1. We seek for the minimum Frobenius-norm distance between $\boldsymbol{\Sigma}$ and $\mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}$:

$$
\min_{\substack{\mathbf{Q} \in \mathbb{R}^{d \times d} \\ \mathbf{Q} \succ \mathbf{0}_d}} \frac{1}{2} \left\| \boldsymbol{\Sigma} - \mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q} \right\|_F^2, \tag{21}
$$

where we required that $\mathbf{Q} \succ \mathbf{0}_d$, i.e., that $\mathbf{Q}$ is positive definite whenever $\boldsymbol{\Sigma}$ is.

We will solve the problem ignoring the constraint and we will prove that the resulting matrix is indeed positive definite whenever $\boldsymbol{\Sigma}$ is.

Lemma B.1. *Let $\boldsymbol{\Sigma} = n\,\mathbb{C}\mathrm{ov}[\mathrm{vec}(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta}))]$, the problem (21) admits a unique solution that is:*

$$
\mathbf{Q} = n\,\mathbb{C}\mathrm{ov}\left[\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi} \mathbb{1}_q \right] = \frac{1}{q^2} \sum_{i=1}^{q} \sum_{j=1}^{q} \boldsymbol{\Sigma}_{iq:(i+1)q, jq:(j+1)q} = \frac{1}{q^2} \left(\mathbb{1}_q \otimes \mathbf{I}_d \right)^T \boldsymbol{\Sigma} \left(\mathbb{1}_q \otimes \mathbf{I}_d \right), \tag{22}
$$

where we denoted with $\boldsymbol{\Sigma}_{i:i',j:j'}$ the submatrix obtained by taking the rows between i and i' and the columns between j and j'. Furthermore, $\mathbf{Q}$ is positive definite whenever $\boldsymbol{\Sigma}$ is.

Proof. Recall that the Kroneker product $\mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}$ constructs a matrix in which $\mathbf{Q}$ is repeated $q \times q$ times, arranged in a square matrix. Thus, it follows that we can rewrite the norm as:

$$
\frac{1}{2} \left\| \boldsymbol{\Sigma} - \mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q} \right\|_F^2 = \frac{1}{2} \sum_{i=1}^{q} \sum_{j=1}^{q} \left\| \mathbf{Q} - \boldsymbol{\Sigma}_{iq:(i+1)q, jq:(j+1)q} \right\|_F^2.
$$

This is a least-squares problem, that can be solved in closed form, yelding to a matrix $\mathbf{Q}$ which is the mean of the blocks $\boldsymbol{\Sigma}_{iq:(i+1)q, jq:(j+1)q}$.

To get the first expression we observe that each block can be rewritten as:

$$
\boldsymbol{\Sigma}_{iq:(i+1)q, jq:(j+1)q} = \mathbb{C}\mathrm{ov}\left[\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_i, \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_j \right].
$$

Given the linearity of the covariance we have:

$$
\begin{aligned}
\frac{1}{q^2} \sum_{i=1}^{q} \sum_{j=1}^{q} \boldsymbol{\Sigma}_{iq:(i+1)q, jq:(j+1)q} &= \frac{1}{q^2} \sum_{i=1}^{q} \sum_{j=1}^{q} \mathbb{C}\mathrm{ov}\left[\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_i, \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_j \right] \\
&= \frac{1}{q^2} \mathbb{C}\mathrm{ov}\left[\sum_{i=1}^{q} \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_i, \sum_{j=1}^{q} \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi}_j \right] \\
&= \frac{1}{q^2} \mathbb{C}\mathrm{ov}\left[\nabla_{\boldsymbol{\theta}} \boldsymbol{\psi} \mathbb{1}_q, \nabla_{\boldsymbol{\theta}} \boldsymbol{\psi} \mathbb{1}_q \right].
\end{aligned}
$$

The last equality follows from the properties of the Kroneker product.

We now prove that matrix $\mathbf{Q}$ is positive definite whenever $\boldsymbol{\Sigma}$ is. $\mathbf{Q}$ is positive definite if and only if:

$$
\inf_{\mathbf{x} \in \mathbb{R}^d : \mathbf{x} \neq \mathbf{0}} \mathbf{x}^T \mathbf{Q} \mathbf{x} > 0.
$$

Let us now consider the following derivation:

$$\begin{aligned}\inf_{\mathbf{x}\in\mathbb{R}^d:\mathbf{x}\neq\mathbf{0}} \mathbf{x}^T\mathbf{Q}\mathbf{x} &= \frac{1}{q^2}\inf_{\mathbf{x}\in\mathbb{R}^d:\mathbf{x}\neq\mathbf{0}} \mathbf{x}^T\left(\mathbb{1}_q\otimes\mathbf{I}_d\right)^T\boldsymbol{\Sigma}\left(\mathbb{1}_q\otimes\mathbf{I}_d\right)\mathbf{x}\\ &\ge \frac{1}{q^2}\inf_{\mathbf{x}\in\mathbb{R}^{dq}:\mathbf{x}\neq\mathbf{0}} \mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x} > 0.\end{aligned}$$

having observed that $\left(\mathbb{1}_q\otimes\mathbf{I}_d\right)\mathbf{x}$ is never null unless $\mathbf{x}$ is null. □

Note that this is equivalent to take a specific choice for the weights $\boldsymbol{\omega} = \frac{1}{q}\mathbb{1}_q$.

B.2 Analysis of the Gap

In this appendix, we upper bound the gap on the objective function value attained by the optimum when we consider either matrix $\boldsymbol{\Sigma}$ or a matrix of the form $\mathbb{1}_q\mathbb{1}_q^T\otimes\mathbf{Q}$.

First of all, let us denote with $l_{\mathbf{A}}(\boldsymbol{\omega})$ the objective function in problem (6), when using $\mathbf{A}$ as covariance model. Let $\mathbf{A}$ and $\mathbf{B}$ be two covariance matrices and let $\boldsymbol{\omega}_{\mathbf{A}}$ and $\boldsymbol{\omega}_{\mathbf{B}}$ be any of the optimal weight for the corresponding covariances. Supposing that $\mathbf{A}$ is the true covariance, we want to bound $0 \le l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{A}}) - l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{B}})$. Using a standard argument from empirical risk minimization:

$$\begin{aligned}l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{A}}) - l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{B}}) &= l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{A}}) - l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{B}}) \pm l_{\mathbf{B}}(\boldsymbol{\omega}_{\mathbf{B}})\\ &\ge l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{A}}) - l_{\mathbf{B}}(\boldsymbol{\omega}_{\mathbf{A}}) + l_{\mathbf{B}}(\boldsymbol{\omega}_{\mathbf{B}}) - l_{\mathbf{A}}(\boldsymbol{\omega}_{\mathbf{B}})\\ &\ge -2\sup_{\boldsymbol{\omega}}\left|l_{\mathbf{A}}(\boldsymbol{\omega}) - l_{\mathbf{B}}(\boldsymbol{\omega})\right|,\end{aligned}$$

where we exploited the fact that $l_{\mathbf{B}}(\boldsymbol{\omega}_{\mathbf{B}}) \le l_{\mathbf{B}}(\boldsymbol{\omega}_{\mathbf{A}})$. Thus, it sufficies to prove an upper bound on $|l_{\mathbf{A}}(\boldsymbol{\omega}) - l_{\mathbf{B}}(\boldsymbol{\omega})|$ that is uniform over $\boldsymbol{\omega}$.

Lemma B.2. *Let $\mathbf{A}\in\mathbb{R}^{d\times dq}$ and $\mathbf{B}\in\mathbb{R}^{dq\times dq}$ symmetric positive definite. Then, it holds that:*

$$\left\|\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\right\|_F \le \frac{\sqrt{d}}{s_{\min}(\mathbf{B})}. \tag{23}$$

Proof. First recall that for a symmetric positive definite matrix the following identity involving the square root holds:

$$\left(\mathbf{B}^{\frac{1}{2}}\right)^T = \left(\mathbf{B}^T\right)^{\frac{1}{2}} = \mathbf{B}^{\frac{1}{2}}.$$

Consider now the following derivation:

$$\begin{aligned}\left\|\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\right\|_F &= \left\|\mathbf{B}^{-\frac{1}{2}}\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\mathbf{B}^{-\frac{1}{2}}\right\|_F\\ &\le \left\|\mathbf{B}^{-\frac{1}{2}}\right\|_2^2\left\|\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\right\|_F\\ &= \frac{1}{s_{\min}(\mathbf{B})}\left\|\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\right\|_F,\end{aligned}$$

where we exploited the inequality $\|\mathbf{X}\mathbf{Y}\|_F \le \|\mathbf{X}\|_2\|\mathbf{Y}\|_F$ and fact that $\left\|\mathbf{B}^{-\frac{1}{2}}\right\|_2 \frac{1}{\sqrt{s_{\min}(\mathbf{B})}}$. Let us bound the second term.

$$\begin{aligned}\left\|\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\right\|_F^2 &= \operatorname{tr}\left(\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\mathbf{B}^{\frac{1}{2}}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\right)\\ &= \operatorname{tr}\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{B}\mathbf{A}^T\left(\mathbf{A}\mathbf{B}\mathbf{A}^T\right)^{-1}\right)\\ &= \operatorname{tr}\left(\mathbf{I}_d\mathbf{I}_d\right) = d,\end{aligned}$$

where we exploited the identity $\|\mathbf{X}\|_F^2 = \operatorname{tr}\left(\mathbf{X}^T\mathbf{X}\right)$ and the cyclic property of the trace. □

Lemma B.3. *Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{dq \times dq}$ be two symmetric positive semidefinite matrices. Then, for any $\boldsymbol{\omega} \in \mathbb{R}^q_+$ it holds that:*

$$\left|l_{\mathbf{A}}(\boldsymbol{\omega}) - l_{\;s_{\min}(\mathbf{Q})vB}(\boldsymbol{\omega})\right| \le \left\|(\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T\right\|_F$$
$$\times \left\|(\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T\right\|_F \left\|\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right\|_F^2 \|\mathbf{B} - \mathbf{A}\|_F .$$

Proof. We explicitly write down the expression of $l_{\mathbf{A}}(\boldsymbol{\omega})$ and $l_{\mathbf{B}}(\boldsymbol{\omega})$ and perform a sequence of algebric manipulations:

$$\begin{aligned}
l_{\mathbf{A}}(\boldsymbol{\omega}) - l_{\mathbf{B}}(\boldsymbol{\omega}) &= \left\|\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_{\left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1}} - \left\|\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right\|^2_{\left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1}} \\
&= \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} - \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left\{\left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} - \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1}\right\} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \left\{\mathbf{I}_{dq} - \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right] \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1}\right\} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \left\{(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d) - (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right\} \\
&\quad \times \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \{\mathbf{B} - \mathbf{A}\} (\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} \\
&= \operatorname{tr}\left(\boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \{\mathbf{B} - \mathbf{A}\} (\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right) \\
&= \operatorname{tr}\left((\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \{\mathbf{B} - \mathbf{A}\}\right) \\
&= \operatorname{vec}\left((\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T\right) \\
&\quad \times \operatorname{vec}(\mathbf{B} - \mathbf{A}) \\
&\le \left\|\operatorname{vec}\left((\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T\right)\right\|_2 \\
&\quad \times \|\operatorname{vec}(\mathbf{B} - \mathbf{A})\|_2 \\
&= \left\|(\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{B} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\boldsymbol{\omega}^T \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})^T \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \mathbf{A} (\boldsymbol{\omega} \otimes \mathbf{I}_d)\right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T\right\|_F \|\mathbf{B} - \mathbf{A}\|_F ,
\end{aligned}$$

where we applied the trace since the quantity is scalar, we exploited the cyclic property of the trace, we used the inequality $\operatorname{tr}(\mathbf{X}^T\mathbf{Y}) = \operatorname{vec}(\mathbf{X})^T \operatorname{vec}(\mathbf{Y})$, Cauchy-Swartz inequality and finally observed that $\|\operatorname{vec}(\mathbf{X})\|_2 = \|\mathbf{X}\|_F$. To conclude consider:

$$\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega} = \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right) = \operatorname{vec}\left(\mathbf{I}_d \widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\boldsymbol{\omega}\right) = \left(\boldsymbol{\omega}^T \otimes \mathbf{I}_d\right) \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right) = (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \operatorname{vec}\left(\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right).$$

Using the properties of the Frobenious norm, the result follows. ☐

Theorem B.1. *Let $\boldsymbol{\Sigma}$ be the true covariance matrix and $\mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}$ be its approximation. Then, it holds that:*

$$\text{gap} \le \frac{2dq}{s_{\min}(\boldsymbol{\Sigma})^2} \left\|\widehat{\nabla}_{\boldsymbol{\theta}} \boldsymbol{\psi}(\boldsymbol{\theta})\right\|_F^2 \left\|\boldsymbol{\Sigma} - \mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}\right\|_F . \tag{24}$$

Proof. We instantiate Lemma B.3 with $\mathbf{A} = \Sigma$ and $\mathbf{B} = \mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q}$. Let us now consider the following identity:

$$\mathbf{I}_d = \frac{1}{q} (\mathbb{1}_q \otimes \mathbf{I}_d)^T (\mathbb{1}_q \otimes \mathbf{I}_d) . \tag{25}$$

For the norm involving $\mathbf{B}$ we employ Lemma B.2 and for the other we directly derive:

$$\begin{aligned}
&\left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) \left[(\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \left(\mathbb{1}_q \mathbb{1}_q^T \otimes \mathbf{Q} \right) (\boldsymbol{\omega} \otimes \mathbf{I}_d) \right]^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \right\|_F = \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) \mathbf{Q}^{-1} (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \right\|_F \\
&\qquad = \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) \mathbf{I}_d \mathbf{Q}^{-1} \mathbf{I}_d (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \right\|_F \\
&\qquad = \frac{1}{q^2} \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) (\mathbb{1}_q \otimes \mathbf{I}_d)^T (\mathbb{1}_q \otimes \mathbf{I}_d) \mathbf{Q}^{-1} (\mathbb{1}_q \otimes \mathbf{I}_d)^T (\mathbb{1}_q \otimes \mathbf{I}_d) (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \right\|_F \\
&\qquad = \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) (\mathbb{1}_q \otimes \mathbf{I}_d)^T (\mathbb{1}_q \otimes \mathbf{I}_d) \left[(\mathbb{1}_q \otimes \mathbf{I}_d)^T \boldsymbol{\Sigma} (\mathbb{1}_q \otimes \mathbf{I}_d) \right]^{-1} (\mathbb{1}_q \otimes \mathbf{I}_d)^T (\mathbb{1}_q \otimes \mathbf{I}_d) (\boldsymbol{\omega} \otimes \mathbf{I}_d)^T \right\|_F \\
&\qquad \le \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) (\mathbb{1}_q \otimes \mathbf{I}_d)^T \right\|_2^2 \left\| (\mathbb{1}_q \otimes \mathbf{I}_d) \left[(\mathbb{1}_q \otimes \mathbf{I}_d)^T \boldsymbol{\Sigma} (\mathbb{1}_q \otimes \mathbf{I}_d) \right]^{-1} (\mathbb{1}_q \otimes \mathbf{I}_d)^T \right\|_F \\
&\qquad \le \left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) (\mathbb{1}_q \otimes \mathbf{I}_d)^T \right\|_2^2 \frac{\sqrt{d}}{s_{\min}(\boldsymbol{\Sigma})},
\end{aligned}$$

where we exploited Lemma B.3. To bound the remaining term we have:

$$\begin{aligned}
\left\| (\boldsymbol{\omega} \otimes \mathbf{I}_d) (\mathbb{1}_q \otimes \mathbf{I}_d)^T \right\|_2 &\le \| \boldsymbol{\omega} \otimes \mathbf{I}_d \|_2 \| \mathbb{1}_q \otimes \mathbf{I}_d \|_2 \\
&\le \| \boldsymbol{\omega} \|_2 \| \mathbf{I}_d \|_2 \| \mathbb{1}_q \|_2 \| \mathbf{I}_d \|_2 \\
&\le \| \boldsymbol{\omega} \|_1 \sqrt{q} \le \sqrt{q}.
\end{aligned}$$

□

C Computational Cost

In this appendix, we present the computational cost of the proposed algorithm Σ-GIRL. The computational cost of the Jacobian estimation is linear in the number of policy parameters d, reward parameters q, samples N and horizon H. The computational cost of the covariance is quadratic in the number of policy parameters d and number of reward parameters q, and linear in samples N and horizon H. For a given ω, evaluating the objective in Equation 5 costs $O(d^3 + d^2 q^2)$. The cost of an expectation maximization step is $O(MkC_{\text{opt}})$, where M and k are the number of agents and clusters and C_{opt} is the cost of optimizing function Q, which depends on the optimizer.

D Additional Experiments

In this appendix, we report some additional experimental details together with some details on the optimization of the objective function employed by Σ-GIRL.

D.1 Optimization of Σ-GIRL objective function

The objective function optimized by Σ-GIRL is, in the general case, non-convex. In the experiments, we optimize this function using the implementation of SLSQP (Sequential Least SQuares Programming) from scipy Python package (`https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html`). We used the deafault parameters and tolerance value $1e-8$. We took the best of 25 in the LQG experiment and 5 in the Gridworld experiment different random initializations.

D.2 Single-IRL

Σ-GIRL Comparison In this section, we compare different choices of the covariance matrix used in Σ-GIRL in the LQG environment. Apart from the full sample covariance matrix and the matrix of Corollary 4.1 (which reduces our algorithm to GIRL), we consider also using a diagonal sample covariance matrix and the identity matrix. The former considers only the uncertainty in each of the entries of the Jacobian matrix, while the latter does not consider the uncertainty.

Figure D.2 shows the results in the LQG environment. We can see that using the uncertainty of the gradient estimation clearly achieves better performance. In the environment considered, using the full sample covariance matrix, offers only a slight improvement when considering few trajectories, compared to the diagonal case. In larger problems, where estimating the full covariance matrix might be prohibitive, using a diagonal covariance model offers improvements compared to not using the uncertainty at all in the gradient estimation.

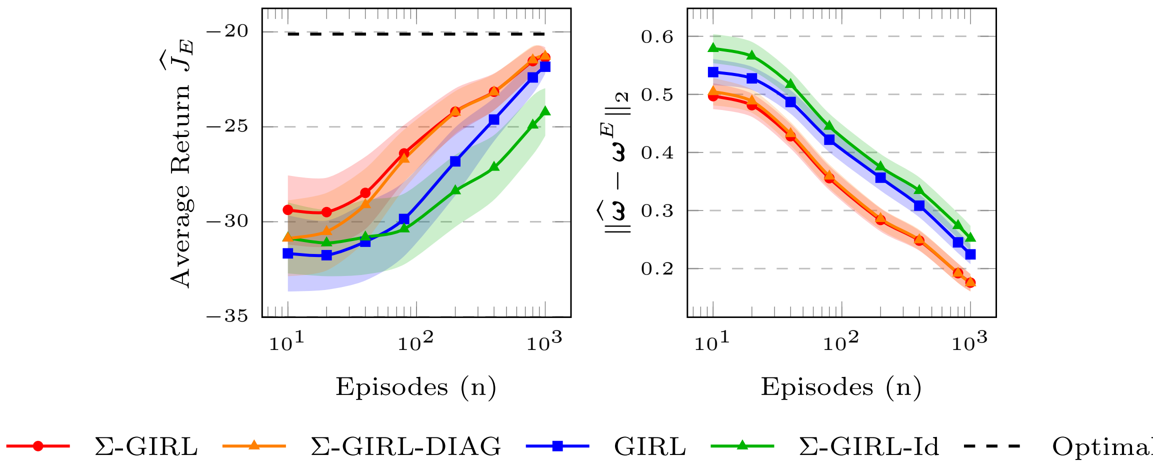

Figure 7: Comparison on the LQG experiment with different choices of covariance model. 100 runs, 95% c.i.

D.3 Multiple-IRL

Gridworld In this section, we perform an empirical analysis the algorithm MLIRL (Babes et al., 2011). The results in Section 7 show that the algorithm is not able to correctly cluster the agents. We perform two experiments. In the first one, we have two agents with two different intentions. In the second one, we have two agents with same intention but different optimal policies, and one agent with a different intention. As show in Figure 8 in the first experiment MLIRL succeeds in the clustering task (left). When we add trajectories performed by two agents with same intentions but different optimal policies the algorithm decreases its performance (right). This behavior explains the results of Section 7 where we have a dataset with three agents sharing the same intention (but different optimal policies) and two agent with different intention.

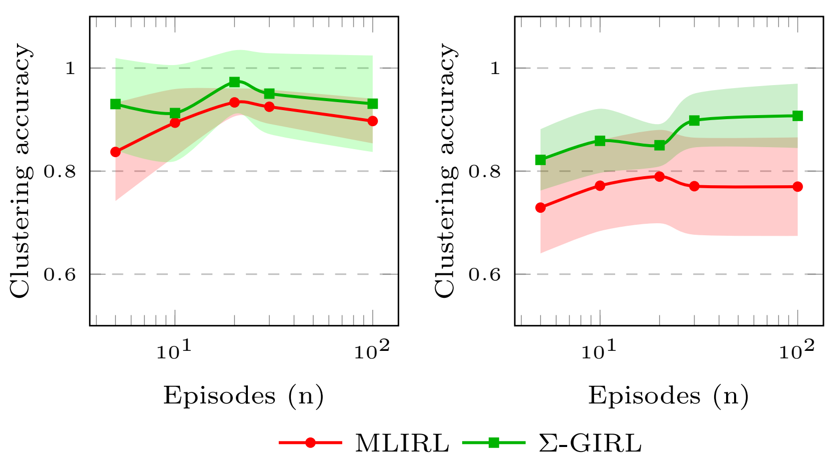

Figure 8: Clustering accuracy in the case of two agents and two clusters (left) and in the case of three agents and two clusters (right). 20 runs 98 % c.i.