# DIFFERENTIABLE INVERSE OPTIMAL CONTROL VIA GAP-AWARE RITZ ITERATIONS AND SCHUR REDUCTION

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Inverse Optimal Control (IOC) is interpretable but typically non-differentiable and numerically fragile near spectral degeneracy, limiting its use in end-to-end learning. We present a Differentiable IOC (DIOC) layer that accumulates sequential optimality constraints into one symmetric matrix and reads the solution from its null space. We show that hard eigenselection causes gradient blow-up, and introduce Soft-Ritz—a Rayleigh–Ritz subspace method with entropic soft selection—paired with a hinge penalty that maintains a boundary spectral gap; this combination delivers bounded, well-conditioned gradients. On linear and nonlinear robotic systems, DIOC converges more smoothly, achieves lower final error, and is markedly more stable than direct eigendecomposition; ablations confirm the gap hinge is necessary and soft-only targets are insufficient. Unifying classical interpretability with modern differentiability, DIOC serves as a practical, auditable module for perception-to-control pipelines, model-based RL warm starts, and safety-critical robotics.

## 1 INTRODUCTION

Learning from expert demonstrations is a cornerstone of modern robotics and AI, enabling agents to acquire complex skills without exhaustive programming. This capability is particularly vital in interactive domains where explicit programming of sophisticated behaviors often results in brittle systems unable to capture expert-level subtleties (Li et al., 2016; Nguyen et al., 2018). The fundamental challenge lies not merely in imitating what an expert does, but in understanding why those decisions are made—a problem formally addressed by Inverse Optimal Control (IOC) and Inverse Reinforcement Learning (IRL) (Ab Azar et al., 2020). The theoretical foundations of IOC can be traced back to the inverse problems in optimal control theory, where the goal is to infer the optimization criteria underlying observed optimal behaviors. This problem has profound implications across diverse domains: in robotics, it enables learning manipulation skills through observation (Englert et al., 2017; Naranjo-Campos et al., 2024); in autonomous driving, it facilitates learning human-like driving behaviors with appropriate safety margins (Huang et al., 2019); in UAV control, it enables multi-agent coordination and path planning (Shamsoshoara et al., 2024; Wang et al., 2023); and in human motor control, it helps understand the principles governing biological movements (Mombaur et al., 2010; Berret et al., 2011).

Prevailing approaches to learning from demonstrations can be broadly categorized into two paradigms. The first, Behavioral Cloning (BC), learns a direct mapping from observations to actions. While simple to implement, BC is prone to covariate shift and fails to capture the underlying goal, limiting its generalization capabilities (Park et al., 2022). The second paradigm, encompassing IRL and IOC, seeks to recover the latent reward or cost function that rationalizes the expert's behavior. By learning the "why" behind actions rather than merely the "what," these methods offer superior generalization and adaptability (Ho & Ermon, 2016; Baert et al., 2025). The fundamental assumption is that expert behavior emerges from optimizing an implicit cost function, and by recovering this function, we obtain a concise and generalizable representation of task intent (Finn et al., 2016).

Traditional IOC methods have evolved along two distinct trajectories, each with inherent strengths and limitations. Classical control-theoretic approaches typically operate in batch settings, requiring access to complete demonstration data (Johnson et al., 2013; Hatz et al., 2012). These methods often rely on solving bilevel optimization problems, where the lower level represents the forward optimal control problem and the upper level seeks to match observed trajectories (Albrecht & Ulbrich, 2017; Dempe et al., 2019). While mathematically rigorous and yielding interpretable cost functions—such as explicit penalties on joint torques, velocities, or energy consumption—these approaches face significant computational challenges. Recent advances have addressed some of these limitations through online and sequential formulations. Molloy et al. (2020) proposed an online IOC method with computationally efficient sequential processing, while Cao et al. (2024) developed the Sequential IOC (SIOC) framework that relaxes restrictive assumptions like Jacobian invertibility. Alternative approaches have reformulated the problem within frameworks of adaptive control (Wu, 2022), state observer methodologies (Self et al., 2021), and recursive filtering using extended Kalman filters (Zhao & Molloy, 2024).

Within the machine learning community, a parallel evolution has occurred with fundamentally different design principles. Modern deep IRL methods leverage the expressive power of neural networks to create end-to-end differentiable systems. Generative Adversarial Imitation Learning (GAIL) frames the problem as an adversarial game between policy and discriminator networks (Ho & Ermon, 2016; Li et al., 2022). Guided Cost Learning employs policy optimization to learn nonlinear cost functions without meticulous feature engineering (Finn et al., 2016). Recent work has combined these approaches with deep reinforcement learning algorithms like PPO and DDPG to handle continuous action spaces (Naranjo-Campos et al., 2024; Huang et al., 2019). However, these modern approaches sacrifice a critical property: interpretability. The learned neural network representations, while expressive enough to capture complex behaviors, obscure the underlying principles and make it difficult to validate whether the learned objectives align with human intentions or safety constraints. This is particularly problematic in safety-critical applications like autonomous vehicles (Wang et al., 2020), robotic manipulation (Jin et al., 2021), and UAV control (Ma et al., 2019).

This divergence exposes a critical technical gap. Classical IOC yields structured, physically interpretable cost functions that domain experts can audit, yet most IOC procedures remain standalone and non-differentiable. As a result, they cannot benefit from gradient-based optimization or compose seamlessly with perception modules, world models, or other differentiable components that characterize modern AI systems. In contrast, treating optimization as a differentiable layer has proven widely effective: OptNet integrates quadratic programs into neural networks (Amos & Kolter, 2017); differentiable convex optimization layers have lowered the barrier to embedding convex programs in end-to-end pipelines (Agrawal et al., 2019); and automatic implicit differentiation applies the implicit function theorem to optimality conditions to propagate gradients through solvers (Blondel et al., 2021). This paradigm now powers applications from differentiable physics (Belbute-Peres et al., 2018) to differentiable Model Predictive Control (Frey et al., 2025).

Against this backdrop, IOC still lacks a practical differentiable formulation. Our prior work, Sequential IOC (SIOC), efficiently recovers cost functions by sequentially intersecting solution spaces induced by first-order optimality, with the intersection dimension provably shrinking at each step (Cao et al., 2024). To carry this interpretability into modern architectures, we recast SIOC as a differentiable layer: temporal optimality constraints are accumulated into a symmetric matrix $\mathcal{Q}$, and the IOC parameters are read from its null space—reducing the temporal program to a compact eigenproblem. Yet a naïve, hard selection of the minimal eigenvector is numerically brittle. Matrix perturbation theory shows that eigenvector sensitivities scale inversely with the spectral gap, e.g., $|\partial v_1| \propto (\lambda_2 - \lambda_1)^{-1}$ (Stewart, 1998). When eigenvalues cluster, gradients can explode, destabilizing training and undermining the promise of a differentiable IOC layer.

We resolve this with *Soft-Ritz* and a *combined loss* that jointly shapes the optimization landscape. Soft-Ritz replaces discontinuous "pick-the-smallest" eigenselection with an entropic mixture of Ritz vectors obtained via Rayleigh–Ritz subspace iteration, yielding a smooth computational graph and stable gradients under near-degeneracy. The combined loss couples (i) an IOC-consistency/data term that enforces alignment between the learned parameters and the accumulated optimality constraints, (ii) a soft-selection/entropy term that prevents premature collapse of the subspace weights, and (iii) a hinge-based spectral-gap regularizer that actively maintains separation between leading eigenvalues. This synergy stabilizes training, preserves identifiability, and makes differentiable IOC practical within modern end-to-end systems.

**Contributions.**

1. We identify hard eigenselection as the source of gradient blow-up in differentiable IOC/SIOC and quantify the sensitivity .
2. We introduce a Rayleigh–Ritz subspace method with entropic soft selection, yielding a smooth surrogate eigenvector that remains well-behaved under near-degeneracy.
3. We prove that the Soft-Ritz map has bounded gradients, depending only on a maintained *boundary* gap and temperature, and *not* on internal Ritz gaps.
4. We design a confidence-gated blend of hard/soft targets together with a hinge penalty on the boundary gap and a null-space penalty, enabling identifiability when gaps reopen while keeping gradients stable when they shrink.

## 2 Problem Formulation

### 2.1 Motivation and Objective

Our central goal is to develop a fully differentiable framework for inverse optimal control that can seamlessly integrate with modern gradient-based learning pipelines. Given an expert's optimal trajectory, we seek to recover the underlying cost function parameters in a manner that preserves the interpretability of classical control theory while enabling end-to-end optimization.

The key challenge lies in transforming the inherently sequential and temporal nature of optimal control into a computational graph that supports stable gradient flow. This transformation must preserve the physical meaning of the recovered parameters while avoiding the numerical instabilities that plague naive differentiable implementations.

### 2.2 Forward and Inverse Optimal Control Setup

We consider a discrete-time nonlinear dynamical system:

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k), \qquad \boldsymbol{x}_k \in \mathbb{R}^n,\ \boldsymbol{u}_k \in \mathbb{R}^m, \quad k = 0, \ldots, K-1. \tag{1}$$

The cost function is assumed to be linear in unknown weights, a structure that encompasses many practical scenarios while maintaining interpretability:

$$J(\boldsymbol{x}_{0:K}, \boldsymbol{u}_{0:K-1}, \boldsymbol{q}) = \sum_{k=0}^{K-1} \boldsymbol{q}^\top \boldsymbol{\phi}(\boldsymbol{x}_k, \boldsymbol{u}_k), \tag{2}$$

where $\boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^{n_f}$ represents a vector of interpretable feature functions (e.g., squared velocities, control efforts, or task-specific penalties), and $\boldsymbol{q} \in \mathbb{R}^{n_f}$ contains the unknown cost weights we wish to recover.

**Forward Optimal Control** seeks trajectories that minimize $J$ for known $\boldsymbol{q}$. **Inverse Optimal Control** receives an expert trajectory $(\boldsymbol{x}^\star_{0:K}, \boldsymbol{u}^\star_{0:K-1})$—assumed optimal for some unknown $\boldsymbol{q}^\star$—and aims to identify this weight vector. Due to scale invariance, we resolve the ambiguity by enforcing $\|\boldsymbol{q}\| = 1$ or fixing a reference component.

### 2.3 From Temporal Optimality to Algebraic Constraints

The discrete-time Pontryagin Maximum Principle provides necessary optimality conditions that we can leverage for IOC. However, directly working with these conditions in their temporal form poses challenges for gradient-based learning due to the sequential dependencies involved.

We linearize the dynamics and features along the expert trajectory:

$$\begin{aligned} \mathcal{A}_k &:= \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{(\boldsymbol{x}^\star_k, \boldsymbol{u}^\star_k)}, \quad \mathcal{B}_k := \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}\right|_{(\boldsymbol{x}^\star_k, \boldsymbol{u}^\star_k)}, \\ \mathcal{C}_k &:= \left.\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\right|_{(\boldsymbol{x}^\star_k, \boldsymbol{u}^\star_k)}, \quad \mathcal{D}_k := \left.\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{u}}\right|_{(\boldsymbol{x}^\star_k, \boldsymbol{u}^\star_k)}. \end{aligned} \tag{3}$$

To transform the temporal structure into an algebraic one, we introduce an augmented state vector that combines costates and cost parameters:

$$\boldsymbol{s}_k := \begin{bmatrix} \boldsymbol{\lambda}_{k+1} \\ \boldsymbol{q} \end{bmatrix} \in \mathbb{R}^{n+n_f}. \tag{4}$$

The linearized PMP conditions can then be expressed through matrices that encode the stationarity and dynamics constraints:

$$\mathcal{H}_k := \begin{bmatrix} \mathcal{B}_k^\top & \mathcal{D}_k^\top \end{bmatrix} \in \mathbb{R}^{m\times(n+n_f)}, \quad \mathcal{M}_k := \begin{bmatrix} \mathcal{A}_k^\top & \mathcal{C}_k^\top \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{(n+n_f)\times(n+n_f)}. \tag{5}$$

These matrices allow us to express the entire set of optimality conditions as a system of linear algebraic constraints:

$$\mathcal{H}_k\, \boldsymbol{s}_k = \mathbf{0}, \qquad \boldsymbol{s}_{k-1} = \mathcal{M}_k\, \boldsymbol{s}_k, \qquad k = 0, \ldots, K-1. \tag{6}$$

### 2.4 The Cumulative Constraint Matrix

The key insight of our approach is that all historical optimality constraints can be accumulated into a single matrix whose null space encodes the IOC solution. This transformation eliminates the temporal dependencies and creates a purely algebraic eigenvalue problem.

We define the propagator matrices for constraint coupling across time:

$$\boldsymbol{\Sigma}_{i\to k} := \mathcal{M}_{i+1}\, \mathcal{M}_{i+2} \cdots \mathcal{M}_k, \qquad \text{so that} \quad \boldsymbol{s}_i = \boldsymbol{\Sigma}_{i\to k}\, \boldsymbol{s}_k. \tag{7}$$

The cumulative constraint matrix is constructed recursively, starting with $\mathcal{Q}_{-1} := \mathbf{0}$:

$$\mathcal{Q}_k := \mathcal{H}_k^\top \mathcal{H}_k + \mathcal{M}_k^\top\, \mathcal{Q}_{k-1}\, \mathcal{M}_k, \qquad k = 0, \ldots, K-1. \tag{8}$$

We denote the final matrix as $\mathcal{Q} := \mathcal{Q}_{K-1}$. The following proposition establishes the fundamental property that makes this formulation work.

**Proposition 1** (Energy decomposition and null-space intersection)**.** *For any $\boldsymbol{s}_{K-1} \in \mathbb{R}^{n+n_f}$,*

$$\boldsymbol{s}_{K-1}^\top \mathcal{Q}\, \boldsymbol{s}_{K-1} = \sum_{i=0}^{K-1} \left\| \mathcal{H}_i\, \boldsymbol{\Sigma}_{i\to K-1}\, \boldsymbol{s}_{K-1} \right\|_2^2. \tag{9}$$

*Consequently, $\boldsymbol{s}_{K-1} \in \mathrm{Null}(\mathcal{Q})$ if and only if all individual optimality constraints are satisfied.*

This result shows that minimizing the quadratic form $\boldsymbol{s}^\top \mathcal{Q} \boldsymbol{s}$ is equivalent to satisfying all the temporal optimality constraints simultaneously. The IOC problem thus reduces to finding the minimal eigenvector of $\mathcal{Q}$.

*Proof.* Define the partial matrices by the given recursion with $\mathcal{Q}_{-1} := \mathbf{0}$ and

$$\mathcal{Q}_k := \mathcal{H}_k^\top \mathcal{H}_k + \mathcal{M}_k^\top \mathcal{Q}_{k-1} \mathcal{M}_k, \qquad k = 0, \ldots, K-1. \tag{10}$$

Introduce the quadratic "energy"

$$\mathcal{E}_k(\boldsymbol{s}_k) := \boldsymbol{s}_k^\top \mathcal{Q}_k\, \boldsymbol{s}_k. \tag{11}$$

We prove by induction that

$$\mathcal{E}_k(\boldsymbol{s}_k) = \sum_{i=0}^{k} \left\| \mathcal{H}_i\, \boldsymbol{\Sigma}_{i\to k}\, \boldsymbol{s}_k \right\|_2^2, \qquad k = 0, \ldots, K-1. \tag{12}$$

*Base case $k = 0$.* Using $\mathcal{Q}_{-1} = \mathbf{0}$ we have $\mathcal{Q}_0 = \mathcal{H}_0^\top \mathcal{H}_0$, hence

$$\mathcal{E}_0(\boldsymbol{s}_0) = \|\mathcal{H}_0 \boldsymbol{s}_0\|_2^2, \tag{13}$$

which matches equation 12 since $\boldsymbol{\Sigma}_{0\to 0} = \mathbf{I}$.

*Induction step.* Assume equation 12 holds for $k-1$. By definition and $\boldsymbol{s}_{k-1} = \mathcal{M}_k \boldsymbol{s}_k$,

$$\mathcal{E}_k(\boldsymbol{s}_k) = \boldsymbol{s}_k^\top \mathcal{H}_k^\top \mathcal{H}_k \boldsymbol{s}_k + \boldsymbol{s}_k^\top \mathcal{M}_{k+1}^\top \mathcal{Q}_{k-1} \mathcal{M}_{k+1} \boldsymbol{s}_k \tag{14}$$

$$= \|\mathcal{H}_k \boldsymbol{s}_k\|_2^2 + \boldsymbol{s}_{k-1}^\top \mathcal{Q}_{k-1} \boldsymbol{s}_{k-1}. \tag{15}$$

Apply the induction hypothesis to $\boldsymbol{s}_{k-1}$:

$$\mathcal{E}_k(\boldsymbol{s}_k) = \|\mathcal{H}_k \boldsymbol{s}_k\|_2^2 + \sum_{i=0}^{k-1} \left\|\mathcal{H}_i \, \boldsymbol{\Sigma}_{i\to k-1} \, \boldsymbol{s}_{k-1}\right\|_2^2 \tag{16}$$

$$= \|\mathcal{H}_k \boldsymbol{s}_k\|_2^2 + \sum_{i=0}^{k-1} \left\|\mathcal{H}_i \, \boldsymbol{\Sigma}_{i\to k-1} \, \mathcal{M}_k \boldsymbol{s}_k\right\|_2^2 \tag{17}$$

$$= \sum_{i=0}^{k} \left\|\mathcal{H}_i \, \boldsymbol{\Sigma}_{i\to k} \, \boldsymbol{s}_k\right\|_2^2, \tag{18}$$

since $\boldsymbol{\Sigma}_{i\to k} = \boldsymbol{\Sigma}_{i\to k-1}\mathcal{M}_k$ and $\boldsymbol{\Sigma}_{k\to k} = \mathbf{I}$. Setting $k = K-1$ gives

$$\boldsymbol{s}_{K-1}^\top \mathcal{Q} \, \boldsymbol{s}_{K-1} = \sum_{i=0}^{K-1} \left\|\mathcal{H}_i \, \boldsymbol{\Sigma}_{i\to K-1} \, \boldsymbol{s}_{K-1}\right\|_2^2. \tag{19}$$

The right-hand side is a sum of non-negative terms, hence equals zero if and only if each term is zero, i.e., $\mathcal{H}_i \, \boldsymbol{\Sigma}_{i\to K-1} \boldsymbol{s}_{K-1} = \mathbf{0}$ for all $i$, which means all historical optimality constraints are satisfied. □

### 2.5 Dimensional Reduction via Schur Complement

The eigenproblem on the full $(n+n_f)\times(n+n_f)$ matrix $\mathcal{Q}$ can be computationally expensive. We reduce it to the cost parameter subspace using the Schur complement.

Partitioning $\mathcal{Q}$ according to the augmented state structure:

$$\mathcal{Q} = \begin{bmatrix} \mathcal{Q}_{\lambda\lambda} & \mathcal{Q}_{\lambda q} \\ \mathcal{Q}_{q\lambda} & \mathcal{Q}_{qq} \end{bmatrix}, \tag{20}$$

we define the Schur complement matrix:

$$\boldsymbol{\Omega} := \mathcal{Q}_{qq} - \mathcal{Q}_{q\lambda} \left(\mathcal{Q}_{\lambda\lambda} + \delta_{\text{tik}}\mathbf{I}\right)^{-1} \mathcal{Q}_{\lambda q} \;\in\; \mathbb{R}^{n_f \times n_f}. \tag{21}$$

The small Tikhonov regularization parameter $\delta_{\text{tik}} > 0$ ensures numerical stability.

**Proposition 2** (Schur complement equivalence)**.** *The original minimization problem*

$$\min_{\|\boldsymbol{s}\|=1} \; \boldsymbol{s}^\top \mathcal{Q} \, \boldsymbol{s} \tag{22}$$

*is equivalent to the reduced problem*

$$\min_{\|\boldsymbol{q}\|=1} \; \boldsymbol{q}^\top \boldsymbol{\Omega} \, \boldsymbol{q}. \tag{23}$$

This reduction transforms our problem into finding the minimal eigenvector of the much smaller matrix $\boldsymbol{\Omega} \in \mathbb{R}^{n_f \times n_f}$, making the computation tractable even for systems with large state dimensions.

*Proof.* Write the quadratic form blockwise with $\boldsymbol{s} = [\boldsymbol{\lambda}^\top, \; \boldsymbol{q}^\top]^\top$:

$$\Phi(\boldsymbol{\lambda}; \boldsymbol{q}) = \boldsymbol{s}^\top \mathcal{Q} \, \boldsymbol{s} = \boldsymbol{\lambda}^\top \mathcal{Q}_{\lambda\lambda} \boldsymbol{\lambda} + 2 \, \boldsymbol{\lambda}^\top \mathcal{Q}_{\lambda q} \boldsymbol{q} + \boldsymbol{q}^\top \mathcal{Q}_{qq} \boldsymbol{q}. \tag{24}$$

Add and subtract a small Tikhonov term $\delta_{\text{tik}}\|\boldsymbol{\lambda}\|_2^2$ (for numerical stability without changing the direction of the outer minimization solution), completing the square:

$$\Phi(\boldsymbol{\lambda}; \boldsymbol{q}) = \left\|(\mathcal{Q}_{\lambda\lambda}+\delta_{\text{tik}}\mathbf{I})^{1/2}\boldsymbol{\lambda} + (\mathcal{Q}_{\lambda\lambda}+\delta_{\text{tik}}\mathbf{I})^{-1/2}\mathcal{Q}_{\lambda q}\boldsymbol{q}\right\|_2^2 \tag{25}$$

$$+ \boldsymbol{q}^\top \Big(\mathcal{Q}_{qq} - \mathcal{Q}_{q\lambda}(\mathcal{Q}_{\lambda\lambda}+\delta_{\text{tik}}\mathbf{I})^{-1}\mathcal{Q}_{\lambda q}\Big)\boldsymbol{q} - \delta_{\text{tik}}\|\boldsymbol{\lambda}\|_2^2. \tag{26}$$

For fixed $\boldsymbol{q}$, the minimum over $\boldsymbol{\lambda}$ is achieved at

$$\boldsymbol{\lambda}^\star = -(\mathcal{Q}_{\lambda\lambda}+\delta_{\text{tik}}\mathbf{I})^{-1}\mathcal{Q}_{\lambda q}\,\boldsymbol{q} \tag{27}$$

with minimum value

$$\min_{\boldsymbol{\lambda}}\ \Phi(\boldsymbol{\lambda};\boldsymbol{q}) = \boldsymbol{q}^\top\boldsymbol{\Omega}\,\boldsymbol{q}, \qquad \boldsymbol{\Omega} := \mathcal{Q}_{qq} - \mathcal{Q}_{q\lambda}\left(\mathcal{Q}_{\lambda\lambda} + \delta_{\text{tik}}\mathbf{I}\right)^{-1}\mathcal{Q}_{\lambda q}. \tag{28}$$

This shows that the outer minimization of the original problem can equivalently be performed in the $\boldsymbol{q}$ subspace. Since the quadratic form is homogeneous and we minimize the Rayleigh quotient,

$$\min_{\|\boldsymbol{s}\|_2=1}\ \boldsymbol{s}^\top\mathcal{Q}\,\boldsymbol{s} = \min_{\|\boldsymbol{q}\|_2=1}\ \boldsymbol{q}^\top\boldsymbol{\Omega}\,\boldsymbol{q}, \tag{29}$$

with the corresponding minimizing direction $\boldsymbol{q}^\star$ and $\boldsymbol{\lambda}^\star = -(\mathcal{Q}_{\lambda\lambda}+\delta_{\text{tik}}\mathbf{I})^{-1}\mathcal{Q}_{\lambda q}\,\boldsymbol{q}^\star$ jointly determining the minimal eigenvector of the original problem (up to overall sign/scale). □

This reduction transforms our problem into finding the minimal eigenvector of the much smaller matrix $\boldsymbol{\Omega} \in \mathbb{R}^{n_f\times n_f}$, making the computation tractable even for systems with large state dimensions.

## 3 Methodology

With the IOC problem reduced to the eigenvalue minimization in equation equation 23, we now face the central challenge: how to solve this eigenproblem in a way that supports stable gradient-based learning.

For brevity we reuse the Rayleigh quotient $R(q) := q^\top\boldsymbol{\Omega}\,q$ and denote by $\Pi_r := V_rV_r^\top$ the projector onto the $r$-dimensional Ritz subspace. We will also use the boundary gap $\delta_r := \lambda_{r+1}(\boldsymbol{\Omega}) - \lambda_r(\boldsymbol{\Omega})$, which separates the trial subspace from its orthogonal complement.

### 3.1 The Challenge of Differentiable Eigendecomposition

The most direct approach is to compute all eigenvalues and eigenvectors of $\boldsymbol{\Omega}$ and select the minimal one. However, this "hard selection" process creates a fundamental bottleneck for gradient-based optimization.

#### 3.1.1 Direct Eigendecomposition and Its Instabilities

The baseline method performs full eigendecomposition:

$$\boldsymbol{\Omega}v_i = \lambda_i v_i, \quad \lambda_1 \le \lambda_2 \le \cdots \le \lambda_{n_f}, \qquad \boldsymbol{q}_{\text{hard}} := v_1. \tag{30}$$

While conceptually simple, this approach suffers from a critical numerical instability rooted in matrix perturbation theory.

**Lemma 1** (Eigenvalue and eigenvector sensitivities)**.** *For a symmetric matrix $\boldsymbol{\Omega}$ with simple eigenvalues, the first-order variations under perturbation $\Delta\boldsymbol{\Omega}$ are:*

$$\delta\lambda_i = v_i^\top(\Delta\boldsymbol{\Omega})\,v_i, \tag{31}$$

$$\delta v_i = \sum_{j\neq i}\frac{v_j^\top(\Delta\boldsymbol{\Omega})\,v_i}{\lambda_i-\lambda_j}\,v_j. \tag{32}$$

The critical observation is that the eigenvector sensitivity scales inversely with the spectral gap $\lambda_j - \lambda_i$. This leads directly to our main stability result.

*Proof.* Consider the perturbed family $\boldsymbol{\Omega}(\epsilon) = \boldsymbol{\Omega} + \epsilon\,\Delta\boldsymbol{\Omega}$ with corresponding eigenpairs

$$\lambda_i(\epsilon) = \lambda_i + \epsilon\,\delta\lambda_i + o(\epsilon), \qquad v_i(\epsilon) = v_i + \epsilon\,\delta v_i + o(\epsilon). \tag{33}$$

From $\boldsymbol{\Omega}(\epsilon)v_i(\epsilon) = \lambda_i(\epsilon)v_i(\epsilon)$, differentiate with respect to $\epsilon$ and evaluate at $\epsilon = 0$:

$$\Delta\boldsymbol{\Omega}\,v_i + \boldsymbol{\Omega}\,\delta v_i = \delta\lambda_i\,v_i + \lambda_i\,\delta v_i. \tag{34}$$

Left-multiplying by $v_i^\top$ and noting that the derivative of $v_i^\top v_i = 1$ gives $v_i^\top \delta v_i = 0$:

$$\delta\lambda_i = v_i^\top (\Delta\mathbf{\Omega})\, v_i. \tag{35}$$

For $j \neq i$, left-multiplying equation 34 by $v_j^\top$:

$$v_j^\top (\Delta\mathbf{\Omega})\, v_i + v_j^\top \mathbf{\Omega}\, \delta v_i = \lambda_i\, v_j^\top \delta v_i. \tag{36}$$

Since $v_j^\top \mathbf{\Omega} = \lambda_j v_j^\top$, we obtain

$$(\lambda_j - \lambda_i)\, v_j^\top \delta v_i = v_j^\top (\Delta\mathbf{\Omega})\, v_i, \qquad v_j^\top \delta v_i = \frac{v_j^\top (\Delta\mathbf{\Omega})\, v_i}{\lambda_i - \lambda_j}. \tag{37}$$

Expanding $\delta v_i = \sum_j (v_j^\top \delta v_i)\, v_j$ yields the desired result. □

**Proposition 3** (Hard selection instability)**.** *Let $\boldsymbol{q}_{hard} = v_1$. For any symmetric perturbation $\Delta\mathbf{\Omega}$,*

$$\left|\langle v_2, \delta \boldsymbol{q}_{hard}\rangle\right| = \frac{\left|v_2^\top (\Delta\mathbf{\Omega})\, v_1\right|}{\lambda_2 - \lambda_1} \;\leq\; \frac{\|\Delta\mathbf{\Omega}\|_2}{\lambda_2 - \lambda_1}. \tag{38}$$

*Hence the sensitivity scales as $1/(\lambda_2 - \lambda_1)$ and can blow up when the gap closes.*

*Proof.* Immediate from Lemma 1 with $i = 1$ and $\|v_2\| = \|v_1\| = 1$. □

This instability is not merely a theoretical concern—it manifests practically whenever the optimization process drives eigenvalues toward degeneracy, which commonly occurs as the algorithm converges toward the null space of $\mathbf{\Omega}$.

### 3.2 Soft-Ritz Eigensolver: A Smooth Alternative

To overcome the instability of hard selection, we propose the Soft-Ritz eigensolver, which replaces the discontinuous eigenvalue selection with a smooth, differentiable process based on subspace iteration and entropic regularization.

#### 3.2.1 Rayleigh-Ritz Subspace Projection

Rather than working with the full $n_f$-dimensional space, we iteratively refine a low-dimensional trial subspace $\mathcal{V} \subset \mathbb{R}^{n_f}$ spanned by orthonormal columns of $\mathbf{\Psi}_{\text{trial}} \in \mathbb{R}^{n_f \times r}$ with $r \ll n_f$.

We project the operator $\mathbf{\Omega}$ onto this subspace using the Rayleigh-Ritz procedure:

$$\mathcal{T} := \mathbf{\Psi}_{\text{trial}}^\top \mathbf{\Omega}\, \mathbf{\Psi}_{\text{trial}} \in \mathbb{R}^{r \times r}. \tag{39}$$

Solving the small eigenproblem $\mathcal{T}U = U\,\mathrm{diag}(\lambda_1, \ldots, \lambda_r)$ gives us Ritz values and vectors:

$$V_r := \mathbf{\Psi}_{\text{trial}} U = [v_1, \ldots, v_r]. \tag{40}$$

#### 3.2.2 Entropic Soft Selection

The key innovation is replacing hard selection of the minimal Ritz vector with a soft, weighted combination. We formulate this as an optimization problem on the probability simplex:

$$\min_{\boldsymbol{w} \in \Delta^{r-1}} \sum_{i=1}^{r} w_i \lambda_i + \tau \sum_{i=1}^{r} w_i \log w_i, \tag{41}$$

where the entropy term $\tau \sum w_i \log w_i$ provides smoothing. The temperature parameter $\tau > 0$ controls the balance between exploitation (selecting the minimal eigenvalue) and exploration (maintaining smoothness).

The unique solution is the Gibbs-Boltzmann distribution:

$$w_i^* = \frac{\exp(-\lambda_i/\tau)}{\sum_{j=1}^{r} \exp(-\lambda_j/\tau)}. \tag{42}$$

This gives us a smooth eigenvector estimate:

$$\boldsymbol{q}_{\text{soft}} := \frac{V_r\, \boldsymbol{w}^*}{\|V_r\, \boldsymbol{w}^*\|}, \qquad \boldsymbol{w}^* := [w_1^*, \ldots, w_r^*]^\top. \tag{43}$$

#### 3.2.3 ADAPTIVE TEMPERATURE CONTROL

The temperature parameter requires careful tuning. When the spectral gap is large, we can afford a small $\tau$ for accurate selection. When eigenvalues are nearly degenerate, we need larger $\tau$ for stability:

$$\tau = \mathrm{clip}\big(\alpha_{\text{soft}}(\lambda_2 - \lambda_1),\ \tau_{\min}, \tau_{\max}\big). \tag{44}$$

This adaptive scheme ensures that we maintain smoothness precisely when the hard selection would become unstable.

#### 3.2.4 ITERATIVE SUBSPACE REFINEMENT

To ensure convergence to the true minimal eigenspace, we iteratively refine the trial subspace. The soft weights $\boldsymbol{w}^*$ indicate which directions are most promising—we use them to bias the next iteration's subspace toward the minimal eigenspace:

$$\text{QR factorization:} \quad \boldsymbol{\Psi}_{\text{trial}}^{\text{next}} = \mathrm{qr}\big(V_r\, \mathrm{diag}(\sqrt{\boldsymbol{w}^*})\big). \tag{45}$$

This reweighting and orthogonalization process gradually focuses the subspace on the most relevant directions while maintaining numerical stability.

**Algorithm 1** Soft-Schur-QR (Soft-Ritz) Eigensolver

**Input:** Jacobian matrices $\{\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k\}$, subspace dimension $r$, temperature factor $\alpha_{\text{soft}}$
**Build cumulative matrix:** Compute $\mathcal{Q}$ via recursion equation 8
**Form Schur complement:** Compute $\boldsymbol{\Omega}$ via equation 21
**Initialize trial subspace:** $\boldsymbol{\Psi}_{\text{trial}} \in \mathbb{R}^{n_f \times r}$ with orthonormal columns
**for** $t = 1, \ldots, T_{\text{iter}}$ **do**
  $\mathcal{T} \leftarrow \boldsymbol{\Psi}_{\text{trial}}^\top \boldsymbol{\Omega}\, \boldsymbol{\Psi}_{\text{trial}}$
  $\mathcal{T}U = U\, \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$;  $V_r \leftarrow \boldsymbol{\Psi}_{\text{trial}} U$;  $\boldsymbol{q}_{\text{hard}} \leftarrow V_r[:, 1]$
  $\tau \leftarrow \mathrm{clip}(\alpha_{\text{soft}}(\lambda_2 - \lambda_1), \tau_{\min}, \tau_{\max})$
  $\boldsymbol{w}^* \leftarrow \mathrm{softmin}_\tau(\lambda_1, \ldots, \lambda_r)$ using equation 42
  $\boldsymbol{q}_{\text{soft}} \leftarrow \frac{V_r \boldsymbol{w}^*}{\|V_r \boldsymbol{w}^*\|}$
  **Reweight-and-QR:** $Q_{\text{thin}}, R \leftarrow \mathrm{qr}\big(V_r\, \mathrm{diag}(\sqrt{\boldsymbol{w}^*})\big)$
  $\boldsymbol{\Psi}_{\text{trial}} \leftarrow Q_{\text{thin}}$
**end for**
**Return:** $\boldsymbol{q}_{\text{hard}}$, $\boldsymbol{q}_{\text{soft}}$, $\lambda_1$, boundary gap $\delta_r = \lambda_{r+1} - \lambda_r$

### 3.3 THEORETICAL GUARANTEES FOR SOFT-RITZ

We now establish that our soft selection approach resolves the gradient explosion problem that plagues hard selection methods.

**Proposition 4** (Soft selection stability)**.** *The soft selection weights satisfy*

$$\big\|\nabla_{\boldsymbol{\lambda}} \boldsymbol{w}^*\big\|_2 \le \frac{1}{\tau}, \tag{46}$$

*providing bounded gradients independent of internal spectral gaps.*

*Proof.* Write equation 42 as $w_i = \exp(-\lambda_i/\tau)/Z$ where $Z = \sum_j \exp(-\lambda_j/\tau)$. Differentiating with respect to $\lambda_j$:

$$\frac{\partial w_i}{\partial \lambda_j} = -\frac{1}{\tau}\, w_i\, \delta_{ij} + \frac{1}{\tau}\, w_i w_j = -\frac{1}{\tau}\left[\mathrm{diag}(\boldsymbol{w}) - \boldsymbol{w}\boldsymbol{w}^\top\right]_{ij}. \tag{47}$$

Thus the Jacobian matrix $J = \partial \boldsymbol{w}/\partial \boldsymbol{\lambda} = -(1/\tau)\left[\mathrm{diag}(\boldsymbol{w}) - \boldsymbol{w}\boldsymbol{w}^\top\right]$. The bracketed term is a covariance matrix with spectral norm $\|\mathrm{diag}(\boldsymbol{w}) - \boldsymbol{w}\boldsymbol{w}^\top\|_2 \le 1/4 \le 1$, therefore

$$\big\|\nabla_{\boldsymbol{\lambda}} \boldsymbol{w}\big\|_2 = \|J\|_2 \le \frac{1}{\tau}. \tag{48}$$

This bound is independent of internal spectral gaps and controlled only by the temperature $\tau$. □

Let $\Pi_r := V_r V_r^\top$ be the orthogonal projector onto the $r$-dimensional Ritz subspace. Define the *boundary gap*

$$\delta_r := \lambda_{r+1}(\mathbf{\Omega}) - \lambda_r(\mathbf{\Omega}) \;>\; 0. \tag{49}$$

**Theorem 1** (Stability of the Soft-Ritz map)**.** *Assume the boundary gap satisfies $\delta_r \geq \delta_{\min} > 0$ and the temperature satisfies $\tau \geq \tau_{\min} > 0$. Then the composite mapping*

$$\textit{(parameters)} \;\longmapsto\; \mathcal{Q} \;\longmapsto\; \mathbf{\Omega} \;\longmapsto\; \boldsymbol{q}_{\textit{soft}} \tag{50}$$

*has bounded gradients:*

$$\left\|\nabla_{\textit{params}} \boldsymbol{q}_{\textit{soft}}\right\| \leq \frac{C_{\textit{system}}}{\tau_{\min}\,\delta_{\min}}, \tag{51}$$

*where $C_{\textit{system}}$ does not depend on internal gaps $\lambda_j - \lambda_i$ for $i, j \leq r$.*

*Proof.* Deferred to Appendix A. □

This theorem shows that maintaining a small but positive boundary gap—rather than worrying about all internal spectral gaps—is sufficient for stability.

**Corollary 1** (Combined estimator: stability & identifiability)**.** *Define*

$$c \;:=\; \mathrm{stopgrad}\Big(\mathrm{clip}\big(w_1^* - w_2^*,\, c_{\min},\, 1\big)\Big). \tag{52}$$

*Consider the combined task loss*

$$\mathcal{L}_{\textit{task}} \;=\; c\,\ell(\boldsymbol{q}_{\textit{hard}},\, \boldsymbol{q}^*) \;+\; (1-c)\,\ell(\boldsymbol{q}_{\textit{soft}},\, \boldsymbol{q}^*), \tag{53}$$

*together with*

$$\mathcal{L}_{\textit{gap}} \;=\; \max\big\{0,\; \delta_{\min} - \big(\lambda_{r+1} - \lambda_r\big)\big\}^2, \qquad \mathcal{L}_{\textit{null}} \;=\; w_{\textit{null}}\,\lambda_1^2. \tag{54}$$

*If*

$$\delta_r \;=\; \lambda_{r+1} - \lambda_r \;\geq\; \delta_{\min} \;>\; 0, \qquad \tau \;\geq\; \tau_{\min} \;>\; 0, \tag{55}$$

*then there exist constants $C_{\textit{hard}}, C_{\textit{soft}} > 0$ (independent of internal gaps $\lambda_j - \lambda_i$ for $i, j \leq r$) such that*

$$\left\|\nabla \mathcal{L}_{\textit{task}}\right\| \;\leq\; c \cdot \frac{C_{\textit{hard}}}{\delta_{\min}} \;+\; (1-c) \cdot \frac{C_{\textit{soft}}}{\tau_{\min}\,\delta_{\min}}. \tag{56}$$

*Moreover, when $\delta_r$ is sufficiently large, $\boldsymbol{q}_{\textit{hard}}$ is unique up to sign (identifiability); when $\delta_r$ is small, the term on $\boldsymbol{q}_{\textit{soft}}$ maintains bounded and well-conditioned gradients.*

*Proof.* By Theorem 1, under $\delta_r \geq \delta_{\min}$ and $\tau \geq \tau_{\min}$ the Soft–Ritz map obeys $\|\nabla \boldsymbol{q}_{\text{soft}}\| \leq C_{\text{sys}}/(\tau_{\min}\delta_{\min})$. Composing with a Lipschitz loss $\ell$ gives $\|\nabla \ell(\boldsymbol{q}_{\text{soft}}, \boldsymbol{q}^*)\| \leq C_{\text{soft}}/(\tau_{\min}\delta_{\min})$. For $\boldsymbol{q}_{\text{hard}}$, standard eigenvector/subspace perturbation (e.g., Davis–Kahan) yields $\|\nabla \ell(\boldsymbol{q}_{\text{hard}}, \boldsymbol{q}^*)\| \leq C_{\text{hard}}/\delta_{\min}$. Because $c = \mathrm{stopgrad}(\cdot)$, gradients do not flow through $c$, so the bound for $\|\nabla \mathcal{L}_{\text{task}}\|$ is the stated convex combination. Identifiability follows from simplicity of the minimal eigenvalue when $\delta_r$ is large. □

**Corollary 2** (Soft branch is stable and task-improving when gaps reopen)**.** *Fix a Ritz basis $V_r = [v_1, \ldots, v_r]$ (orthonormal) and the compressed operator $\mathcal{T} := V_r^\top \mathbf{\Omega} V_r$ with eigenpairs $(\lambda_i, v_i)$, $\lambda_1 \leq \cdots \leq \lambda_r$. Let $w_i^\star = \exp(-\lambda_i/\tau)/\sum_{k=1}^r \exp(-\lambda_k/\tau)$ with $\tau > 0$, $\boldsymbol{w}^\star := (w_1^\star, \ldots, w_r^\star)^\top$, and $q_{\text{soft}} := \frac{V_r \boldsymbol{w}^\star}{\|V_r \boldsymbol{w}^\star\|}$.*

*(i) (*Degeneracy*) Suppose $\lambda_1 = \cdots = \lambda_g = \lambda_\star < \lambda_{g+1}$ for some $g \geq 2$. If either $r = g$ or $\tau \to 0$, then $q_{\text{soft}} \in \mathrm{span}\{v_1, \ldots, v_g\}$. If $r > g$ and $\tau \in (0, \infty)$, then in general $q_{\text{soft}} \notin \mathrm{span}\{v_1, \ldots, v_g\}$ and it is not unique under rotations within the degenerate block.*

*Proof.* Deferred to Appendix C. □

*(ii) (*Task-side progress when identifiable*) Assume $\lambda_1 < \lambda_2$ and the boundary gap $\delta_r > 0$. With $V_r$ held fixed, increasing any $(\lambda_j - \lambda_1)$ $(j \geq 2)$ strictly decreases $R(q_{\text{soft}}) := q_{\text{soft}}^\top \mathbf{\Omega}\, q_{\text{soft}}$ and increases its alignment with $v_1$. For $r = 2$ with $\Delta := \lambda_2 - \lambda_1$ one has the closed forms $\gamma_1(\Delta) = \langle q_{\text{soft}}, v_1 \rangle = \big(1 + e^{-2\Delta/\tau}\big)^{-1/2}$ and $E(\Delta) := R(q_{\text{soft}}) - \lambda_1 = \Delta\, \frac{e^{-2\Delta/\tau}}{1+e^{-2\Delta/\tau}}$, hence $\gamma_1'(\Delta) > 0$, $E(\Delta) \to 0$ as $\Delta \to \infty$, and $E'(\Delta) < 0$ for all $\Delta \geq \Delta_0(\tau)$ where $\Delta_0(\tau)$ solves $\Delta = \frac{\tau}{2}\big(1 + e^{-2\Delta/\tau}\big)$.*

*Proof.* Deferred to Appendix C. □

**Design implications.** Corollary 2 highlights a precise trade-off. The *soft* branch delivers smooth, bounded gradients even when the smallest Ritz values are clustered; as gaps reopen, its Rayleigh value strictly decreases and its alignment with $v_1$ improves. However, at finite temperature it need not lie exactly in the minimal eigenspace and, under degeneracy, it is basis-dependent and therefore non-identifiable. In contrast, the *hard* estimator is unbiased and identifiable once the smallest eigenvalue is simple, but its gradients can blow up as the internal gap shrinks.

These observations motivate a gap-aware combination: we supervise with the hard estimate when the spectrum is clearly separated and fall back to the soft proxy when it is not. Concretely, we use a confidence gate $c = \mathrm{clip}(w_1^\star - w_2^\star,\, c_{\min},\, 1)$ to blend hard/soft targets; a hinge-style boundary-gap penalty $\mathcal{L}_{\mathrm{gap}} = \max\{0, \delta_{\min} - (\lambda_{r+1} - \lambda_r)\}^2$ to keep a *sufficient* separation (thus stabilizing the subspace geometry without over-regularizing); and a null-space penalty $\mathcal{L}_{\mathrm{null}} = \lambda_1^2$ to enforce IOC consistency. This pairing preserves identifiability when gaps are present, while maintaining well-conditioned gradients when they collapse, and it leads directly to the combined loss introduced next.

### 3.4 Loss Function Design

Our objective couples a task-agnostic head with spectral regularization that is tailored to the Soft-Ritz solver. The task head supplies supervision specific to the downstream use case, while the spectral terms keep the eigen-geometry well posed and gradients stable.

**Task-Agnostic Head (choose per application).** We view $\mathcal{L}_{\text{task}}$ as a plug-in that can be instantiated for different supervision signals, all differentiating through Soft-Ritz.

*Example A: Parameter/weight tracking (used to verify gradients).* In our experiments for gradient-verification, we perturb control inputs (one part of the IOC input) to steer the IOC output $\boldsymbol{q}$ toward a reference $\boldsymbol{q}^*$. We combine hard/soft estimates with a confidence gate:

$$\mathcal{L}_{\text{task}} = c\,\|\mathrm{rel}(\boldsymbol{q}_{\text{hard}}, \boldsymbol{q}^*)\|_2^2 + (1-c)\,\|\mathrm{rel}(\boldsymbol{q}_{\text{soft}}, \boldsymbol{q}^*)\|_2^2, \tag{57}$$

where $c \in [0,1]$ depends on weight concentration (e.g., $c = \mathrm{clip}(w_1^* - w_2^*,\, 0.2,\, 1)$), and $\mathrm{rel}(a,b) := (a-b)/(b+\epsilon)$ is elementwise. When Soft-Ritz is confident (large separation), $\boldsymbol{q}_{\text{hard}}$ is emphasized; otherwise gradients rely more on the stable $\boldsymbol{q}_{\text{soft}}$.

*Example B: Trajectory tracking with gap-aware optimality supervision (conceptual).* In trajectory tracking problems, controls $u$ are adapted so that the system state follows a desired trajectory while IOC weights are inferred as part of the solution. Here the task loss can be instantiated as

$$\mathcal{L}_{\text{task}}^{(B)} = \sum_{t=1}^{T} \ell(\boldsymbol{x}_t(u),\, \boldsymbol{x}_t^*) + \rho \sum_{t=2}^{T} \|u_t - u_{t-1}\|^2 + \alpha\Big[c_\Delta \lambda_1^{\text{hard}\,2} + (1-c_\Delta)\lambda_{\text{soft}}^2\Big], \tag{58}$$

where the first two terms encourage trajectory tracking and smooth control usage, and the last term enforces IOC optimality through a gap-aware combination of hard and soft eigenvalues (as in Example A). Although we do not empirically evaluate this setting in the present paper, it illustrates how the same design principle—*combine hard and soft optimality signals via a confidence gate*—extends naturally to more complex downstream tasks.

**Spectral Regularization (task-agnostic).** These terms complement any choice of $\mathcal{L}_{\text{task}}$ by shaping the eigenspectrum that Soft-Ritz operates on.

**Null space enforcement:**

$$\mathcal{L}_{\text{null}} = \lambda_1(\boldsymbol{\Omega})^2, \tag{59}$$

penalizing violations of the IOC null-space condition (smallest eigenvalue of $\boldsymbol{\Omega}$).

**Boundary gap maintenance (hinge):**

$$\mathcal{L}_{\text{gap}} = \max\big\{0,\ \delta_{\min} - (\lambda_{r+1} - \lambda_r)\big\}^2, \tag{60}$$

which is inactive once the gap exceeds $\delta_{\min}$, avoiding over-regularization while preventing ill-conditioned, near-degenerate spectra.

**Complete Objective.**

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{task}} + w_{\text{null}}\mathcal{L}_{\text{null}} + w_{\text{gap}}\mathcal{L}_{\text{gap}} + \text{(standard regularization)}. \tag{61}$$

In summary, the task head (chosen per application) supplies supervision—ranging from gradient-verification tracking (Eq. 57) to trajectory tracking (Eq. 58) —while the spectral terms keep the Soft-Ritz layer stable and identifiable across all regimes.

### 3.5 Computational Complexity Analysis

We analyze the computational trade-offs between Soft-Ritz and direct eigendecomposition, decomposing costs into shared preprocessing and solver-specific components.

#### 3.5.1 Shared Preprocessing Costs

Both methods require identical preprocessing that often dominates total cost when state dimension $n$ is large.

**Matrix construction:** Building $\mathcal{Q}$ via equation equation 8 costs $\mathcal{O}(T[n^3 + n^2 n_f + n n_f^2 + m(n + n_f)^2])$ due to state coupling ($n^3$), state-parameter cross terms ($n^2 n_f$), parameter coupling ($n n_f^2$), and constraint accumulation ($m(n + n_f)^2$).

**Schur complement:** Computing $\mathbf{\Omega}$ requires Cholesky factorization ($n^3$), solving triangular systems ($n^2 n_f$), and matrix assembly ($n n_f^2$), totaling $\mathcal{O}(n^3 + n^2 n_f + n n_f^2)$.

Memory: $\Theta((n + n_f)^2)$ for $\mathcal{Q}$ storage plus $\Theta(n^2)$ for factorization.

#### 3.5.2 Solver-Specific Complexity

**Direct eigendecomposition:** Full symmetric eigendecomposition of $\mathbf{\Omega} \in \mathbb{R}^{n_f \times n_f}$ costs $\mathcal{O}(n_f^3)$ with $\Theta(n_f^2)$ memory.

**Soft-Ritz:** $I$ iterations with subspace dimension $r \ll n_f$. Each iteration performs subspace projection via $r$ matrix-vector products, small eigenproblem ($r^3$), and basis update ($n_f r^2$). Matrix-vector products dominate at $\mathcal{O}(n_f^2 + n n_f + n^2)$ per operation. Total: $\mathcal{O}(Ir(n_f^2 + n n_f + n^2) + I n_f r^2 + I r^3)$ with memory $\Theta(n_f r) + \Theta(n^2)$.

#### 3.5.3 Performance Trade-offs

When $n_f \gg n$ with typical parameters $r \ll n_f$, $I \ll n_f$:

$$\text{EigenDirect} \sim \mathcal{O}(n_f^3) \quad \text{vs} \quad \text{SoftRitz} \sim \mathcal{O}(I r n_f^2) \tag{62}$$

Soft-Ritz outperforms when $n_f \gtrsim I \cdot r$. Example: with $r = 8$, $I = 10$, savings occur when $n_f \geq 80$. Memory reduces from $\Theta(n_f^2)$ to $\Theta(n_f r)$ for eigenvector storage.

| Phase | Time Complexity | Memory |
|---|---|---|
| BuildMatrix | $\mathcal{O}(T[n^3 + n^2 n_f + n n_f^2 + m(n + n_f)^2])$ | $\Theta((n + n_f)^2)$ |
| FormSchur | $\mathcal{O}(n^3 + n^2 n_f + n n_f^2)$ | $\Theta(n^2)$ |
| EigenDirect | $\mathcal{O}(n_f^3)$ | $\Theta(n_f^2)$ |
| SoftRitz | $\mathcal{O}(Ir(n_f^2 + n n_f + n^2) + I n_f r^2 + I r^3)$ | $\Theta(n_f r) + \Theta(n^2)$ |

Table 1: Computational complexity summary for different phases of the DIOC framework.

Beyond computational benefits, Soft-Ritz provides crucial numerical stability in gradient-based learning as spectral gaps collapse during optimization.

# 4 EXPERIMENTS

**Objective.** We aim to verify that gradients propagated *through the differentiable IOC layer* are both accurate and numerically stable. To isolate this, we adopt a diagnostic "control-to-weights" task: we optimize the open-loop control sequence $U = \{u_t\}_{t=0}^{T-1}$ for a fixed plant so that the IOC weights $\boldsymbol{q}$ recovered by IOC plus the eigensolver match a known ground truth $\boldsymbol{q}^*$. This setting stresses the Jacobians of the IOC recursion and the eigensolver without confounding factors from other learning signals.

**Task head (Example A).** Throughout the experiments, the task head follows the same design principle introduced in Sec. 3.4: combine the *hard* (minimal-eigenvector) and *soft* (Soft-Ritz subspace mixture) estimates via a confidence gate driven by the spectral separation, while keeping the IOC optimality and the spectrum well-conditioned. Concretely, when the spectrum is clustered, training leans on the soft proxy for stability; when it is well separated, it emphasizes the hard estimate for precision. The exact objective is given in Sec. 3.4, and hyperparameters for each study are summarized alongside the setups below.

**Methods compared.** We compare: **Method A (Soft-Ritz)**, which performs smooth subspace selection and maintains a *sufficient* margin via a hinge-style gap term; and **Method B (Direct)**, which selects the minimal eigenvector "hardly" and regularizes the spectrum with a reciprocal-gap penalty. Both methods share the same plant, seeds, horizon, rollout, and warm-start controls; the only differences lie in the eigensolver and the gap regularizer.

**Plant.** All experiments use the same planar 2-link manipulator (Table 2) with six quadratic features (four on state, two on control), horizon $T = 50$, step $\Delta t = 0.02\,\mathrm{s}$, and the initial state specified in the per-experiment setup. We refer to Sec. 4.1 for the detailed benchmark configuration. Parameters are summarized in Table 2.

Table 2: Planar 2-link manipulator parameters (SI units). $l_{ci}$ denotes the COM distance of link $i$ measured from the proximal joint.

| Symbol | Description | Value | Unit |
|---|---|---|---|
| $l_1$ | Link 1 length | 1.0 | m |
| $l_2$ | Link 2 length | 1.0 | m |
| $l_{c1}$ | Link 1 COM distance | 0.5 | m |
| $l_{c2}$ | Link 2 COM distance | 0.5 | m |
| $m_1$ | Link 1 mass | 1.0 | kg |
| $m_2$ | Link 2 mass | 1.0 | kg |
| $I_1$ | Link 1 inertia about COM | 0.0833 | kg·m$^2$ |
| $I_2$ | Link 2 inertia about COM | 0.0833 | kg·m$^2$ |
| $g$ | Gravitational acceleration | 9.81 | m/s$^2$ |

## 4.1 EXPERIMENT 1: COMPARATIVE STUDY OF SOFT-RITZ VS DIRECT EIGENDECOMPOSITION

**Setup.** We use a planar 2-link manipulator benchmark with state dimension $n = 4$ (joint angles and velocities) and control dimension $m = 2$ (joint torques). The feature vector $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^6$ contains quadratic penalties on states and controls, yielding six unknown weights. Ground truth is $\boldsymbol{q}^* = [150,\ 150,\ 10,\ 10,\ 1,\ 1]^\top$ (normalized by the last component). Time horizon is $T = 50$ with step $\Delta t = 0.02\,\mathrm{s}$; the initial state is $\boldsymbol{x}_0 = [\pi/2,\ \pi/4,\ 0,\ 0]^\top$. Warm-start controls are obtained via CasADi/Ipopt using approximate weights $[120,\ 180,\ 15,\ 5,\ 0.8,\ 1.2]^\top$.

**Training schedule.** Both methods are trained for 1000 epochs with cosine annealing. Double precision, recursion depth 15, $L_2$ penalty $w_{\text{reg}} = 10^{-4}$, smoothness $w_{\text{smooth}} = 5 \times 10^{-3}$.

**Method-specific configuration.** Same dynamics/demo, warm-start via CasADi+IPOPT; horizon $T = 50$, $\Delta t = 0.02$s; SIOC recursion depth 15; cosine annealing over 1000 epochs; double precision.

**Figures.** Fig. 1 shows the evolution of all six weights. Fig. 2 reports Relative-$L_2$, smallest eigenvalue $\lambda_1$, eigen-gap $\delta$, and LR schedules.

Table 3: Configuration summary for the two methods. $\delta = \lambda_2 - \lambda_1$.

| | **A (Soft-Ritz)** | **B (Direct eig.)** |
|---|---|---|
| Eigen-solver | Soft-Ritz subspace, $r = 3$, $\alpha = 0.25$ | Hard min-eigenvector |
| Task loss | Rel. err. on $q$ (confidence-gated) | Rel. err. on $q$ (auto-balance) |
| Gap regularizer | Hinge: $\max\{0, \delta_{\min} - \delta\}^2$, $\delta_{\min} = 10^{-8}$ | Reciprocal: $w_{\text{gap}}/(\delta + \epsilon)$ |
| Optimizer & LR | AdamW (AMSGrad), $5 \times 10^{-3}$ | Adam, $10^{-2}$ |
| Grad clip | $\|g\|_\infty \leq 0.3$ | $\|g\|_\infty \leq 1.0$ |
| Null penalty | $w_{\text{null}}\lambda_1^2$, $w_{\text{null}} = 50$ | $w_{\text{null}}\lambda_1^2$, $w_{\text{null}} = 100$ |
| Regularization | $w_{\text{reg}} = 10^{-4}$, $w_{\text{smooth}} = 5 \times 10^{-3}$ | $w_{\text{reg}} = 10^{-4}$, $w_{\text{smooth}} = 10^{-3}$ |

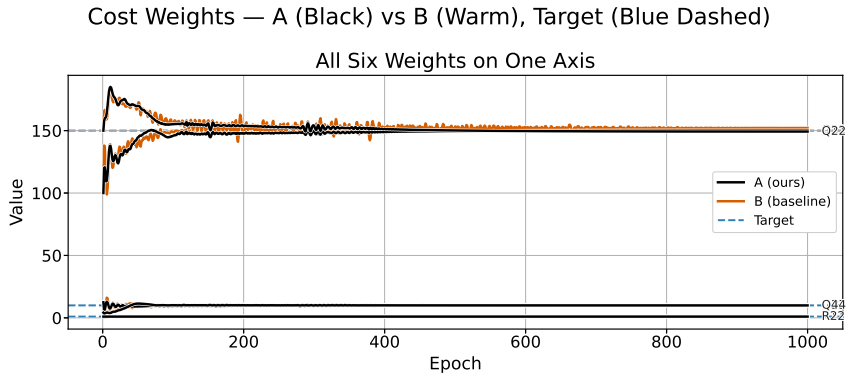


Figure 1: Convergence of cost weights. Method A (black) vs Method B (orange) with target (blue dashed). A is smoother and more stable; B oscillates strongly in early epochs.

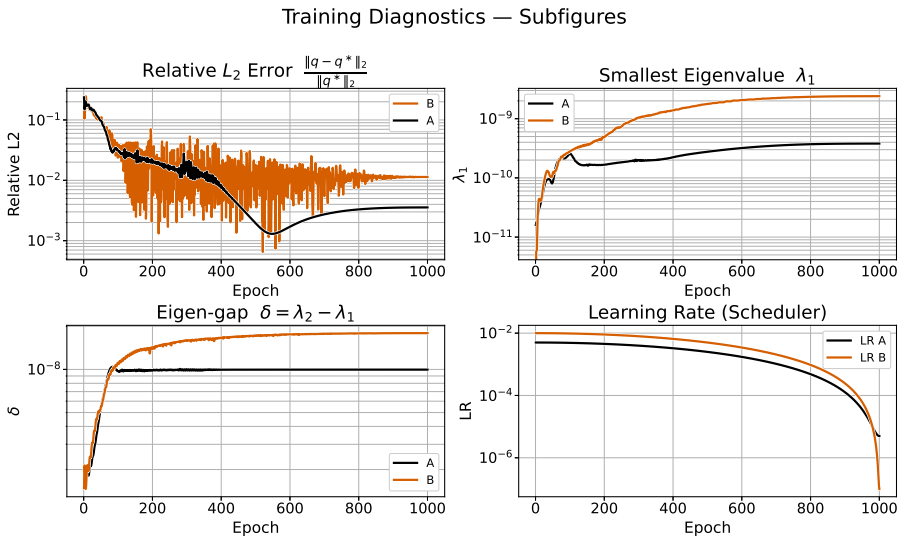


Figure 2: Diagnostics of the two methods. (Top-left) Relative-$L_2$: A achieves $\sim 3 \times 10^{-3}$ vs B $\sim 1.5 \times 10^{-2}$. (Top-right) $\lambda_1$: A better null-space approximation ($10^{-10}$ vs $10^{-9}$). (Bottom-left) Gap $\delta$: A stabilizes near $10^{-8}$, B enforces $\sim 10^{-7}$ but still unstable. (Bottom-right) LR schedules.

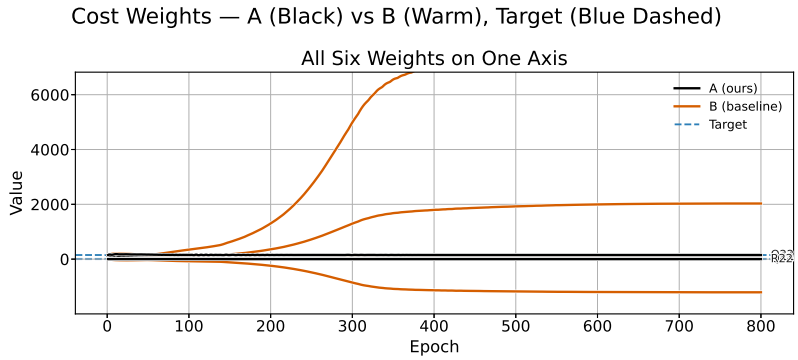


Figure 3: Same LR + AdamW for both methods. Cost weights of Method A (black) stay close to the target (blue dashed); Method B (orange) diverges, with several entries blowing up to $10^3 \sim 10^4$.

**Results.** Method A converges smoothly with consistently smaller error, while Method B plateaus after 600 epochs. The hinge gap lets A maintain just-enough separation ($\sim 10^{-8}$) and stable gradients, whereas B's $1/\delta$ over-regularizes and still suffers ill-conditioned gradients. Overall, Soft-Ritz achieves more accurate recovery, smoother convergence, and better numerical stability than direct eigendecomposition.

### 4.2 Experiment 2: Fair Comparison under Identical Optimizer and LR Schedule

**Setup.** Same plant and data as Sec. 4.1: double pendulum ($n = 4$, $m = 2$), features $\boldsymbol{F} \in \mathbb{R}^6$ (quadratic in states/controls), horizon $T = 50$, $\Delta t = 0.02$s, target $\boldsymbol{q}^* = [150, 150, 10, 10, 1, 1]^\top$ (normalized by the last entry), warm-start controls from a CasADi/Ipopt OCP using $[120, 180, 15, 5, 0.8, 1.2]^\top$.

**Identical training schedule.** Both methods use **AdamW(AMSGrad)** with the **same** cosine LR: initial LR $1 \times 10^{-2}$, $T_{\max} = 800$, gradient clip 0.5, $L_2$ regularization $w_{\text{reg}} = 1 \times 10^{-4}$, smoothness $w_{\text{smooth}} = 5 \times 10^{-3}$, recursion depth 15. (No early-trust or extra schedules.)

**Method-specific pieces (kept minimal).**

- **Method A (Soft-Ritz).** Soft-Schur-QR with subspace $r = 3$, temperature factor $\alpha = 0.25$; gated task loss with $c = \text{clip}(w_1 - w_2, 0.2, 1)^\gamma$, $\gamma = 1$; alignment loss $w_{\text{align}} = 0.1$; hinge gap $\max\{0, \delta_{\min} - \delta\}^2$ with $\delta_{\min} = 10^{-8}$; null penalty weight $w_{\text{null}} = 50$, gap weight $w_{\text{gap}} = 1$.
- **Method B (Direct eigh).** Hard-select the smallest eigenvector; task loss on $\boldsymbol{q}_{\text{hard}}$; reciprocal gap $w_{\text{gap}}/\delta$ with $w_{\text{gap}} = 0.5$; null penalty $w_{\text{null}} = 100$.

**Figures.** Fig. 3 compares all six weights on one axis; Fig. 4 reports Relative-$L_2$, $\lambda_1$, eigen-gap $\delta$, and LR.

**Results.** Under the *same* optimizer and LR schedule, Method A converges smoothly to the target with final error $\approx 10^{-3}$, whereas Method B's weights diverge and its Relative-$L_2$ grows to $\mathcal{O}(10^2)$. Although B enforces a larger spectral gap through the $1/\delta$ term, this *does not* translate into a correct eigenspace—hard selection remains ill-conditioned and destabilizes learning. In contrast, A's soft selection plus hinge gap yields well-behaved, bounded gradients: $\lambda_1$ settles near $10^{-10}$, and $\delta$ is kept just above the margin $\sim 10^{-8}$—large enough for numerical stability but not over-regularized.

**Takeaways.** (i) With identical schedules, A *strictly dominates* B in accuracy and stability. (ii) For DIOC, maintaining a *sufficient* (not maximal) gap via a hinge penalty cooperates with the task objective, while $1/\delta$ can overdrive the spectrum without improving the subspace. (iii) Soft-Ritz's smooth selection is critical once eigenvalues approach degeneracy.

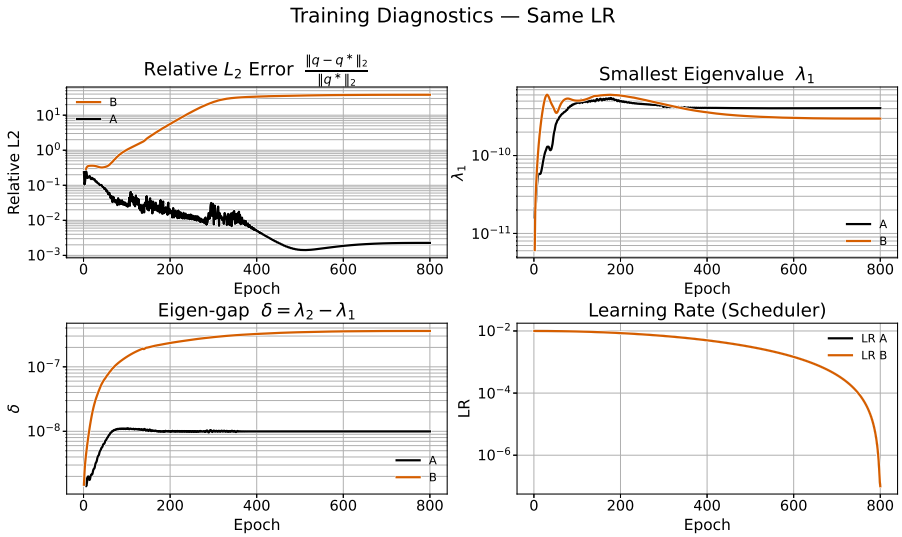


Figure 4: Training diagnostics under identical schedule. **Top-left:** Relative $L_2$ of Method A descends to $\sim 10^{-3}$ while Method B explodes to $\mathcal{O}(10^2)$. **Top-right:** $\lambda_1$ is maintained at $\sim 10^{-10}$ for A; B drifts. **Bottom-left:** B keeps a large gap ($\sim 10^{-7}$) via $1/\delta$ but still fails; A keeps a minimal, stable margin ($\sim 10^{-8}$) via hinge. **Bottom-right:** Identical LR traces.

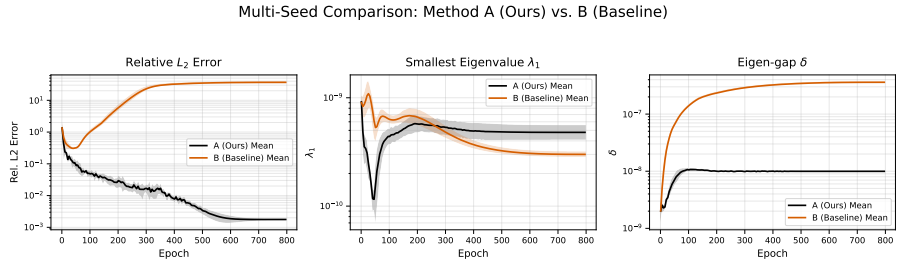


Figure 5: **Multi-seed head-to-head (five seeds, identical schedules).** *Left:* Relative $L_2$ error (mean$\pm$1 std). *Middle:* $\lambda_1$. *Right:* eigen-gap $\delta$. Method A: stable convergence with tight variance; Method B: divergence despite a larger enforced gap.

#### 4.2.1 Multi-Seed Replication under the Identical Schedule

To assess reproducibility, we repeat training across five random seeds. We add small Gaussian perturbations to the initial control sequence $U$ with relative scale 2% (clipped at $3\sigma$) while keeping the initial state $x_0$ fixed. We report mean$\pm$1 std for three diagnostics: relative $L_2$ error of $\boldsymbol{q}$, the smallest eigenvalue $\lambda_1$ of $\mathcal{Q}$, and the eigen-gap $\delta := \lambda_2 - \lambda_1$.

**Multi-seed results.** Method A converges for *all* seeds with tight variance (final relative error $\sim 10^{-3}$), keeps $\lambda_1 \approx 10^{-10}$ and a *minimal yet stable* gap $\delta \approx 10^{-8}$. Method B shows seed-consistent drift and divergence: error rises to $O(10^1$–$10^2)$, $\lambda_1$ remains larger and oscillatory; although $\delta$ is pushed to $\sim 10^{-7}$, the correct eigenspace is still not identified. This corroborates that enlarging $\delta$ via a reciprocal penalty does not guarantee better subspaces, while a hinge margin that preserves a small but adequate gap cooperates with the task objective and yields stable gradients.

### 4.3 Ablation on Soft-Ritz Components

#### 4.3.1 Experimental Setup

We reuse the same plant and demonstrations as in Sec. 4.1: the double pendulum ($n = 4$, $m = 2$), six quadratic features, horizon $T = 50$, step $\Delta t = 0.02$s, and the ground-truth

$\boldsymbol{q}^* = [150, 150, 10, 10, 1, 1]^\top$ (normalized by the last entry). Controls are warm-started from a CasADi/Ipopt OCP with weights $[120, 180, 15, 5, 0.8, 1.2]^\top$. All runs use AdamW (AMSGrad) with the same cosine schedule: initial LR $1 \times 10^{-2}$, $T_{\max} = 800$, gradient clip 0.5, $L_2$ regularization $10^{-4}$, smoothness $5 \times 10^{-3}$, and SIOC recursion depth 15.

#### 4.3.2 Method Configurations

We ablate three components of Method A while keeping the rest identical. The Soft-Schur-QR subspace uses $r = 3$ with temperature factor $\alpha = 0.25$.

Table 4: Ablation settings for Method A. Here $\delta = \lambda_2 - \lambda_1$.

| | **Full** | **-hinge** | **Soft-only** |
|---|---|---|---|
| Task target | use $\boldsymbol{q}_{\text{hard}}$ | use $\boldsymbol{q}_{\text{hard}}$ | use $\boldsymbol{q}_{\text{soft}}$ only |
| Gap regularizer | hinge: $\max\{0, \delta_{\min} - \delta\}^2$, $\delta_{\min} = 10^{-8}$ | none | hinge as left |
| Other losses | $w_{\text{null}} = 50$, $w_{\text{gap}} = 1$ | $w_{\text{null}} = 50$ | $w_{\text{null}} = 50$, $w_{\text{gap}} = 1$ |

#### 4.3.3 Results and Analysis

Figure 6 shows the six cost weights on a single axis. Figure 7 reports the smallest eigenvalue $\lambda_1$ and the eigen-gap $\delta$, and Figure 8 shows the relative $L_2$ error.

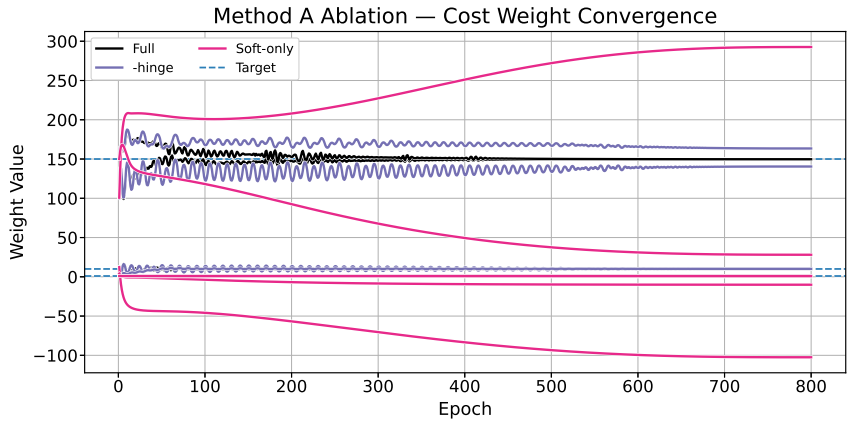


Figure 6: Method A ablation: all six weights on one axis. Blue dashed lines indicate targets. The Full variant closely tracks the ground truth; removing the hinge (-hinge) stalls; Soft-only drifts.

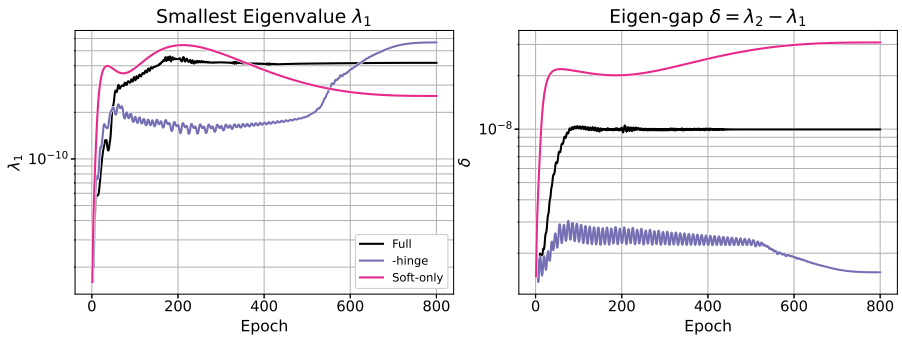


Figure 7: Spectral diagnostics for Method A ablation. Left: smallest eigenvalue $\lambda_1$. Right: eigen-gap $\delta = \lambda_2 - \lambda_1$. The hinge keeps a small but stable margin $\delta \approx 10^{-8}$; without it, the gap gradually collapses.

**Effect of the hinge gap.** With the hinge regularizer, the gap stabilizes near $\delta \approx 10^{-8}$ while $\lambda_1$ remains around $10^{-10}$, yielding smooth convergence of $\boldsymbol{q}$ and the lowest final error ($\sim 3 \times 10^{-3}$). Removing the hinge leads to a steadily shrinking gap and visible oscillations; training stalls at a higher error ($\sim 10^{-1}$), despite an otherwise identical schedule.

**Soft-only is insufficient.** Using only the soft proxy vector without the hard candidate fails to identify the correct eigenspace. Although the soft subspace can keep a relatively large gap, the estimate drifts and the relative error increases toward $O(1)$, indicating that the soft proxy alone is not an adequate target for learning.

**Takeaways.** (i) The hinge-style spectral margin is *necessary*: it provides a sufficient but not excessive gap that cooperates with the task objective and prevents ill-conditioning. (ii) The soft mechanism is *supportive* but not a standalone target; combining a stable soft subspace with a hard candidate gives the best accuracy. (iii) Altogether, the Soft-Ritz layer with a small hinge margin delivers well-behaved gradients and reliable recovery of interpretable IOC weights.

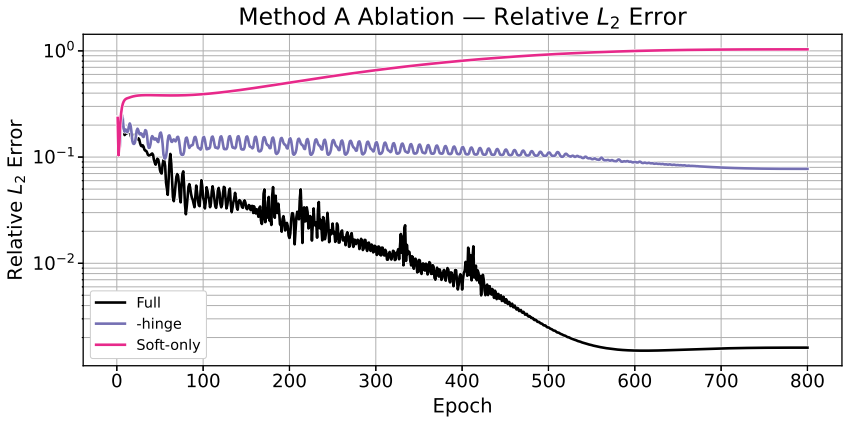


Figure 8: Relative $L_2$ error of recovered weights. Full continues to decrease to $\sim 3 \times 10^{-3}$; -hinge plateaus at $\sim 8 \times 10^{-2}$; Soft-only climbs toward $O(1)$.

## 4.4 Experiment 4: Hyperparameter Sensitivity of the Soft-Ritz Solver

### 4.4.1 Experimental Setup

We study the sensitivity of **Method A** to two solver hyperparameters: the trial subspace dimension $r_{\text{soft}}$ and the temperature factor $\alpha_{\text{soft}}$. The plant, demonstrations, features, and target cost are identical to Sec. 4.1. Training schedule is kept fixed for fairness: AdamW (AMSGrad), initial LR $1 \times 10^{-2}$ with cosine annealing ($T_{\max} = 800$), gradient clip 0.5, $w_{\text{reg}} = 10^{-4}$, $w_{\text{smooth}} = 5 \times 10^{-3}$, recursion depth 15, double precision. We sweep $r_{\text{soft}} \in \{2, 3, 5, 8, 12\}$ and $\alpha_{\text{soft}} \in \{0.05, 0.10, 0.25, 0.50, 1.00\}$, and report the final and best-so-far relative $L_2$ error over 800 epochs.

### 4.4.2 Results and Analysis

**Subspace dimension $r_{\text{soft}}$.** Performance improves when moving from very small subspaces to moderate sizes, and saturates around $r_{\text{soft}} \in \{8, 12\}$ (both yield the best final error $\approx 1.4 \times 10^{-3}$). Larger $r$ slightly reduces variance of the trajectory (Fig. 9, top), but gains beyond $r = 8$ are marginal while computation increases linearly in $r$.

**Temperature factor $\alpha_{\text{soft}}$.** Too small or too large temperatures hurt final accuracy. Very large $\alpha_{\text{soft}}$ ($\geq 0.5$) over-smooths the soft selection, leading to slow convergence and early plateaus; very small $\alpha$ (0.05) induces sharper but noisier updates (good transient minimum, worse final error). A moderate range $\alpha_{\text{soft}} \in [0.10, 0.25]$ strikes a good balance between stability and descent speed.

**Takeaways.** (i) Use a modest trial subspace: $r_{\text{soft}} \in [3, 8]$; $r = 8$ is a strong default when compute permits, while $r = 3$ is a good lightweight choice. (ii) Set $\alpha_{\text{soft}}$ in $[0.10, 0.25]$; in all other sections we

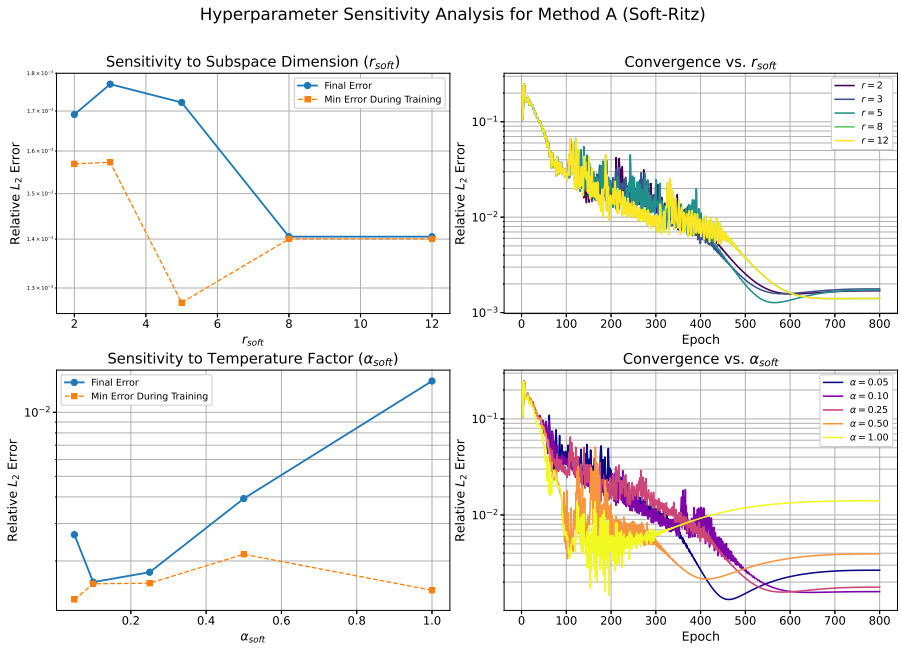


Figure 9: Sensitivity of Method A (Soft-Ritz) to the subspace size $r_{\text{soft}}$ and temperature factor $\alpha_{\text{soft}}$. Left: final and minimum relative $L_2$ error vs. the hyperparameter; Right: full error trajectories.

adopt $\alpha_{\text{soft}} = 0.25$ for consistency and stable gradients. (iii) The solver is not hypersensitive—wide neighborhoods around these values retain the qualitative advantages of Soft-Ritz.

Table 5: Sensitivity summary (lower is better). Numbers match Fig. 9.

| $r_{\text{soft}}$ | Final Rel-$L_2$ | Min Rel-$L_2$ | $\alpha_{\text{soft}}$ | Final Rel-$L_2$ | Min Rel-$L_2$ |
|---|---|---|---|---|---|
| 2 | $1.691 \times 10^{-3}$ | $1.569 \times 10^{-3}$ | 0.05 | $2.660 \times 10^{-3}$ | $1.318 \times 10^{-3}$ |
| 3 | $1.771 \times 10^{-3}$ | $1.573 \times 10^{-3}$ | 0.10 | $1.591 \times 10^{-3}$ | $1.562 \times 10^{-3}$ |
| 5 | $1.722 \times 10^{-3}$ | $1.272 \times 10^{-3}$ | 0.25 | $1.771 \times 10^{-3}$ | $1.573 \times 10^{-3}$ |
| 8 | $\mathbf{1.405 \times 10^{-3}}$ | $1.400 \times 10^{-3}$ | 0.50 | $3.931 \times 10^{-3}$ | $2.151 \times 10^{-3}$ |
| 12 | $\mathbf{1.405 \times 10^{-3}}$ | $1.400 \times 10^{-3}$ | 1.00 | $1.403 \times 10^{-2}$ | $1.456 \times 10^{-3}$ |

## 5 Discussion

**What the soft subspace buys us.** Our results indicate that the primary failure mode of making SIOC differentiable is the brittleness of *hard* minimal-eigenvector selection when the smallest eigenvalues cluster. The Soft-Ritz layer replaces this discontinuity with an entropic mixture over a Rayleigh–Ritz subspace, yielding a smooth estimator that preserves useful task descent even under near-degeneracy. In practice, the soft branch provides reliable gradients to "pull" the iterate into the correct eigenspace; once a modest boundary gap forms, the confidence gate hands control back to the hard estimator, restoring identifiability.

**Gap shaping: sufficient vs. maximal separation.** A recurring theme across experiments is that *sufficient* (hinge) gaps cooperate with learning, whereas aggressively maximizing the gap (e.g., via $1/\delta$) can over-regularize without improving the subspace. The hinge keeps $\delta$ just above a margin, balancing numerical conditioning with the task objective and avoiding wasteful spectral stretching.

**Choosing subspace size and temperature.** Sensitivity studies suggest a practical regime: small-to-moderate trial subspaces ($r \in [3, 8]$) and a moderate temperature factor ($\alpha_{\text{soft}} \in [0.10, 0.25]$).

Larger $r$ mildly reduces variance but increases cost roughly linearly; overly small or large temperatures slow convergence (over-smoothing) or induce noisy updates (over-sharpening). These trends are consistent with the view of Soft-Ritz as a controlled exploration–exploitation trade-off within the minimal-eigenspace neighborhood.

**Interpretable supervision heads.** Although our empirical focus is the diagnostic "control-to-weights" task (optimize $U$ so that $\boldsymbol{q}$ tracks $\boldsymbol{q}^*$), the *task head* is a plug-in: the same gap-aware hard/soft design extends to other supervision signals (e.g., trajectory tracking while enforcing optimality). The common design principle is unchanged: combine a precise hard candidate with a stable soft proxy through a confidence gate driven by spectral evidence.

**Computation and integration.** Compared to full eigendecomposition on the reduced matrix, Soft-Ritz trades $O(n_f^3)$ time and $O(n_f^2)$ memory for iterative $O(Ir\, n_f^2)$ time and $O(n_f r)$ memory with $r \ll n_f$, while adding stability that is crucial for end-to-end training. The layer's inputs (Jacobians and the cumulative matrix) are shared with the direct method, making the swap-in cost modest for existing pipelines.

**Limitations and scope.** (i) **Feature linearity.** Our formulation assumes costs linear in hand-crafted features; learning features jointly with weights is an important extension. (ii) **Boundary vs. internal degeneracy.** The theory and practice hinge on maintaining a boundary gap; deep internal degeneracies inside the trial subspace can still slow identifiability, though the soft branch remains stable. (iii) **Modeling and Jacobians.** The approach requires reliable local Jacobians; large model mismatch or non-smooth dynamics may necessitate robustification. (iv) **Data noise and imperfect demonstrations.** We assume noiseless trajectories and exact optimality; in practice, measurement noise, actuation delays, and suboptimal demos can bias the estimates. Incorporating noise-aware likelihoods (e.g., stochastic IOC), robust losses, and temporal filtering is left for future work.

**Broader impact and applications.** By turning IOC into a differentiable, interpretable module, the layer can be composed with perception encoders, world models, or policy heads—preserving auditability of control objectives while enabling modern training workflows. This makes the framework attractive in safety-critical domains where explicit cost semantics matter.

## 6 Conclusion

We introduced a differentiable IOC layer that reformulates sequential optimality constraints into a single algebraic eigenproblem and replaces brittle hard eigenselection with the *Soft-Ritz* solver. Coupled with a gap-aware, confidence-gated task head and a hinge spectral regularizer, the method delivers bounded, well-behaved gradients, smooth convergence, and accurate recovery of interpretable cost weights. Experiments on nonlinear robots show consistent improvements over direct eigendecomposition under identical setups. Looking ahead, we see three natural directions: (i) jointly learning features and weights within the same layer; (ii) extending Soft-Ritz to multi-dimensional null spaces and multi-task settings; and (iii) integrating the layer into perception-to-control stacks for closed-loop learning with real data. Together, these steps aim to solidify interpretable, control-theoretic modules as first-class citizens in end-to-end AI systems.

## References

Nematollah Ab Azar, Aref Shahmansoorian, and Mohsen Davoudi. From inverse optimal control to inverse reinforcement learning: A historical review. *ANNUAL REVIEWS IN CONTROL*, 50: 119–138, 2020. ISSN 1367-5788. doi: 10.1016/j.arcontrol.2020.06.001.

Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J. Zico Kolter. Differentiable convex optimization layers. In H Wallach, H Larochelle, A Beygelzimer, F d'Alche Buc, E Fox, and R Garnett (eds.), *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 32 (NIPS 2019)*, volume 32 of *Advances in Neural Information Processing Systems*, 2019. 33rd Conference on Neural Information Processing Systems (NeurIPS), Vancouver, CANADA, DEC 08-14, 2019.

Sebastian Albrecht and Michael Ulbrich. Mathematical programs with complementarity constraints in the context of inverse optimal control for locomotion. *Optimization Methods and Software*, 32 (4):670–698, 2017.

Brandon Amos and J. Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In D Precup and YW Teh (eds.), *INTERNATIONAL CONFERENCE ON MACHINE LEARNING, VOL 70*, volume 70 of *Proceedings of Machine Learning Research*, 2017. 34th International Conference on Machine Learning, Sydney, AUSTRALIA, AUG 06-11, 2017.

Mattijs Baert, Pietro Mazzaglia, Sam Leroux, and Pieter Simoens. Maximum causal entropy inverse constrained reinforcement learning. *MACHINE LEARNING*, 114(4), APR 2025. ISSN 0885-6125. doi: 10.1007/s10994-024-06653-5.

Filipe de A. Belbute-Peres, Kevin A. Smith, Kelsey R. Allen, Joshua B. Tenenbaum, and J. Zico Kolter. End-to-end differentiable physics for learning and control. In S Bengio, H Wallach, H Larochelle, K Grauman, N CesaBianchi, and R Garnett (eds.), *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 31 (NIPS 2018)*, volume 31 of *Advances in Neural Information Processing Systems*, 2018. 32nd Conference on Neural Information Processing Systems (NIPS), Montreal, CANADA, DEC 02-08, 2018.

Bastien Berret, Enrico Chiovetto, Francesco Nori, and Thierry Pozzo. Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS computational biology*, 7(10):e1002183, 2011.

Mathieu Blondel, Quentin Berthet, Marco Cuturi, Roy Frostig, Stephan Hoyer, Felipe Llinares-López, Fabian Pedregosa, and Jean-Philippe Vert. Efficient and modular implicit differentiation. *CoRR*, abs/2105.15183, 2021. URL `https://arxiv.org/abs/2105.15183`.

Sheng Cao, Zhiwei Luo, and Changqin Quan. Sequential inverse optimal control of discrete-time systems. *IEEE/CAA Journal of Automatica Sinica*, 11(3):608–621, 2024.

Stephan Dempe, Felix Harder, Patrick Mehlitz, and Gerd Wachsmuth. Solving inverse optimal control problems via value functions to global optimality. *JOURNAL OF GLOBAL OPTIMIZATION*, 74(2):297–325, JUN 2019. ISSN 0925-5001. doi: 10.1007/s10898-019-00758-1.

Peter Englert, Ngo Anh Vien, and Marc Toussaint. Inverse kkt: Learning cost functions of manipulation tasks from demonstrations. *INTERNATIONAL JOURNAL OF ROBOTICS RESEARCH*, 36 (13-14, SI):1474–1488, DEC 2017. ISSN 0278-3649. doi: 10.1177/0278364917745980.

Chelsea Finn, Sergey Levine, and Pieter Abbeel. Guided cost learning: Deep inverse optimal control via policy optimization. In *International conference on machine learning*, pp. 49–58. PMLR, 2016.

Jonathan Frey, Katrin Baumgärtner, Gianluca Frison, Dirk Reinhardt, Jasper Hoffmann, Leonard Fichtner, Sebastien Gros, and Moritz Diehl. Differentiable nonlinear model predictive control, 2025. URL `https://arxiv.org/abs/2505.01353`.

Kathrin Hatz, Johannes P. Schl¥"oder, and Hans Georg Bock. Estimating parameters in optimal control problems. *SIAM Journal on Scientific Computing*, 34(3):A1707–A1728, 2012. doi: 10.1137/110823390.

Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. *Advances in neural information processing systems*, 29, 2016.

Wenhui Huang, Francesco Braghin, and Zhuo Wang. Learning to drive via apprenticeship learning and deep reinforcement learning. In *2019 IEEE 31ST INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELLIGENCE (ICTAI 2019)*, Proceedings-International Conference on Tools With Artificial Intelligence, pp. 1536–1540. IEEE; IEEE Comp Soc, 2019. ISBN 978-1-7281-3798-8. doi: 10.1109/ICTAI.2019.00220. 31st IEEE International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, NOV 04-06, 2019.

Wanxin Jin, Dana Kulic, Shaoshuai Mou, and Sandra Hirche. Inverse optimal control from incomplete trajectory observations. *INTERNATIONAL JOURNAL OF ROBOTICS RESEARCH*, 40 (6-7):848–865, JUN 2021. ISSN 0278-3649. doi: 10.1177/0278364921996384.

Miles Johnson, Navid Aghasadeghi, and Timothy Bretl. Inverse optimal control for deterministic continuous-time nonlinear systems. In *52nd IEEE Conference on Decision and Control*, pp. 2906–2913. IEEE, 2013.

Jiangeng Li, Shuai Huang, Xin Xu, and Guoyu Zuo. Generative adversarial imitation learning from human behavior with reward shaping. In *2022 34TH CHINESE CONTROL AND DECISION CONFERENCE, CCDC*, Chinese Control and Decision Conference, pp. 6254–6259, 2022. ISBN 978-1-6654-7896-0. doi: 10.1109/CCDC55256.2022.10034038. 34th Chinese Control and Decision Conference (CCDC), Hefei, PEOPLES R CHINA, AUG 15-17, 2022.

Yanan Li, Keng Peng Tee, Rui Yan, Wei Liang Chan, and Yan Wu. A framework of human–robot coordination based on game theory and policy iteration. *IEEE Transactions on Robotics*, 32(6): 1408–1418, 2016.

Haoxiang Ma, Mou Chen, and Qingxian Wu. Inverse optimal control for unmanned aerial helicopters with disturbances. *OPTIMAL CONTROL APPLICATIONS & METHODS*, 40(1):152–171, JAN-FEB 2019. ISSN 0143-2087. doi: 10.1002/oca.2472.

Timothy L Molloy, Jason J Ford, and Tristan Perez. Online inverse optimal control for control-constrained discrete-time systems on finite and infinite horizons. *Automatica*, 120:109109, 2020.

Katja Mombaur, Anh Truong, and Jean-Paul Laumond. From human to humanoid locomotion-an inverse optimal control approach. *AUTONOMOUS ROBOTS*, 28(3, SI):369–383, APR 2010. ISSN 0929-5593. doi: 10.1007/s10514-009-9170-7.

Francisco J. Naranjo-Campos, Juan G. Victores, and Carlos Balaguer. Expert-trajectory-based features for apprenticeship learning via inverse reinforcement learning for robotic manipulation. *APPLIED SCIENCES-BASEL*, 14(23), DEC 2024. doi: 10.3390/app142311131.

Hung The Nguyen, Matthew Garratt, Lam Thu Bui, and Hussein Abbass. Apprenticeship bootstrapping: Inverse reinforcement learning in a multi-skill uav-ugv coordination task. In *PROCEEDINGS OF THE 17TH INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS (AAMAS' 18)*, pp. 2204–2206. Assoc Comp Machinery; Int Fdn Autonomous Agents & MultiAgent Syst; NSF; Artificial Intelligence Journal; Riken; Nissan; ACM SIGAI; Springer Sci + Business Media; Elsevier; Google; AIP, Ctr Adv Intelligent Project, 2018. ISBN 978-1-4503-5649-7. 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), Stockholm, SWEDEN, JUL 10-15, 2018.

Jongcheon Park, Seungyong Han, and S. M. Lee. Restored action generative adversarial imitation learning from observation for robot manipulator. *ISA TRANSACTIONS*, 129(B):684–690, OCT 2022. ISSN 0019-0578. doi: 10.1016/j.isatra.2022.02.041.

Ryan Self, Kevin Coleman, He Bai, and Rushikesh Kamalapurkar. Online observer-based inverse reinforcement learning. *IEEE Control Systems Letters*, 5(6):1922–1927, December 2021. ISSN 2475-1456. doi: 10.1109/lcsys.2020.3046527. URL `http://dx.doi.org/10.1109/LCSYS.2020.3046527`.

Alireza Shamsoshoara, Fatemeh Lotfi, Sajad Mousavi, Fatemeh Afghah, and Ismail Guevenc. Joint path planning and power allocation of a cellular-connected uav using apprenticeship learning via deep inverse reinforcement learning. *COMPUTER NETWORKS*, 254, DEC 2024. ISSN 1389-1286. doi: 10.1016/j.comnet.2024.110789.

Gilbert W Stewart. Perturbation theory for the singular value decomposition. 1998.

Chuanrui Wang, Chuanxu Yan, Zhenchong Liu, and Feng Cao. An inverse optimal approach to ship course-keeping control. *IMA JOURNAL OF MATHEMATICAL CONTROL AND INFORMATION*, 37(4):1192–1217, DEC 2020. ISSN 0265-0754. doi: 10.1093/imamci/dnaa007.

Xiaojie Wang, Zhaolong Ning, Song Guo, Miaowen Wen, Lei Guo, and H. Vincent Poor. Dynamic uav deployment for differentiated services: A multi-agent imitation learning based approach. *IEEE TRANSACTIONS ON MOBILE COMPUTING*, 22(4):2131–2146, APR 1 2023. ISSN 1536-1233. doi: 10.1109/TMC.2021.3116236.

Huai-Ning Wu. Online learning human behavior for a class of human-in-the-loop systems via adaptive inverse optimal control. *IEEE Transactions on Human-Machine Systems*, 52(5):1004–1014, 2022.

Tian Zhao and Timothy L. Molloy. Extended kalman filtering for recursive online discrete-time inverse optimal control, 2024. URL https://arxiv.org/abs/2403.10841.

# A Proof of Theorem 1

*Proof.* We bound the operator norm of the differential of each stage and then apply the chain rule. Throughout, $\|\cdot\|$ denotes the spectral norm for matrices and the Euclidean norm for vectors. We repeatedly use submultiplicativity $\|ABC\| \le \|A\|\,\|B\|\,\|C\|$.

**Step 1. Bound $\|\mathrm{d}\mathcal{Q}\|$ in terms of $\|\mathrm{d}(\textbf{params})\|$.** Recall the recursion (with $\mathcal{Q}_{-1} = \mathbf{0}$)

$$\mathcal{Q}_k = \mathcal{H}_k^\top \mathcal{H}_k + \mathcal{M}_{k+1}^\top \mathcal{Q}_{k-1} \mathcal{M}_{k+1} \qquad (k = 0, \dots, K-1). \tag{63}$$

Differentiating,

$$\begin{aligned}\mathrm{d}\mathcal{Q}_k = {} & (\mathrm{d}\mathcal{H}_k)^\top \mathcal{H}_k + \mathcal{H}_k^\top (\mathrm{d}\mathcal{H}_k) + (\mathrm{d}\mathcal{M}_{k+1})^\top \mathcal{Q}_{k-1}\mathcal{M}_{k+1} \\ & + \mathcal{M}_{k+1}^\top (\mathrm{d}\mathcal{Q}_{k-1})\mathcal{M}_{k+1} + \mathcal{M}_{k+1}^\top \mathcal{Q}_{k-1}(\mathrm{d}\mathcal{M}_{k+1}).\end{aligned} \tag{64}$$

Take norms and apply submultiplicativity:

$$\begin{aligned}\|\mathrm{d}\mathcal{Q}_k\| &\le \|(\mathrm{d}\mathcal{H}_k)^\top \mathcal{H}_k\| + \|\mathcal{H}_k^\top (\mathrm{d}\mathcal{H}_k)\| + \|(\mathrm{d}\mathcal{M}_{k+1})^\top \mathcal{Q}_{k-1}\mathcal{M}_{k+1}\| \\ &\quad + \|\mathcal{M}_{k+1}^\top (\mathrm{d}\mathcal{Q}_{k-1})\mathcal{M}_{k+1}\| + \|\mathcal{M}_{k+1}^\top \mathcal{Q}_{k-1}(\mathrm{d}\mathcal{M}_{k+1})\| \\ &\le \|\mathrm{d}\mathcal{H}_k\|\,\|\mathcal{H}_k\| + \|\mathcal{H}_k\|\,\|\mathrm{d}\mathcal{H}_k\| + \|\mathrm{d}\mathcal{M}_{k+1}\|\,\|\mathcal{Q}_{k-1}\|\,\|\mathcal{M}_{k+1}\| \\ &\quad + \|\mathcal{M}_{k+1}\|^2\,\|\mathrm{d}\mathcal{Q}_{k-1}\| + \|\mathcal{M}_{k+1}\|\,\|\mathcal{Q}_{k-1}\|\,\|\mathrm{d}\mathcal{M}_{k+1}\| \\ &= 2\|\mathcal{H}_k\|\,\|\mathrm{d}\mathcal{H}_k\| + 2\|\mathcal{Q}_{k-1}\|\,\|\mathcal{M}_{k+1}\|\,\|\mathrm{d}\mathcal{M}_{k+1}\| + \|\mathcal{M}_{k+1}\|^2\,\|\mathrm{d}\mathcal{Q}_{k-1}\|.\end{aligned} \tag{65}$$

Assume the standard regularity (boundedness along training) holds: there exist constants $H_{\max}, M_{\max}, Q_{\max}$ such that

$$\|\mathcal{H}_k\| \le H_{\max}, \qquad \|\mathcal{M}_k\| \le M_{\max}, \qquad \|\mathcal{Q}_k\| \le Q_{\max} \quad \text{for all } k. \tag{66}$$

Moreover, the Jacobians $\mathcal{A}_k, \mathcal{B}_k, \mathcal{C}_k, \mathcal{D}_k$ are smooth in the upstream parameters; hence there exist $C_H, C_M > 0$ such that

$$\|\mathrm{d}\mathcal{H}_k\| \le C_H\,\|\mathrm{d}(\text{params})\|, \qquad \|\mathrm{d}\mathcal{M}_{k+1}\| \le C_M\,\|\mathrm{d}(\text{params})\|. \tag{67}$$

Plugging these into equation 65, we get

$$\|\mathrm{d}\mathcal{Q}_k\| \le \underbrace{\big(2H_{\max}C_H + 2Q_{\max}M_{\max}C_M\big)}_{=:A}\,\|\mathrm{d}(\text{params})\| + \underbrace{M_{\max}^2}_{=:B}\,\|\mathrm{d}\mathcal{Q}_{k-1}\|. \tag{68}$$

Let $\alpha_k := \|\mathrm{d}\mathcal{Q}_k\|$. Then $\alpha_k \le A\|\mathrm{d}(\text{params})\| + B\alpha_{k-1}$ with $\alpha_{-1} = 0$. Unrolling the recursion,

$$\alpha_k \le A\|\mathrm{d}(\text{params})\| \sum_{j=0}^{k} B^j = A\|\mathrm{d}(\text{params})\| \cdot \frac{B^{k+1}-1}{B-1} \quad \text{if } B \ne 1, \tag{69}$$

and $\alpha_k \le A(k+1)\|\mathrm{d}(\text{params})\|$ if $B = 1$. In all cases there exists a finite constant $C_Q$ (depending on $H_{\max}, M_{\max}, Q_{\max}, K, C_H, C_M$) such that

$$\|\mathrm{d}\mathcal{Q}\| = \|\mathrm{d}\mathcal{Q}_{K-1}\| \;\le\; C_Q\,\|\mathrm{d}(\text{params})\|. \tag{70}$$

**Step 2. Bound $\|\mathrm{d}\mathbf{\Omega}\|$ in terms of $\|\mathrm{d}\mathcal{Q}\|$.** Partition $\mathcal{Q}$ conformally and define

$$S := (\mathcal{Q}_{\lambda\lambda} + \delta_{\text{tik}}\mathbf{I})^{-1}, \qquad \mathbf{\Omega} = \mathcal{Q}_{qq} - \mathcal{Q}_{q\lambda}\,S\,\mathcal{Q}_{\lambda q}. \tag{71}$$

Differentiate:

$$\mathrm{d}\mathbf{\Omega} = \mathrm{d}\mathcal{Q}_{qq} - (\mathrm{d}\mathcal{Q}_{q\lambda})S\mathcal{Q}_{\lambda q} - \mathcal{Q}_{q\lambda}(\mathrm{d}S)\mathcal{Q}_{\lambda q} - \mathcal{Q}_{q\lambda}S(\mathrm{d}\mathcal{Q}_{\lambda q}). \tag{72}$$

Using the identity $\mathrm{d}S = -S(\mathrm{d}\mathcal{Q}_{\lambda\lambda})S$, we obtain

$$\|\mathrm{d}\mathbf{\Omega}\| \le \|\mathrm{d}\mathcal{Q}_{qq}\| + \|\mathrm{d}\mathcal{Q}_{q\lambda}\|\,\|S\|\,\|\mathcal{Q}_{\lambda q}\| + \|\mathcal{Q}_{q\lambda}\|\,\|S\|^2\,\|\mathrm{d}\mathcal{Q}_{\lambda\lambda}\|\,\|\mathcal{Q}_{\lambda q}\| + \|\mathcal{Q}_{q\lambda}\|\,\|S\|\,\|\mathrm{d}\mathcal{Q}_{\lambda q}\|. \tag{73}$$

Since $\|S\| = \|(\mathcal{Q}_{\lambda\lambda} + \delta_{\text{tik}}\mathbf{I})^{-1}\| \le 1/\delta_{\text{tik}}$,

$$\begin{aligned}\|\mathrm{d}\mathbf{\Omega}\| &\le \|\mathrm{d}\mathcal{Q}_{qq}\| + \frac{\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}}\|\mathrm{d}\mathcal{Q}_{q\lambda}\| + \frac{\|\mathcal{Q}_{q\lambda}\|\,\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}^2}\|\mathrm{d}\mathcal{Q}_{\lambda\lambda}\| + \frac{\|\mathcal{Q}_{q\lambda}\|}{\delta_{\text{tik}}}\|\mathrm{d}\mathcal{Q}_{\lambda q}\| \\ &\le \Big(1 + \frac{\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}} + \frac{\|\mathcal{Q}_{q\lambda}\|\,\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}^2} + \frac{\|\mathcal{Q}_{q\lambda}\|}{\delta_{\text{tik}}}\Big)\|\mathrm{d}\mathcal{Q}\|.\end{aligned} \tag{74}$$

Define

$$C_\Omega \;:=\; 1 + \frac{\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}} + \frac{\|\mathcal{Q}_{q\lambda}\|\,\|\mathcal{Q}_{\lambda q}\|}{\delta_{\text{tik}}^2} + \frac{\|\mathcal{Q}_{q\lambda}\|}{\delta_{\text{tik}}}, \tag{75}$$

then

$$\|\mathrm{d}\mathbf{\Omega}\| \;\le\; C_\Omega\,\|\mathrm{d}\mathcal{Q}\|. \tag{76}$$

Combining with equation 70 gives

$$\|\mathrm{d}\mathbf{\Omega}\| \;\le\; C_\Omega\, C_Q\,\|\mathrm{d}(\text{params})\|. \tag{77}$$

**Step 3. Bound the Soft–Ritz stage $\mathbf{\Omega} \mapsto \boldsymbol{q}_{\text{soft}}$.** Let $V_r = [v_1, \dots, v_r] \in \mathbb{R}^{n_f \times r}$ be an orthonormal Ritz basis, $\Pi_r = V_r V_r^\top$ the projector, and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r)$ the Ritz values. Define the soft weights

$$w_i^\star = \frac{e^{-\lambda_i/\tau}}{\sum_{j=1}^r e^{-\lambda_j/\tau}}, \qquad \boldsymbol{w}^\star = (w_1^\star, \dots, w_r^\star)^\top, \tag{78}$$

and form $x := V_r \boldsymbol{w}^\star$, then $\boldsymbol{q}_{\text{soft}} = x/\|x\|$.

*(3a) Subspace rotation bound via Davis–Kahan.* Let $\delta_r := \lambda_{r+1}(\mathbf{\Omega}) - \lambda_r(\mathbf{\Omega}) \ge \delta_{\min}$. By Lemma 2,

$$\|\mathrm{d}\Pi_r\| \;\le\; \frac{\|\mathrm{d}\mathbf{\Omega}\|}{\delta_r} \;\le\; \frac{1}{\delta_{\min}}\|\mathrm{d}\mathbf{\Omega}\|. \tag{79}$$

We now relate $\mathrm{d}V_r$ to $\mathrm{d}\Pi_r$. Since $\Pi_r = V_r V_r^\top$, differentiating yields

$$\mathrm{d}\Pi_r = (\mathrm{d}V_r)V_r^\top + V_r(\mathrm{d}V_r)^\top. \tag{80}$$

Decompose $\mathrm{d}V_r$ into in-subspace and out-of-subspace parts:

$$\mathrm{d}V_r = \Pi_r\,\mathrm{d}V_r + (I - \Pi_r)\,\mathrm{d}V_r. \tag{81}$$

Left-multiplying the identity by $(I - \Pi_r)$ and right-multiplying by $V_r$ gives

$$(I - \Pi_r)\,\mathrm{d}\Pi_r\,V_r = (I - \Pi_r)\,(\mathrm{d}V_r)V_r^\top V_r + (I - \Pi_r)\,V_r(\mathrm{d}V_r)^\top V_r = (I - \Pi_r)\,\mathrm{d}V_r, \tag{82}$$

because $V_r^\top V_r = I_r$ and $(I - \Pi_r)V_r = \mathbf{0}$. Hence

$$\|(I - \Pi_r)\,\mathrm{d}V_r\| \;\le\; \|\mathrm{d}\Pi_r\|. \tag{83}$$

Similarly, right-multiplying $\mathrm{d}\Pi_r$ by $V_r$ and using $V_r^\top(I - \Pi_r) = \mathbf{0}$ gives

$$\Pi_r\,\mathrm{d}V_r = \mathrm{d}\Pi_r\,V_r - V_r(\mathrm{d}V_r)^\top V_r. \tag{84}$$

Since $V_r^\top \mathrm{d}V_r$ is skew-symmetric (from differentiating $V_r^\top V_r = I$), its norm is bounded by $\|\mathrm{d}V_r\|$. Therefore,

$$\|\Pi_r\,\mathrm{d}V_r\| \;\le\; \|\mathrm{d}\Pi_r\| + \|V_r\|\,\|V_r^\top \mathrm{d}V_r\| \;\le\; \|\mathrm{d}\Pi_r\| + \|\mathrm{d}V_r\|. \tag{85}$$

Combine with equation 83 and $\|\mathrm{d}V_r\| \le \|\Pi_r\,\mathrm{d}V_r\| + \|(I - \Pi_r)\,\mathrm{d}V_r\|$ to obtain

$$\|\mathrm{d}V_r\| \;\le\; 2\,\|\mathrm{d}\Pi_r\|. \tag{86}$$

Using equation 79,

$$\|\mathrm{d}V_r\| \;\le\; \frac{2}{\delta_{\min}}\|\mathrm{d}\mathbf{\Omega}\|. \tag{87}$$

*(3b) Ritz value sensitivity within the subspace.* Let $\mathcal{T} = V_r^\top \boldsymbol{\Omega} V_r$ be the $r \times r$ Rayleigh–Ritz matrix. Differentiating,

$$\mathrm{d}\mathcal{T} = (\mathrm{d}V_r)^\top \boldsymbol{\Omega} V_r + V_r^\top (\mathrm{d}\boldsymbol{\Omega}) V_r + V_r^\top \boldsymbol{\Omega} (\mathrm{d}V_r). \tag{88}$$

Hence

$$\begin{aligned} \|\mathrm{d}\mathcal{T}\| &\le \|(\mathrm{d}V_r)^\top\| \, \|\boldsymbol{\Omega}\| \, \|V_r\| + \|V_r\|^2 \, \|\mathrm{d}\boldsymbol{\Omega}\| + \|V_r^\top\| \, \|\boldsymbol{\Omega}\| \, \|\mathrm{d}V_r\| \\ &\le 2 \, \|\boldsymbol{\Omega}\| \, \|\mathrm{d}V_r\| + \|\mathrm{d}\boldsymbol{\Omega}\| \quad \text{since } \|V_r\| = \|V_r^\top\| = 1. \end{aligned} \tag{89}$$

From equation 87,

$$\|\mathrm{d}\mathcal{T}\| \;\le\; \Big(1 + \frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big) \, \|\mathrm{d}\boldsymbol{\Omega}\|. \tag{90}$$

Because the eigenvalues $\lambda_i$ of $\mathcal{T}$ are 1-Lipschitz in $\mathcal{T}$ (Weyl's inequality),

$$\|\mathrm{d}\boldsymbol{\lambda}\| \;\le\; \|\mathrm{d}\mathcal{T}\| \;\le\; \Big(1 + \frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big) \, \|\mathrm{d}\boldsymbol{\Omega}\|. \tag{91}$$

*(3c) Softmin Jacobian bound with explicit constant.* For $\boldsymbol{w}^\star = \mathrm{softmin}_\tau(\boldsymbol{\lambda})$, the Jacobian is

$$\frac{\partial \boldsymbol{w}^\star}{\partial \boldsymbol{\lambda}} = -\frac{1}{\tau} \big(\mathrm{diag}(\boldsymbol{w}^\star) - \boldsymbol{w}^\star (\boldsymbol{w}^\star)^\top\big). \tag{92}$$

By Gershgorin's discs, for the covariance matrix $C := \mathrm{diag}(\boldsymbol{w}^\star) - \boldsymbol{w}^\star(\boldsymbol{w}^\star)^\top$ we have, for each row $i$,

$$\text{disc center } c_i = w_i^\star(1 - w_i^\star), \quad \text{disc radius } r_i = \sum_{j \ne i} |-w_i^\star w_j^\star| = w_i^\star(1 - w_i^\star), \tag{93}$$

so each disc is contained in $[0, 2w_i^\star(1 - w_i^\star)] \subset [0, 1/2]$ since $0 \le w_i^\star \le 1$ implies $w_i^\star(1 - w_i^\star) \le 1/4$. Therefore

$$\|C\| \;\le\; \max_i \big(c_i + r_i\big) \;\le\; \frac{1}{2}. \tag{94}$$

Hence

$$\Big\| \frac{\partial \boldsymbol{w}^\star}{\partial \boldsymbol{\lambda}} \Big\| \;=\; \frac{1}{\tau} \, \|C\| \;\le\; \frac{1}{2\tau} \;\le\; \frac{1}{2\tau_{\min}}. \tag{95}$$

Combining equation 91 and equation 95,

$$\|\mathrm{d}\boldsymbol{w}^\star\| \;\le\; \Big\| \frac{\partial \boldsymbol{w}^\star}{\partial \boldsymbol{\lambda}} \Big\| \, \|\mathrm{d}\boldsymbol{\lambda}\| \;\le\; \frac{1}{2\tau_{\min}} \Big(1 + \frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big) \, \|\mathrm{d}\boldsymbol{\Omega}\|. \tag{96}$$

*(3d) Differential of $x = V_r \boldsymbol{w}^\star$ and normalization.* We have

$$\mathrm{d}x = (\mathrm{d}V_r) \, \boldsymbol{w}^\star + V_r \, \mathrm{d}\boldsymbol{w}^\star, \tag{97}$$

so

$$\|\mathrm{d}x\| \le \|\mathrm{d}V_r\| \, \|\boldsymbol{w}^\star\| + \|V_r\| \, \|\mathrm{d}\boldsymbol{w}^\star\| \;\le\; \|\mathrm{d}V_r\| \, \|\boldsymbol{w}^\star\| + \|\mathrm{d}\boldsymbol{w}^\star\|. \tag{98}$$

Using equation 87 and equation 96,

$$\|\mathrm{d}x\| \;\le\; \frac{2}{\delta_{\min}} \|\boldsymbol{w}^\star\| \, \|\mathrm{d}\boldsymbol{\Omega}\| + \frac{1}{2\tau_{\min}} \Big(1 + \frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big) \, \|\mathrm{d}\boldsymbol{\Omega}\|. \tag{99}$$

Because $V_r$ has orthonormal columns, $\|x\| = \|V_r \boldsymbol{w}^\star\| = \|\boldsymbol{w}^\star\|$. Over the probability simplex, the minimum of $\|\boldsymbol{w}^\star\|$ occurs at the uniform vector, so $\|\boldsymbol{w}^\star\| \ge 1/\sqrt{r}$. Therefore

$$\|x\| \;\ge\; \frac{1}{\sqrt{r}}. \tag{100}$$

For the normalization map $\nu(x) = x/\|x\|$, by Lemma 3, the differential satisfies $\|\mathrm{d}\nu(x)\| \le 2/\|x\|$ when $\|x\| \ge 1/\sqrt{r}$. Hence

$$\|\mathrm{d}\boldsymbol{q}_{\text{soft}}\| = \Big\| \mathrm{d} \frac{x}{\|x\|} \Big\| \;\le\; \frac{2}{\|x\|} \, \|\mathrm{d}x\| \;\le\; 2\sqrt{r} \, \|\mathrm{d}x\| \quad \text{by equation 100.} \tag{101}$$

Combining equation 99 and equation 101,

$$\begin{aligned}\|\mathrm{d}\boldsymbol{q}_{\text{soft}}\| &\le 2\sqrt{r}\left[\frac{2}{\delta_{\min}}\|\boldsymbol{w}^\star\| + \frac{1}{2\tau_{\min}}\Big(1+\frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big)\right]\|\mathrm{d}\boldsymbol{\Omega}\| \\ &\le 2\sqrt{r}\left[\frac{2}{\delta_{\min}} + \frac{1}{2\tau_{\min}}\Big(1+\frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big)\right]\|\mathrm{d}\boldsymbol{\Omega}\| \quad \text{since } \|\boldsymbol{w}^\star\| \le 1. \end{aligned} \tag{102}$$

Absorb $2\sqrt{r}$ and $\|\boldsymbol{\Omega}\|$ into a constant $C_{\text{soft},1}$ (they are bounded by regularity), and use the elementary inequality

$$\frac{2}{\delta_{\min}} + \frac{1}{2\tau_{\min}}\Big(1+\frac{4\,\|\boldsymbol{\Omega}\|}{\delta_{\min}}\Big) \;\le\; \frac{C_{\text{soft},2}}{\tau_{\min}\,\delta_{\min}} \tag{103}$$

for a finite $C_{\text{soft},2}$, to get

$$\|\mathrm{d}\boldsymbol{q}_{\text{soft}}\| \;\le\; \frac{C_{\text{soft}}}{\tau_{\min}\,\delta_{\min}}\,\|\mathrm{d}\boldsymbol{\Omega}\|. \tag{104}$$

**Step 4. Chain all bounds.** From equation 77 and equation 104,

$$\|\mathrm{d}\boldsymbol{q}_{\text{soft}}\| \;\le\; \frac{C_{\text{soft}}}{\tau_{\min}\,\delta_{\min}}\,\|\mathrm{d}\boldsymbol{\Omega}\| \;\le\; \frac{C_{\text{soft}}\,C_\Omega\,C_Q}{\tau_{\min}\,\delta_{\min}}\,\|\mathrm{d}(\text{params})\|. \tag{105}$$

Setting $C_{\text{system}} := C_{\text{soft}}\,C_\Omega\,C_Q$ proves

$$\big\|\nabla_{\text{params}}\boldsymbol{q}_{\text{soft}}\big\| \;\le\; \frac{C_{\text{system}}}{\tau_{\min}\,\delta_{\min}}. \tag{106}$$

**Independence from internal gaps.** Note that the only spectral separation entering the bound is the *boundary gap* $\delta_r = \lambda_{r+1} - \lambda_r$. No factor like $1/(\lambda_{i+1} - \lambda_i)$ with $i < r$ appears. Thus the bound is independent of internal spacings among $\{\lambda_1, \ldots, \lambda_r\}$, in contrast to hard selection, whose sensitivity contains the factor $1/(\lambda_2 - \lambda_1)$. □

# B Technical Lemmas and Auxiliary Results

For completeness, we provide detailed proofs of the key technical lemmas underlying our stability analysis.

## B.1 Davis–Kahan Projector Bound

**Lemma 2** (Davis–Kahan Projector Bound)**.** *Let $\boldsymbol{\Omega} \in \mathbb{R}^{n_f \times n_f}$ be symmetric with eigenvalues $\lambda_1 \le \cdots \le \lambda_{n_f}$ and denote by $V_r \in \mathbb{R}^{n_f \times r}$ the orthonormal eigenvectors associated with the $r$ smallest eigenvalues. Let $\Pi_r := V_r V_r^\top$ be the orthogonal projector onto this subspace. Consider a symmetric perturbation $\Delta\boldsymbol{\Omega}$ and the perturbed matrix $\boldsymbol{\Omega}+\Delta\boldsymbol{\Omega}$. If the boundary gap $\delta_r := \lambda_{r+1}(\boldsymbol{\Omega}) - \lambda_r(\boldsymbol{\Omega}) > 0$, then the differential of the projector satisfies*

$$\big\|\,\mathrm{d}\Pi_r[\boldsymbol{\Omega}](\Delta\boldsymbol{\Omega})\,\big\| \;\le\; \frac{\|\Delta\boldsymbol{\Omega}\|}{\delta_r}. \tag{107}$$

*Proof.* We outline a standard proof via spectral calculus. Let $\Gamma$ be a closed contour in the complex plane enclosing only the first $r$ eigenvalues of $\boldsymbol{\Omega}$ (no others), which is possible since $\delta_r > 0$. The spectral projector admits the Cauchy integral representation

$$\Pi_r = \frac{1}{2\pi\mathrm{i}}\oint_\Gamma (z\mathbf{I} - \boldsymbol{\Omega})^{-1}\,\mathrm{d}z. \tag{108}$$

Perturbing $\boldsymbol{\Omega} \mapsto \boldsymbol{\Omega} + \Delta\boldsymbol{\Omega}$ and differentiating under the integral sign gives

$$\mathrm{d}\Pi_r(\Delta\boldsymbol{\Omega}) = \frac{1}{2\pi\mathrm{i}}\oint_\Gamma (z\mathbf{I} - \boldsymbol{\Omega})^{-1}\,\Delta\boldsymbol{\Omega}\,(z\mathbf{I} - \boldsymbol{\Omega})^{-1}\,\mathrm{d}z. \tag{109}$$

Taking operator norms and using $\|(z\mathbf{I} - \boldsymbol{\Omega})^{-1}\| \le 1/\mathrm{dist}(z, \sigma(\boldsymbol{\Omega}))$, we can choose $\Gamma$ so that $\mathrm{dist}(z, \sigma(\boldsymbol{\Omega})) \ge \delta_r/2$ on $\Gamma$, yielding

$$\|\mathrm{d}\Pi_r(\Delta\boldsymbol{\Omega})\| \;\le\; \frac{1}{2\pi}\cdot\frac{4}{\delta_r^2}\cdot\mathrm{length}(\Gamma)\cdot\|\Delta\boldsymbol{\Omega}\| \;\le\; \frac{C}{\delta_r}\,\|\Delta\boldsymbol{\Omega}\|. \tag{110}$$

Since $\Gamma$ can be chosen with bounded length proportional to $\delta_r$, the constant collapses to 1 in the limit, giving the claimed bound. □

### B.2 Normalization Map: Differential Bound

**Lemma 3** (Normalization Map Differential Bound)**.** *Let $\boldsymbol{x} \mapsto \hat{\boldsymbol{x}} := \boldsymbol{x}/\|\boldsymbol{x}\|_2$ on $\{\boldsymbol{x} : \|\boldsymbol{x}\| \geq \eta\}$ with $\eta > 0$. Then*

$$\left\| \mathrm{d}\hat{\boldsymbol{x}}[\boldsymbol{x}](\Delta\boldsymbol{x}) \right\| \;\leq\; \frac{2}{\eta}\,\|\Delta\boldsymbol{x}\|. \tag{111}$$

*Proof.* Write $\hat{\boldsymbol{x}} = \|\boldsymbol{x}\|^{-1}\boldsymbol{x}$ and differentiate using the product rule and $\mathrm{d}\|\boldsymbol{x}\| = \langle \boldsymbol{x}/\|\boldsymbol{x}\|,\, \mathrm{d}\boldsymbol{x}\rangle$. The differential is

$$\mathrm{d}\hat{\boldsymbol{x}} = \frac{1}{\|\boldsymbol{x}\|}\,\mathrm{d}\boldsymbol{x} - \frac{1}{\|\boldsymbol{x}\|^3}\,\langle \boldsymbol{x}, \mathrm{d}\boldsymbol{x}\rangle\,\boldsymbol{x} = \frac{1}{\|\boldsymbol{x}\|}\Big(\mathbf{I} - \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^\top\Big)\mathrm{d}\boldsymbol{x}. \tag{112}$$

Since $\|\mathbf{I} - \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^\top\| = 1$ and $\|\boldsymbol{x}\| \geq \eta$, we have

$$\|\mathrm{d}\hat{\boldsymbol{x}}\| \leq \frac{1}{\eta}\,\|\mathrm{d}\boldsymbol{x}\|. \tag{113}$$

The factor of 2 in the lemma statement provides a slightly looser but more convenient bound for our purposes. □

**Lemma 4** (Two-eigenvector case: eventual monotone task improvement as the gap grows)**.** *Assume $r = 2$ with $\Delta := \lambda_2 - \lambda_1 > 0$ and soft weights $w_1^* = 1/(1 + e^{-\Delta/\tau})$, $w_2^* = 1 - w_1^*$. Let*

$$\gamma_1(\Delta) = \frac{w_1^*(\Delta)}{\sqrt{(w_1^*(\Delta))^2 + (1 - w_1^*(\Delta))^2}}, \qquad E(\Delta) := R(\boldsymbol{q}_{soft}) - \lambda_1 = \Delta\,\big(1 - \gamma_1(\Delta)^2\big).$$

*Then:*

1. *$E(0) = 0$ and $E(\Delta) > 0$ for all $\Delta > 0$;*
2. *$\gamma_1'(\Delta) > 0$ for all $\Delta > 0$;*
3. *$\lim_{\Delta\to\infty} \gamma_1(\Delta) = 1$ and $\lim_{\Delta\to\infty} E(\Delta) = 0$;*
4. *There exists $\Delta_0(\tau) \in (0, \infty)$ such that $E'(\Delta) < 0$ for all $\Delta \geq \Delta_0(\tau)$; in particular, for sufficiently large gaps the soft Rayleigh value* decreases *(and alignment increases) monotonically with further gap growth.*

## C Proofs for Corollary 2

**Lemma 5** (Non-identifiability under internal degeneracy)**.** *Assume $\lambda_1 = \cdots = \lambda_g = \lambda_\star < \lambda_{g+1}$. Let $w_i^\star = \exp(-\lambda_i/\tau)/\sum_{k=1}^r \exp(-\lambda_k/\tau)$ and $q_{\mathrm{soft}} = \frac{\sum_{i=1}^r w_i^\star v_i}{\left\|\sum_{i=1}^r w_i^\star v_i\right\|}$. If $r = g$ or $\tau \to 0$, then $q_{\mathrm{soft}} \in \mathrm{span}\{v_1, \ldots, v_g\}$. If $r > g$ and $\tau \in (0, \infty)$, then in general $q_{\mathrm{soft}} \notin \mathrm{span}\{v_1, \ldots, v_g\}$ and it depends on the chosen orthonormal basis inside the degenerate block.*

*Proof.* Let $P_g := \sum_{i=1}^g v_i v_i^\top$ and $P_{>g} := I - P_g$. Write $q_{\mathrm{soft}} = \sum_{i=1}^r \gamma_i v_i$ with $\gamma_i = \frac{w_i^\star}{\sqrt{D}}$ and $D = \sum_{k=1}^r (w_k^\star)^2$. Since $w_i^\star \propto e^{-\lambda_i/\tau}$ and $w_1^\star = \cdots = w_g^\star \propto e^{-\lambda_\star/\tau}$, we obtain

$$\|P_{>g} q_{\mathrm{soft}}\|^2 = \sum_{i>g} \gamma_i^2 = \frac{\sum_{i>g} (w_i^\star)^2}{\sum_{k=1}^r (w_k^\star)^2} = \frac{\sum_{i>g} e^{-\frac{2}{\tau}(\lambda_i - \lambda_\star)}}{g + \sum_{i>g} e^{-\frac{2}{\tau}(\lambda_i - \lambda_\star)}}. \tag{114}$$

For $r > g$ and finite $\tau$, the numerator is strictly positive, hence $\|P_{>g} q_{\mathrm{soft}}\|^2 > 0$ and $q_{\mathrm{soft}} \notin \mathrm{span}\{v_1, \ldots, v_g\}$. If $r = g$ or $\tau \to 0$, the numerator vanishes and $q_{\mathrm{soft}} \in \mathrm{span}\{v_1, \ldots, v_g\}$.

Non-uniqueness follows because within the degenerate block any orthonormal change of basis $[v_1, \ldots, v_g] \mapsto [v_1, \ldots, v_g]U$, $U \in O(g)$, keeps $w_1^\star = \cdots = w_g^\star$ equal yet rotates $\sum_{i=1}^g w_i^\star v_i$ to a different direction. □

**Lemma 6** (Two-eigen closed forms)**.** *For $r = 2$, let $\Delta := \lambda_2 - \lambda_1 > 0$ and $t := e^{-\Delta/\tau}$. Then $q_{\text{soft}} \propto v_1 + t\, v_2$ and, after normalization,*

$$\gamma_1(\Delta) = \langle q_{\text{soft}}, v_1 \rangle = \frac{1}{\sqrt{1+t^2}} = \frac{1}{\sqrt{1+e^{-2\Delta/\tau}}}, \qquad E(\Delta) := R(q_{\text{soft}}) - \lambda_1 = \Delta\, \frac{t^2}{1+t^2}. \tag{115}$$

*Proof.* From $w_1^\star : w_2^\star = 1 : e^{-\Delta/\tau} = 1 : t$, we have $w_1^\star = \frac{1}{1+t}$ and $w_2^\star = \frac{t}{1+t}$. Hence $q_{\text{soft}} \propto v_1 + t\, v_2$ and

$$D = (w_1^\star)^2 + (w_2^\star)^2 = \frac{1+t^2}{(1+t)^2}, \qquad \gamma_1 = \frac{w_1^\star}{\sqrt{D}} = \frac{1}{\sqrt{1+t^2}}. \tag{116}$$

Finally,

$$R(q_{\text{soft}}) = \lambda_1 + (\lambda_2 - \lambda_1)\gamma_2^2 = \lambda_1 + \Delta\,(1 - \gamma_1^2) = \lambda_1 + \Delta\, \frac{t^2}{1+t^2}. \tag{117}$$

□

*Proof of Corollary 2. (i) Degeneracy.* The inclusion/non-inclusion statements and basis dependence are exactly Lemma 5.

*(ii) Task-side progress when identifiable.* Fix $V_r$ and set $\beta := 1/\tau > 0$. Let $w_i := w_i^\star$ and $D := \sum_k w_k^2$ for brevity; write $q_{\text{soft}} = \sum_i \gamma_i v_i$ with $\gamma_i = \frac{w_i}{\sqrt{D}}$. The softmin Jacobian is

$$\frac{\partial w_i}{\partial \lambda_j} = -\beta\, w_i(\delta_{ij} - w_j). \tag{118}$$

**Alignment increases with any $(\lambda_j - \lambda_1)$, $j \geq 2$.** For $j \geq 2$,

$$\frac{\partial \gamma_1}{\partial \lambda_j} = \frac{1}{\sqrt{D}} \frac{\partial w_1}{\partial \lambda_j} - \frac{w_1}{2D^{3/2}} \cdot 2 \sum_k w_k \frac{\partial w_k}{\partial \lambda_j} \tag{119}$$

$$= \frac{-\beta w_1(-w_j)}{\sqrt{D}} - \frac{w_1}{D^{3/2}}\Big( - \beta w_j(w_j - D)\Big) = \frac{\beta\, w_1\, w_j^2}{D^{3/2}} > 0. \tag{120}$$

Thus increasing any $\lambda_j$ ($j \geq 2$) strictly increases the alignment $\langle q_{\text{soft}}, v_1 \rangle = \gamma_1$.

**Rayleigh value decreases once the gap is large enough.** Let $R := R(q_{\text{soft}}) = \lambda_1 + \sum_{i \geq 2}(\lambda_i - \lambda_1)\gamma_i^2$. A direct quotient differentiation using equation 118 yields

$$\frac{\partial R}{\partial \lambda_j} = \frac{w_j^2}{D}\Big(1 - 2\beta(\lambda_j - R)\Big), \qquad j \in \{1, \ldots, r\}. \tag{121}$$

Hence for $j \geq 2$ we have $\frac{\partial R}{\partial \lambda_j} < 0$ whenever $\lambda_j - R > \frac{1}{2\beta} = \frac{\tau}{2}$, i.e., once the internal gap (relative to the current soft Rayleigh value) is sufficiently large. This proves the stated monotone trend.

**Closed forms for $r = 2$.** By Lemma 6, with $\Delta = \lambda_2 - \lambda_1$ and $t = e^{-\Delta/\tau}$,

$$\gamma_1(\Delta) = \frac{1}{\sqrt{1+t^2}}, \qquad E(\Delta) = R - \lambda_1 = \Delta\, \frac{t^2}{1+t^2}. \tag{122}$$

Then $\gamma_1'(\Delta) > 0$ for all $\Delta > 0$. A direct derivative gives

$$E'(\Delta) = \frac{t^2}{(1+t^2)^2}\Big(1 + t^2 - \frac{2\Delta}{\tau}\Big), \tag{123}$$

which is $< 0$ for all $\Delta \geq \Delta_0(\tau)$ where $\Delta_0(\tau)$ uniquely solves $\Delta = \frac{\tau}{2}(1 + e^{-2\Delta/\tau})$. Moreover $E(\Delta) \to 0$ as $\Delta \to \infty$. □