Gradient Descent on Two ReLU Neurons: Global Landscape and Bifurcation Dynamics

author names withheld

Under Review for the Workshop on High-dimensional Learning Dynamics, 2026

Abstract

Understanding how gradient descent learns features remains a central challenge in neural network theory. We study this question in a minimal multi-index setting that already exhibits rich multi-phase gradient dynamics: a well-specified two-neuron ReLU teacher-student model under isotropic Gaussian inputs, trained by population gradient flow from small random initialization. We show that the dynamics are organized by two directions: the "easy" bisector direction, which carries the leading signal, and the "hard" splitting direction, which governs specialization. To characterize the loss structure and learning behavior, we analyze the population squared-loss landscape and show that every nonzero critical point is either a global minimum or a saddle. We then track the gradient-flow trajectory, showing that student neurons first collapse toward a bisector saddle before escaping and specializing to teacher neurons. Our results provide both a landscape and dynamics account of the multi-phase symmetry-breaking behavior that arises in a simple multi-index model and under standard algorithmic and architectural choices.

1. Introduction

A central promise of gradient-based training of neural networks is *feature learning*: rather than fitting the target function on a fixed feature space, neural networks can extract features by aligning their parameters with the latent structure of the data or target function. A recent line of theoretical works has made this intuition precise by analyzing teacher-student models under Gaussian inputs, where the target function is a *multi-index model* [1, 4, 16] which depends on a small number of input directions (e.g., a shallow neural network with $O_d(1)$ neurons). In these settings, it has been shown that gradient descent training of a shallow "student" neural network discovers the relevant subspace, and achieves a sample complexity superior to kernel methods (or neural networks in the "lazy" regime [15]).

However, most existing theory becomes tractable only after imposing simplifying assumptions that deviate from standard algorithmic and architectural choices. Examples include (i) near-orthogonal teacher neurons [24, 29], (ii) local convergence conditions [43, 44], (iii) activation functions with high *information exponent* [27, 30, 35], and (iv) modified algorithms such as loss preprocessing [16, 42] and spherical/Stiefel gradients [2, 5, 6]. While these assumptions enable clean analyses, they do not capture the complex *multi-phase* learning dynamics present in standard ReLU networks [12, 38]. A notable multi-index example exhibiting such nontrivial dynamics is the "well-specified" setting in which both the student and teacher are two-layer neural networks with matching activation functions and more than one neuron: for commonly used nonlinearities such as ReLU, the optimization trajectory exhibits symmetry-breaking and bifurcation behavior, with neurons initially moving toward an "average" direction before specializing to different teacher directions. While this plateau-followed-by-specialization phenomenon has appeared in the statistical physics literature [8, 22, 32], it is not captured by analyses that rely on the above simplifications. To our knowledge, a rigorous and quantitative understanding of the loss landscape and multi-phase training dynamics remains incomplete and largely open.

1.1. Our contributions

We study arguably the simplest multi-index model setting that exhibits this bifurcation phenomenon: learning a two-neuron ReLU network under isotropic Gaussian inputs via population gradient flow, starting from small initialization. Specifically, the teacher model is

$$f^*(x) = \mathrm{ReLU}(\langle v_1, x\rangle) + \mathrm{ReLU}(\langle v_2, x\rangle), \qquad x \sim \mathcal{N}(0, I_d), \tag{1}$$

where $\mathrm{ReLU}(z) = \max\{z, 0\}$, $\|v_1\|_2 = \|v_2\|_2 = 1$ and $\angle(v_1, v_2) = \alpha \in (0, \frac{\pi}{2}]$, and the student has the same architecture with trainable weights w_1, w_2.

As we will see, the geometry of this model is organized by two directions in the teacher plane: the bisector $(v_1 + v_2)$, which carries the dominant signal, and the difference $(v_1 - v_2)$, which governs specialization. In the language of leap complexity for Gaussian multi-index models [2, 10], the target function (1) has *leap exponent* 2, where the bisector and difference vectors specify the "easy" leap-1 direction and the "hard" leap-2 direction, respectively. This suggests a two-stage picture for gradient-based learning: the student neurons are first attracted toward the leap-1 bisector, and only later separate toward the individual teacher directions along the leap-2 direction. Our goal is to make this collapse-then-splitting picture quantitative by studying both the population squared loss landscape and the gradient-flow dynamics from small initialization. Specifically,

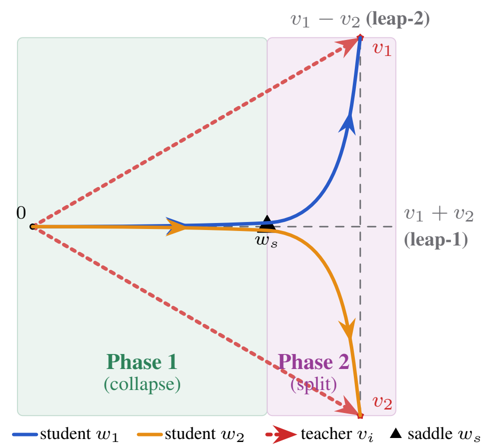

Figure 1: Illustration of bifurcation dynamics.

- **Landscape analysis.** Section 3 studies the loss geometry and the critical-point structure. We reduce the stationarity conditions to a three-dimensional problem and isolate the remaining in-plane and out-of-plane critical families. We establish that every nonzero critical point of the population loss is either a global minimum or a saddle, so the squared loss has no spurious nonzero local minima. This resolves a two-neuron conjecture from Tian [37] in the acute angle ($\alpha \le \pi/2$) setting. This identifies the saddle geometry that controls the later high-dimensional escape phase.
- **Small-initialization dynamics.** Section 4 studies the gradient-flow trajectory from small initialization and characterizes the symmetry breaking and specialization behavior — see Figure 1. We show that for sufficiently small initialization, both student directions stay close to the bisector for fixed finite time and approach the symmetric saddle. We then analyze escape from that saddle, including global convergence to a minimum for typical initializations, as well as a quantitative escape-time bound under a structured initialization that is logarithmic in the ambient dimension d. Thus the model exposes a concrete dimension-dependent time scale for feature specialization.

1.2. Further related work

Learning multi-index models. Recent works have shown shallow neural networks can efficiently learn low-dimensional latent structure via gradient descent [1, 4, 9, 16, 17, 21]. Specifically, the sample complexity has been shown to be superior to that of kernel methods and to be governed by the *information exponent* [5, 18] for single-index models and the *leap exponent* [2, 10] for multi-index models. Our target function (1) is an example of a leap-2 multi-index model, but, crucially, it also contains a leap-1 subspace, which complicates the SGD dynamics. In fact, most prior works studying gradient-based feature learning either directly assume link functions with information exponent 2 [7, 24, 25, 29] or higher [27, 30, 35], or remove the low-leap components through algorithmic modifications [2, 16, 20, 42]. Consequently, these analyses do not fully characterize the learning dynamics of two ReLU neurons with standard gradient-based training.

Training dynamics of ReLU networks. Restricting to shallow neural networks with ReLU activations, prior works have studied landscape properties such as the existence of spurious minima [33, 34]; the concurrent work of [23] gives a summary-statistics characterization of the population minima and proposes certain ansatz to simplify the fixed-point equations. For SGD dynamics, while the single-neuron case is relatively well-understood [36, 37, 40, 41], when the target network contains more than one neuron, existing dynamical analyses either assume local convergence, thereby bypassing the bifurcation behavior [43, 44], or require exponential width to invoke mean-field global convergence [3, 14]. The two-phase trajectory illustrated in Figure 1 also relates to the "plateau" phenomenon observed in the statistical mechanics literature [8, 19]. Specifically, the specialization phenomenon has been investigated in [22, 28] using the ODE analysis originated from [31, 32]. However, to our knowledge, a complete understanding of the landscape and rigorous characterization of the learning dynamics have not yet been established even in the two-neuron setting (1).

2. Setup and Leap Intuition

We use the teacher f^* defined in (1). The student has trainable weights $w_1, w_2 \in \mathbb{R}^d$ and output

$$f(x) = \mathrm{ReLU}(\langle w_1, x\rangle) + \mathrm{ReLU}(\langle w_2, x\rangle). \tag{2}$$

The student is trained by gradient flow on the squared loss under isotropic Gaussian inputs:

$$L(w_1, w_2) = \tfrac{1}{2}\mathbb{E}_{x\sim\mathcal{N}(0,I_d)}[(f(x) - f^*(x))^2], \qquad \tfrac{\mathrm{d}w_i}{\mathrm{d}t} = -\nabla_{w_i} L(w_1, w_2),\ i \in \{1, 2\}, \tag{3}$$

starting from initialization $w(0) = (w_1(0), w_2(0))$ with $\|w(0)\|_2 = \sigma\sqrt{d}$. Within the teacher plane $\mathrm{span}\{v_1, v_2\}$, we highlight two important directions: the *bisector* and the *splitting direction*,

$$u_1 = \bar{v} = \tfrac{v_1+v_2}{\|v_1+v_2\|_2}, \qquad u_2 = v_\Delta = \tfrac{v_1-v_2}{\|v_1-v_2\|_2}. \tag{4}$$

The pair (u_1, u_2) forms an orthonormal basis of the teacher plane. Throughout the paper, we work in the small-initialization regime where $\sigma > 0$ is sufficiently small.

For the two-ReLU teacher network defined in (1), let $z_1 := \langle u_1, x\rangle$ and $z_2 := \langle u_2, x\rangle$ be the projections of the input onto the bisector and splitting directions defined in (4). Note that $f^*(x)$ depends on x only through these two coordinates, and hence we can write $f^*(x) = g(z_1, z_2)$.

Proposition 1 (Teacher leap complexity) *The two-ReLU network f^* in* (1) *has leap exponent 2, where the bisector coordinate z_1 has leap 1 and the splitting coordinate z_2 has leap 2.*

This proposition is the teacher-side explanation of Figure 1: the leading correlation pulls both students toward the bisector, while the splitting coordinate becomes decisive only after the full squared loss creates an escape direction away from the collapsed state.

3. Loss landscape

We first isolate the leading correlation term. For a width-k ReLU teacher with atoms $v_1, \dots, v_k$, define the correlation loss $L_{\mathrm{corr}}(w_1, \dots, w_m) = -\sum_{j=1}^{m} \mathbb{E}[f^*(x)\, \mathrm{ReLU}(\langle w_j, x\rangle)]$ for a width-m student. By ReLU homogeneity, this is the first-order dominant term in the squared loss near small student weights. The setup in Section 2 corresponds to the special case $k = m = 2$.

Theorem 2 (Warm up: landscape of the correlation loss) *Let $v_1, \dots, v_k$ be unit vectors with $\angle(v_i, v_j) = \alpha \in (0, \pi/2]$ for all $i \neq j$, and let $\bar{v} = (\sum_i v_i)/\|\sum_i v_i\|_2$. Then $(\bar{v}, \dots, \bar{v})$ is the unique global minimum of L_{corr} among unit-norm student weights.*

In the two-neuron model, this averaged direction is exactly the bisector. Thus the correlation loss explains the initial collapse, but not specialization. That separation comes from the full squared loss.

Theorem 3 (Landscape of the squared loss) *Let L be the population squared loss in (3). Every nonzero stationary point of L is, up to permutation of the two student neurons, one of the following:*

1. *a global minimum, $w_1 = v_1$ and $w_2 = v_2$;*
2. *an in-plane saddle supported on the bisector directions $\bar{v}$ and $-\bar{v}$;*
3. *an out-of-plane saddle of the symmetric form $w_1 = r(\cos\kappa\, \bar{v} + \sin\kappa\, \xi)$, $w_2 = r(\cos\kappa\, \bar{v} - \sin\kappa\, \xi)$, where $r > 0$, $\kappa \in (0, \pi/2)$, and $\xi \perp \mathrm{span}\{v_1, v_2\}$ is a unit vector.*

Consequently, the squared loss has no spurious local minima among nonzero student weights.

The proof reduces stationarity to three dimensions and solves the remaining in-plane and out-of-plane cases. For the dynamics, the key point is that the positive-bisector family contains the collapsed saddle reached after the leap-1 feature is learned, before the leap-2 splitting feature has emerged.

Remark 4 *Tian [37] conjectured the unconstrained population squared loss for $k = 2$ ReLU teacher-student under isotropic Gaussian inputs has no spurious local minima. Wu et al. [39] made progress by proving the absence of spurious minima for* orthogonal *teachers on the* unit-norm manifold. *Our Theorem 3 resolves this open problem in the acute angle case: it provides an unconditional, complete classification of every nonzero critical point of the unconstrained squared loss for general $\alpha \in (0, \pi/2]$, showing that every nonzero critical point is either the global minimum or a saddle.*

4. Dynamics and Escape Time Scaling

Let $r_+ = \frac{2}{\pi}((\pi - \alpha/2)\cos(\alpha/2) + \sin(\alpha/2))$, set $r_s = r_+/2$, and write $w_s = r_s \bar{v}$ and $w_{\mathrm{saddle}} = (w_s, w_s)$. We now track gradient flow from small initialization.

Lemma 5 (Phase 1: approaching the bisector saddle) *Fix $\varepsilon, \rho > 0$ and $0 < T_1 < T < \infty$. Assume $\|w_i(0)\|_2 \asymp \sigma$, $\angle(\bar{w}_i(0), \bar{v}) \leq 2\pi/3, i = 1, 2$, with constants independent of $\sigma\sqrt{d}$. For all sufficiently small σ, the gradient-flow trajectory satisfies $\max_{i=1,2} \|\bar{w}_i(t) - \bar{v}\| < \varepsilon$ for all $t \in [T_1, T]$, and there is a time $t_0 = t_0(\alpha, \rho)$ such that $\inf\{t > 0 : \|w(t) - w_{\mathrm{saddle}}\| \leq \rho\} \leq t_0$.*

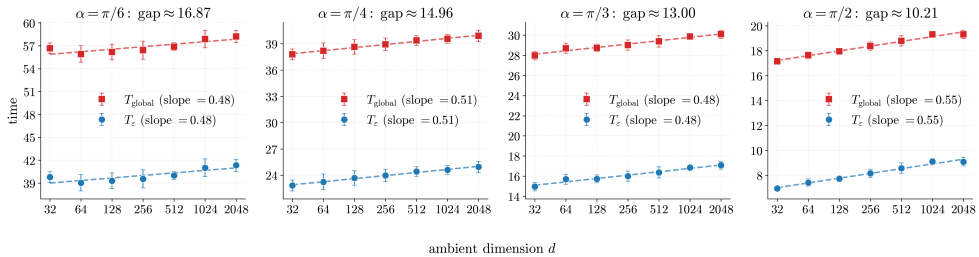

Figure 2: Escape and convergence times vs. dimension at fixed $\sigma = 10^{-4}$, averaged over 5 seeds. Dashed fits show the predicted $\log d$ dependence of T_ε; the later descent time is nearly independent of d.

In words, when the weights are small, the projected teacher force dominates and points toward the bisector. Once the directions are close to $\bar{v}$, radial dynamics carries the norms toward r_s. Complete the directions $u_1 = \bar{v}, u_2 = v_\Delta$ from (4) to an orthonormal basis $u_1, \ldots, u_d$ of $\mathbb{R}^d$.

Near w_{saddle}, the Hessian separates the exits. The planar antisymmetric mode $(v_\Delta, -v_\Delta)$ is unstable, with twice the magnitude of each out-of-plane antisymmetric mode; the symmetric modes are stable. Thus the local geometry favors the in-plane split in Figure 1. Generic saddle avoidance then implies convergence to a global minimum for almost every initialization, in view of theorem 3 (see Theorem 54). The remaining question is quantitative: how long does the trajectory spend near the collapsed saddle before the splitting coordinate becomes visible?

Theorem 6 (Logarithmic escape under structured initialization) *Let $\varepsilon = \varepsilon(\alpha) > 0$ be a small constant. For all sufficiently large d and every initialization scale σ satisfying $\sigma\sqrt{d} \ll 1$, decompose each student as $w_i(t) = \sum_{j=1}^{d} z_{i,j}(t) u_j$, where $u_1 = \bar{v}$ and $u_2 = v_\Delta$. Assume $z_{1,1}(0) + z_{2,1}(0) \asymp \sigma\sqrt{d}, z_{1,2}(0) - z_{2,2}(0) \asymp \sigma, z_{1,2}(0) > 0 > z_{2,2}(0)$, with $|z_{1,1}(0) - z_{2,1}(0)| \lesssim \sigma$ and $\sum_{i=1}^{2} \sum_{j=3}^{d} z_{i,j}(0)^2 \lesssim \sigma^2$. Let T_ε be the first time at which $\sum_{i=1}^{2} \sum_{j=2}^{d} z_{i,j}(t)^2 \geq \varepsilon^2$. Then $T_\varepsilon \leq (\frac{1}{\mu_\alpha} - 1) \log \frac{1}{\sigma} + \frac{1}{2} \log d + T_0$ for constants μ_α and T_0 depending on α.*

Notably, the dependence of this escape time on the ambient dimension is *logarithmic in* d, which is consistent with the "search phase" behavior in learning information/leap exponent 2 functions [5, 13, 26]. Figure 2 shows the same scaling for random initialization. The escape time grows linearly in $\log d$, while the later descent time is essentially dimension independent, supporting the view that the high-dimensional cost lies in the leap-2 symmetry-breaking phase.

5. Conclusion

We studied population gradient flow for a two-neuron ReLU teacher-student model, a minimal multi-index setting that already shows staged high-dimensional feature learning. The loss has no nonzero spurious local minima, but it contains a bisector saddle that organizes training. From small initialization, gradient flow first learns the easy bisector direction, then escapes along the harder splitting direction. The logarithmic escape time identifies where dimension enters: in the leap-2 symmetry-breaking phase, rather than in the final descent after specialization begins. Several directions remain open, including: (i) obtaining an end-to-end escape-time analysis from random initialization; (ii) extending the landscape and dynamical characterization to other teacher-student settings and high-leap multi-index functions.

References

[1] Emmanuel Abbe, Enric Boix Adsera, and Theodor Misiakiewicz. The merged-staircase property: a necessary and nearly sufficient condition for sgd learning of sparse functions on two-layer neural networks. In *Conference on Learning Theory*, pages 4782–4887. PMLR, 2022.

[2] Emmanuel Abbe, Enric Boix Adsera, and Theodor Misiakiewicz. Sgd learning on neural networks: leap complexity and saddle-to-saddle dynamics. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 2552–2623. PMLR, 2023.

[3] Shunta Akiyama and Taiji Suzuki. On learnability via gradient method for two-layer ReLU neural networks in teacher-student setting. In *International Conference on Machine Learning*, 2021.

[4] Jimmy Ba, Murat A Erdogdu, Taiji Suzuki, Zhichao Wang, Denny Wu, and Greg Yang. High-dimensional asymptotics of feature learning: How one gradient step improves the representation. *Advances in Neural Information Processing Systems*, 35:37932–37946, 2022.

[5] Gerard Ben Arous, Reza Gheissari, and Aukosh Jagannath. Online stochastic gradient descent on non-convex losses from high-dimensional inference. *Journal of Machine Learning Research*, 22(106):1–51, 2021.

[6] Gérard Ben Arous, Cédric Gerbelot, and Vanessa Piccolo. Stochastic gradient descent in high dimensions for multi-spiked tensor pca. *arXiv preprint arXiv:2410.18162*, 2024.

[7] Gérard Ben Arous, Murat A Erdogdu, N Mert Vural, and Denny Wu. Learning quadratic neural networks in high dimensions: Sgd dynamics and scaling laws. *arXiv preprint arXiv:2508.03688*, 2025.

[8] Michael Biehl, Peter Riegler, and Christian Wöhler. Transient dynamics of on-line learning in two-layered neural networks. *Journal of Physics A: Mathematical and General*, 29(16): 4769–4780, 1996.

[9] Alberto Bietti, Joan Bruna, Clayton Sanford, and Min Jae Song. Learning single-index models with shallow neural networks. *Advances in neural information processing systems*, 35:9768–9783, 2022.

[10] Alberto Bietti, Joan Bruna, and Loucas Pillaud-Vivien. On learning gaussian multi-index models with gradient flow. *arXiv preprint arXiv:2310.19793*, 2023.

[11] Jérôme Bolte, Aris Daniilidis, and Adrian Lewis. The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *SIAM Journal on Optimization*, 17(4):1205–1223, 2007.

[12] Etienne Boursier, Loucas Pillaud-Vivien, and Nicolas Flammarion. Gradient flow dynamics of shallow relu networks for square loss and orthogonal inputs. *Advances in Neural Information Processing Systems*, 35:20105–20118, 2022.

[13] Yuxin Chen, Yuejie Chi, Jianqing Fan, and Cong Ma. Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval. *Mathematical Programming*, 176 (1):5–37, 2019.

[14] Lenaic Chizat. Sparse optimization on measures with over-parameterized gradient descent. *Mathematical Programming*, 194(1):487–532, 2022.

[15] Lenaic Chizat, Edouard Oyallon, and Francis Bach. On lazy training in differentiable programming. *Advances in neural information processing systems*, 32, 2019.

[16] Alexandru Damian, Jason Lee, and Mahdi Soltanolkotabi. Neural networks can learn representations with gradient descent. In *Conference on Learning Theory*, pages 5413–5452. PMLR, 2022.

[17] Yatin Dandi, Florent Krzakala, Bruno Loureiro, Luca Pesce, and Ludovic Stephan. How two-layer neural networks learn, one (giant) step at a time. *Journal of Machine Learning Research*, 25(349):1–65, 2024.

[18] Rishabh Dudeja and Daniel Hsu. Learning single-index models in gaussian space. In *Conference On Learning Theory*, pages 1887–1930. PMLR, 2018.

[19] Kenji Fukumizu and Shun-ichi Amari. Local minima and plateaus in hierarchical structures of multilayer perceptrons. *Neural networks*, 13(3):317–327, 2000.

[20] Rong Ge, Yunwei Ren, Xiang Wang, and Mo Zhou. Understanding deflation process in over-parametrized tensor decomposition. *Advances in Neural Information Processing Systems*, 34: 1299–1311, 2021.

[21] Margalit Glasgow. Sgd finds then tunes features in two-layer neural networks with near-optimal sample complexity: A case study in the xor problem. *arXiv preprint arXiv:2309.15111*, 2023.

[22] Sebastian Goldt, Madhu S. Advani, Andrew M. Saxe, Florent Krzakala, and Lenka Zdeborová. Dynamics of stochastic gradient descent for two-layer neural networks in the teacher-student setup. *Journal of Statistical Mechanics: Theory and Experiment*, page 124010, 2020. doi: 10.1088/1742-5468/abc61e.

[23] Jie Huang, Bruno Loureiro, and Stefano Sarao Mannelli. Sharp description of local minima in the loss landscape of high-dimensional two-layer relu neural networks. *arXiv preprint arXiv:2604.09412*, 2026.

[24] Yuanzhi Li, Tengyu Ma, and Hongyang R Zhang. Learning over-parametrized two-layer neural networks beyond ntk. In *Conference on learning theory*, pages 2613–2682. PMLR, 2020.

[25] Simon Martin, Francis Bach, and Giulio Biroli. On the impact of overparameterization on the training of a shallow neural network in high dimensions. In *International Conference on Artificial Intelligence and Statistics*, pages 3655–3663. PMLR, 2024.

[26] Andrea Montanari and Zihao Wang. Phase transitions for feature learning in neural networks. *arXiv preprint arXiv:2602.01434*, 2026.

[27] Kazusato Oko, Yujin Song, Taiji Suzuki, and Denny Wu. Learning sum of diverse features: computational hardness and efficient gradient-based training for ridge combinations. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 4009–4081. PMLR, 2024.

[28] Elisa Oostwal, Michiel Straat, and Michael Biehl. Hidden unit specialization in layered neural networks: ReLU vs. sigmoidal activation. *Physica A: Statistical Mechanics and its Applications*, 564:125517, 2021. doi: 10.1016/j.physa.2020.125517.

[29] Yunwei Ren and Jason D Lee. Learning orthogonal multi-index models: A fine-grained information exponent analysis. *arXiv preprint arXiv:2410.09678*, 2024.

[30] Yunwei Ren, Eshaan Nichani, Denny Wu, and Jason D Lee. Emergence and scaling laws in sgd learning of shallow neural networks. *arXiv preprint arXiv:2504.19983*, 2025.

[31] Peter Riegler and Michael Biehl. On-line backpropagation in two-layered neural networks. *Journal of Physics A: Mathematical and General*, 28(20):L507–L513, 1995.

[32] David Saad and Sara A. Solla. On-line learning in soft committee machines. *Physical Review E*, 52(4):4225–4243, 1995.

[33] Itay Safran and Ohad Shamir. Spurious local minima are common in two-layer relu neural networks. In *International conference on machine learning*, pages 4433–4441. PMLR, 2018.

[34] Itay M Safran, Gilad Yehudai, and Ohad Shamir. The effects of mild over-parameterization on the optimization landscape of shallow relu neural networks. In *Conference on Learning Theory*, pages 3889–3934. PMLR, 2021.

[35] Berfin Şimşek, Amire Bendjeddou, and Daniel Hsu. Learning Gaussian multi-index models with gradient flow: Time complexity and directional convergence. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pages 4204–4212. PMLR, 2025.

[36] Mahdi Soltanolkotabi. Learning relus via gradient descent. *Advances in neural information processing systems*, 30, 2017.

[37] Yuandong Tian. An analytical formula of population gradient for two-layered ReLU network and its applications in convergence and critical point analysis. In *International Conference on Machine Learning*, 2017.

[38] Mingze Wang and Chao Ma. Understanding multi-phase optimization dynamics and rich nonlinear behaviors of relu networks. *Advances in Neural Information Processing Systems*, 36: 35654–35747, 2023.

[39] Chenwei Wu, Jiajun Luo, and Jason D Lee. No spurious local minima in a two hidden unit relu network. 2018.

[40] Weihang Xu and Simon Du. Over-parameterization exponentially slows down gradient descent for learning a single neuron. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 1155–1198. PMLR, 2023.

[41] Gilad Yehudai and Ohad Shamir. Learning a single neuron with gradient methods. In *Conference on Learning Theory*, pages 3756–3786. PMLR, 2020.

[42] Bohan Zhang, Zihao Wang, Hengyu Fu, and Jason D Lee. Neural networks learn generic multi-index models near information-theoretic limit. *arXiv preprint arXiv:2511.15120*, 2025.

[43] Mo Zhou, Rong Ge, and Chi Jin. A local convergence theory for mildly over-parameterized two-layer neural network. In *Conference on Learning Theory*, pages 4577–4632. PMLR, 2021.

[44] Zhenyu Zhu, Fanghui Liu, and Volkan Cevher. How gradient descent balances features: A dynamical analysis for two-layer neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

AI Assistance Disclosure

The authors used AI assistance for editing, formatting, and submission preparation. All scientific content, claims, proofs, and final text were reviewed and approved by the authors.

Supplementary Material

Appendix A. Some useful lemmas

A.1. Closed-form expressions for loss functions

From [33], we have the following two useful lemmas.

Lemma 7

$$\mathbb{E}_{x\sim\mathcal{N}(0,I_d)}[\phi(\langle u,x\rangle)\phi(\langle v,x\rangle)] = \frac{\|u\|_2\|v\|_2}{2\pi}\Big(\sin\theta_{u,v} + (\pi-\theta_{u,v})\cos\theta_{u,v}\Big)$$

Remark 8 *For simplicity, we denote* $l(\theta) := \frac{(\pi-\theta)\cos\theta+\sin\theta}{2\pi}, \theta\in[0,\pi]$.

Lemma 9 $l(\theta)$ *is monotonically decreasing and* $0\le l(\theta)\le \frac{1}{2}$ *when* $\theta\in[0,\pi]$.

Proof $l'(\theta) = -\frac{(\pi-\theta)\sin\theta}{2\pi} \le 0$. Thus, for all $0\le\theta\le\pi, 0 = l(\pi)\le l(\theta)\le l(0) = \frac{1}{2}$. ■

Lemma 10 *For any* $i\in\{1,2\}$, *we have*

$$2l\left(\pi-\frac{\alpha}{2}\right) \le l(\theta_{i1}) + l(\theta_{i2}) \le 2l\left(\frac{\alpha}{2}\right).$$

Moreover, $l(\theta_{i1})+l(\theta_{i2}) = 2l(\frac{\alpha}{2})$ *if and only if* $\theta_{i1}=\theta_{i2}=\frac{\alpha}{2}$.

Proof We first prove the lower bound. It suffices to consider $i=1$. Let

$$\vartheta_j := \pi-\theta_{1j} = \angle(-\bar{w}_1, v_j), \quad j=1,2.$$

The triangle inequality applied to $-\bar{w}_1, v_1, v_2$ gives

$$\vartheta_1+\vartheta_2\ge\alpha.$$

Since

$$l(\pi-x) = \frac{\sin x - x\cos x}{2\pi}, \quad x\in[0,\pi],$$

define $h(x) := \sin x - x\cos x$ on $[0,\pi]$. Then $h'(x) = x\sin x \ge 0$, so h is increasing. Choose $a,b\in[0,\alpha]$ such that

$$a+b=\alpha, \quad a\le\vartheta_1, \quad b\le\vartheta_2.$$

This is possible because $\vartheta_1 + \vartheta_2 \geq \alpha$; for instance, take $a = \min\{\vartheta_1, \alpha\}$ and $b = \alpha - a$. Hence

$$h(\vartheta_1) + h(\vartheta_2) \geq h(a) + h(b).$$

Since $\alpha \leq \pi/2$, we have $h''(x) = \sin x + x\cos x \geq 0$ for $x \in [0, \alpha]$, so h is convex on $[0, \alpha]$. Jensen's inequality gives

$$h(a) + h(b) \geq 2h\left(\frac{a+b}{2}\right) = 2h\left(\frac{\alpha}{2}\right).$$

Therefore,

$$l(\theta_{11}) + l(\theta_{12}) = l(\pi - \vartheta_1) + l(\pi - \vartheta_2) \geq 2l\left(\pi - \frac{\alpha}{2}\right).$$

It remains to prove the upper bound and characterize the equality case.

Case 1: $\theta_{11}, \theta_{12} \in [0, \alpha]$.

We only need to prove $l(\theta_{11}) + l(\theta_{12}) \leq 2l(\frac{\alpha}{2})$. By the triangular inequality, we have $\theta_{12} \geq \alpha - \theta_{11}$, then by lemma 9, we have $l(\theta_{11}) + l(\theta_{12}) \leq l(\theta_{11}) + l(\alpha - \theta_{11})$. Let $l_1(x) = l(x) + l(\alpha - x), x \in [0, \alpha]$. We claim that l_1 is strictly concave on $[0, \alpha]$. Indeed, since

$$l'(x) = -\frac{(\pi - x)\sin x}{2\pi},$$

we have

$$\begin{aligned} l_1'(x) &= l'(x) - l'(\alpha - x) \\ &= -\frac{(\pi - x)\sin x}{2\pi} + \frac{(\pi - \alpha + x)\sin(\alpha - x)}{2\pi}. \end{aligned}$$

Differentiating once more gives

$$\begin{aligned} 2\pi l_1''(x) &= \sin x - (\pi - x)\cos x + \sin(\alpha - x) - (\pi - \alpha + x)\cos(\alpha - x) \\ &= \big(\sin x + \sin(\alpha - x)\big) - \big((\pi - x)\cos x + (\pi - \alpha + x)\cos(\alpha - x)\big). \end{aligned}$$

Using

$$\sin x + \sin(\alpha - x) = 2\sin\frac{\alpha}{2}\cos\left(x - \frac{\alpha}{2}\right)$$

and

$$\begin{aligned} &(\pi - x)\cos x + (\pi - \alpha + x)\cos(\alpha - x) \\ &\quad = 2\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\cos\left(x - \frac{\alpha}{2}\right) + 2\left(x - \frac{\alpha}{2}\right)\sin\frac{\alpha}{2}\sin\left(x - \frac{\alpha}{2}\right), \end{aligned}$$

we get

$$\begin{aligned} 2\pi l_1''(x) = {} & 2\left(\sin\frac{\alpha}{2} - \left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\right)\cos\left(x - \frac{\alpha}{2}\right) \\ & - 2\left(x - \frac{\alpha}{2}\right)\sin\frac{\alpha}{2}\sin\left(x - \frac{\alpha}{2}\right). \end{aligned}$$

Since $0 < \alpha \leq \pi/2$, we have

$$\tan\frac{\alpha}{2} < \pi - \frac{\alpha}{2},$$

and hence

$$\sin\frac{\alpha}{2} - \left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} < 0.$$

Moreover, $\left|x - \frac{\alpha}{2}\right| \leq \frac{\alpha}{2} \leq \frac{\pi}{4}$, so

$$\cos\left(x - \frac{\alpha}{2}\right) > 0, \qquad \left(x - \frac{\alpha}{2}\right)\sin\left(x - \frac{\alpha}{2}\right) \geq 0.$$

Therefore $l_1''(x) < 0$ on $[0, \alpha]$, so l_1 is strictly concave. Since $l_1(x) = l_1(\alpha - x)$, its maximum is attained at $x = \alpha/2$. Thus, we get $l_1(x) \leq l_1(\frac{\alpha}{2})$, which implies

$$l(\theta_{11}) + l(\theta_{12}) \leq l(\theta_{11}) + l(\alpha - \theta_{11}) \leq 2l(\frac{\alpha}{2}).$$

If equality holds, we have $\theta_{11} = \frac{\alpha}{2}$ and $\theta_{11} + \theta_{12} = \alpha$, which is equivalent to $\theta_{11} = \theta_{12} = \frac{\alpha}{2}$.

Case 2: One of θ_{11}, θ_{12} lies in $[\alpha, \pi]$ and the other lies in $[0, \alpha]$. By symmetry, assume $\theta_{11} \in [\alpha, \pi]$ and $\theta_{12} \in [0, \alpha]$.

In this case, $\theta_{11} - \theta_{12} \leq \alpha$. And thus $l(\theta_{11}) + l(\theta_{12}) \leq l(\theta_{11}) + l(\theta_{11} - \alpha) \leq l(0) + l(\alpha) \leq 2l(\frac{\alpha}{2})$.

Case 3: $\theta_{11}, \theta_{12} \in [\alpha, \pi]$.

We have $l(\theta_{11}) + l(\theta_{12}) \leq 2l(\alpha) \leq 2l(\frac{\alpha}{2})$. ■

Lemma 11

$$\begin{aligned} u \neq 0, \frac{\partial}{\partial u}\mathbb{E}_{x\sim\mathcal{N}(0,I_d)}[\phi(\langle u, x\rangle)\phi(\langle v, x\rangle)] &= \mathbb{E}_{x\sim\mathcal{N}(0,I_d)}[1_{u^\top x \geq 0} \cdot \phi(\langle v, x\rangle) \cdot x] \\ &= \frac{1}{2\pi}\|v\|_2 \bar{u}\sin(\theta_{u,v}) + \frac{1}{2\pi}(\pi - \theta_{u,v})v. \end{aligned}$$

By above two lemmas, we can get

$$\begin{aligned} L(w_1, w_2) &= \frac{1}{2}\mathbb{E}_{x\sim\mathcal{N}(0,I_d)}\left[\phi(\langle w_1, x\rangle) + \phi(\langle w_2, x\rangle) - \phi(\langle v_1, x\rangle) - \phi(\langle v_2, x\rangle)^2\right] \\ &= \frac{1}{2}\Bigg(\frac{1}{2}(\|w_1\|_2^2 + \|w_2\|_2^2 + 1 + 1) + 2\|w_1\|_2\|w_2\|_2 l(\gamma) \\ &\qquad - 2\sum_{i=1}^{2}\|w_i\|_2\sum_{j=1}^{2}l(\theta_{ij}) + 2l(\alpha)\Bigg) \\ &= \frac{1}{4}\left(\|w_1\|_2^2 + \|w_2\|_2^2\right) + \|w_1\|_2\|w_2\|_2 l(\gamma) \\ &\qquad - \sum_{i=1}^{2}\|w_i\|_2\sum_{j=1}^{2}l(\theta_{ij}) + l(\alpha) + \frac{1}{2}. \end{aligned}$$

A.2. Closed-form expressions for gradient flow

We will use the gradient flow $\frac{\mathrm{d}w_i}{\mathrm{d}t} = -\frac{\mathrm{d}L}{\mathrm{d}w_i}$ to analyze the trajctory. We can compute the related gradients as follows

$$\begin{aligned}\frac{\mathrm{d}w_1}{\mathrm{d}t} &= \frac{1}{2\pi}\Big((\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2 + (\sin\theta_{11}+\sin\theta_{12})\bar{w}_1\Big)\\ &\quad - \left(\frac{1}{2} + \frac{\sin\gamma}{2\pi}\cdot\frac{\|w_2\|_2}{\|w_1\|_2}\right)w_1 - \frac{\pi-\gamma}{2\pi}w_2, \qquad (5)\\ \frac{\mathrm{d}w_2}{\mathrm{d}t} &= \frac{1}{2\pi}\Big((\pi-\theta_{21})v_1 + (\pi-\theta_{22})v_2 + (\sin\theta_{21}+\sin\theta_{22})\bar{w}_2\Big)\\ &\quad - \frac{\pi-\gamma}{2\pi}w_1 - \left(\frac{1}{2} + \frac{\sin\gamma}{2\pi}\cdot\frac{\|w_1\|_2}{\|w_2\|_2}\right)w_2. \qquad (6)\end{aligned}$$

Lemma 12 *For the gradient flow of the squared loss, we have calculated the related terms as follows:*

$$\frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t} = \frac{1}{2\pi\|w_1\|_2}(I - \bar{w}_1\bar{w}_1^\top)\Big((\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2 - (\pi-\gamma)w_2\Big), \qquad (7)$$

$$\frac{\mathrm{d}\bar{w}_2}{\mathrm{d}t} = \frac{1}{2\pi\|w_2\|_2}(I - \bar{w}_2\bar{w}_2^\top)\Big((\pi-\theta_{21})v_1 + (\pi-\theta_{22})v_2 - (\pi-\gamma)w_1\Big), \qquad (8)$$

$$\begin{aligned}\frac{\mathrm{d}\|w_1\|_2}{\mathrm{d}t} &= \frac{1}{2\pi}\Big((\sin\theta_{11}+\sin\theta_{12}) + (\pi-\theta_{11})\cos\theta_{11} + (\pi-\theta_{12})\cos\theta_{12}\Big)\\ &\quad - \frac{1}{2}\|w_1\|_2 - \frac{\sin\gamma + (\pi-\gamma)\cos\gamma}{2\pi}\|w_2\|_2, \qquad (9)\\ \frac{\mathrm{d}\|w_2\|_2}{\mathrm{d}t} &= \frac{1}{2\pi}\Big((\sin\theta_{21}+\sin\theta_{22}) + (\pi-\theta_{21})\cos\theta_{21} + (\pi-\theta_{22})\cos\theta_{22}\Big)\\ &\quad - \frac{\sin\gamma + (\pi-\gamma)\cos\gamma}{2\pi}\|w_1\|_2 - \frac{1}{2}\|w_2\|_2. \qquad (10)\end{aligned}$$

And the related derivative of the angles are given by

$$\begin{aligned}\frac{\mathrm{d}\theta_{11}}{\mathrm{d}t} &= \left\langle \frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t}, v_1 \right\rangle\\ &= \frac{1}{2\pi\|w_1\|_2}\cdot\Big((\pi-\theta_{11}) + (\pi-\theta_{12})\cos\alpha - (\pi-\gamma)\|w_2\|_2\cos\theta_{21}\\ &\qquad - (\pi-\theta_{11})\cos^2\theta_{11} - (\pi-\theta_{12})\cos\theta_{12}\cos\theta_{11}\\ &\qquad + (\pi-\gamma)\|w_2\|_2\cos\gamma\cos\theta_{11}\Big)\\ &= \frac{1}{2\pi\|w_1\|_2}\cdot\Big((\pi-\theta_{11})\sin^2\theta_{11} + (\pi-\theta_{12})(\cos\alpha - \cos\theta_{12}\cos\theta_{11}) \qquad (11)\\ &\qquad - (\pi-\gamma)\|w_2\|_2(\cos\theta_{21} - \cos\gamma\cos\theta_{11})\Big), \qquad (12)\end{aligned}$$

$$\frac{\mathrm{d}\theta_{12}}{\mathrm{d}t} = \left\langle \frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t}, v_2 \right\rangle$$

$$
\begin{aligned}
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11}) \cos\alpha + (\pi - \theta_{12}) - (\pi - \gamma)\|w_2\|_2 \cos\theta_{22} \\
&\qquad - (\pi - \theta_{11}) \cos\theta_{11} \cos\theta_{12} - (\pi - \theta_{12}) \cos^2\theta_{12} \\
&\qquad + (\pi - \gamma)\|w_2\|_2 \cos\gamma \cos\theta_{12} \Big) \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11})(\cos\alpha - \cos\theta_{11} \cos\theta_{12}) + (\pi - \theta_{12}) \sin^2\theta_{12} \qquad (13) \\
&\qquad - (\pi - \gamma)\|w_2\|_2 (\cos\theta_{22} - \cos\gamma \cos\theta_{12}) \Big), \qquad (14)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\theta_{21}}{\mathrm{d}t} &= \left\langle \frac{\mathrm{d}\bar{w}_2}{\mathrm{d}t}, v_1 \right\rangle \\
&= \frac{1}{2\pi\|w_2\|_2} \cdot \Big((\pi - \theta_{21}) + (\pi - \theta_{22}) \cos\alpha - (\pi - \gamma)\|w_1\|_2 \cos\theta_{11} \\
&\qquad - (\pi - \theta_{21}) \cos^2\theta_{21} - (\pi - \theta_{22}) \cos\theta_{22} \cos\theta_{21} \\
&\qquad + (\pi - \gamma)\|w_1\|_2 \cos\gamma \cos\theta_{21} \Big) \\
&= \frac{1}{2\pi\|w_2\|_2} \cdot \Big((\pi - \theta_{21}) \sin^2\theta_{21} + (\pi - \theta_{22})(\cos\alpha - \cos\theta_{22} \cos\theta_{21}) \qquad (15) \\
&\qquad - (\pi - \gamma)\|w_1\|_2 (\cos\theta_{11} - \cos\gamma \cos\theta_{21}) \Big), \qquad (16)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\theta_{22}}{\mathrm{d}t} &= \left\langle \frac{\mathrm{d}\bar{w}_2}{\mathrm{d}t}, v_2 \right\rangle \\
&= \frac{1}{2\pi\|w_2\|_2} \cdot \Big((\pi - \theta_{21}) \cos\alpha + (\pi - \theta_{22}) - (\pi - \gamma)\|w_1\|_2 \cos\theta_{12} \\
&\qquad - (\pi - \theta_{21}) \cos\theta_{21} \cos\theta_{22} - (\pi - \theta_{22}) \cos^2\theta_{22} \\
&\qquad + (\pi - \gamma)\|w_1\|_2 \cos\gamma \cos\theta_{22} \Big) \\
&= \frac{1}{2\pi\|w_2\|_2} \cdot \Big((\pi - \theta_{21})(\cos\alpha - \cos\theta_{21} \cos\theta_{22}) + (\pi - \theta_{22}) \sin^2\theta_{22} \qquad (17) \\
&\qquad - (\pi - \gamma)\|w_1\|_2 (\cos\theta_{12} - \cos\gamma \cos\theta_{22}) \Big). \qquad (18)
\end{aligned}
$$

The related derivative of the direction of the weights are given by

$$
\begin{aligned}
\left\langle \frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t}, \bar{w}_2 \right\rangle &= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11}) \cos\theta_{21} + (\pi - \theta_{12}) \cos\theta_{22} - (\pi - \gamma)\|w_2\|_2 \\
&\qquad - (\pi - \theta_{11}) \cos\theta_{11} \cos\gamma - (\pi - \theta_{12}) \cos\theta_{12} \cos\gamma \\
&\qquad + (\pi - \gamma)\|w_2\|_2 \cos^2\gamma \Big) \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11})(\cos\theta_{21} - \cos\theta_{11} \cos\gamma) \\
&\qquad + (\pi - \theta_{12})(\cos\theta_{22} - \cos\theta_{12} \cos\gamma)
\end{aligned}
$$

$$- (\pi - \gamma)\|w_2\|_2 \sin^2\gamma\Big) \tag{19}$$

Hermite polynomials We denote the normalized probabilist's Hermite polynomials by $\{h_k(\cdot)\}_{k\geq 0}$, where each

$$h_k(x) \coloneqq \frac{(-1)^k}{\sqrt{k!}} \cdot e^{x^2/2} \cdot \frac{\mathrm{d}^k}{\mathrm{d}x^k} e^{-x^2/2}.$$

For example,

$$h_0(x) = 1, \quad h_1(x) = x, \quad h_2(x) = \frac{x^2-1}{\sqrt{2}}, \quad h_3(x) = \frac{x^3-3x}{\sqrt{3!}}.$$

A.3. Hermite-leap convention

Let $Z = (Z_1, \dots, Z_p) \sim \mathcal{N}(0, I_p)$ and let $g \in L^2(\mathcal{N}(0, I_p))$ have normalized Hermite expansion

$$g(Z) = \sum_{\beta \in \mathbb{N}_0^p} \widehat{g}_\beta h_\beta(Z), \qquad h_\beta(Z) \coloneqq \prod_{j=1}^{p} h_{\beta_j}(Z_j).$$

For a set of already revealed coordinates $A \subseteq [p]$, the leap out of A is

$$\kappa_g(A) \coloneqq \inf\left\{ \sum_{j\notin A} \beta_j : \widehat{g}_\beta \neq 0, \quad \sum_{j \notin A} \beta_j > 0 \right\}.$$

Thus $\kappa_g(A)$ is the smallest Hermite degree involving at least one coordinate outside A. A function is leap-k if, along some ordering of its relevant coordinates, the largest such leap is k.

A.4. Proof of proposition 1

Proof [Proof of proposition 1] Let

$$u \coloneqq \bar{v} = \frac{v_1 + v_2}{\|v_1 + v_2\|_2}, \qquad w \coloneqq \frac{v_1 - v_2}{\|v_1 - v_2\|_2}, \qquad a \coloneqq \cos\frac{\alpha}{2}, \qquad b \coloneqq \sin\frac{\alpha}{2}.$$

Since $\|v_1\|_2 = \|v_2\|_2 = 1$, the vectors u and w are orthonormal and

$$v_1 = au + bw, \qquad v_2 = au - bw.$$

Define

$$U \coloneqq \langle u, X\rangle, \qquad W \coloneqq \langle w, X\rangle, \qquad Z \coloneqq X - Uu - Ww.$$

Because $X \sim \mathcal{N}(0, I_d)$ and u, w are orthonormal, (U, W) is a pair of independent standard Gaussians, $Z \perp\!\!\!\perp (U, W)$, and

$$f^*(X) = g(U, W) \coloneqq \phi(aU + bW) + \phi(aU - bW).$$

Let

$$\widehat{\phi}_k \coloneqq \mathbb{E}_{G\sim\mathcal{N}(0,1)}[\phi(G)h_k(G)].$$

For $G_\pm = aU \pm bW$, the Gaussian-Hermite identity gives

$$\mathbb{E}[\phi(G_\pm)h_m(U)h_n(W)] = \widehat{\phi}_{m+n}\sqrt{\frac{(m+n)!}{m!\,n!}}\,a^m(\pm b)^n.$$

Therefore, if

$$g(U,W) = \sum_{m,n\geq 0} \widehat{g}_{m,n} h_m(U)h_n(W),$$

then

$$\widehat{g}_{m,n} = \widehat{\phi}_{m+n}\sqrt{\frac{(m+n)!}{m!\,n!}}\,a^m b^n\big(1+(-1)^n\big).$$

In particular, $\widehat{g}_{m,1} = 0$ for every $m \geq 0$.

It remains to check that the claimed nonzero coefficients are indeed nonzero. Since $h_1(t) = t$,

$$\widehat{\phi}_1 = \mathbb{E}[G^2 \mathbb{1}_{\{G>0\}}] = \frac{1}{2},$$

and since $h_2(t) = (t^2-1)/\sqrt{2}$,

$$\widehat{\phi}_2 = \frac{1}{\sqrt{2}}\mathbb{E}[(G^3 - G)\mathbb{1}_{\{G>0\}}] = \frac{1}{2\sqrt{\pi}}.$$

Thus

$$\widehat{g}_{1,0} = 2\widehat{\phi}_1 a = a \neq 0, \qquad \widehat{g}_{0,2} = 2\widehat{\phi}_2 b^2 = \frac{b^2}{\sqrt{\pi}} \neq 0,$$

where $a > 0$ and $b > 0$ because $\alpha \in (0, \pi/2]$.

Finally, $f^*(X)$ is measurable with respect to (U,W), so no Hermite coefficient involving a direction outside $\mathrm{span}\{u,w\}$ can be nonzero. The only nonzero first-order Hermite signal is along $u = \bar{v}$, while all degree-one terms in the $w = v_\Delta$ coordinate vanish and a degree-two term in w is nonzero. Hence the first learned subspace in the Hermite-leap sense is $\mathrm{span}\{\bar{v}\}$, and the leap from $\mathrm{span}\{\bar{v}\}$ to $\mathrm{span}\{\bar{v}, v_\Delta\}$ is 2. Therefore the teacher is a leap-2 Gaussian multi-index function. ■

Remark 13 *The symmetry argument above uses the equal coefficients in the teacher. If instead*

$$Y = c_1\phi(\langle \hat{v}_1, X\rangle) + c_2\phi(\langle \hat{v}_2, X\rangle)$$

with $c_1 \neq c_2$, then after the same (u,w) rotation we obtain

$$Y = c_1\phi(aU + bW) + c_2\phi(aU - bW),$$

and

$$\mathbb{E}[WY] = \frac{b}{2}(c_1 - c_2) \neq 0.$$

So the splitting direction is already visible at order 1, and the leap-2 hierarchy above is special to the symmetric teacher studied in this paper.

A.5. Other useful lemmas

Lemma 14 *Recall that* $n_{v,w} := \bar{v} - \cos\angle(v,w)\bar{w}, \bar{n}_{v,w} = \frac{n_{w,v}}{\|n_{w,v}\|_2}.$ *The Hessian of* $L(w_1, w_2)$ *at point* $w(t) = (w_1(t), w_2(t))$ *is given on the main diagonals*

$$\frac{\partial^2 L}{\partial w_i^2} = \frac{1}{2} I_d + h_1(w_i, w_{3-i}) - h_1(w_i, v_1) - h_1(w_i, v_2), \quad i = 1, 2, \tag{20}$$

and on the off-diagonals by

$$\frac{\partial^2 L}{\partial w_i \partial w_{3-i}} = h_2(w_i, w_{3-i}), \quad i = 1, 2, \tag{21}$$

where

$$h_1(w, v) = \frac{\sin\angle(w,v)\|v\|_2}{2\pi\|w\|_2}\left(I - \bar{w}\bar{w}^\top + \bar{n}_{v,w}\bar{n}_{v,w}^\top\right),$$

and

$$h_2(w, v) = \frac{1}{2\pi}\left(\big(\pi - \angle(w,v)\big)I_d + \bar{n}_{v,w}\bar{w}^\top + \bar{n}_{w,v}\bar{v}^\top\right).$$

Proof details are in Theorem 5 of [33].

We define $t_4 = \sup_{t \geq t_3}\{t : |w(t_4) - w_{\text{saddle}}| = \varepsilon_{\text{escape saddle point}}\}$.

Lemma 15 *There exists* $c > 0$ *such that for any finite time* $t > t_4$, *we have* $\|w_i(t)\|_2 \geq c, i = 1, 2.$

Proof First we prove that $L\big(w_1(t_4), w_2(t_4)\big)$ is strictly less than $L(w_s, w_s)$. Notice that

$$\begin{aligned} L\big(w_1(t_4), w_2(t_4)\big) - L(w_s, w_s) = &\big(w(t_4) - w_{\text{saddle}}\big)^\top H(w_{\text{saddle}})\big(w(t_4) - w_{\text{saddle}}\big) \\ &+ o(\|w(t_4) - w_{\text{saddle}}\|_2^2), \end{aligned}$$

and thus for small enough $\epsilon_{\text{escape saddle point}}$, if $L(w_1(t_4), w_2(t_4) > L(w_s, w_s)$ we have

$$\big(w(t_4) - w_{\text{saddle}}\big)^\top H(w_{\text{saddle}})\big(w(t_4) - w_{\text{saddle}}\big) > 0.$$

However, on the other hand, for small enough $\epsilon_{\text{escape saddle point}}$, by (201),

$$\begin{aligned} \frac{\mathrm{d}\|w - w_s\|_2^2}{\mathrm{d}t} &= (w - w_s)^\top \frac{\mathrm{d}w}{\mathrm{d}t} \\ &= (w - w_{\text{saddle}})^\top\big(-H(w_{\text{saddle}})\big)(w - w_{\text{saddle}})^\top + o(\|w - w_{\text{saddle}}\|_2^2), \end{aligned}$$

we have $\frac{\mathrm{d}\|w - w_{\text{saddle}}\|_2^2}{\mathrm{d}t} < 0$. And hence w is still approaching to w_{saddle}, instead of escaping, which contradicts with our assumption that $t_4 = \sup_{t \geq t_3}\{t : \|w(t) - w_{\text{saddle}}\| = \varepsilon_{\text{escape saddle point}}\}$. In the next steps we assume $L(w_s, w_s) - L\big(w_1(t_4), w_2(t_4)\big) > \varepsilon'$, for some $\epsilon' = \epsilon'(\epsilon_{\text{escape saddle point}}) > 0$.

By the uniform continuity of loss in w_1, there exists $\delta > 0$ such that for any $\|w_1\|_2 < \delta$,$|L(0, w_2) - L(w_1, w_2)| < \frac{\varepsilon'}{2}$. We have

$$\begin{aligned}
L(w_s, w_s) &= \frac{1}{4}(r_s^2 + r_s^2) + r_s^2 \cdot l(0) - 2l(\frac{\alpha}{2}) \cdot 2r_s + l(\alpha) + \frac{1}{2} \\
&= l(\alpha) + \frac{1}{2} - \frac{1}{\pi^2}\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)^2, \\
L(0, w_2) &= \frac{1}{2}\mathbb{E}_{\mathcal{N}(0,I_d)}\Big(\mathrm{ReLU}(\langle w_2, x\rangle) - \mathrm{ReLU}(\langle v_1, x\rangle) - \mathrm{ReLU}(\langle v_2, x\rangle)\Big)^2 \\
&= \frac{1}{4}\|w_2\|_2^2 + \frac{1}{4} + \frac{1}{4} - \Big(l(\theta_{21}) + l(\theta_{22})\Big)\|w_2\|_2 + l(\alpha) \\
&\overset{(*)}{\geq} \frac{1}{4}\|w_2\|_2^2 - 2l(\frac{\alpha}{2})\|w_2\|_2 + l(\alpha) + \frac{1}{2}, \\
&= \Big(\frac{1}{2}\|w_2\|_2 - \frac{1}{\pi}\big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\big)\Big)^2 \\
&\quad + l(\alpha) + \frac{1}{2} - \frac{1}{\pi^2}\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)^2 \\
&\geq l(\alpha) + \frac{1}{2} - \frac{1}{\pi^2}\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)^2 \\
&= L(w_s, w_s), \qquad (22)
\end{aligned}$$

where the $(*)$ comes from lemma 10. If there exists $t > t_4$ such that $\|w_1(t)\|_2 < \delta$, then we have

$$L(w_1(t), w_2(t)) > L\big(0, w_2(t)\big) - \frac{\varepsilon'}{2} \geq L(w_s, w_s) - \frac{\varepsilon'}{2}, \qquad (23)$$

which contradicts with $L(w_s, w_s) - L\big(w_1(t), w_2(t)\big) \geq L(w_s, w_s) - L\big(w_1(t_4), w_2(t_4)\big) > \varepsilon'$. Then, we have $\|w_1(t)\|_2 \geq \delta$. Similarly, we have $\|w_2(t)\|_2 \geq \delta, \forall t > t_4$. Let $c = \delta$ and we got the lower bound of weights' norm. ■

Lemma 16 *Among all the saddle points in the plane spanned by* $\{v_1, v_2\}$, *the loss at* $w_{\mathrm{saddle}} = (w_s, w_s)$ *is the smallest.*

Proof Case 1: $\bar{w}_1 = \bar{w}_2 = \bar{v}$, $\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)$.

Let $w_s = \frac{1}{\pi}\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)\bar{v} =: r_s\bar{v}$, then we have

$$\begin{aligned}
L &= L(w_s, w_s) \\
&= \frac{1}{4}\Big(r_s^2 + r_s^2\Big) + r_s^2 \cdot l(0) - 2l(\frac{\alpha}{2}) \cdot 2r_s + l(\alpha) + \frac{1}{2} \\
&= -r_s^2 + l(\alpha) + \frac{1}{2}.
\end{aligned}$$

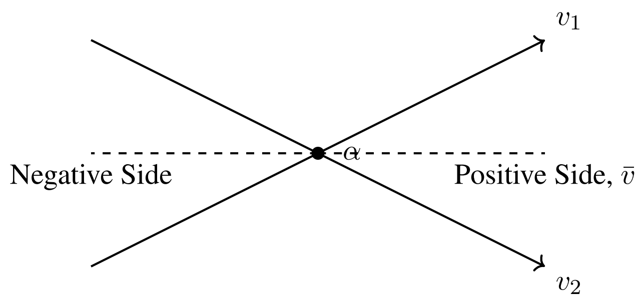

Figure 3: The positive and negative sides of the angular bisector.

Case 2: $\bar{w}_1 = \bar{w}_2 = -\bar{v}$, $\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$.

Let $w'_s = -\frac{1}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right)\bar{v} =: -r'_s\bar{v}$, then we have

$$\begin{aligned} L &= L(w'_s, w'_s) \\ &= \frac{1}{4}\left(r'^2_s + r'^2_s\right) + r'^2_s \cdot l(0) - 2l(\pi - \frac{\alpha}{2}) \cdot 2r'_s + l(\alpha) + \frac{1}{2} \\ &= -r'^2_s + l(\alpha) + \frac{1}{2}. \end{aligned}$$

Case 3: $\bar{w}_1 = \bar{v}$, $\bar{w}_2 = -\bar{v}$, $\|w_1\|_2 = \frac{2}{\pi}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right)$, $\|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$ and its permutation.

We have

$$\begin{aligned} L &= L(2w_s, -2w'_s) \\ &= \frac{1}{4}(4r^2_s + 4r'^2_s) - 2r_s l(\frac{\alpha}{2}) - 2r'_s l(\pi - \frac{\alpha}{2}) + l(\alpha) + \frac{1}{2} \\ &= l(\alpha) + \frac{1}{2}. \end{aligned}$$

To sum up, since $r_s > r'_s$, it's very easy to see lemma 16 holds. ∎

Lemma 17 *If* $\min\{\theta_{11} + \theta_{12}, \theta_{21} + \theta_{22}\} > \pi$*, then* $L(w_1, w_2) > L(w_s, w_s)$*, where we use the notation* $w_s = \frac{1}{\pi}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right)\bar{v}$.

Proof We have that

$$L(w_1, w_2) = \frac{1}{4}\left(\|w_1\|^2_2 + \|w_2\|^2_2\right) + \|w_1\|_2\|w_2\|_2 l(\gamma)$$

$$
\begin{aligned}
&- \sum_{i=1}^{2} \|w_i\|_2 \sum_{j=1}^{2} l(\theta_{ij}) + l(\alpha) + \frac{1}{2} \\
\geq\ & \frac{1}{4}\big(\|w_1\|_2^2 + \|w_2\|_2^2\big) - \big(l(\theta_{11}) + l(\theta_{12})\big)\|w_1\|_2 \\
&- \big(l(\theta_{21}) + l(\theta_{22})\big)\|w_2\|_2 + l(\alpha) + \frac{1}{2} \\
\geq\ & \frac{1}{4}\big(\|w_1\|_2^2 + \|w_2\|_2^2\big) - \big(l(\theta_{11}) + l(\pi - \theta_{11})\big)\|w_1\|_2 \\
&- \big(l(\theta_{21}) + l(\pi - \theta_{21})\big)\|w_2\|_2 + l(\alpha) + \frac{1}{2} \\
\geq\ & \frac{1}{4}\big(\|w_1\|_2^2 + \|w_2\|_2^2\big) - \frac{1}{2}(\|w_1\|_2 + \|w_2\|_2) + l(\alpha) + \frac{1}{2} \\
\geq\ & -\frac{1}{2} + l(\alpha) + \frac{1}{2}.
\end{aligned}
$$

Since $r_s = 2l(\frac{\alpha}{2}) \geq 2l(\frac{\pi}{4}) = \frac{\frac{3\pi}{4}+1}{\sqrt{2}\pi} > \frac{1}{\sqrt{2}}$, and thus

$$
L(w_1, w_2) > -r_s^2 + l(\alpha) + \frac{1}{2} = L(w_s, w_s).
$$

■

Appendix B. Proofs for landscape

We use the angle shorthand

$$
\theta_{ij} = \angle(w_i, v_j), \qquad \gamma = \angle(w_1, w_2). \tag{24}
$$

B.1. Proofs for correlation loss

B.1.1. PROOF OF THEOREM 2

Proof Since the correlation loss is separable across student neurons, it is enough to analyze the single-neuron objective. Accordingly, throughout the proof we write

$$
L_{\text{corr}}(w) := -\mathbb{E}_{x\sim\mathcal{N}(0,I_d)}\big[f^*(x)\phi(\langle w, x\rangle)\big].
$$

Let $c = \cos\alpha$, and define

$$
l(\theta) := \frac{(\pi - \theta)\cos\theta + \sin\theta}{2\pi}, \qquad \Phi(w) = \sum_{i=1}^{k} l(\angle(w, v_i)).
$$

Then the closed-form ReLU correlation gives $L_{\text{corr}}(w) = -\Phi(w)$, so it is enough to prove that Φ is maximized at $\bar{v}$. Define $g(t) = l(\arccos t)$ for $t \in [-1, 1]$. Since

$$
l(\theta) = \frac{(\pi - \theta)\cos\theta + \sin\theta}{2\pi},
$$

we have

$$g(t) = \frac{t(\pi - \arccos t) + \sqrt{1-t^2}}{2\pi}.$$

A direct computation gives

$$g'(t) = \frac{\pi - \arccos t}{2\pi}, \qquad g''(t) = \frac{1}{2\pi\sqrt{1-t^2}} > 0 \quad (t \in (-1,1)). \tag{25}$$

Thus $\Phi(w) = \sum_{i=1}^k g(w \cdot v_i)$.

First compute $\bar{v} \cdot v_i$. Since $v_i \cdot v_j = c$ for $i \neq j$,

$$\left\| \sum_{i=1}^k v_i \right\|_2^2 = k + k(k-1)c = k(1+(k-1)c), \qquad v_i \cdot \sum_{j=1}^k v_j = 1 + (k-1)c.$$

Therefore $\bar{v} \cdot v_i = a$, where $a := \sqrt{(1+(k-1)c)/k}$. Since $\alpha \in (0, \pi/2]$, we have $c \in [0,1)$, hence $1/k \le a^2 < 1$ and $0 < a < 1$. Let $\gamma := \arccos a \in (0, \pi/2)$. Then $\gamma = \angle(\bar{v}, v_i)$ for every i, and $\Phi(\bar{v}) = \sum_{i=1}^k g(\bar{v} \cdot v_i) = kg(a)$.

We first show that a maximizer may be assumed to lie in $\operatorname{span}\{v_1, \dots, v_k\}$. Let w_* be a maximizer of Φ on the unit sphere. By the Lagrange multiplier condition,

$$\nabla\Phi(w_*) = \lambda w_*.$$

Since $\nabla\Phi(w) = \sum_{i=1}^k g'(w \cdot v_i) v_i$, we have

$$\sum_{i=1}^k g'(w_* \cdot v_i) v_i = \lambda w_*.$$

The left-hand side lies in $\operatorname{span}\{v_i\}$. We claim that $\lambda \neq 0$. If $\lambda = 0$, then $\sum_{i=1}^k g'(w_* \cdot v_i) v_i = 0$. Taking the inner product with $\sum_{j=1}^k v_j$, we obtain

$$0 = \sum_{i=1}^k g'(w_* \cdot v_i) \left(v_i \cdot \sum_{j=1}^k v_j \right) = (1+(k-1)c) \sum_{i=1}^k g'(w_* \cdot v_i).$$

Because $1 + (k-1)c > 0$ and $g'(t) = (\pi - \arccos t)/(2\pi) \ge 0$, this forces $g'(w_* \cdot v_i) = 0$ for every i. Since $g'(t) = 0$ if and only if $t = -1$, we get $w_* \cdot v_i = -1$ for all i. As w_* and v_i are unit vectors, this means $w_* = -v_i$ for all i, forcing all v_i to be equal, contradicting $\alpha > 0$. Therefore $\lambda \neq 0$, and hence $w_* \in \operatorname{span}\{v_i\}$. It is therefore enough to prove the desired inequality for $w \in \operatorname{span}\{v_i\}$.

We split the proof into two cases.

Case 1: $a \le 1/\sqrt{2}$**.**

Since $a = \cos\gamma$ and $\gamma \in (0, \pi/2)$, the assumption $a \le 1/\sqrt{2}$ is equivalent to $\gamma \in [\pi/4, \pi/2)$.

We prove the scalar inequality

$$g(y) \le g(a) + \frac{g'(a)}{2a}(y^2 - a^2), \qquad y \in [-1,1]. \tag{26}$$

Write $y = \cos\theta$, where $\theta \in [0, \pi]$, and recall that $a = \cos\gamma$. The desired inequality is equivalent to

$$l(\theta) \le l(\gamma) + \frac{\pi - \gamma}{4\pi\cos\gamma}\big(\cos^2\theta - \cos^2\gamma\big).$$

For fixed γ, define

$$R_\gamma(\theta) = l(\gamma) + \frac{\pi - \gamma}{4\pi\cos\gamma}(\cos^2\theta - \cos^2\gamma) - l(\theta).$$

Clearly $R_\gamma(\gamma) = 0$. Since $l'(\theta) = -(\pi - \theta)\sin\theta/(2\pi)$, we get

$$R'_\gamma(\theta) = \frac{\sin\theta}{2\pi}\left[(\pi - \theta) - \frac{\pi - \gamma}{\cos\gamma}\cos\theta\right]. \tag{27}$$

For this fixed γ, write $H_\gamma(\theta) = (\pi - \theta) - \frac{\pi-\gamma}{\cos\gamma}\cos\theta$. Then (27) says $R'_\gamma(\theta) = \sin\theta\, H_\gamma(\theta)/(2\pi)$.

We now determine the sign of H_γ. First, $H_\gamma(\gamma) = 0$, and $H_\gamma(0) = \pi - \frac{\pi-\gamma}{\cos\gamma}$. The inequality $H_\gamma(0) < 0$ is equivalent to $\pi\cos\gamma < \pi - \gamma$. On $\gamma \in [\pi/4, \pi/2)$, the function $\gamma \mapsto \pi - \gamma - \pi\cos\gamma$ satisfies

$$\pi - \frac{\pi}{4} - \pi\cos\frac{\pi}{4} = \frac{3\pi}{4} - \frac{\pi}{\sqrt{2}} > 0,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\big(\pi - \gamma - \pi\cos\gamma\big) = -1 + \pi\sin\gamma \ge -1 + \frac{\pi}{\sqrt{2}} > 0.$$

Thus $\pi - \gamma - \pi\cos\gamma > 0$, proving $H_\gamma(0) < 0$.

Next, $H'_\gamma(\theta) = -1 + \frac{\pi-\gamma}{\cos\gamma}\sin\theta$ and $H''_\gamma(\theta) = \frac{\pi-\gamma}{\cos\gamma}\cos\theta$. For $\theta \in [0, \gamma]$, we have $\cos\theta \ge 0$, because $\gamma < \pi/2$. Hence $H''_\gamma(\theta) \ge 0$ on $[0, \gamma]$, so H_γ is convex there. Since $H_\gamma(0) < 0$ and $H_\gamma(\gamma) = 0$, convexity gives $H_\gamma(\theta) < 0$ for $0 < \theta < \gamma$.

Now consider $\theta \in (\gamma, \pi/2]$. We show $H_\gamma(\theta) > 0$. At $\theta = \gamma$,

$$H'_\gamma(\gamma) = -1 + (\pi - \gamma)\tan\gamma.$$

For $\gamma \ge \pi/4$,

$$\left(\pi - \frac{\pi}{4}\right)\tan\frac{\pi}{4} = \frac{3\pi}{4} > 1.$$

Also,

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\big((\pi - \gamma)\tan\gamma\big) = -\tan\gamma + (\pi - \gamma)\sec^2\gamma = \frac{(\pi - \gamma) - \sin\gamma\cos\gamma}{\cos^2\gamma} > 0,$$

because $\pi - \gamma > \pi/2$ and $\sin\gamma\cos\gamma \le 1/2$. Thus $(\pi - \gamma)\tan\gamma > 1$, so $H'_\gamma(\gamma) > 0$. Since $H''_\gamma(\theta) \ge 0$ on $[\gamma, \pi/2]$, it follows that $H'_\gamma(\theta) > 0$ there, and hence

$$H_\gamma(\theta) > H_\gamma(\gamma) = 0, \qquad \gamma < \theta \le \frac{\pi}{2}.$$

Finally, if $\theta \in [\pi/2, \pi]$, then $\cos\theta \le 0$, so $-\frac{\pi-\gamma}{\cos\gamma}\cos\theta \ge 0$, and also $\pi - \theta \ge 0$. Therefore $H_\gamma(\theta) \ge 0$ on $[\pi/2, \pi]$.

Combining these sign results, we have

$$H_\gamma(\theta) < 0 \quad (0 < \theta < \gamma), \qquad H_\gamma(\theta) > 0 \quad (\gamma < \theta < \pi). \tag{28}$$

Since $R'_\gamma(\theta) = \sin\theta\, H_\gamma(\theta)/(2\pi)$, $\sin\theta \geq 0$ on $[0, \pi]$, and the sign of H_γ is given by (28), the function R_γ decreases on $(0, \gamma)$ and increases on (γ, π). Hence R_γ attains its global minimum at $\theta = \gamma$, where $R_\gamma(\gamma) = 0$. Thus $R_\gamma(\theta) \geq 0$, which proves (26).

Now let $y_i = w \cdot v_i$. Summing (26) yields

$$\Phi(w) = \sum_i g(y_i) \leq kg(a) + \frac{g'(a)}{2a}\left(\sum_i y_i^2 - ka^2\right). \tag{29}$$

It remains to show $\sum_i y_i^2 \leq ka^2$. Indeed,

$$\sum_i y_i^2 = w^\top \left(\sum_i v_i v_i^\top\right) w.$$

The operator $\sum_i v_i v_i^\top$ has eigenvalue $1 + (k-1)c = ka^2$ in the direction $\bar{v}$, and eigenvalue $1 - c$ on the orthogonal complement of $\bar{v}$ inside $\operatorname{span}\{v_i\}$. Since $c \geq 0$, $ka^2 = 1 + (k-1)c \geq 1 - c$. Thus the largest eigenvalue of $\sum_i v_i v_i^\top$ on $\operatorname{span}\{v_i\}$ is ka^2, and so $\sum_i y_i^2 \leq ka^2$. The summed estimate (29) gives $\Phi(w) \leq kg(a) = \Phi(\bar{v})$.

Moreover, this argument identifies the equality case. The preceding sign analysis of R shows that the scalar inequality is strict unless $y = a$. Let

$$D_i = g(a) + \frac{g'(a)}{2a}(y_i^2 - a^2) - g(y_i) \geq 0.$$

Then

$$kg(a) - \Phi(w) = \frac{g'(a)}{2a}\left(ka^2 - \sum_i y_i^2\right) + \sum_i D_i.$$

Since $a > 0$ and

$$g'(a) = \frac{\pi - \arccos a}{2\pi} > 0,$$

we have $g'(a)/(2a) > 0$. Thus both terms on the right-hand side are nonnegative. Hence equality $\Phi(w) = \Phi(\bar{v})$ forces $D_i = 0$ for every i, and therefore

$$w \cdot v_i = y_i = a = \bar{v} \cdot v_i \qquad \text{for all } i.$$

Since $w, \bar{v} \in \operatorname{span}\{v_i\}$ and the Gram matrix

$$(v_i \cdot v_j)_{i,j} = (1-c)I + c\mathbf{1}\mathbf{1}^\top$$

is positive definite for $c \in [0, 1)$, these identities imply

$$w = \bar{v}.$$

Thus equality in the first case occurs only at $\bar{v}$. This proves the first case.

Case 2: $a \geq 1/\sqrt{2}$**.**

Recall that $\Phi(w) = \sum_{i=1}^{k} l(\theta_i)$ with $\theta_i = \angle(w, v_i)$ and $l(\theta) = ((\pi - \theta)\cos\theta + \sin\theta)/(2\pi)$, equivalently $\Phi(w) = \sum_{i=1}^{k} g(w \cdot v_i)$ with $g(t) = l(\arccos t)$. The global claim is $\Phi(w) \leq \Phi(\bar{v})$ for every unit w.

In this case $a = \cos\gamma$ and $\gamma \in (0, \pi/4]$. Let $b = \sin\gamma$, and define $u_i = (v_i - a\bar{v})/b$. Then

$$v_i = a\bar{v} + bu_i, \qquad u_i \perp \bar{v}, \qquad \|u_i\|_2 = 1, \qquad \sum_i u_i = 0, \qquad u_i \cdot u_j = -\frac{1}{k-1} \quad (i \neq j). \tag{30}$$

Because a maximizer may be assumed to lie in $\operatorname{span}\{v_i\}$, write

$$w = \cos\varphi\,\bar{v} + \sin\varphi\,u, \tag{31}$$

where $\varphi \in [0, \pi]$, $u \perp \bar{v}$, $\|u\|_2 = 1$, and $u \in \operatorname{span}\{u_i\}$. When $\sin\varphi = 0$, the choice of u does not change w, so we choose such a unit u. Set $t_i = u \cdot u_i$.

The identities in (30) imply

$$\sum_i t_i = 0, \qquad \sum_i t_i^2 = \frac{k}{k-1}, \qquad -1 \leq t_i \leq 1. \tag{32}$$

Indeed, $\sum_i t_i = u \cdot \sum_i u_i = 0$, and the regular-simplex frame identity $\sum_i u_i u_i^\top = \frac{k}{k-1} I_{\operatorname{span}\{u_i\}}$ gives $\sum_i t_i^2 = u^\top (\sum_i u_i u_i^\top) u = k/(k-1)$. Moreover, $w \cdot v_i = a\cos\varphi + bt_i \sin\varphi$. For fixed u, define

$$F(\varphi; u) = \sum_{i=1}^{k} g(a\cos\varphi + bt_i \sin\varphi).$$

This equals $\Phi(w)$ for the vector w in (31).

We first handle $\varphi \geq \gamma$. For fixed u, write

$$z_i(\varphi; u) = a\cos\varphi + bt_i \sin\varphi.$$

Then

$$\frac{\partial}{\partial\varphi} F(\varphi; u) = -a\sin\varphi \sum_i g'\big(z_i(\varphi; u)\big) + b\cos\varphi \sum_i t_i g'\big(z_i(\varphi; u)\big). \tag{33}$$

Since g' is increasing by (25), and $z_i(\varphi; u)$ is an increasing affine function of t_i, we have

$$\sum_i t_i g'\big(z_i(\varphi; u)\big) = \frac{1}{k} \sum_{i<j} (t_i - t_j)\left[g'\big(z_i(\varphi; u)\big) - g'\big(z_j(\varphi; u)\big)\right] \geq 0, \tag{34}$$

because $\sum_i t_i = 0$.

If $\varphi \geq \pi/2$, then $\cos\varphi \leq 0$. The first term on the right-hand side of (33) is nonpositive because $g' \geq 0$, and the second term is nonpositive by (34). Hence $\partial_\varphi F(\varphi; u) \leq 0$.

If $\gamma \leq \varphi < \pi/2$, then $(a/b)\tan\varphi \geq (a/b)\tan\gamma = 1$. In this range $\cos\varphi > 0$, and we can rewrite

$$\frac{\partial}{\partial\varphi} F(\varphi; u) = b\cos\varphi \sum_i \left(t_i - \frac{a}{b}\tan\varphi\right) g'\big(z_i(\varphi; u)\big).$$

Since $t_i \le 1 \le (a/b)\tan\varphi$ and $g' \ge 0$, we again have $\partial_\varphi F(\varphi; u) \le 0$. Thus $\varphi \ge \gamma$ implies $\partial_\varphi F(\varphi; u) \le 0$. Consequently, on the interval $[\gamma, \pi]$, the function $F(\cdot; u)$ is nonincreasing, and hence

$$F(\varphi; u) \le F(\gamma; u), \qquad \varphi \ge \gamma.$$

Thus this range cannot give a value larger than the value attained closer to the angle bisector. To prove $F(\varphi; u) \le F(0; u) = \Phi(\bar{v})$ for all φ, it remains to consider $0 \le \varphi \le \gamma$. We prove the next estimate first for $0 \le \varphi < \gamma$; the endpoint $\varphi = \gamma$ then follows by continuity. Since $\gamma \le \pi/4$, we have $\gamma + \varphi \le 2\gamma \le \pi/2$. Thus $a\cos\varphi - b\sin\varphi = \cos(\gamma + \varphi) \ge 0$, and therefore $a\cos\varphi + bt\sin\varphi \ge 0$ for every $t \in [-1, 1]$. For this fixed φ, define

$$h_\varphi(t) = g(a\cos\varphi + bt\sin\varphi).$$

Since $g'''(s) = s/(2\pi(1 - s^2)^{3/2})$, the chain rule gives

$$h_\varphi'''(t) = (b\sin\varphi)^3 g'''\big(a\cos\varphi + bt\sin\varphi\big) = (b\sin\varphi)^3 \frac{1}{2\pi} \frac{a\cos\varphi + bt\sin\varphi}{\big(1 - (a\cos\varphi + bt\sin\varphi)^2\big)^{3/2}}. \tag{35}$$

Here $b\sin\varphi \ge 0$, and the numerator $a\cos\varphi + bt\sin\varphi$ in (35) is nonnegative. Thus $h_\varphi'''(t) \ge 0$ on $t \in [-1, 1]$.

Let $s_0 = -1/(k-1)$. Let $Q_\varphi(t)$ be the quadratic polynomial satisfying

$$Q_\varphi(1) = h_\varphi(1), \qquad Q_\varphi(s_0) = h_\varphi(s_0), \qquad Q_\varphi'(s_0) = h_\varphi'(s_0).$$

Now fix $t \in [-1, 1]$, which is the relevant domain because each $t_i = u \cdot u_i$ lies in $[-1, 1]$. Also $s_0 \in [-1, 0]$, so both interpolation nodes s_0 and 1 lie in the same interval. Applying the Hermite interpolation remainder with a double node at s_0 and a simple node at 1 gives

$$h_\varphi(t) - Q_\varphi(t) = \frac{h_\varphi'''(\xi)}{6}(t - s_0)^2(t - 1), \tag{36}$$

where ξ lies between t, s_0, and 1. Since $t \in [-1, 1]$, we have $h_\varphi'''(\xi) \ge 0$, $(t - s_0)^2 \ge 0$, and $t - 1 \le 0$. Thus the remainder in (36) is nonpositive, so $h_\varphi(t) \le Q_\varphi(t)$ for all $t \in [-1, 1]$, and therefore $\sum_i h_\varphi(t_i) \le \sum_i Q_\varphi(t_i)$.

Since Q_φ is quadratic, (32) and the identities $1 + (k-1)s_0 = 0$, $1 + (k-1)s_0^2 = k/(k-1)$ show that the two collections $(t_1, \dots, t_k)$ and $(1, s_0, \dots, s_0)$ have the same zeroth, first, and second moments. To see the substitution explicitly, write $Q_\varphi(t) = q_0 + q_1 t + q_2 t^2$. Then

$$\begin{aligned} \sum_i Q_\varphi(t_i) &= kq_0 + q_1 \sum_i t_i + q_2 \sum_i t_i^2 \\ &= kq_0 + q_2 \frac{k}{k-1}, \end{aligned}$$

whereas

$$\begin{aligned} Q_\varphi(1) + (k-1)Q_\varphi(s_0) &= q_0 + q_1 + q_2 + (k-1)\big(q_0 + q_1 s_0 + q_2 s_0^2\big) \\ &= kq_0 + q_1\big(1 + (k-1)s_0\big) + q_2\big(1 + (k-1)s_0^2\big) \end{aligned}$$

$$= kq_0 + q_2 \frac{k}{k-1}.$$

Therefore $\sum_i Q_\varphi(t_i) = Q_\varphi(1) + (k-1)Q_\varphi(s_0)$. Using the defining interpolation conditions of Q_φ, this becomes

$$\sum_i h_\varphi(t_i) \le h_\varphi(1) + (k-1)h_\varphi(s_0).$$

Hence $F(\varphi; u) \le J(\varphi)$, where, with a, b, k fixed throughout this case,

$$J(\varphi) = g(a\cos\varphi + b\sin\varphi) + (k-1)g\left(a\cos\varphi - \frac{b}{k-1}\sin\varphi\right). \tag{37}$$

It remains to prove $J(\varphi) \le J(0)$. Since $a = \cos\gamma$ and $b = \sin\gamma$, we have $a\cos\varphi + b\sin\varphi = \cos(\gamma - \varphi)$. Set $z_-(\varphi) = a\cos\varphi - \frac{b}{k-1}\sin\varphi$. Since $0 \le \varphi \le \gamma \le \pi/4$, we have

$$0 \le \cos(\gamma+\varphi) = a\cos\varphi - b\sin\varphi \le z_-(\varphi) \le a\cos\varphi \le a = \cos\gamma \le 1. \tag{38}$$

Thus $z_-(\varphi) \in [0,1]$. Define $\eta(\varphi) = \arccos z_-(\varphi) \in [0,\pi]$, so that $z_-(\varphi) = \cos\eta(\varphi)$. Since the arccosine is decreasing, (38) gives

$$\gamma \le \eta(\varphi) \le \gamma + \varphi. \tag{39}$$

Write $q(\varphi) = \eta(\varphi) - \gamma$. Then (39) gives $0 \le q(\varphi) \le \varphi$. Also,

$$g'\big(\cos(\gamma-\varphi)\big) = \frac{\pi - \gamma + \varphi}{2\pi}, \qquad g'\big(z_-(\varphi)\big) = \frac{\pi - \eta(\varphi)}{2\pi} = \frac{\pi - \gamma - q(\varphi)}{2\pi}.$$

Using these two derivative values, together with

$$\frac{\mathrm{d}}{\mathrm{d}\varphi}\cos(\gamma-\varphi) = b\cos\varphi - a\sin\varphi, \qquad \frac{\mathrm{d}}{\mathrm{d}\varphi}z_-(\varphi) = -a\sin\varphi - \frac{b}{k-1}\cos\varphi,$$

and differentiating (37), we obtain

$$\begin{aligned} J'(\varphi) &= \frac{1}{2\pi}\Big[(\pi-\gamma+\varphi)\big(b\cos\varphi - a\sin\varphi\big) \\ &\qquad + (k-1)\big(\pi-\gamma-q(\varphi)\big)\left(-a\sin\varphi - \frac{b}{k-1}\cos\varphi\right)\Big] \\ &= \frac{1}{2\pi}\Big[-a\sin\varphi\,[k(\pi-\gamma) + \varphi - (k-1)q(\varphi)] \\ &\qquad + b\cos\varphi\big(\varphi + q(\varphi)\big)\Big]. \end{aligned} \tag{40}$$

We want to show that $J'(\varphi) \le 0$ on $0 \le \varphi \le \gamma$; then $J(\varphi) \le J(0)$, which is the remaining desired bound. The expression inside the brackets in (40), viewed as a function of $q(\varphi)$ with φ fixed, has derivative $a(k-1)\sin\varphi + b\cos\varphi > 0$. Since $q(\varphi) \le \varphi$, (40) gives

$$J'(\varphi) \le \frac{1}{2\pi}\left[-a\sin\varphi\,[k(\pi-\gamma) - (k-2)\varphi] + 2b\varphi\cos\varphi\right]. \tag{41}$$

At $\varphi = 0$, the exact formula (40) gives $J'(0) = 0$. For $0 < \varphi \le \gamma$, it suffices to prove

$$2b\varphi\cos\varphi \le a\sin\varphi\,[k(\pi-\gamma) - (k-2)\varphi].$$

Dividing by $a \sin\varphi > 0$, this is equivalent to

$$2\tan\gamma \cdot \varphi\cot\varphi \le k(\pi-\gamma) - (k-2)\varphi. \tag{42}$$

Since $0 < \varphi \le \gamma \le \pi/4$, we have $\varphi\cot\varphi \le 1$. Hence

$$2\tan\gamma \cdot \varphi\cot\varphi \le 2\tan\gamma \le 2.$$

On the other hand,

$$k(\pi-\gamma) - (k-2)\varphi \ge k(\pi-\gamma) - (k-2)\gamma = k(\pi-\gamma) - (k-2)\gamma = k\pi - 2(k-1)\gamma.$$

Since $\gamma \le \pi/4$, we have

$$k\pi - 2(k-1)\gamma \ge k\pi - \frac{(k-1)\pi}{2} = \frac{(k+1)\pi}{2}.$$

As $k \ge 2$, $(k+1)\pi/2 \ge 3\pi/2 > 2$. Therefore

$$2\tan\gamma \cdot \varphi\cot\varphi \le k(\pi-\gamma) - (k-2)\varphi.$$

Thus (42) holds, and (41) gives $J'(\varphi) \le 0$. Consequently $J(\varphi) \le J(0)$. But $J(0) = kg(a) = \Phi(\bar{v})$, hence $F(\varphi; u) \le J(\varphi) \le \Phi(\bar{v})$.

The same estimates also give the equality case. For $0 < \varphi \le \gamma \le \frac{\pi}{4}$, the final bound used above is strict, because

$$2\tan\gamma \cdot \varphi\cot\varphi \le 2 < \frac{(k+1)\pi}{2} \le k(\pi-\gamma) - (k-2)\varphi.$$

Thus $J'(\varphi) < 0$ on $(0, \gamma]$, and hence

$$J(\varphi) < J(0) = \Phi(\bar{v}), \qquad 0 < \varphi \le \gamma.$$

For $\varphi \ge \gamma$, the monotonicity estimate gives

$$F(\varphi) \le F(\gamma) \le J(\gamma) < J(0) = \Phi(\bar{v}).$$

Therefore equality in the second case is possible only when $\varphi = 0$, namely $w = \bar{v}$. This proves the second case.

Combining the two cases $a \le 1/\sqrt{2}$ and $a \ge 1/\sqrt{2}$, we conclude that for every unit vector $w, \Phi(w) \le \Phi(\bar{v})$. If equality holds, then w is a maximizer of Φ. By the reduction at the beginning of the proof, any maximizer lies in $\operatorname{span}\{v_i\}$. The equality discussion in the two cases therefore applies and shows that equality holds if and only if $w = \bar{v}$. Therefore $\Phi(\bar{v}) = \max_{\|w\|=1} \Phi(w)$, and $\bar{v}$ is the only maximizer. Since $L_{\text{corr}}(w) = -\Phi(w)$, this is equivalent to $L_{\text{corr}}(w) \ge L_{\text{corr}}(\bar{v})$ for every unit vector w, with equality if and only if $w = \bar{v}$. Hence $\bar{v}$ is the unique global minimum of the correlation loss on the unit sphere. ∎

B.2. Proofs for squared loss

Lemma 18 (Characterization of stationary points) *For nonzero student weights w_1, w_2, define*

$a_1 = \sin\theta_{11} + \sin\theta_{12} - \pi\|w_1\|_2 - \sin\gamma\|w_2\|_2, \quad a_2 = \sin\theta_{21} + \sin\theta_{22} - \pi\|w_2\|_2 - \sin\gamma\|w_1\|_2.$

Then (w_1, w_2) is a stationary point of L if and only if

$$\begin{aligned}(\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2 + a_1\bar{w}_1 - (\pi-\gamma)\|w_2\|_2\bar{w}_2 = 0,\\ (\pi-\theta_{21})v_1 + (\pi-\theta_{22})v_2 + a_2\bar{w}_2 - (\pi-\gamma)\|w_1\|_2\bar{w}_1 = 0.\end{aligned} \tag{43}$$

B.2.1. PROOF OF LEMMA 18

Proof By Equations (5) and (6), stationarity is equivalent to the right-hand sides of both gradient-flow equations being zero. The first equation becomes

$$(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2+(\sin\theta_{11}+\sin\theta_{12})\bar{w}_1-\pi\|w_1\|_2\bar{w}_1-\sin\gamma\|w_2\|_2\bar{w}_1-(\pi-\gamma)\|w_2\|_2\bar{w}_2 = 0$$

after multiplying by 2π. The second identity follows in the same way from Equation (6). Conversely, these two displayed identities are exactly the two gradient-flow equations after multiplication by 2π, so they are equivalent to stationarity. ■

Proposition 19 (Reduction to three dimensions) *Assume $w_1, w_2 \neq 0$ satisfy (43). For $i = 1, 2$, write $\bar{w}_i = \bar{w}_i^{\|} + \bar{w}_i^{\perp}$, where $\bar{w}_i^{\|} \in \operatorname{span}\{v_1, v_2\}$ and $\bar{w}_i^{\perp} \perp \operatorname{span}\{v_1, v_2\}$. Then $\bar{w}_1^{\perp}$ and $\bar{w}_2^{\perp}$ are linearly dependent. In particular,*

$$\dim\operatorname{span}\{v_1, v_2, \bar{w}_1, \bar{w}_2\} \leq 3.$$

B.2.2. PROOF OF PROPOSITION 19

Note that we have for critical point

$$\begin{aligned}\frac{dw_1}{dt} = 0 = \frac{1}{2\pi}\left[(\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2\right] + \frac{\hat{w}_1}{2\pi}(\sin\theta_{11}+\sin\theta_{12})\\ -\left[\frac{1}{2}w_1 + \frac{\|w_2\|_2}{\|w_1\|_2}\left(\frac{\sin\gamma}{2\pi}w_1\right) + \left(\frac{\pi-\gamma}{2\pi}w_2\right)\right]\end{aligned} \tag{44}$$

$$\begin{aligned}\frac{dw_2}{dt} = 0 = \frac{1}{2\pi}\left[(\pi-\theta_{21})v_1 + (\pi-\theta_{22})v_2\right] + \frac{\hat{w}_2}{2\pi}(\sin\theta_{21}+\sin\theta_{22})\\ -\left[\frac{1}{2}w_2 + \frac{\|w_1\|_2}{\|w_2\|_2}\left(\frac{\sin\gamma}{2\pi}w_2\right) + \left(\frac{\pi-\gamma}{2\pi}w_1\right)\right],\end{aligned} \tag{45}$$

where γ is the angle between w_1 and w_2. If we write $\hat{w}_1 = \hat{w}_1^{\perp} + \hat{w}_1^{\|}$ and $\hat{w}_2 = \hat{w}_2^{\perp} + \hat{w}_2^{\|}$ where $\hat{w}_i^{\|} \in \operatorname{span}\{v_1, v_2\}$ and $\hat{w}_i^{\perp} \perp \operatorname{span}\{v_1, v_2\}$, then by (44) and (45), we have $\hat{w}_1^{\perp} = c\hat{w}_2^{\perp}$ for some constant $c \neq 0$, and thus $\hat{w}_1^{\perp} \mathbin{//} \hat{w}_2^{\perp}$ or $\hat{w}_1^{\perp} = \hat{w}_2^{\perp} = 0$.

In the reduced three-dimensional problem we use the basis

$$u_1 = \bar{v}, \qquad u_2 = v_\Delta, \qquad u_3 \perp \operatorname{span}\{v_1, v_2\}, \tag{46}$$

and write student directions as

$$\bar{w}_i = (\cos\phi_i\sin\beta_i, \sin\phi_i\sin\beta_i, \cos\beta_i), \qquad i = 1, 2. \tag{47}$$

B.2.3. AUXILIARY ANGLE-SUM ESTIMATE FOR LEMMA 21

Lemma 20 *Assume* (w_1, w_2) *satisfies* (43)*, with* β_i, ϕ_i *defined by* (47)*. If* $(\beta_1, \beta_2) \neq (\frac{\pi}{2}, \frac{\pi}{2})$*, then*

$$\min\{\theta_{11} + \theta_{12}, \theta_{21} + \theta_{22}\} < \pi.$$

Proof In the coordinates used in lemma 21, the first coordinate of

$$(\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2$$

is

$$\cos\frac{\alpha}{2}\,(2\pi - \theta_{i1} - \theta_{i2}),$$

which is strictly positive. Indeed, $\cos(\alpha/2) > 0$, and equality in the second factor would force $\theta_{i1} = \theta_{i2} = \pi$, impossible because $v_1 \neq v_2$.

Because $(\beta_1, \beta_2) \neq (\pi/2, \pi/2)$, at least one of $\cos\beta_1, \cos\beta_2$ is nonzero. We first exclude $\gamma = \pi$. If $\gamma = \pi$, then $\bar{w}_2 = -\bar{w}_1$ and $\pi - \gamma = 0$. Taking the third coordinate in whichever equation has $\cos\beta_i \neq 0$ forces the coefficient of $\bar{w}_i$ in that equation to vanish. The corresponding stationarity equation then reduces to

$$(\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2 = 0,$$

contradicting the positive first coordinate above. Hence $\gamma \neq \pi$, so $\pi - \gamma > 0$.

Next, both $\cos\beta_1$ and $\cos\beta_2$ must be nonzero. If, for example, $\cos\beta_1 = 0$, then the third coordinate of the first equation in (43) gives $(\pi - \gamma)\|w_2\|_2 \cos\beta_2 = 0$, hence $\cos\beta_2 = 0$, a contradiction. The other case is identical.

Taking third coordinates in (43), and substituting the result back into the two vector equations, gives

$$(\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2 = (\pi - \gamma)\|w_2\|_2 \left(\bar{w}_2 - \frac{\cos\beta_2}{\cos\beta_1}\bar{w}_1\right),$$

$$(\pi - \theta_{21})v_1 + (\pi - \theta_{22})v_2 = (\pi - \gamma)\|w_1\|_2 \left(\bar{w}_1 - \frac{\cos\beta_1}{\cos\beta_2}\bar{w}_2\right).$$

Comparing first coordinates and using their positivity yields

$$\sin\beta_2\cos\phi_2 - \frac{\cos\beta_2}{\cos\beta_1}\sin\beta_1\cos\phi_1 > 0,$$

$$\sin\beta_1\cos\phi_1 - \frac{\cos\beta_1}{\cos\beta_2}\sin\beta_2\cos\phi_2 > 0.$$

If $\cos\beta_1$ and $\cos\beta_2$ had the same sign, these two strict inequalities would contradict each other. Therefore they have opposite signs. Using this in the first inequality, at least one of $\sin\beta_1\cos\phi_1$ and $\sin\beta_2\cos\phi_2$ must be positive.

For each i,

$$\cos\theta_{i1} + \cos\theta_{i2} = \bar{w}_i \cdot (v_1 + v_2) = 2\cos\frac{\alpha}{2}\sin\beta_i\cos\phi_i.$$

On the other hand,

$$\cos\theta_{i1} + \cos\theta_{i2} = 2\cos\frac{\theta_{i1} + \theta_{i2}}{2}\cos\frac{\theta_{i1} - \theta_{i2}}{2}.$$

The triangle inequality gives

$$|\theta_{i1} - \theta_{i2}| \le \angle(v_1, v_2) = \alpha \le \frac{\pi}{2},$$

so $\cos((\theta_{i1} - \theta_{i2})/2) > 0$. Since $\cos(\alpha/2) > 0$, the sign of $\sin\beta_i \cos\phi_i$ is the sign of $\cos((\theta_{i1} + \theta_{i2})/2)$. Thus

$$\sin\beta_i \cos\phi_i > 0 \quad \Longleftrightarrow \quad \theta_{i1} + \theta_{i2} < \pi.$$

Since at least one of $\sin\beta_i \cos\phi_i$ is positive, the desired strict angle-sum condition follows. ■

Lemma 21 (Three-dimensional symmetry) *Let (w_1, w_2) be a stationary point satisfying (43), and let β_i, ϕ_i be defined by (47). If $(\beta_1, \beta_2) \neq (\pi/2, \pi/2)$, then $\phi_1 = \phi_2 = 0$.*

B.2.4. PROOF OF LEMMA 21

By lemma 20, after possibly swapping the two student neurons, we may assume

$$\theta_{11} + \theta_{12} < \pi.$$

In this case, by (44) and (45), we have that the perpendicular components of w_1 with respect to the plane spanned by v_1 and v_2 is parallel to the perpendicular components of w_2 with respect to the plane spanned by v_1 and v_2, i.e., if we write $\hat{w}_1 = \hat{w}_1^\perp + \hat{w}_1^\parallel$ and $\hat{w}_2 = \hat{w}_2^\perp + \hat{w}_2^\parallel$ where $\hat{w}_i^\parallel \in \text{span}\{v_1, v_2\}$ and $\hat{w}_i^\perp \perp \text{span}\{v_1, v_2\}$, then $\hat{w}_1^\perp \,//\, \hat{w}_2^\perp$. In addition, we have that

$$\frac{\pi - \theta_{11}}{\pi - \theta_{12}} = \frac{\pi - \theta_{21}}{\pi - \theta_{22}}. \tag{48}$$

Let x-axis be the angular bisector of v_1 and v_2, and y-axis be the perpendicular direction of $\hat{v}$, and z-axis be orthogonal to the plane spanned by v_1 and v_2.

Let β_i be the angle between w_i and z-axis, let ϕ_i be the angle between w_i and $\hat{v}$ with positive x-axis. The sign of ϕ_i is positive if w_i is clockwise to $\hat{v}$, and negative otherwise.

$$\begin{aligned}
v_1 &= (\cos(\alpha/2), \sin(\alpha/2), 0) \\
v_2 &= (\cos(\alpha/2), -\sin(\alpha/2), 0) \\
\hat{v} &= \frac{v_1 + v_2}{\|v_1 + v_2\|_2} = (\cos(\alpha/2), 0, 0)
\end{aligned}$$

We have that

$$\begin{aligned}
\bar{w}_1 &= (\cos(\phi_1)\sin(\beta_1), \sin(\phi_1)\sin(\beta_1), \cos(\beta_1)) \\
\bar{w}_2 &= (\cos(\phi_2)\sin(\beta_2), \sin(\phi_2)\sin(\beta_2), \cos(\beta_2)) \\
\bar{\beta}_1 &= (\cos(\phi_1)\cos(\beta_1), \sin(\phi_1)\cos(\beta_1), -\sin(\beta_1)) \\
\bar{\beta}_2 &= (\cos(\phi_2)\cos(\beta_2), \sin(\phi_2)\cos(\beta_2), -\sin(\beta_2)) \\
\bar{\phi}_1 &= (-\sin(\phi_1), \cos(\phi_1), 0) \\
\bar{\phi}_2 &= (-\sin(\phi_2), \cos(\phi_2), 0)
\end{aligned}$$

$$
\begin{aligned}
(\pi-\theta_{i1})v_1+(\pi-\theta_{i2})v_2 &= (\cos(\alpha/2)(2\pi-\theta_{i1}-\theta_{i2}),\sin(\alpha/2)(\theta_{i2}-\theta_{i1}),0)\\
\cos\theta_{i1} &= \cos(\alpha/2-\phi_i)\sin\beta_i\\
\cos\theta_{i2} &= \cos(\alpha/2+\phi_i)\sin\beta_i.
\end{aligned}
$$

For simplicity, we let

$$
\begin{aligned}
C_i &= \cos(\alpha/2)[2\pi-(\theta_{i1}+\theta_{i2})]\\
S_i &= \sin(\alpha/2)[\theta_{i2}-\theta_{i1}].
\end{aligned}
$$

From (44) and (45), we have that

$$C_1\sin\beta_1\cos\phi_1+S_1\sin\beta_1\sin\phi_1+\sin\theta_{11}+\sin\theta_{12}-\pi r_1-r_2\sin\gamma-(\pi-\gamma)r_2\cos\gamma=0 \tag{49}$$

$$C_2\sin\beta_2\cos\phi_2+S_2\sin\beta_2\sin\phi_2+\sin\theta_{21}+\sin\theta_{22}-\pi r_2-r_1\sin\gamma-(\pi-\gamma)r_1\cos\gamma=0 \tag{50}$$

$$C_1\cos\beta_1\cos\phi_1+S_1\cos\beta_1\sin\phi_1-(\pi-\gamma)r_2(\sin\beta_2\cos\beta_1\cos(\phi_1-\phi_2)-\cos\beta_2\sin\beta_1)=0 \tag{51}$$

$$C_2\cos\beta_2\cos\phi_2+S_2\cos\beta_2\sin\phi_2-(\pi-\gamma)r_1(\sin\beta_1\cos\beta_2\cos(\phi_2-\phi_1)-\cos\beta_1\sin\beta_2)=0 \tag{52}$$

$$-C_1\sin\phi_1+S_1\cos\phi_1-(\pi-\gamma)r_2\sin\beta_2\sin(\phi_2-\phi_1)=0 \tag{53}$$

$$-C_2\sin\phi_2+S_2\cos\phi_2-(\pi-\gamma)r_1\sin\beta_1\sin(\phi_1-\phi_2)=0 \tag{54}$$

We have that

$$\gamma=\arccos(\sin\beta_1\sin\beta_2\cos(\phi_1-\phi_2)+\cos\beta_1\cos\beta_2),$$

and that

$$
\begin{aligned}
\cos\theta_{i1} &= \sin\beta_i\cos(\phi_i-\alpha/2),\\
\cos\theta_{i2} &= \sin\beta_i\cos(\phi_i+\alpha/2).
\end{aligned}
$$

Multiplying (51) by $\sin\phi_1$, we have that

$$C_1\sin\phi_1\cos\beta_1\cos\phi_1+S_1\sin\phi_1^2\cos\beta_1=(\pi-\gamma)r_2\sin\phi_1 \tag{55}$$

$$\times(\sin\beta_2\cos\beta_1\cos(\phi_1-\phi_2)-\cos\beta_2\sin\beta_1)) \tag{56}$$

Multiplying (53) by $\cos\beta_1\cos\phi_1$, we have that

$$-C_1\sin\phi_1\cos\beta_1\cos\phi_1+S_1\cos\phi_1^2\cos\beta_1=(\pi-\gamma)r_2\cos\beta_1\cos\phi_1\sin\beta_2\sin(\phi_2-\phi_1) \tag{57}$$

Adding (56) and (57), we have that

$$S_1 \cos\beta_1 = (\pi-\gamma) r_2 \left[(\sin\beta_2\cos\beta_1\cos(\phi_1-\phi_2) - \cos\beta_2\sin\beta_1)\sin\phi_1 \right. \\ \left. + \sin\beta_2\sin(\phi_2-\phi_1)\cos\beta_1\cos\phi_1\right],$$

where we have that

$$\begin{aligned} &(\sin\beta_2\cos\beta_1\cos(\phi_1-\phi_2) - \cos\beta_2\sin\beta_1)\sin\phi_1 \\ &\quad + \sin\beta_2\sin(\phi_2-\phi_1)\cos\beta_1\cos\phi_1 \\ =& \underline{\sin\beta_2\cos\beta_1\cos\phi_1\cos\phi_2\sin\phi_1} + \sin\beta_2\cos\beta_1\sin^2\phi_1\sin\phi_2 - \cos\beta_2\sin\beta_1\sin\phi_1 \\ &- \underline{\sin\beta_2\sin\phi_1\cos\phi_2\cos\beta_1\cos\phi_1} + \sin\beta_2\cos\phi_1^2\sin\phi_2\cos\beta_1 \\ =& \sin\beta_2\cos\beta_1\sin\phi_2 - \cos\beta_2\sin\beta_1\sin\phi_1. \end{aligned}$$

And we have that

$$S_1\cos\beta_1 = (\pi-\gamma)r_2\left[\sin\beta_2\cos\beta_1\sin\phi_2 - \cos\beta_2\sin\beta_1\sin\phi_1\right]. \tag{58}$$

Similarly, we have that

$$S_2\cos\beta_2 = (\pi-\gamma)r_1\left[\sin\beta_1\cos\beta_2\sin\phi_1 - \cos\beta_1\sin\beta_2\sin\phi_2\right], \tag{59}$$
$$C_1\cos\beta_1 = (\pi-\gamma)r_2\left[\cos\beta_1\sin\beta_2\cos\phi_2 - \cos\beta_2\sin\beta_1\cos\phi_1\right], \tag{60}$$
$$C_2\cos\beta_2 = (\pi-\gamma)r_1\left[\cos\beta_2\sin\beta_1\cos\phi_1 - \cos\beta_1\sin\beta_2\cos\phi_2\right]. \tag{61}$$

Noth that we have

$$\frac{\pi-\theta_{11}}{\pi-\theta_{12}} = \frac{\pi-\theta_{21}}{\pi-\theta_{22}},$$

Case 1: $\theta_{11}+\theta_{12}<\pi$**,** $\theta_{21}+\theta_{22}<\pi$: Without loss of generality, we assume $\phi_1,\phi_2\in(0,\frac{\alpha}{2})$. Now we are going to prove that $-C_i\sin\phi_i + S_i\cos\phi_i \le 0$, for all $\beta_i\in(0,\pi)$, $i=1,2$ and $\alpha\in(0,\frac{\pi}{2})$. We assume that $\beta_1\in(0,\pi/2)$ and we are going to prove that $-C_1\sin\phi_1+S_1\cos\phi_1\le 0$ for $\phi_1\in(0,\frac{\alpha}{2})$.

First, we are going to prove, when $\phi_1\in(0,\pi/2)$ is fixed, $\theta_{12}-\theta_{11}$ takes the maximum value when $\beta_1=\pi/2$. Since

$$\theta_{12}-\theta_{11} = \arccos\left(\sin\beta_1\cos(\alpha/2+\phi_1)\right) - \arccos\left(\sin\beta_1\cos(\alpha/2-\phi_1)\right),$$

we have that

$$\begin{aligned} \frac{\partial}{\partial\beta_1}(\theta_{12}-\theta_{11}) &= -\frac{\cos\beta_1\cos(\alpha/2+\phi_1)}{\sqrt{1-\sin^2\beta_1\cos^2(\alpha/2+\phi_1)}} + \frac{\cos\beta_1\cos(\alpha/2-\phi_1)}{\sqrt{1-\sin^2\beta_1\cos^2(\alpha/2-\phi_1)}} \\ &= \cos\beta_1\left[\frac{\cos(\alpha/2-\phi_1)}{\sqrt{1-\sin^2\beta_1\cos^2(\alpha/2-\phi_1)}} - \frac{\cos(\alpha/2+\phi_1)}{\sqrt{1-\sin^2\beta_1\cos^2(\alpha/2+\phi_1)}}\right]. \end{aligned}$$

When $\phi_1\in(0,\frac{\alpha}{2})$ is fixed, we have that $\cos(\alpha/2-\phi_1)>\cos(\alpha/2+\phi_1)$. And thus we have that $\frac{d}{d\beta_1}(\theta_{12}-\theta_{11})>0$, which implies that $\theta_{12}-\theta_{11}\le\theta_{12}(\beta_1=\pi/2)-\theta_{11}(\beta_1=\pi/2)=2\phi_1$.

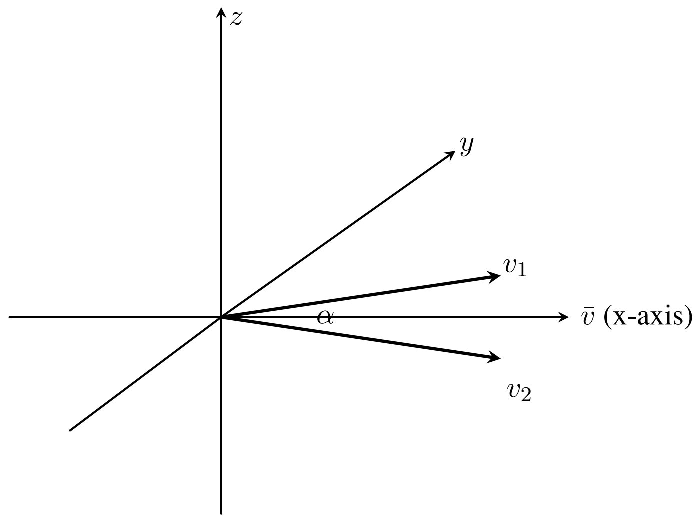

Figure 4: A graphical explanation of the 3D problem.

And thus we have that

$$\begin{aligned}\frac{\theta_{12}-\theta_{11}}{\tan\phi_1} &\le \frac{2\phi_1}{\tan\phi_1}\\ &\le 2 < \frac{\pi}{\tan\frac{\pi}{4}}\\ &\le \frac{\pi}{\tan\frac{\alpha}{2}} \le \frac{2\pi-\theta_{11}-\theta_{12}}{\tan\frac{\alpha}{2}},\end{aligned}$$

where the last inequality comes from the assumption $\theta_{11}+\theta_{12}\le\pi$. And thus we have that

$$\sin\frac{\alpha}{2}\left(\theta_{12}-\theta_{11}\right)\cos\phi_1 < \cos\frac{\alpha}{2}\left(2\pi-\theta_{11}-\theta_{12}\right)\sin\phi_1.$$

This implies that

$$-C_1\sin\phi_1 + S_1\cos\phi_1 \le 0.$$

Similarly, we have that

$$-C_2\sin\phi_2 + S_2\cos\phi_2 < 0.$$

Note that by (53) and (54), we have that

$$\frac{-C_1\sin\phi_1+S_1\cos\phi_1}{r_2\sin\beta_2} + \frac{-C_2\sin\phi_2+S_2\cos\phi_2}{r_1\sin\beta_1} = 0,$$

which implies that (since the two terms on the left hand side have the same sign)

$$-C_1\sin\phi_1 + S_1\cos\phi_1 = -C_2\sin\phi_2 + S_2\cos\phi_2 = 0. \tag{62}$$

Note that by (58)∼(61), we have that

$$\frac{S_1}{C_1} = \frac{S_2}{C_2}.$$

And thus we have $\phi_1 = \phi_2$.

By (62), we have that

$$\begin{aligned}\tan\phi_i &= \frac{\tan\frac{\alpha}{2}\left(\theta_{i2}-\theta_{i1}\right)}{2\pi-\theta_{i1}-\theta_{i2}}\\ &\le \frac{\tan\frac{\alpha}{2}2\phi_i}{pi}\\ &\le \frac{2}{\pi}\tan\phi_i.\end{aligned}$$

And thus $\phi_i = 0$.

Case 2: $\theta_{11}+\theta_{12}<\pi, \theta_{21}+\theta_{22}>\pi$. Note that $\frac{\pi-\theta_{11}}{\pi-\theta_{12}} = \frac{\pi-\theta_{21}}{\pi-\theta_{22}}$. Thus w_1 projection in the $x-y$ plane and w_2 projection in the $x-y$ plane are at the same side of the x-axis. Note that we have

$$\begin{aligned}0 &= \frac{dw_1}{dt}\\ &= \frac{1}{2\pi}\left[(\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2 + (\sin\theta_{11}+\sin\theta_{12})\bar{w}_1\right]\\ &\quad - \left(\frac{1}{2}+\frac{\sin\gamma}{2\pi}\cdot\frac{\|w_2\|_2}{\|w_1\|_2}\right)w_1 - \frac{\pi-\gamma}{2\pi}w_2\\ &= \frac{1}{2\pi}\left[(\pi-\theta_{11})v_1 + (\pi-\theta_{12})v_2\right]\\ &\quad - \frac{1}{2\pi}\left(\pi\|w_1\|_2 + \|w_2\|_2\sin\gamma - \sin\theta_{11} - \sin\theta_{12}\right)\bar{w}_1\\ &\quad - \frac{\pi-\gamma}{2\pi}\|w_2\|_2\bar{w}_2.\end{aligned}$$

Similarly, we have that

$$\begin{aligned}0 &= \frac{dw_2}{dt}\\ &= \frac{1}{2\pi}\left[(\pi-\theta_{21})v_1 + (\pi-\theta_{22})v_2\right]\\ &\quad - \frac{1}{2\pi}\left(\pi\|w_2\|_2 + \|w_1\|_2\sin\gamma - \sin\theta_{21} - \sin\theta_{22}\right)\bar{w}_2\\ &\quad - \frac{\pi-\gamma}{2\pi}\|w_1\|_2\bar{w}_1.\end{aligned}$$

Note that we have $(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2$ is parallel to $(\pi-\theta_{21})v_1+(\pi-\theta_{22})v_2$, and one should note that if $x\mathbf{a}+y\mathbf{b}+z\mathbf{c}=0$, $p\mathbf{a}+q\mathbf{b}+r\mathbf{c}=0$ for some $p,q,r\in\mathbb{R}$ and $\mathbf{a},\mathbf{b},\mathbf{c}$ are not parallel to each other, then $\frac{x}{p}=\frac{y}{q}=\frac{z}{r}$.

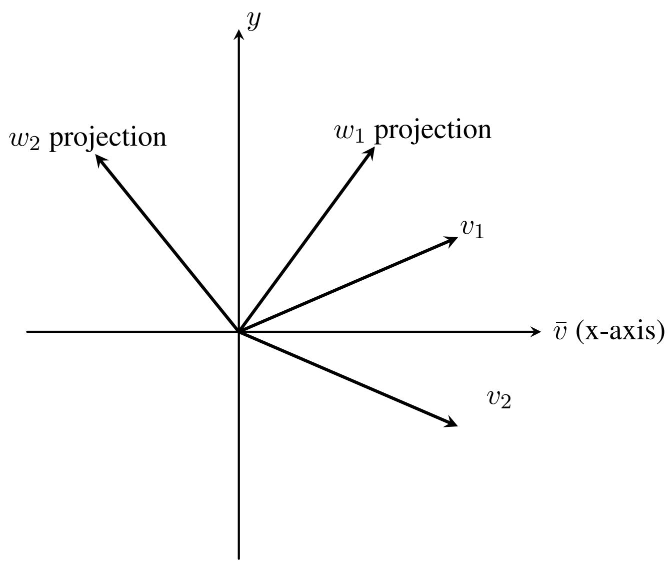

Figure 5: v_1, v_2, and the projections of w_1 and w_2 in the $x - y$ plane.

This implies that

$$\begin{aligned}\frac{\pi-\theta_{11}}{\pi-\theta_{21}} &= \frac{\pi-\theta_{12}}{\pi-\theta_{22}} \\ &= \frac{\pi\|w_1\|_2+\|w_2\|_2\sin\gamma-\sin\theta_{11}-\sin\theta_{12}}{\big(\pi-\gamma\big)\|w_1\|_2} \\ &= \frac{\big(\pi-\gamma\big)\|w_2\|_2}{\pi\|w_2\|_2+\|w_1\|_2\sin\gamma-\sin\theta_{21}-\sin\theta_{22}}.\end{aligned}$$

This implies that $\pi\|w_1\|_2+\|w_2\|_2\sin\gamma-\sin\theta_{11}-\sin\theta_{12}>0$ and $\pi\|w_2\|_2+\|w_1\|_2\sin\gamma-\sin\theta_{21}-\sin\theta_{22}>0$.

If the projection of w_1 lies outside the angle between v_1 and v_2 as shown in Figure 6. And thus we have $a\Bigg[(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2\Bigg]-b\bar{w}_1-c\bar{w}_2=0$, for some $a,b,c\in\mathbb{R}_+$.

However, the projections of $-w_1$, $-w_2$, and $(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2$ all lie on the same side of the vector v_1. Therefore, the planar coordinates of $\frac{\mathrm{d}w_1}{dt}$ cannot be 0. And thus the projection of w_1 lies inside the angle between v_1 and v_2.

When the projection of w_1 lies inside the angle between v_1 and v_2, if we prove the vector $((\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2)$ lies inside the angle between w_1 and $\bar{v}$ as shown in Figure 7, then $\frac{\mathrm{d}w_1}{dt}$ can't be 0 since $a((\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2)-b\bar{w}_1-c\bar{w}_2$ can't be 0 for any $a,b,c\in\mathbb{R}_+$ in such case. Next we are going to prove that the vector $((\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2)$ lie inside the angle between w_1 and $\bar{v}$.

Lemma 22 *The vector $(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2$ lies inside the angle between w_1 and $\bar{v}$.*

Proof Fix $\phi_1\in(0,\frac{\alpha}{2})$. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be orthogonal planer unit vectors such that $\mathbf{e}_1$ is parallel to the x-axis and $\mathbf{e}_2$ is parallel to the y-axis. For simplicity, we let $u=(\pi-\theta_{11})v_1+(\pi-\theta_{12})v_2$. We

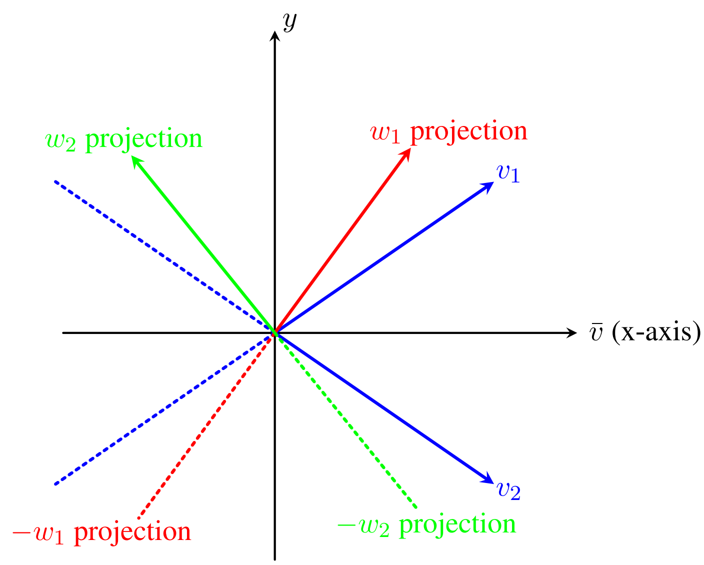

Figure 6: v_1, v_2, and the projections of w_1 and w_2 in the $x - y$ plane.

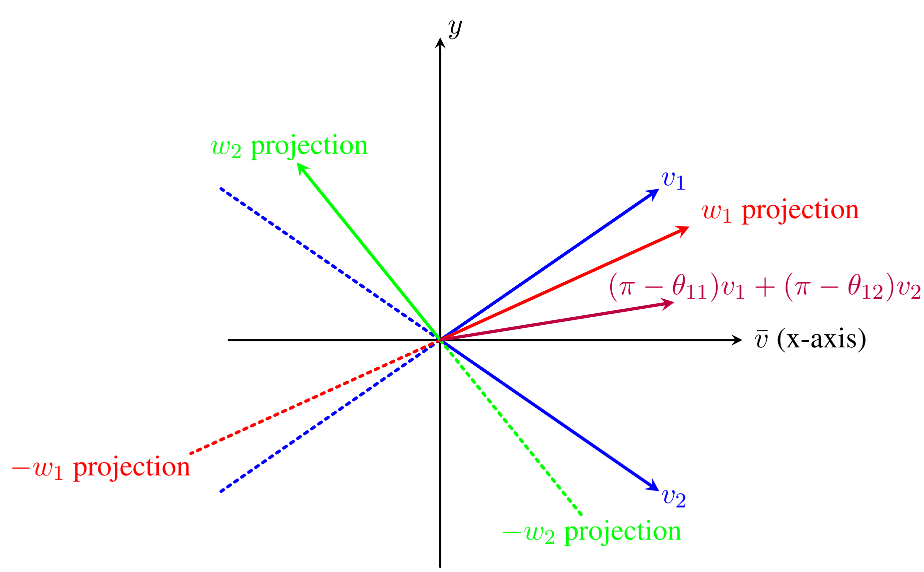

Figure 7: v_1, v_2, and the projections of w_1 and w_2 in the $x - y$ plane.

only need to prove that $\angle(u, v_1) \geq \angle(w_1, v_1), v_2 = \cos\alpha \cdot v_1 + \sin\frac{\alpha}{2} v_1^{\perp}$. Note that

$$\tan\angle(u, v_1) = \frac{\big(\pi - \theta_{12}\big)\sin\alpha}{\big(\pi - \theta_{11}\big) + \big(\pi - \theta_{12}\big)\cos\alpha},$$

we only need to prove $\tan\angle(u, v_1) \geq \tan\angle(w_1, v_1) = \tan(\frac{\alpha}{2} - \phi_1)$. We only need to prove

$$\big(\pi - \theta_{12}\big)\sin\alpha\cos\left(\frac{\alpha}{2} - \phi_1\right) \geq \Big[\big(\pi - \theta_{11}\big) + \big(\pi - \theta_{12}\big)\cos\alpha\Big]\sin\left(\frac{\alpha}{2} - \phi_1\right).$$

Move $\big(\pi - \theta_{12}\big)$ term to the left hand side, we have only need to prove

$$\big(\pi - \theta_{12}\big)\sin\left(\frac{\alpha}{2} + \phi_1\right) \geq \big(\pi - \theta_{11}\big)\sin\left(\frac{\alpha}{2} - \phi_1\right). \tag{63}$$

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\phi_1}\Big[\big(\pi - \theta_{12}\big)\sin\Big(\frac{\alpha}{2} + \phi_1\Big) - \big(\pi - \theta_{11}\big)\sin\Big(\frac{\alpha}{2} - \phi_1\Big)\Big] \\
&= \Big(\pi - \theta_{12}\Big)\cos\Big(\frac{\alpha}{2} + \phi_1\Big) - \frac{\sin\beta_1\sin^2\left(\frac{\alpha}{2} + \phi_1\right)}{\sin\theta_{12}} \\
&\quad + \Big(\pi - \theta_{11}\Big)\cos\Big(\frac{\alpha}{2} - \phi_1\Big) - \frac{\sin\beta_1\sin^2\left(\frac{\alpha}{2} - \phi_1\right)}{\sin\theta_{11}}.
\end{aligned}$$

Since we have

$$\sin^2\theta_{12} = 1 - \cos^2\theta_{12} = 1 - \sin^2\beta_1\cos^2\left(\frac{\alpha}{2} + \phi_1\right) \geq \sin^2\beta_1\sin^2\left(\frac{\alpha}{2} + \phi_1\right),$$

we have that

$$\sin\theta_{12} \geq \sin\beta_1\sin\left(\frac{\alpha}{2} + \phi_1\right).$$

Similarly, we have that

$$\sin\theta_{11} \geq \sin\beta_1\sin\left(\frac{\alpha}{2} - \phi_1\right).$$

And thus we have that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\phi_1}\Big[\big(\pi - \theta_{12}\big)\sin\Big(\frac{\alpha}{2} + \phi_1\Big) - \big(\pi - \theta_{11}\big)\sin\Big(\frac{\alpha}{2} - \phi_1\Big)\Big] \\
&\geq \Big(\pi - \theta_{12}\Big)\cos\Big(\frac{\alpha}{2} + \phi_1\Big) - \sin\Big(\frac{\alpha}{2} + \phi_1\Big) + \Big(\pi - \theta_{11}\Big)\cos\Big(\frac{\alpha}{2} - \phi_1\Big) - \sin\Big(\frac{\alpha}{2} - \phi_1\Big) \\
&\geq \frac{\pi}{2}\cos\Big(\frac{\alpha}{2} + \phi_1\Big) - \sin\Big(\frac{\alpha}{2} + \phi_1\Big) + \frac{\pi}{2}\cos\Big(\frac{\alpha}{2} - \phi_1\Big) - \sin\Big(\frac{\alpha}{2} - \phi_1\Big).
\end{aligned}$$

Note that we have that

$$\begin{aligned}
\cos\Big(\frac{\alpha}{2} + \phi_1\Big) + \cos\Big(\frac{\alpha}{2} - \phi_1\Big) &= 2\cos\Big(\frac{\alpha}{2}\Big)\cos\phi_1, \\
\sin\Big(\frac{\alpha}{2} + \phi_1\Big) + \sin\Big(\frac{\alpha}{2} - \phi_1\Big) &= 2\sin\Big(\frac{\alpha}{2}\Big)\cos\phi_1,
\end{aligned}$$

And thus

$$\frac{\mathrm{d}}{\mathrm{d}\phi_1}\Big[\big(\pi - \theta_{12}\big)\sin\Big(\frac{\alpha}{2} + \phi_1\Big) - \big(\pi - \theta_{11}\big)\sin\Big(\frac{\alpha}{2} - \phi_1\Big)\Big]$$

$$\geq 2\cos\phi_1\Big(\frac{\pi}{2}\cos\left(\frac{\alpha}{2}\right) - \sin\left(\frac{\alpha}{2}\right)\Big)$$
$$\geq 2\cos\phi_1\Big(\cos\left(\frac{\alpha}{2}\right) - \sin\left(\frac{\alpha}{2}\right)\Big) \geq 0.$$

The last inequality follows from the fact that $\phi_1 \in \big(0, \frac{\alpha}{2}\big)$ and $\alpha \in (0, \pi/2)$. ■

By Lemma 22, we have that the three vectors $((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2)$, $-w_1$, and $-w_2$ lies exactly on the same side of the vector w_1. And thus the planer coordinate of $\frac{\mathrm{d}w_1}{dt}$ can't be 0 as shown in Figure 7.

B.2.5. SYMMETRY REDUCTION FOR THE OUT-OF-PLANE CASE

Lemma 23 (A scalar separation lemma) *Let*

$$0 < t < m < \frac{\pi}{2}.$$

Define

$$A_-(m) := \frac{2(\pi - 2m)}{2m - \sin 2m}, \qquad A_+(m) := \frac{2(\pi - 2m)}{2m + \sin 2m},$$

and

$$X = X(t;m) := t + A_-(m)\big(t - \cos^2 m\, \tan t\big), \qquad Y = Y(t;m) := t + A_+(m)\big(t + \sin^2 m\, \cot t\big).$$

If $\mathcal{A} > 0$ satisfies

$$\mathcal{A}X = \frac{\pi}{2} + \arctan\left(\frac{\mathcal{A}\tan t}{\tan^2 m}\right),$$

then

$$\mathcal{A}X > \frac{Y}{\mathcal{A}}.$$

Proof With m fixed, put $a = m - \sin m \cos m$, $b = m + \sin m \cos m$,

$$P(t) = t - \cos^2 m \tan t, \qquad Q(t) = t + \sin^2 m \cot t.$$

The quantities $a, b, P(t), Q(t)$ recur throughout the proof. Then $a, b > 0$,

$$A_-(m) = \frac{\pi - 2m}{a}, \qquad A_+(m) = \frac{\pi - 2m}{b},$$
$$X = t + (\pi - 2m)\frac{P(t)}{a}, \qquad Y = t + (\pi - 2m)\frac{Q(t)}{b}.$$

Also $X > 0$ and $Y > 0$, since

$$P(t) > t - \cos^2 t \tan t = t - \sin t \cos t > 0, \qquad Q(t) > 0.$$

We first record the elementary comparisons needed below. Since

$$P'(t) = 1 - \cos^2 m \sec^2 t > 0, \qquad Q'(t) = 1 - \sin^2 m \csc^2 t < 0$$

on $0 < t < m$, and since $P(m) = a$, $Q(m) = b$, we have $0 < P(t) < a$ and $Q(t) > b$. Next, differentiating $P(t)Q(t)$ gives

$$\begin{aligned}\frac{d}{dt}\big(P(t)Q(t)\big) &= \frac{\tan^2 m - \tan^2 t}{1+\tan^2 m}\left[Q(t) - \frac{P(t)}{\tan^2 t}\right] \\ &= \frac{\tan^2 m - \tan^2 t}{1+\tan^2 m}\,\frac{\tan t - t + t\tan^2 t}{\tan^2 t} > 0.\end{aligned}$$

Thus $P(t)Q(t) < ab$. We shall also use $t(P(t)/a + Q(t)/b) < 2m$. To prove it, differentiate $t(P(t)/a + Q(t)/b)$:

$$\frac{d}{dt}\left[t\left(\frac{P(t)}{a} + \frac{Q(t)}{b}\right)\right] = \frac{2t - \cos^2 m(\tan t + t\sec^2 t)}{a} + \frac{2t + \sin^2 m(\cot t - t\csc^2 t)}{b}. \tag{64}$$

The numerator $2t - \cos^2 m(\tan t + t\sec^2 t)$ in (64) is positive because it is larger than $t - \sin t\cos t > 0$. For the numerator $2t + \sin^2 m(\cot t - t\csc^2 t)$, note that $\cot t - t\csc^2 t < 0$, and therefore

$$2t + \sin^2 m(\cot t - t\csc^2 t) > 2t + \cot t - t\csc^2 t = \frac{\sin t\cos t - t\cos 2t}{\sin^2 t} > 0. \tag{65}$$

The positivity of the two numerators in (64), with the second one bounded in (65), shows that the derivative in (64) is positive. Comparing $t(P(t)/a + Q(t)/b)$ with its value at $t = m$ gives $t(P(t)/a + Q(t)/b) < 2m$.

The inequalities $P(t)Q(t) < ab$ and $t(P(t)/a + Q(t)/b) < 2m$ imply

$$\begin{aligned}(\pi - m)^2 - XY &= (\pi - m)^2 - \left(t + (\pi - 2m)\frac{P(t)}{a}\right) \\ &\qquad \cdot \left(t + (\pi - 2m)\frac{Q(t)}{b}\right) \\ &= m^2 - t^2 \\ &\quad + (\pi - 2m)\left[2m - t\left(\frac{P(t)}{a} + \frac{Q(t)}{b}\right)\right] \\ &\quad + (\pi - 2m)^2\left(1 - \frac{P(t)Q(t)}{ab}\right) > 0.\end{aligned}$$

Therefore $\sqrt{XY} < \pi - m$.

We need one sharper comparison. Since $ab - P(t)Q(t) > 0$ and $aQ(t) - bP(t) > ab - bP(t) = b(a - P(t)) > 0$, we claim that

$$\frac{t(aQ(t) - bP(t))}{ab - P(t)Q(t)} < \frac{\pi}{2}. \tag{66}$$

Because $b = m + \frac{1}{2}\sin 2m < \pi/2$, it is enough to prove

$$\frac{t(aQ(t) - bP(t))}{ab - P(t)Q(t)} < b,$$

or equivalently

$$b(ab - P(t)Q(t)) - t(aQ(t) - bP(t)) > 0. \tag{67}$$

Expanding $b(ab - P(t)Q(t)) - t(aQ(t) - bP(t))$ in (67) gives

$$b(ab - P(t)Q(t)) - t(aQ(t) - bP(t)) = bm^2 - at^2 - 2m\sin^2 m\, t \cot t. \tag{68}$$

At $t = m$, the right-hand side of (68) is zero. Its derivative with respect to t is

$$2t\left[m\sin^2 m\,\frac{t - \sin t\cos t}{t\sin^2 t} - a\right]. \tag{69}$$

For $0 < u < \pi/2$, the function $u \mapsto (u - \sin u\cos u)/(u\sin^2 u)$ is strictly increasing, since

$$\frac{d}{du}\left[\frac{u - \sin u\cos u}{u\sin^2 u}\right] = \frac{\cos u}{\sin^3 u}\left[\frac{\tan u}{u} + \left(\frac{\sin u}{u}\right)^2 - 2\right] > 0$$

by Wilker's inequality. Since $t < m$, the bracketed factor $m\sin^2 m\,(t - \sin t\cos t)/(t\sin^2 t) - a$ in (69) is strictly less than $m\sin^2 m\,\frac{m - \sin m\cos m}{m\sin^2 m} - a = 0$. Thus the derivative in (69) is negative on $(0, m)$. Since the right-hand side of (68) equals zero at $t = m$, it is positive for $0 < t < m$. Hence (66) follows.

We now prove

$$\sqrt{XY} + \arctan\left(\frac{\tan^2 m}{\tan t}\sqrt{\frac{X}{Y}}\right) < \pi. \tag{70}$$

If $\sqrt{XY} \le \pi/2$, (70) follows because the arctangent term is strictly smaller than $\pi/2$. Hence assume $\sqrt{XY} > \pi/2$. From $\sqrt{XY} < \pi - m$, we get $m < \pi - \sqrt{XY} < \pi/2$. For any $u \in (0, \pi/2)$ with $u \neq m$, strict convexity of $\tan$ gives the strict tangent-line bound

$$\tan u > \tan m + \sec^2 m\,(u - m).$$

Taking $u = \pi - \sqrt{XY}$ and using $\sec^2 m = 1 + \tan^2 m$ gives

$$\tan(\pi - \sqrt{XY}) > \tan m + (1 + \tan^2 m)(\pi - \sqrt{XY} - m).$$

It remains to prove

$$\frac{\tan^2 m}{\tan t}\sqrt{\frac{X}{Y}} \le \tan m + (1 + \tan^2 m)(\pi - \sqrt{XY} - m). \tag{71}$$

After multiplying (71) by $\cos^2 m$ and using $\sqrt{XY} = Y\sqrt{X/Y}$, (71) is equivalent to

$$\sqrt{\frac{X}{Y}}\big(Y + \sin^2 m\cot t\big) \le \pi - m + \sin m\cos m.$$

But $Y + \sin^2 m\cot t = \frac{\pi - m + \sin m\cos m}{m + \sin m\cos m}(t + \sin^2 m\cot t)$, so it suffices to show

$$X\big(t + \sin^2 m\cot t\big)^2 \le Y(m + \sin m\cos m)^2. \tag{72}$$

We prove (72) in the present case $\sqrt{XY} > \pi/2$. With t, m fixed, for $\mu \ge \pi - 2m$, set

$$X_\mu = t + \mu\frac{P(t)}{a}, \qquad Y_\mu = t + \mu\frac{Q(t)}{b}.$$

At $\mu = \pi - 2m$, these are exactly the original quantities X and Y. Hence (72) is the same as

$$Y_{\pi-2m}b^2 - X_{\pi-2m}Q(t)^2 \geq 0.$$

The expression $Y_\mu b^2 - X_\mu Q(t)^2$ is affine in μ, and

$$\frac{d}{d\mu}\left(Y_\mu b^2 - X_\mu Q(t)^2\right) = \frac{Q(t)(ab - P(t)Q(t))}{a} > 0.$$

Thus this affine expression is strictly increasing in μ. If (72) were false, then its value at $\mu = \pi - 2m$ would satisfy

$$Y_{\pi-2m}b^2 - X_{\pi-2m}Q(t)^2 < 0.$$

Since the same affine expression tends to $+\infty$ as $\mu \to \infty$, it would have exactly one zero after $\pi - 2m$. Denote this zero by μ_*. Then $\mu_* > \pi - 2m$, and

$$X_{\mu_*}Q(t)^2 = Y_{\mu_*}b^2.$$

Solving this equality gives

$$\mu_* = \frac{ta}{Q(t)}\frac{Q(t)^2 - b^2}{ab - P(t)Q(t)}.$$

At this value,

$$\sqrt{X_{\mu_*}Y_{\mu_*}} = \frac{b}{Q(t)}Y_{\mu_*} = \frac{tb}{Q(t)} + \mu_* = t\,\frac{aQ(t) - bP(t)}{ab - P(t)Q(t)}.$$

The bound (66) says that $t(aQ(t) - bP(t))/(ab - P(t)Q(t)) < \pi/2$. Since both X_μ and Y_μ increase with μ, the inequality $\mu_* > \pi - 2m$ would imply

$$\begin{aligned}\sqrt{XY} &< \sqrt{X_{\mu_*}Y_{\mu_*}}\\ &< \frac{\pi}{2},\end{aligned}$$

contradicting the present case. Therefore (72) holds, equivalently $XQ(t)^2 \leq Yb^2$, and hence

$$\frac{\tan^2 m}{\tan t}\sqrt{\frac{X}{Y}} < \tan(\pi - \sqrt{XY}).$$

Because $0 < \pi - \sqrt{XY} < \pi/2$, the preceding strict tangent inequality gives

$$\sqrt{XY} + \arctan\left(\frac{\tan^2 m}{\tan t}\sqrt{\frac{X}{Y}}\right) < \pi.$$

Using $\arctan r + \arctan(1/r) = \pi/2$ for $r > 0$, the inequality (70) is equivalent to

$$\sqrt{XY} < \frac{\pi}{2} + \arctan\left(\frac{\tan t}{\tan^2 m}\sqrt{\frac{Y}{X}}\right).$$

Therefore

$$\frac{\frac{\pi}{2} + \arctan\left(\frac{\tan t}{\tan^2 m}\sqrt{\frac{Y}{X}}\right)}{\sqrt{Y/X}} > X. \tag{73}$$

For fixed t, m, the function

$$r \mapsto \frac{\frac{\pi}{2} + \arctan(r \tan t / \tan^2 m)}{r}, \qquad r > 0,$$

is strictly decreasing, because its derivative is

$$\frac{\frac{r \tan t/\tan^2 m}{1+(r\tan t/\tan^2 m)^2} - \left(\frac{\pi}{2} + \arctan(r\tan t/\tan^2 m)\right)}{r^2} < 0.$$

The hypothesis of the lemma says that the value of this function at $r = \mathcal{A}$ equals X, while (73) says that its value at $r = \sqrt{Y/X}$ is larger than X. Hence $\mathcal{A} > \sqrt{Y/X}$. Consequently, $\mathcal{A}X > \sqrt{XY} > Y/\mathcal{A}$, which proves the claim.

■

Lemma 24 (Out-of-plane symmetry) *Let u_1, u_2, u_3 be the basis in (46). Assume $w_1, w_2 \neq 0$ satisfy (43), lie in $\operatorname{span}\{u_1, u_3\}$, and do not both lie in the teacher plane. Then, after possibly replacing u_3 by $-u_3$, there exist $r > 0$ and $\kappa \in (0, \pi/2)$ such that*

$$w_1 = r(\cos\kappa\, u_1 + \sin\kappa\, u_3), \qquad w_2 = r(\cos\kappa\, u_1 - \sin\kappa\, u_3).$$

Equivalently, in the coordinates (47), one has $\beta_1 + \beta_2 = \pi$, $\phi_1 = \phi_2 = 0$, and $\|w_1\|_2 = \|w_2\|_2$.

Proof [Proof of lemma 24]

Choose coordinates so that the first coordinate is the angular bisector direction, the second coordinate lies in the teacher plane, and the third coordinate is orthogonal to the teacher plane. Then

$$v_1 = \big(\cos(\alpha/2), \sin(\alpha/2), 0\big),$$
$$v_2 = \big(\cos(\alpha/2), -\sin(\alpha/2), 0\big).$$

Since w_1, w_2 lie in the perpendicular plane of the angular bisector, write $\bar{w}_i = (x_i, 0, z_i)$, $r_i = \|w_i\|_2$, and $\gamma = \angle(\bar{w}_1, \bar{w}_2)$. Then $\bar{w}_i \cdot v_1 = \bar{w}_i \cdot v_2$, so $\theta_{i1} = \theta_{i2}$; denote this common angle by θ_i. The system (43) reduces to

$$(\pi - \theta_i)(v_1 + v_2) + \big(2\sin\theta_i - \pi r_i - \sin\gamma\, r_j\big)\bar{w}_i - (\pi - \gamma) r_j \bar{w}_j = 0, \qquad \{i, j\} = \{1, 2\}. \tag{74}$$

The case $z_1 = z_2 = 0$ is excluded by the hypothesis that the two directions do not both lie in the teacher plane. We next rule out the one-zero cases. Suppose $z_1 = 0$ and $z_2 \neq 0$. Then $\gamma \notin \{0, \pi\}$, since parallel or antiparallel unit directions would force $z_2 = 0$. Taking the third coordinate of (74) for $i = 1$ gives

$$-(\pi - \gamma) r_2 z_2 = 0,$$

a contradiction. The case $z_2 = 0$ and $z_1 \neq 0$ is identical. Hence $z_1 z_2 \neq 0$, and we are in the genuinely out-of-plane case.

First $0 < \gamma < \pi$. If γ were 0 or π, then $\bar{w}_1$ and $\bar{w}_2$ would be parallel or antiparallel. Since $z_1 z_2 \neq 0$, this direction is not the bisector direction, and the $v_1 + v_2$ component in (74) could not be cancelled. Thus $0 < \gamma < \pi$.

Taking the two-dimensional determinant in the first and third coordinates with $\bar{w}_1$ in (74) for $i = 1$, and with $\bar{w}_2$ in (74) for $i = 2$, gives

$$
\begin{aligned}
&2\cos\frac{\alpha}{2}(\pi - \theta_1)z_1 + (\pi - \gamma)r_2(x_1 z_2 - z_1 x_2) = 0,\\
&2\cos\frac{\alpha}{2}(\pi - \theta_2)z_2 - (\pi - \gamma)r_1(x_1 z_2 - z_1 x_2) = 0.
\end{aligned} \tag{75}
$$

All coefficients except the displayed signed factors are positive, so z_1 and z_2 have opposite signs. Choose the sign of the third coordinate so that $z_1 > 0$ and $z_2 < 0$. Then for some $p, q \in (0, \pi)$,

$$
\bar{w}_1 = (\cos p, 0, \sin p), \qquad \bar{w}_2 = (\cos q, 0, -\sin q).
$$

Under this parametrization, the signed determinant $x_1 z_2 - z_1 x_2$ in (75) equals $-\sin(p+q)$. The first identity in (75) shows that $x_1 z_2 - z_1 x_2 < 0$. Hence $p + q \in (0, \pi)$, and therefore $\gamma = p + q$.

The angles with the two teachers satisfy $\cos\theta_1 = \cos\frac{\alpha}{2}\cos p$ and $\cos\theta_2 = \cos\frac{\alpha}{2}\cos q$. Taking inner products of (74) with $\bar{w}_1$ when $i = 1$, and with $\bar{w}_2$ when $i = 2$, gives

$$
\begin{aligned}
&\pi r_1 + \big(\sin(p+q) + (\pi - p - q)\cos(p+q)\big)r_2 = 2\big(\sin\theta_1 + (\pi - \theta_1)\cos\theta_1\big),\\
&\big(\sin(p+q) + (\pi - p - q)\cos(p+q)\big)r_1 + \pi r_2 = 2\big(\sin\theta_2 + (\pi - \theta_2)\cos\theta_2\big).
\end{aligned} \tag{76}
$$

Projecting (74) onto the tangent directions of $\bar{w}_1$ and $\bar{w}_2$ gives

$$
\begin{aligned}
&(\pi - p - q)\sin(p+q)r_2 = 2\cos\frac{\alpha}{2}(\pi - \theta_1)\sin p,\\
&(\pi - p - q)\sin(p+q)r_1 = 2\cos\frac{\alpha}{2}(\pi - \theta_2)\sin q.
\end{aligned} \tag{77}
$$

Eliminating r_1, r_2 from (76) and (77) gives

$$
\begin{aligned}
&(\pi - p - q)\sin(p+q)\big(\sin\theta_1 + (\pi - \theta_1)\cos\theta_1\big)\\
&\quad = \big(\sin(p+q) + (\pi - p - q)\cos(p+q)\big)\cos\frac{\alpha}{2}(\pi - \theta_1)\sin p\\
&\qquad + \pi\cos\frac{\alpha}{2}(\pi - \theta_2)\sin q,
\end{aligned} \tag{78}
$$

$$
\begin{aligned}
&(\pi - p - q)\sin(p+q)\big(\sin\theta_2 + (\pi - \theta_2)\cos\theta_2\big)\\
&\quad = \pi\cos\frac{\alpha}{2}(\pi - \theta_1)\sin p\\
&\qquad + \big(\sin(p+q) + (\pi - p - q)\cos(p+q)\big)\cos\frac{\alpha}{2}(\pi - \theta_2)\sin q.
\end{aligned} \tag{79}
$$

We claim that (78) and (79) force $p = q$. Suppose instead that $p \neq q$, and by symmetry assume $p > q$. Then

$$
0 < \frac{p-q}{2} < \frac{p+q}{2} < \frac{\pi}{2}.
$$

Also
$$0 < \frac{\theta_1 - \theta_2}{2} < \frac{p-q}{2},$$
because the map $r \mapsto \arccos(\cos(\alpha/2)\cos r)$ has derivative strictly between 0 and 1 on $(0,\pi)$. Since $\cos\theta_1 + \cos\theta_2 = \cos(\alpha/2)(\cos p + \cos q) > 0$, we have $(\theta_1+\theta_2)/2 < \pi/2$.

Adding and subtracting the identities $\cos\theta_1 = \cos(\alpha/2)\cos p$ and $\cos\theta_2 = \cos(\alpha/2)\cos q$, we obtain
$$\begin{aligned} \cos\frac{\theta_1+\theta_2}{2}\cos\frac{\theta_1-\theta_2}{2} &= \cos\frac{\alpha}{2}\cos\frac{p+q}{2}\cos\frac{p-q}{2}, \\ \sin\frac{\theta_1+\theta_2}{2}\sin\frac{\theta_1-\theta_2}{2} &= \cos\frac{\alpha}{2}\sin\frac{p+q}{2}\sin\frac{p-q}{2}. \end{aligned} \tag{80}$$
Hence
$$\tan\frac{\theta_1+\theta_2}{2}\tan\frac{\theta_1-\theta_2}{2} = \tan\frac{p+q}{2}\tan\frac{p-q}{2}.$$
Since $0 < (\theta_1+\theta_2)/2 < \pi/2$, this gives
$$\pi - \frac{\theta_1+\theta_2}{2} = \frac{\pi}{2} + \arctan\left(\frac{\tan\frac{\theta_1-\theta_2}{2}}{\tan\frac{p+q}{2}\tan\frac{p-q}{2}}\right). \tag{81}$$

Subtracting (79) from (78), and then substituting
$$\begin{aligned} p &= \frac{p+q}{2} + \frac{p-q}{2}, & q &= \frac{p+q}{2} - \frac{p-q}{2}, \\ \theta_1 &= \frac{\theta_1+\theta_2}{2} + \frac{\theta_1-\theta_2}{2}, & \theta_2 &= \frac{\theta_1+\theta_2}{2} - \frac{\theta_1-\theta_2}{2}, \end{aligned}$$
gives the following identity:
$$\begin{aligned} &(\pi-p-q)\sin(p+q)\Big[2\cos\frac{\theta_1+\theta_2}{2}\sin\frac{\theta_1-\theta_2}{2} \\ &\quad - 2\left(\pi - \frac{\theta_1+\theta_2}{2}\right)\sin\frac{\theta_1+\theta_2}{2}\sin\frac{\theta_1-\theta_2}{2} \\ &\quad - (\theta_1-\theta_2)\cos\frac{\theta_1+\theta_2}{2}\cos\frac{\theta_1-\theta_2}{2}\Big] \\ &= \cos\frac{\alpha}{2}\big(\sin(p+q) + (\pi-p-q)\cos(p+q) - \pi\big)\Big[2\left(\pi - \frac{\theta_1+\theta_2}{2}\right)\cos\frac{p+q}{2}\sin\frac{p-q}{2} \\ &\quad - (\theta_1-\theta_2)\sin\frac{p+q}{2}\cos\frac{p-q}{2}\Big]. \end{aligned}$$
Using (80) in this identity and collecting the terms containing $\pi - (\theta_1+\theta_2)/2$, we get
$$\begin{aligned} &\left(\pi - \frac{\theta_1+\theta_2}{2}\right)\big(p+q-\sin(p+q)\big)\cos\frac{p+q}{2}\tan\frac{p-q}{2} \\ &= \sin\frac{p+q}{2}\Big[\big(p+q-\sin(p+q)\big)\frac{\theta_1-\theta_2}{2} \\ &\qquad + 2(\pi-p-q)\left(\frac{\theta_1-\theta_2}{2} - \cos^2\frac{p+q}{2}\tan\frac{\theta_1-\theta_2}{2}\right)\Big]. \end{aligned}$$

Therefore

$$\pi - \frac{\theta_1+\theta_2}{2} = \tan\frac{p+q}{2}\cot\frac{p-q}{2}\left[\frac{\theta_1-\theta_2}{2} + \frac{2(\pi-p-q)}{p+q-\sin(p+q)} \times\left(\frac{\theta_1-\theta_2}{2} - \cos^2\frac{p+q}{2}\tan\frac{\theta_1-\theta_2}{2}\right)\right]. \tag{82}$$

Adding (78) and (79), and making the same substitution, gives

$$\begin{aligned}
&(\pi-p-q)\sin(p+q)\Big[2\sin\frac{\theta_1+\theta_2}{2}\cos\frac{\theta_1-\theta_2}{2} \\
&\quad + 2\left(\pi - \frac{\theta_1+\theta_2}{2}\right)\cos\frac{\theta_1+\theta_2}{2}\cos\frac{\theta_1-\theta_2}{2} \\
&\quad + (\theta_1-\theta_2)\sin\frac{\theta_1+\theta_2}{2}\sin\frac{\theta_1-\theta_2}{2}\Big] \\
&= \cos\frac{\alpha}{2}\big(\sin(p+q) + (\pi-p-q)\cos(p+q)+\pi\big)\Big[2\left(\pi-\frac{\theta_1+\theta_2}{2}\right)\sin\frac{p+q}{2}\cos\frac{p-q}{2} \\
&\quad - (\theta_1-\theta_2)\cos\frac{p+q}{2}\sin\frac{p-q}{2}\Big].
\end{aligned}$$

Using (80) again and collecting $\pi - (\theta_1+\theta_2)/2$, we obtain

$$\begin{aligned}
&\left(\pi - \frac{\theta_1+\theta_2}{2}\right)\big(p+q+\sin(p+q)\big)\sin\frac{p+q}{2}\cot\frac{p-q}{2} \\
&= \cos\frac{p+q}{2}\Big[\big(p+q+\sin(p+q)\big)\frac{\theta_1-\theta_2}{2} \\
&\qquad + 2(\pi-p-q)\left(\frac{\theta_1-\theta_2}{2} + \sin^2\frac{p+q}{2}\cot\frac{\theta_1-\theta_2}{2}\right)\Big].
\end{aligned}$$

Thus

$$\pi - \frac{\theta_1+\theta_2}{2} = \cot\frac{p+q}{2}\tan\frac{p-q}{2}\left[\frac{\theta_1-\theta_2}{2} + \frac{2(\pi-p-q)}{p+q+\sin(p+q)} \times\left(\frac{\theta_1-\theta_2}{2} + \sin^2\frac{p+q}{2}\cot\frac{\theta_1-\theta_2}{2}\right)\right]. \tag{83}$$

Apply lemma 23 with

$$t = \frac{\theta_1-\theta_2}{2}, \qquad m = \frac{p+q}{2}, \qquad \mathcal{A} = \frac{\tan\frac{p+q}{2}}{\tan\frac{p-q}{2}}.$$

The inequalities above verify the range assumptions. Comparing (82) with (81) gives exactly the hypothesis of lemma 23:

$$\begin{aligned}
&\frac{\tan\frac{p+q}{2}}{\tan\frac{p-q}{2}}\left[\frac{\theta_1-\theta_2}{2} + \frac{2(\pi-p-q)}{p+q-\sin(p+q)}\left(\frac{\theta_1-\theta_2}{2} - \cos^2\frac{p+q}{2}\tan\frac{\theta_1-\theta_2}{2}\right)\right] \\
&= \frac{\pi}{2} + \arctan\left(\frac{\left(\tan\frac{p+q}{2}/\tan\frac{p-q}{2}\right)\tan\frac{\theta_1-\theta_2}{2}}{\tan^2\frac{p+q}{2}}\right).
\end{aligned}$$

Therefore the lemma gives

$$\pi - \frac{\theta_1+\theta_2}{2} > \frac{\tan\frac{p-q}{2}}{\tan\frac{p+q}{2}}\left[\frac{\theta_1-\theta_2}{2} + \frac{2(\pi-p-q)}{p+q+\sin(p+q)} \times\left(\frac{\theta_1-\theta_2}{2} + \sin^2\frac{p+q}{2}\cot\frac{\theta_1-\theta_2}{2}\right)\right]$$
$$= \cot\frac{p+q}{2}\tan\frac{p-q}{2}\left[\frac{\theta_1-\theta_2}{2} + \frac{2(\pi-p-q)}{p+q+\sin(p+q)} \times\left(\frac{\theta_1-\theta_2}{2} + \sin^2\frac{p+q}{2}\cot\frac{\theta_1-\theta_2}{2}\right)\right],$$

which contradicts (83). Thus $p > q$ is impossible; by symmetry $q > p$ is impossible as well. Hence $p = q$.

We have shown $\bar{w}_1 = (\cos p, 0, \sin p)$, $\bar{w}_2 = (\cos p, 0, -\sin p)$, and $0 < p < \pi/2$. It remains only to compare the lengths. Now $\theta_1 = \theta_2$ and $\gamma = 2p$. Taking the third coordinate in (74) for $i = 1$ and $i = 2$ gives

$$\begin{aligned} 0 &= 2\sin\theta_1 - \pi r_1 + \big((\pi-\gamma) - \sin\gamma\big) r_2, \\ 0 &= 2\sin\theta_1 - \pi r_2 + \big((\pi-\gamma) - \sin\gamma\big) r_1. \end{aligned} \tag{84}$$

Subtracting the two equations in (84) yields $\big(2\pi - \gamma - \sin\gamma\big)(r_1 - r_2) = 0$. The coefficient is positive for $0 < \gamma < \pi$, so $r_1 = r_2$. Thus w_1 and w_2 are mirror images across the teacher plane.

Finally compare $\bar{w}_1 = (\cos p, 0, \sin p)$ and $\bar{w}_2 = (\cos p, 0, -\sin p)$ with the spherical coordinates used in the main text,

$$\bar{w}_i = (\cos\phi_i \sin\beta_i,\ \sin\phi_i \sin\beta_i,\ \cos\beta_i).$$

Because both directions have zero second coordinate and positive first coordinate, $\phi_1 = \phi_2 = 0$. The third coordinates then give

$$\beta_1 = \frac{\pi}{2} - p, \qquad \beta_2 = \frac{\pi}{2} + p,$$

and hence $\beta_1 + \beta_2 = \pi$. This proves the lemma.

■

Lemma 25 (Out-of-plane stationary points) *Let $w_1, w_2 \neq 0$ lie in $\operatorname{span}\{u_1, u_3\}$ but not both in the teacher plane. Then (w_1, w_2) is a stationary point if and only if, after possibly replacing u_3 by $-u_3$, there exist $r > 0$ and $\kappa \in (0, \pi/2)$ such that*

$$w_1 = r(\cos\kappa\,\bar{v} + \sin\kappa\,u_3), \qquad w_2 = r(\cos\kappa\,\bar{v} - \sin\kappa\,u_3).$$

Every such stationary point is a saddle.

B.2.6. PROOF OF LEMMA 25

Proof By lemma 24, after possibly replacing u_3 by $-u_3$, the out-of-plane solution has the form

$$w_1 = r(\cos\kappa\,u_1 + \sin\kappa\,u_3), \qquad w_2 = r(\cos\kappa\,u_1 - \sin\kappa\,u_3)$$

for some $r > 0$ and $\kappa \in (0, \pi/2)$. Let θ denote the common angle between each student direction and each teacher direction. Then

$$\cos\theta = \cos\kappa \cos\frac{\alpha}{2}. \tag{85}$$

Recalling that $\gamma = 2\kappa$, from (43) we have

$$(\pi - \theta)(v_1 + v_2) + \big(2\sin\theta - \pi r - \sin(2\kappa) r\big)\bar{w}_1 - (\pi - 2\kappa) r \bar{w}_2 = 0. \tag{86}$$

By the above equation, observing that the coordinate orthogonal to the xy-plane (the plane spanned by v_1 and v_2) equals zero, we have

$$\pi r + \sin(2\kappa) r - 2\sin\theta = \pi r - 2\kappa r.$$

Hence

$$r = \frac{2\sin\theta}{\sin 2\kappa + 2\kappa}. \tag{87}$$

With this substitution, (86) becomes

$$(\pi - \theta)(v_1 + v_2) - 2(\pi - 2\kappa)\frac{2\sin\theta}{\sin 2\kappa + 2\kappa}(\bar{w}_1 + \bar{w}_2) = 0.$$

Note that $\|v_1 + v_2\|_2^2 = 4\cos^2\frac{\alpha}{2}$ and $(\bar{w}_1 + \bar{w}_2) \cdot (v_1 + v_2) = 4\cos\frac{\alpha}{2}\cos\kappa$. Taking the dot product of the above equation with $(v_1 + v_2)$, we obtain

$$(\pi - \theta)\cos\frac{\alpha}{2} - (\pi - 2\kappa)\frac{2\sin\theta}{\sin 2\kappa + 2\kappa}\cos\kappa = 0. \tag{88}$$

And next we are going to show that the Hessian of the loss function has a negative eigenvalue at this point.

Let

$$h_1(w, v) = \frac{\sin(\theta_{w,v})\|v\|_2}{2\pi\|w\|_2}\Big(I - \bar{w}\bar{w}^\top + \bar{n}_{v,w}\bar{n}_{v,w}^\top\Big),$$
$$h_2(w, v) = \frac{1}{2\pi}\Big((\pi - \theta_{w,v})I + \bar{n}_{w,v}\bar{v}^\top + \bar{n}_{v,w}\bar{w}^\top\Big),$$

by Theorem 5 in [33], we have

$$\frac{\partial^2 L}{\partial w_1^2} = \frac{1}{2}I + h_1(w_1, w_2) - \sum_{j=1}^{2} h_1(w_1, v_j),$$
$$\frac{\partial^2 L}{\partial w_2^2} = \frac{1}{2}I + h_1(w_2, w_1) - \sum_{j=1}^{2} h_1(w_2, v_j),$$
$$\frac{\partial^2 L}{\partial w_1 \partial w_2} = h_2(w_1, w_2).$$

Let $u := \dfrac{v_1 - v_2}{\|v_1 - v_2\|_2}$ and use $(u, -u)$ as the perturbation direction. We will prove that

$$(u, -u)\Big(\nabla^2 L(w_1, w_2)\Big)(u, -u)^\top < 0.$$

Since $\|v_1 - v_2\|_2 = 2\sin\frac{\alpha}{2}$, we have

$$u \cdot v_1 = \sin\frac{\alpha}{2}, \quad u \cdot v_2 = -\sin\frac{\alpha}{2}.$$

By symmetry, u is orthogonal to the angular bisector and w_1, w_2 lie in the perpendicular plane, so

$$u \cdot \bar{w}_1 = u \cdot \bar{w}_2 = 0, \quad u \cdot \bar{n}_{w_j,w_i} = 0 \quad \text{for } i \neq j.$$

$$\begin{aligned}
u \cdot \bar{n}_{v_1,w_1} &= \frac{\sin(\alpha/2)}{\sin\theta}, \quad u \cdot \bar{n}_{v_2,w_1} = -\frac{\sin(\alpha/2)}{\sin\theta}, \\
u \cdot \bar{n}_{v_1,w_2} &= \frac{\sin(\alpha/2)}{\sin\theta}, \quad u \cdot \bar{n}_{v_2,w_2} = -\frac{\sin(\alpha/2)}{\sin\theta}.
\end{aligned}$$

Next we will calculate the terms of quadratic form $u^\top \nabla^2 L(w_1, w_2) u$ separately.

We start with the identity term.

$$u^\top \Big(\frac{1}{2}I\Big)u + (-u)^\top \Big(\frac{1}{2}I\Big)(-u) = \frac{1}{2}\|u\|_2^2 + \frac{1}{2}\|u\|_2^2 = 1.$$

For the $h_1(w_i, w_j)$ terms with $i \neq j$, using $\gamma = 2\kappa$, $\|w_1\|_2 = \|w_2\|_2$, and $u \perp \bar{w}_i$, $u \perp \bar{n}_{w_j,w_i}$, we obtain

$$u^\top h_1(w_1, w_2)u = \frac{\sin\gamma\|w_2\|_2}{2\pi\|w_1\|_2}\|u\|_2^2 = \frac{\sin(2\kappa)}{2\pi},$$

and similarly $(-u)^\top h_1(w_2, w_1)(-u) = \frac{\sin(2\kappa)}{2\pi}$.

For the $-h_1(w_i, v_j)$ terms, each fixed $i \in \{1, 2\}$ and $j \in \{1, 2\}$ gives

$$u^\top h_1(w_i, v_j)u = \frac{\sin\theta}{2\pi\|w_i\|_2}\Big(\|u\|_2^2 + (u \cdot \bar{n}_{v_j,w_i})^2\Big) = \frac{1}{2\pi\|w_1\|_2}\Big(\sin\theta + \frac{\sin^2(\alpha/2)}{\sin\theta}\Big),$$

For the $h_2(w_i, w_j)$ terms, using $u \perp \bar{w}_i$ and $u \perp \bar{n}_{w_j,w_i}$, we have

$$u^\top h_2(w_1, w_2)(-u) = \frac{1}{2\pi}(\pi - \gamma)u^\top(-u) = -\frac{\pi - 2\kappa}{2\pi},$$

and symmetrically $(-u)^\top h_2(w_2, w_1)u = -\frac{\pi - 2\kappa}{2\pi}$.

In summary, we have

$$\begin{aligned}
(u, -u)\nabla^2 L(w_1, w_2)(u, -u)^\top &= u^\top(\frac{1}{2}I)u + (-u)^\top(\frac{1}{2}I)(-u) \\
&\quad + u^\top h_1(w_1, w_2)u + (-u)^\top h_1(w_2, w_1)(-u) \\
&\quad - u^\top h_1(w_1, v_1)u - (-u)^\top h_1(w_1, v_2)(-u) \\
&\quad - u^\top h_1(w_2, v_1)u - (-u)^\top h_1(w_2, v_2)(-u) \\
&\quad + u^\top h_2(w_1, w_2)(-u) + (-u)^\top h_2(w_2, w_1)u \\
&= 1 + 2 \times \frac{\sin(2\kappa)}{2\pi} - 4 \times \frac{1}{2\pi\|w_1\|_2}\Big(\sin\theta + \frac{\sin^2(\alpha/2)}{\sin\theta}\Big)
\end{aligned}$$

$$- 2 \times \frac{\pi - 2\kappa}{2\pi}.$$

Note that we have

$$\frac{2}{\pi \|w_1\|_2} \sin\theta = \frac{\sin(2\kappa) + 2\kappa}{\pi},$$

and thus we have

$$\begin{aligned} Q &= 1 + 2 \times \frac{\sin(2\kappa)}{2\pi} - \frac{\sin(2\kappa) + 2\kappa}{\pi} - \frac{\sin(2\kappa) + 2\kappa}{\pi} \cdot \frac{\sin^2(\alpha/2)}{\sin^2\theta} \\ &\quad - \frac{\pi - 2\kappa}{\pi} \\ &= -\frac{\sin(2\kappa) + 2\kappa}{\pi} \cdot \frac{\sin^2(\alpha/2)}{\sin^2\theta} \leq 0. \end{aligned}$$

And thus we have $(u, -u)\nabla^2 L(w_1, w_2)(u, -u)^\top < 0$, which implies that there is a negative eigenvalue of the Hessian matrix at the critical point, so the critical point is a saddle point. ■

Lemma 26 (In-plane stationary points) *Let*

$$r_+ = \frac{2}{\pi}\left(\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right), \quad r_- = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right).$$

If $w_1, w_2 \neq 0$ lie in $\mathrm{span}\{v_1, v_2\}$, then the stationary points are exactly the global minima (v_1, v_2) up to permutation and the collapsed saddle configurations along $\pm\bar{v}$ with total radii r_+ or r_-. Every non-global in-plane stationary point is a saddle.

B.2.7. PROOF OF LEMMA 26

Proof By (44) and (45), we have

$$\begin{aligned} &(\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2 \\ &\quad + \left(2\sin\frac{\theta_{11} + \theta_{12}}{2}\cos\frac{\theta_{11} - \theta_{12}}{2} - \pi\|w_1\|_2 - \sin\gamma\|w_2\|_2\right)\bar{w}_1 \\ &\quad - (\pi - \gamma)\|w_2\|_2\bar{w}_2 = 0, \end{aligned} \tag{89}$$

$$\begin{aligned} &(\pi - \theta_{21})v_1 + (\pi - \theta_{22})v_2 \\ &\quad + \left(2\sin\frac{\theta_{21} + \theta_{22}}{2}\cos\frac{\theta_{21} - \theta_{22}}{2} - \pi\|w_2\|_2 - \sin\gamma\|w_1\|_2\right)\bar{w}_2 \\ &\quad - (\pi - \gamma)\|w_1\|_2\bar{w}_1 = 0. \end{aligned} \tag{90}$$

Constructing the 2 dimensional space as $e_1 = \bar{v}$ and $e_2 = \frac{v_1 - v_2}{\|v_1 - v_2\|_2}$. Let

$$\begin{aligned} v_1 &= (\cos(\alpha/2), \sin(\alpha/2)), \\ v_2 &= (\cos(\alpha/2), -\sin(\alpha/2)), \\ \bar{w}_1 &= (\cos\theta_1, \sin\theta_1), \end{aligned}$$

$$\bar{w}_2 = (\cos\theta_2, \sin\theta_2),$$

where $\theta_i \in [0, 2\pi), i = 1, 2$. By (89) and (90), we have

$$\begin{aligned}
&(2\pi - \theta_{11} - \theta_{12}) \cos\frac{\alpha}{2} \\
&\quad + \left(2\sin\frac{\theta_{11}+\theta_{12}}{2}\cos\frac{\theta_{11}-\theta_{12}}{2} - \pi\|w_1\|_2 - \sin\gamma\|w_2\|_2\right)\cos\theta_1 \\
&\quad - (\pi - \gamma)\|w_2\|_2 \cos\theta_2 = 0, && (91)\\
&(\theta_{12} - \theta_{11}) \sin\frac{\alpha}{2} \\
&\quad + \left(2\sin\frac{\theta_{11}+\theta_{12}}{2}\cos\frac{\theta_{11}-\theta_{12}}{2} - \pi\|w_1\|_2 - \sin\gamma\|w_2\|_2\right)\sin\theta_1 \\
&\quad - (\pi - \gamma)\|w_2\|_2 \sin\theta_2 = 0, && (92)\\
&(2\pi - \theta_{21} - \theta_{22}) \cos\frac{\alpha}{2} \\
&\quad + \left(2\sin\frac{\theta_{21}+\theta_{22}}{2}\cos\frac{\theta_{21}-\theta_{22}}{2} - \pi\|w_2\|_2 - \sin\gamma\|w_1\|_2\right)\cos\theta_2 \\
&\quad - (\pi - \gamma)\|w_1\|_2 \cos\theta_1 = 0, && (93)\\
&(\theta_{22} - \theta_{21}) \sin\frac{\alpha}{2} \\
&\quad + \left(2\sin\frac{\theta_{21}+\theta_{22}}{2}\cos\frac{\theta_{21}-\theta_{22}}{2} - \pi\|w_2\|_2 - \sin\gamma\|w_1\|_2\right)\sin\theta_2 \\
&\quad - (\pi - \gamma)\|w_1\|_2 \sin\theta_1 = 0. && (94)
\end{aligned}$$

If $\gamma \neq 0$ and $\gamma \neq \pi$, the four equations are:

$$(\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha = (\pi r_1 + r_2\sin\gamma - \sin\theta_{11} - \sin\theta_{12})\cos\theta_{11} + (\pi - \gamma)\cos\theta_{21} r_2, \tag{95}$$

$$(\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12}) = (\pi r_1 + r_2\sin\gamma - \sin\theta_{11} - \sin\theta_{12})\cos\theta_{12} + (\pi - \gamma)\cos\theta_{22} r_2, \tag{96}$$

$$(\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha = (\pi r_2 + r_1\sin\gamma - \sin\theta_{21} - \sin\theta_{22})\cos\theta_{21} + (\pi - \gamma)\cos\theta_{11} r_1, \tag{97}$$

$$(\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22}) = (\pi r_2 + r_1\sin\gamma - \sin\theta_{21} - \sin\theta_{22})\cos\theta_{22} + (\pi - \gamma)\cos\theta_{12} r_1, \tag{98}$$

with

$$\begin{aligned}
\theta_{11} &= \min(2\pi - |\theta_1 - \alpha/2|, |\theta_1 - \alpha/2|), \\
\theta_{12} &= \min(2\pi - |\theta_1 + \alpha/2|, |\theta_1 + \alpha/2|), \\
\theta_{21} &= \min(2\pi - |\alpha/2 - \theta_2|, |\alpha/2 - \theta_2|), \\
\theta_{22} &= \min(2\pi - |\alpha/2 + \theta_2|, |\alpha/2 + \theta_2|), \\
\gamma &= \min(|\theta_1 - \theta_2|, 2\pi - |\theta_1 - \theta_2|),
\end{aligned}$$

and $\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22}, \gamma \in [0, \pi]$. By combining equations (95) and (96), we can derive a system of linear equations to solve for the unknowns r_1 and r_2. We assume that $\theta_{11} \neq \theta_{12}$; if this condition fails, the two equations become linearly dependent and the system becomes underdetermined.

$$\pi \cos\theta_{11} r_1 + \big(\sin\gamma\cos\theta_{11} + (\pi-\gamma)\cos\theta_{21}\big) r_2 = b_1,$$
$$\pi \cos\theta_{12} r_1 + \big(\sin\gamma\cos\theta_{12} + (\pi-\gamma)\cos\theta_{22}\big) r_2 = b_2,$$

where

$$b_1 = (\pi-\theta_{11}) + (\pi-\theta_{12})\cos\alpha + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{11},$$
$$b_2 = (\pi-\theta_{11})\cos\alpha + (\pi-\theta_{12}) + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{12}.$$

Note that we have

$$\begin{aligned}
&\pi\cos\theta_{11}\big(\sin\gamma\cos\theta_{12} + (\pi-\gamma)\cos\theta_{22}\big) - \pi\cos\theta_{12}\big(\sin\gamma\cos\theta_{11} + (\pi-\gamma)\cos\theta_{21}\big) \\
&= \pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big).
\end{aligned}$$

And solving the linear system, we can get the values of r_1 and r_2:

$$\begin{aligned}
r_1 &= \frac{b_1\big(\sin\gamma\cos\theta_{12} + (\pi-\gamma)\cos\theta_{22}\big) - b_2\big(\sin\gamma\cos\theta_{11} + (\pi-\gamma)\cos\theta_{21}\big)}{\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)}, \\
r_2 &= \frac{\pi\cos\theta_{11} b_2 - \pi\cos\theta_{12} b_1}{\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)} \\
&= \frac{b_1\cos\theta_{12} - b_2\cos\theta_{11}}{(\pi-\gamma)\big(\cos\theta_{21}\cos\theta_{12} - \cos\theta_{22}\cos\theta_{11}\big)}.
\end{aligned}$$

Similarly, we can solve for r_1 and r_2 from equations (97) and (98):

$$\pi \cos\theta_{21} r_2 + \big(\sin\gamma\cos\theta_{21} + (\pi-\gamma)\cos\theta_{11}\big) r_1 = c_1,$$
$$\pi \cos\theta_{22} r_2 + \big(\sin\gamma\cos\theta_{22} + (\pi-\gamma)\cos\theta_{12}\big) r_1 = c_2,$$

where

$$c_1 = (\pi-\theta_{21}) + (\pi-\theta_{22})\cos\alpha + (\sin\theta_{21} + \sin\theta_{22})\cos\theta_{21},$$
$$c_2 = (\pi-\theta_{21})\cos\alpha + (\pi-\theta_{22}) + (\sin\theta_{21} + \sin\theta_{22})\cos\theta_{22}.$$

The determinant again equals $\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)$. Solving the linear system, we can get the values of r_1 and r_2:

$$\begin{aligned}
r_1 &= \frac{\pi\big(c_1\cos\theta_{22} - c_2\cos\theta_{21}\big)}{\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)} \\
&= \frac{\big((\pi-\theta_{21}) + (\pi-\theta_{22})\cos\alpha\big)\cos\theta_{22} - \big((\pi-\theta_{21})\cos\alpha + (\pi-\theta_{22})\big)\cos\theta_{21}}{(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)}, \\
r_2 &= \frac{c_1\big(\sin\gamma\cos\theta_{22} + (\pi-\gamma)\cos\theta_{12}\big) - c_2\big(\sin\gamma\cos\theta_{21} + (\pi-\gamma)\cos\theta_{11}\big)}{\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big)}.
\end{aligned}$$

By equating the two expressions for r_1 obtained from previous steps and multiplying both sides by $\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22}-\cos\theta_{12}\cos\theta_{21}\big)$, we can get the following equation:

$$\begin{aligned}E_1(\theta_1,\theta_2) = &\Big((\pi-\theta_{11})+(\pi-\theta_{12})\cos\alpha+(\sin\theta_{11}+\sin\theta_{12})\cos\theta_{11}\Big)\\ &\cdot\Big(\sin\gamma\cos\theta_{12}+(\pi-\gamma)\cos\theta_{22}\Big)\\ &-\Big((\pi-\theta_{11})\cos\alpha+(\pi-\theta_{12})+(\sin\theta_{11}+\sin\theta_{12})\cos\theta_{12}\Big)\\ &\cdot\Big(\sin\gamma\cos\theta_{11}+(\pi-\gamma)\cos\theta_{21}\Big)\\ &-\pi\Big(\big((\pi-\theta_{21})+(\pi-\theta_{22})\cos\alpha\big)\cos\theta_{22}\\ &\quad-\big((\pi-\theta_{21})\cos\alpha+(\pi-\theta_{22})\big)\cos\theta_{21}\Big)\\ =&\,0.\end{aligned}$$

By equating the two expressions for r_2 obtained from previous steps and multiplying both sides by $\pi(\pi-\gamma)\big(\cos\theta_{11}\cos\theta_{22}-\cos\theta_{12}\cos\theta_{21}\big)$, we can get the following equation:

$$\begin{aligned}E_2(\theta_1,\theta_2) = &\Big(\sin\gamma\cos\theta_{21}+(\pi-\gamma)\cos\theta_{11}\Big)\\ &\cdot\Big((\pi-\theta_{21})\cos\alpha+(\pi-\theta_{22})+(\sin\theta_{21}+\sin\theta_{22})\cos\theta_{22}\Big)\\ &-\Big(\sin\gamma\cos\theta_{22}+(\pi-\gamma)\cos\theta_{12}\Big)\\ &\cdot\Big((\pi-\theta_{21})+(\pi-\theta_{22})\cos\alpha+(\sin\theta_{21}+\sin\theta_{22})\cos\theta_{21}\Big)\\ &+\pi\Big(\big((\pi-\theta_{11})+(\pi-\theta_{12})\cos\alpha\big)\cos\theta_{12}\\ &\quad-\big((\pi-\theta_{11})\cos\alpha+(\pi-\theta_{12})\big)\cos\theta_{11}\Big)\\ =&\,0.\end{aligned}$$

To find critical points in the plane spanned by v_1 and v_2, we solve the system $E_1(\theta_1,\theta_2)=0$ and $E_2(\theta_1,\theta_2)=0$ simultaneously. We plot $E_1(\theta_1,\theta_2)^2+E_2(\theta_1,\theta_2)^2$ as a function of θ_1 and θ_2, marking in red all points where $E_1(\theta_1,\theta_2)^2+E_2(\theta_1,\theta_2)^2\le 10^{-4}$. From our earlier analysis, when $\theta_1=\theta_2$, the two equations become linearly dependent. Therefore, we mark the line $\theta_1-\theta_2=2k\pi$ for $k\in\mathbb{Z}$ in cyan and exclude these points when identifying critical points in the plane spanned by v_1 and v_2. And we will additionally find the critical points when θ_1 and θ_2 theoretically. Hence the only possible critical points when $\theta_1=\theta_2$ are $(\theta_1,\theta_2)=(0,0)$ and $(\theta_1,\theta_2)=(\pi,\pi)$.

Lemma 27 *Let $w_1,w_2\in\operatorname{span}\{v_1,v_2\}$ and (w_1,w_2) is a critical point. If $\gamma=0$, then we have*

$$\bar{w}_1=\bar{w}_2=\bar{v},\quad \|w_1\|_2+\|w_2\|_2=\frac{2}{\pi}\bigg((\pi-\frac{\alpha}{2})\cos\frac{\alpha}{2}+\sin\frac{\alpha}{2}\bigg)$$

or

$$\bar{w}_1=\bar{w}_2=-\bar{v},\quad \|w_1\|_2+\|w_2\|_2=\frac{2}{\pi}\bigg(\sin\frac{\alpha}{2}-\frac{\alpha}{2}\cos\frac{\alpha}{2}\bigg).$$

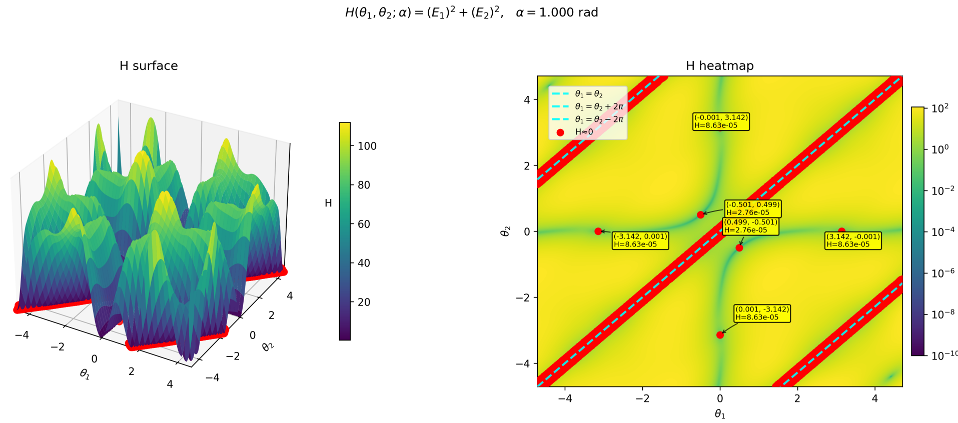

Figure 8: Graphical illustration of the critical points in the plane spanned by v_1 and v_2, where red points show solutions satisfying $E_1(\theta_1, \theta_2)^2 + E_2(\theta_1, \theta_2)^2 \leq 10^{-4}$ with $\alpha = 1$. From our theoretical analysis when $\theta_1 = \theta_2$, the only possible critical points are $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (\pi, \pi)$. The graph shows all critical points except those on the line $\theta_1 - \theta_2 = 2k\pi$ for $k \in \mathbb{Z}$. Therefore, the complete set of critical points includes $(\theta_1, \theta_2) = (0, 0)$, $(\theta_1, \theta_2) = (\pi, \pi)$, $(\theta_1, \theta_2) = (0, \pi)$, $(\theta_1, \theta_2) = (\pi, 0)$, $(\theta_1, \theta_2) = (\frac{\alpha}{2}, -\frac{\alpha}{2})$, and $(\theta_1, \theta_2) = (-\frac{\alpha}{2}, \frac{\alpha}{2})$.

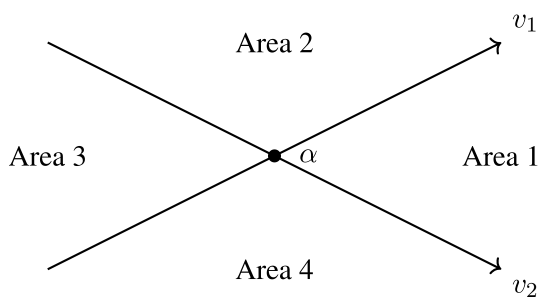

Figure 9: Four areas formed by two intersecting lines with vectors v_1 and v_2 and angle α

If $\gamma = \pi$, we have

$$\bar{w}_1 = \bar{v}, \quad \bar{w}_2 = -\bar{v},$$
$$\|w_1\|_2 = \frac{2}{\pi}\left(\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right),$$
$$\|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right),$$

or

$$\bar{w}_1 = -\bar{v}, \quad \bar{w}_2 = -\bar{v},$$
$$\|w_1\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right),$$

$$\|w_2\|_2 = \frac{2}{\pi}\left(\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right),$$

Proof [Proof of Lemma 27] If $\gamma = 0$, then $\bar{w}_1 = \bar{w}_2, \theta_{11} = \theta_{12}, \theta_{21} = \theta_{22}$. (89) and (90) turn to

$$(\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2 + \left(2\sin\frac{\theta_{11}+\theta_{12}}{2}\cos\frac{\theta_{11}-\theta_{12}}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\bar{w}_1 = 0. \tag{99}$$

Case 1: $\gamma = 0$ and w_1 lies in Area 1 in Figure 9. We have $\theta_{11} + \theta_{12} = \alpha$. Without loss of generosity, we assume $\theta_{11} \le \theta_{12}$. Let $\theta_{12} - \theta_{11} = p$, then $\bar{w}_1 = (\cos\frac{p}{2}, \sin\frac{p}{2})$. We have

$$(2\pi - \alpha)\cos\frac{\alpha}{2} + \left(2\sin\frac{\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\cos\frac{p}{2} = 0, \tag{100}$$

$$p\sin\frac{\alpha}{2} + \left(2\sin\frac{\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\sin\frac{p}{2} = 0. \tag{101}$$

If $p \neq 0$, by (101), we have $2\sin\frac{\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big) \neq 0$. Then, we can get

$$\begin{aligned}\frac{(2\pi - \alpha)\cos\frac{\alpha}{2}}{p\sin\frac{\alpha}{2}} &= \frac{-\left(2\sin\frac{\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\cos\frac{p}{2}}{-\left(2\sin\frac{\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\sin\frac{p}{2}} \\ &= \cot\frac{p}{2}.\end{aligned}$$

Then, we have $\left(\pi - \frac{\alpha}{2}\right)\cot\frac{\alpha}{2} = \frac{p}{2}\cot\frac{p}{2}$. Let $g(x) = x\cot x, x \in (0, \pi)$. Then

$$g'(x) = \frac{(\cos x - x\sin x)\sin x - x\cos x \cdot \cos x}{\sin^2 x} = \frac{\frac{1}{2}\sin 2x - x}{\sin^2 x} < 0.$$

Thus, we have $(\pi - \frac{\alpha}{2})\cot\frac{\alpha}{2} = \frac{p}{2}\cot\frac{p}{2} \le \lim_{x\to 0^+} g(x) = 1$. Then, we have $(\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} \le \sin\frac{\alpha}{2}$. Recall that $\alpha \in (0, \frac{\pi}{2}]$, we have $(\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} \ge (\pi - \frac{\pi}{4})\cos\frac{\pi}{4} > \sin\frac{\pi}{4} \ge \sin\frac{\alpha}{2}$, which contradicts with $(\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} \le \sin\frac{\alpha}{2}$. Hence, we have $p = 0$. Back to (100), we have

$$\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\left(\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right).$$

Case 2: $\gamma = 0$ and w_1 lies in Area $2, 3, 4$ in Figure 9. By (99), we have $\bar{w}_1$ lies in Area 3. Then, we have $\theta_{11} + \theta_{12} = 2\pi - \alpha$. Without loss of generosity, we have $\theta_{11} \le \theta_{12}$. Let $\theta_{12} - \theta_{11} = p$, we have $\bar{w}_1 = (-\cos\frac{p}{2}, \sin\frac{p}{2})$. From (99), we have

$$\alpha\cos\frac{\alpha}{2} - \left(2\sin\frac{2\pi - \alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\cos\frac{p}{2} = 0, \tag{102}$$

$$p\sin\frac{\alpha}{2} + \left(2\sin\frac{2\pi - \alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\right)\sin\frac{p}{2} = 0.$$

If $p \neq 0$, then $2\sin\frac{2\pi-\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big) \neq 0$. Thus, we have

$$\frac{\alpha\cos\frac{\alpha}{2}}{p\sin\frac{\alpha}{2}} = \frac{\Big(2\sin\frac{2\pi-\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\Big)\cos\frac{p}{2}}{-\Big(2\sin\frac{2\pi-\alpha}{2}\cos\frac{p}{2} - \pi\big(\|w_1\|_2 + \|w_2\|_2\big)\Big)\sin\frac{p}{2}} = -\cot\frac{p}{2}. \tag{103}$$

Then, we get $\alpha\cot\frac{\alpha}{2} = -p\cot\frac{p}{2}$, which implies $\frac{p}{2} > \frac{\pi}{2}$. However, $\bar{w}_1$ lies between the $-v_1$ and $-v_2$ and we have $p \leq \alpha \leq \frac{\pi}{2}$, which contradicts with $\frac{p}{2} > \frac{\pi}{2}$. Hence, we have $p = 0$. Back to (102),we have

$$\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right).$$

Then, we consider the case $\gamma \neq 0$ in the following discussion.

Case 3: $\gamma = \pi$ and w_1 is located in Area 1 and w_2 is located in Area 3 in Figure 9. Without loss of generality, we assume $\theta_1 \leq \theta_2$. Then, we have

$$\begin{aligned}
&\theta_2 = \theta_1 + \pi, \quad \theta_{11} = \frac{\alpha}{2} - \theta_1, \quad \theta_{12} = \frac{\alpha}{2} + \theta_1,\\
&\theta_{11} + \theta_{12} = \alpha, \quad \theta_{12} - \theta_{11} = 2\theta_1,\\
&\theta_{21} = \theta_2 - \frac{\alpha}{2} = \theta_1 + \pi - \frac{\alpha}{2},\\
&\theta_{22} = 2\pi - \theta_2 - \frac{\alpha}{2} = \pi - \theta_1 - \frac{\alpha}{2},\\
&\theta_{21} + \theta_{22} = 2\pi - \alpha, \quad \theta_{22} - \theta_{21} = -2\theta_1.
\end{aligned}$$

Then, (91), (92), (93) and (94) reduces to

$$(2\pi-\alpha)\cos\frac{\alpha}{2} + \left(2\sin\frac{\alpha}{2}\cos\theta_1 - \pi\|w_1\|_2\right)\cos\theta_1 = 0, \tag{104}$$

$$2\theta_1\sin\frac{\alpha}{2} + \left(2\sin\frac{\alpha}{2}\cos\theta_1 - \pi\|w_1\|_2\right)\sin\theta_1 = 0, \tag{105}$$

$$\alpha\cos\frac{\alpha}{2} - \left(2\sin\frac{\alpha}{2}\cos\theta_1 - \pi\|w_2\|_2\right)\cos\theta_1 = 0, \tag{106}$$

$$-2\theta_1\sin\frac{\alpha}{2} - \left(2\sin\frac{\alpha}{2}\cos\theta_1 - \pi\|w_2\|_2\right)\sin\theta_1 = 0. \tag{107}$$

If $\theta_1 = 0$, (104) and (106) turn into

$$\begin{aligned}
&(2\pi-\alpha)\cos\frac{\alpha}{2} + \left(2\sin\frac{\alpha}{2} - \pi\|w_1\|_2\right) = 0,\\
&\alpha\cos\frac{\alpha}{2} - \left(2\sin\frac{\alpha}{2} - \pi\|w_2\|_2\right) = 0.
\end{aligned}$$

Thus, we have

Then we consider the case $\theta_1 \neq 0$. (105)+ (107) yields

$$\pi(\|w_2\|_2 - \|w_1\|_2)\sin\theta_1 = 0. \tag{108}$$

Recall that $0 < |\theta_1| \leq \frac{\alpha}{2} \leq \frac{\pi}{4}$, we have $\|w_1\|_2 = \|w_2\|_2$. Then, (104)+ (106) yields

$$0 = (2\pi - \alpha)\cos\frac{\alpha}{2} + \alpha\cos\frac{\alpha}{2} = 2\pi\cos\frac{\alpha}{2},$$

which contradicts with $\alpha \in (0, \frac{\pi}{2}]$.

Case 4: $\gamma = \pi$ and w_1 is located in Area 2 and w_2 is located in Area 4 in Figure 9. We have

$$\begin{aligned}
&\theta_2 = \theta_1 + \pi,\\
&\theta_{11} = \theta_1 - \frac{\alpha}{2}, \quad \theta_{12} = \theta_1 + \frac{\alpha}{2},\\
&\theta_{11} + \theta_{12} = 2\theta_1, \quad \theta_{12} - \theta_{11} = \alpha,\\
&\theta_{21} = 2\pi - \theta_2 + \frac{\alpha}{2} = \pi - \theta_1 + \frac{\alpha}{2},\\
&\theta_{22} = 2\pi - \theta_2 - \frac{\alpha}{2} = \pi - \theta_1 - \frac{\alpha}{2},\\
&\theta_{21} + \theta_{22} = 2\pi - 2\theta_1, \quad \theta_{22} - \theta_{21} = -\alpha.
\end{aligned}$$

Then, (91), (92), (93) and (94) reduces to

$$(2\pi - 2\theta_1)\cos\frac{\alpha}{2} + \left(2\sin\theta_1\cos\frac{\alpha}{2} - \pi\|w_1\|_2\right)\cos\theta_1 = 0, \tag{109}$$

$$\alpha\sin\frac{\alpha}{2} + \left(2\sin\theta_1\cos\frac{\alpha}{2} - \pi\|w_1\|_2\right)\sin\theta_1 = 0, \tag{110}$$

$$2\theta_1\cos\frac{\alpha}{2} - \left(2\sin\theta_1\cos\frac{\alpha}{2} - \pi\|w_2\|_2\right)\cos\theta_1 = 0, \tag{111}$$

$$-\alpha\sin\frac{\alpha}{2} - \left(2\sin\theta_1\cos\frac{\alpha}{2} - \pi\|w_2\|_2\right)\sin\theta_1 = 0. \tag{112}$$

By (110), we have $\theta_1 \neq 0$. (110)+(112) yileds $\pi(\|w_2\|_2 - \|w_1\|_2)\sin\theta_1 = 0$, which implies $\|w_1\|_2 = \|w_2\|_2$. Taking sum of (109) and (111), wehave $2\pi\cos\frac{\alpha}{2} = 0$, which contradicts with $\alpha \in (0, \frac{\pi}{2}]$. ■

Lemma 28 *Let $w_1, w_2 \in \text{span}\{v_1, v_2\}$, and suppose that (w_1, w_2) satisfies one of the following cases:*

- *both w_1 and w_2 are in Area 2;*
- *both w_1 and w_2 are in Area 4;*
- *w_1 is in Area 2 and w_2 is in Area 3;*
- *w_1 is in Area 3 and w_2 is in Area 4.*

Then (w_1, w_2) *cannot be a critical point.*

Proof [Proof of Lemma 28] Without loss of generality, we assume $\theta_1 \le \theta_2$. Let (w_1, w_2) be a critical point. If both w_1 and w_2 lies in Area 2, then we have $\alpha = \theta_{12} - \theta_{11}, \theta_{21} = \gamma + \theta_{11}, \theta_{22} = \gamma + \theta_{12}$. By (12) and (14), we have

$$\begin{aligned}
0 &= \frac{\mathrm{d}\theta_{11}}{\mathrm{d}t} \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11}) \sin^2\theta_{11} + (\pi - \theta_{12})(\cos(\theta_{12} - \theta_{11}) - \cos\theta_{12}\cos\theta_{11}) \\
&\qquad - (\pi - \gamma)\|w_2\|_2 (\cos\left(\gamma + \theta_{11}\right) - \cos\gamma\cos\theta_{11}) \Big) \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11}) \sin^2\theta_{11} + (\pi - \theta_{12})\sin\theta_{11}\sin\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma\sin\theta_{11} \Big),
\end{aligned} \tag{113}$$

$$\begin{aligned}
0 &= \frac{\mathrm{d}\theta_{12}}{\mathrm{d}t} \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11})(\cos(\theta_{12} - \theta_{11}) - \cos\theta_{11}\cos\theta_{12}) + (\pi - \theta_{12}) \sin^2\theta_{12} \\
&\qquad - (\pi - \gamma)\|w_2\|_2 (\cos\left(\gamma + \theta_{12}\right) - \cos\gamma\cos\theta_{12}) \Big) \\
&= \frac{1}{2\pi\|w_1\|_2} \cdot \Big((\pi - \theta_{11})\sin\theta_{11}\sin\theta_{12} + (\pi - \theta_{12}) \sin^2\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma\sin\theta_{12} \Big).
\end{aligned} \tag{114}$$

Thus, we have

$$\begin{aligned}
0 &= (\pi - \theta_{11}) \sin^2\theta_{11} + (\pi - \theta_{12})\sin\theta_{11}\sin\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma\sin\theta_{11} \\
&= \sin\theta_{11} \cdot \Big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma \Big),
\end{aligned} \tag{115}$$

$$\begin{aligned}
0 &= (\pi - \theta_{11})\sin\theta_{11}\sin\theta_{12} + (\pi - \theta_{12}) \sin^2\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma\sin\theta_{12} \\
&= \sin\theta_{12} \cdot \Big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12} + (\pi - \gamma)\|w_2\|_2 \sin\gamma \Big).
\end{aligned} \tag{116}$$

If $\theta_{11} > 0$, then the right side of (115) is positive, which implies $\theta_{11} = 0$. Similarly, by (116), we have $\theta_{12} = 0$, which contradicts with $\alpha > 0$.

If w_1 lies in Area 2 and w_2 lies in Area 3, (115) still holds, which implies $\theta_{11} = 0$. Then we have $\theta_{12} = \alpha, \theta_{22} = 2\pi - \alpha - \gamma$. Then we have $\cos\theta_{22} = \cos(\gamma + \alpha) = \cos(\gamma + \theta_{12})$ and (116) also holds, which implies $\theta_{12} = 0$. We thus arrive at a contradiction.

Using the same argument, we can show that if w_1 lies in Area 4 and w_2 lies in Area 4 or Area 3, then (w_1, w_2) cannot be a critical point. ■

By Lemma 28, only five cases remain to be considered:

- w_1 lies in Area 1 and w_2 lies in Area 1;

- w_1 lies in Area 2 and w_2 lies in Area 1;
- w_1 lies in Area 1 and w_2 lies in Area 3;
- w_1 lies in Area 2 and w_2 lies in Area 4;
- w_1 lies in Area 3 and w_2 lies in Area 3.

Next we are going to discuss each case one by one.

Case 1: $\gamma \neq 0, \pi$ and w_1, w_2 are both lie in Area 1 in Figure 9

If $\gamma \neq 0$ and $\gamma \neq \pi$, by (91) and (92), we have

$$\|w_2\|_2 = \frac{(2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} \sin \theta_1 - (\theta_{12} - \theta_{11}) \sin \frac{\alpha}{2} \cos \theta_1}{(\pi - \gamma) \sin(\theta_1 - \theta_2)} \tag{117}$$

and by (93) and (94), we have

$$\|w_1\|_2 = \frac{(2\pi - \theta_{21} - \theta_{22}) \cos \frac{\alpha}{2} \sin \theta_2 - (\theta_{22} - \theta_{21}) \sin \frac{\alpha}{2} \cos \theta_2}{(\pi - \gamma) \sin(\theta_2 - \theta_1)}. \tag{118}$$

Plugging (118) and (117) to (91) and (93), we have

$$\begin{aligned}
0 = {} & (2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} \cdot (\pi - \gamma) \sin(\theta_1 - \theta_2) \\
& + 2 \sin \frac{\theta_{11} + \theta_{12}}{2} \cos \frac{\theta_{11} - \theta_{12}}{2} \cos \theta_1 \cdot (\pi - \gamma) \sin(\theta_1 - \theta_2) \\
& + \pi \cos \theta_1 \Big((2\pi - \theta_{21} - \theta_{22}) \cos \frac{\alpha}{2} \sin \theta_2 - (\theta_{22} - \theta_{21}) \sin \frac{\alpha}{2} \cos \theta_2 \Big) \\
& - \big(\sin \gamma \cos \theta_1 + (\pi - \gamma) \cos \theta_2 \big) \\
& \quad \cdot \Big((2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} \sin \theta_1 - (\theta_{12} - \theta_{11}) \sin \frac{\alpha}{2} \cos \theta_1 \Big),
\end{aligned} \tag{119}$$

$$\begin{aligned}
0 = {} & (2\pi - \theta_{21} - \theta_{22}) \cos \frac{\alpha}{2} \cdot (\pi - \gamma) \sin(\theta_1 - \theta_2) \\
& + 2 \sin \frac{\theta_{21} + \theta_{22}}{2} \cos \frac{\theta_{21} - \theta_{22}}{2} \cos \theta_2 \cdot (\pi - \gamma) \sin(\theta_1 - \theta_2) \\
& - \pi \cos \theta_2 \Big((2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} \sin \theta_1 - (\theta_{12} - \theta_{11}) \sin \frac{\alpha}{2} \cos \theta_1 \Big) \\
& + \big(\sin \gamma \cos \theta_2 + (\pi - \gamma) \cos \theta_1 \big) \\
& \quad \cdot \Big((2\pi - \theta_{21} - \theta_{22}) \cos \frac{\alpha}{2} \sin \theta_2 - (\theta_{22} - \theta_{21}) \sin \frac{\alpha}{2} \cos \theta_2 \Big).
\end{aligned} \tag{120}$$

Without loss of generality, we assume $\theta_1 \geq \theta_2$. Then, We have $\theta_1 - \theta_2 = \gamma$. Let $\theta_1 + \theta_2 = p$, we have

$$\begin{aligned}
& \theta_1 = \frac{p + \gamma}{2}, \quad \theta_2 = \frac{p - \gamma}{2}, \quad \theta_1 - \theta_2 = \gamma, \\
& \theta_{11} = \frac{\alpha}{2} - \theta_1 = \frac{\alpha - p - \gamma}{2},
\end{aligned}$$

$$
\begin{aligned}
&\theta_{12} = \frac{\alpha}{2} + \theta_1 = \frac{\alpha + p + \gamma}{2}, \\
&\theta_{11} + \theta_{12} = \alpha, \quad \theta_{12} - \theta_{11} = p + \gamma, \\
&\theta_{21} = \gamma + \frac{\alpha}{2} - \theta_1 = \frac{\alpha - p + \gamma}{2}, \\
&\theta_{22} = \frac{\alpha}{2} - \gamma + \theta_1 = \frac{\alpha + p - \gamma}{2}, \\
&\theta_{21} + \theta_{22} = \alpha, \quad \theta_{22} - \theta_{21} = p - \gamma.
\end{aligned} \tag{121}
$$

Let $(\pi - \frac{\alpha}{2}) \cos \frac{\alpha}{2} = A$ and $\sin \frac{\alpha}{2} = B$. Back to (119), we get

$$
\begin{aligned}
0 &= 2(\pi-\gamma)\sin\gamma \cdot A + 2(\pi-\gamma)\sin\gamma\cos^2\frac{p+\gamma}{2}\cdot B \\
&\quad - \pi\cos\frac{p+\gamma}{2}\Big(-2\sin\frac{p-\gamma}{2}\cdot A + (p-\gamma)\cos\frac{p-\gamma}{2}\cdot B\Big) \\
&\quad - \Big(\sin\gamma\cos\frac{p+\gamma}{2} + (\pi-\gamma)\cos\frac{p-\gamma}{2}\Big) \\
&\quad\ \cdot \Big(2\sin\frac{p+\gamma}{2}\cdot A - (p+\gamma)\cos\frac{p+\gamma}{2}\cdot B\Big) \\
&= \Big(2(\pi-\gamma)\sin\gamma + \pi(\sin p - \sin\gamma) - \sin\gamma\sin(p+\gamma) \\
&\quad\ - (\pi-\gamma)(\sin p + \sin\gamma)\Big)\cdot A \\
&\quad + \Big(\big(2(\pi-\gamma) + (p+\gamma)\big)\sin\gamma\cos^2\frac{p+\gamma}{2} \\
&\quad\ + \big(-\pi(p-\gamma) + (\pi-\gamma)(p+\gamma)\big)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\Big)\cdot B \\
&= \Big(2(\pi-\gamma)\sin\gamma - \sin\gamma\sin(p+\gamma) + \gamma\sin p \\
&\quad\ - (2\pi-\gamma)\sin\gamma\Big)\cdot A \\
&\quad + \Big((2\pi+p-\gamma)\sin\gamma\cos^2\frac{p+\gamma}{2} \\
&\qquad + \big(2\pi\gamma - \gamma p - \gamma^2\big)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\Big)\cdot B \\
&= \Big(\gamma(\sin p - \sin\gamma) - \sin\gamma\sin(p+\gamma)\Big)\cdot A \\
&\quad + \Big((2\pi+p-\gamma)\sin\gamma\cdot\frac{1+\cos(p+\gamma)}{2} \\
&\qquad + \big(2\pi\gamma - \gamma p - \gamma^2\big)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\Big)\cdot B \\
&= \Big(\gamma(\sin p - \sin\gamma)
\end{aligned}
$$

$$
\begin{aligned}
&- \sin\gamma(\sin p\cos\gamma + \cos p\sin\gamma)\Big)\cdot A\\
&+\bigg((2\pi+p-\gamma)\sin\gamma\Big(\frac{1+\cos p\cos\gamma}{2}-\frac{\sin p\sin\gamma}{2}\Big)\\
&\quad+\big(2\pi\gamma-\gamma p-\gamma^2\big)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\bigg)\cdot B
\end{aligned}
\tag{122}
$$

Similarly, Back to (120), we have

$$
\begin{aligned}
0 &= 2(\pi-\gamma)\sin\gamma\cdot A + 2(\pi-\gamma)\sin\gamma\cos^2\frac{p-\gamma}{2}\cdot B\\
&\quad-\pi\cos\frac{p-\gamma}{2}\bigg(2\sin\frac{p+\gamma}{2}\cdot A-(p+\gamma)\cos\frac{p+\gamma}{2}\cdot B\bigg)\\
&\quad+\bigg(\sin\gamma\cos\frac{p-\gamma}{2}+(\pi-\gamma)\cos\frac{p+\gamma}{2}\bigg)\\
&\quad\cdot\bigg(2\sin\frac{p-\gamma}{2}\cdot A-(p-\gamma)\cos\frac{p-\gamma}{2}\cdot B\bigg)\\
&=\bigg(-\gamma(\sin\gamma+\sin p)+\sin\gamma(\sin p\cos\gamma-\cos p\sin\gamma)\bigg)\cdot A\\
&\quad+\bigg((2\pi-p-\gamma)\sin\gamma\Big(\frac{1+\cos p\cos\gamma}{2}+\frac{\sin p\sin\gamma}{2}\Big)\\
&\qquad+(2\pi\gamma+\gamma p-\gamma^2)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\bigg)\cdot B
\end{aligned}
\tag{123}
$$

(122)+(123) yields

$$
\begin{aligned}
0 &= (\underbrace{-2\gamma\sin\gamma-2\sin^2\gamma\cos p}_{\text{I}})\cdot A\\
&\quad+\underbrace{\begin{pmatrix}(2\pi-\gamma)\sin\gamma(1+\cos p\cos\gamma)-p\sin p\sin^2\gamma\\ +(4\pi\gamma-2\gamma^2)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\end{pmatrix}}_{\text{II}}\cdot B,
\end{aligned}
\tag{124}
$$

and (122)-(123) yields

$$
\begin{aligned}
0 &= (\underbrace{2\gamma\sin p-2\sin p\sin\gamma\cos\gamma}_{\text{III}})\cdot A\\
&\quad+\underbrace{\begin{pmatrix}p\sin\gamma(1+\cos p\cos\gamma)-(2\pi-\gamma)\sin p\sin^2\gamma\\ -2\gamma p\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\end{pmatrix}}_{\text{IV}}\cdot B.
\end{aligned}
\tag{125}
$$

By (124) and (125), we have $\text{I}\cdot\text{IV}=\text{II}\cdot\text{III}$. By simple calculation, we have

$$
\frac{1}{2}\text{I}\cdot\text{IV}=(2\pi-\gamma)(\gamma\sin^3\gamma\sin p\sin^4\gamma\sin p\cos p)
$$

$$
\begin{aligned}
&- p\Big(\cancel{\gamma\sin^2\gamma} + \sin^3\gamma\cos p \cancel{+\gamma\sin^2\gamma\cos\gamma\cos p} \\
&\quad + \sin^3\gamma\cos\gamma\cos^2 p - \gamma^2\sin\gamma\cos\gamma \cancel{-\gamma\sin^2\gamma\cos\gamma\cos p} \\
&\quad - \gamma^2\sin\gamma\cos p \cancel{-\gamma\sin^2\gamma\cos^2 p}\Big) \\
&= (2\pi-\gamma)(\gamma\sin^3\gamma\sin p + \sin^4\gamma\sin p\cos p) \\
&\quad - p\Big(\gamma\sin^2\gamma\sin^2 p + \sin^3\gamma\cos p + \sin^3\gamma\cos\gamma\cos^2 p \\
&\quad - \gamma^2\sin\gamma(\cos\gamma+\cos p)\Big), \qquad (126)
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2}\mathrm{II}\cdot\mathrm{III} &= (2\pi-\gamma)\Big(\cancel{\gamma\sin\gamma\sin p} - \sin^2\gamma\cos\gamma\sin p \\
&\quad \cancel{+\gamma\sin\gamma\cos\gamma\sin p\cos p} - \sin^2\gamma\cos^2\gamma\sin p\cos p \\
&\quad + \gamma^2\cos\gamma\sin p - \cancel{\gamma\sin\gamma\cos^2\gamma\sin p} \\
&\quad + \gamma^2\sin p\cos p \cancel{-\gamma\sin\gamma\cos\gamma\sin p\cos p}\Big) \\
&\quad + p(-\gamma\sin^2\gamma\sin^2 p + \sin^3\gamma\cos\gamma\sin^2 p) \\
&= (2\pi-\gamma)\Big(\gamma\sin^3\gamma\sin p - \sin^2\gamma\cos\gamma\sin p - \sin^2\gamma\cos^2\gamma\sin p\cos p \\
&\quad + \gamma^2\sin p(\cos\gamma+\cos p)\Big) \\
&\quad + p(-\gamma\sin^2\gamma\sin^2 p + \sin^3\gamma\cos\gamma\sin^2 p). \qquad (127)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
0 =& \frac{1}{2}(\mathrm{II}\cdot\mathrm{III} - \mathrm{I}\cdot\mathrm{IV}) \qquad (128) \\
=& (2\pi-\gamma)\Big(-\sin^2\gamma\cos\gamma\sin p - \sin^2\gamma\sin p\cos p + \gamma^2\sin p(\cos\gamma+\cos p)\Big) \\
&+ p\Big(\sin^3\gamma\cos p + \sin^3\gamma\cos\gamma - \gamma^2\sin\gamma(\cos\gamma+\cos p)\Big) \\
=& (2\pi-\gamma)\Big(\sin p\big(\gamma^2(\cos\gamma+\cos p) - \sin^2\gamma(\cos\gamma+\cos p)\big)\Big) \\
&+ p\Big(\sin\gamma\big(\sin^2\gamma(\cos\gamma+\cos p) - \gamma^2(\cos\gamma+\cos p)\big)\Big) \\
=& (\cos\gamma+\cos p)\Big((2\pi-\gamma)\sin p(\gamma^2-\sin^2\gamma) - p\sin\gamma(\gamma^2-\sin^2\gamma)\Big) \\
=& (\cos\gamma+\cos p)\big((2\pi-\gamma)\sin p - p\sin\gamma\big)(\gamma^2-\sin^2\gamma).
\end{aligned}
$$

If $\cos\gamma+\cos p = 0$, then $|\gamma| = |p| = \frac{\pi}{2}$. Notice that $\frac{\alpha}{2} \geq |p| = \frac{\pi}{2}$, which contradicts with $\alpha < \pi$. By assumption, $\gamma \neq 0$, then $\gamma^2 - \sin^2\gamma \neq 0$. We can get $(2\pi-\gamma)\sin p - p\sin\gamma = 0$. We will show that $p = 0$. If not, then $\sin\gamma = (2\pi-\gamma)\frac{\sin p}{p} > 0$, which implies $\gamma > 0$. Then $|p| = |2\theta-\gamma| \leq \gamma$. Let $g(x) = \frac{\sin x}{x}, x \in (0,\pi]$. we have $g'(x) = \frac{x\cos x - \sin x}{x^2}$. Let $h(x) = x\cos x - \sin x, x \in (0,\pi]$. Then

$h'(x) = -x\sin x \le 0$. Thus $h(x) < h(0) = 0$ and $g'(x) = \frac{h(x)}{x^2} < 0$. Recall $|p| \le \gamma \le \alpha < \pi$, we have

$$\frac{\sin\gamma}{2\pi-\gamma} = \frac{\sin p}{p} = \frac{\sin|p|}{|p|} \ge \frac{\sin\gamma}{\gamma},$$

which contradicts with $2\pi - \gamma > \gamma$ and $\sin\gamma > 0$. Then, we get $p = 0$ and back to (122), we have

$$\begin{aligned}
0 &= -4\big(\pi - \frac{\alpha}{2}\big)\cos\frac{\alpha}{2}\sin\gamma(\gamma+\sin\gamma)\\
&\quad + 4\big(\pi-\frac{\gamma}{2}\big)\sin\frac{\alpha}{2}(\gamma+\sin\gamma)(1+\cos\gamma)\\
&= 4(\gamma+\sin\gamma)\sin\gamma\sin\frac{\alpha}{2}\Big(-\big(\pi-\frac{\alpha}{2}\big)\cot\frac{\alpha}{2} + \big(\pi-\frac{\gamma}{2}\big)\cdot\frac{1+\cos\gamma}{\sin\gamma}\Big)\\
&= 4(\gamma+\sin\gamma)\sin\gamma\sin\frac{\alpha}{2}\big(u(\frac{\gamma}{2}) - u(\frac{\alpha}{2})\big),
\end{aligned}$$

where $u(x) = (\pi - x)\cot x$. Noticing $u(\frac{\gamma}{2}) = u(\frac{\alpha}{2}) > 0$, which impies $\gamma > 0$. When $x \in (0, \frac{\pi}{2})$, we have $u'(x) = \frac{-(\pi-x)-\sin x\cos x}{\sin^2 x} < 0$, which implies $\gamma = \alpha$. Hence, $\bar{w}_1 = v_1$ and $\bar{w}_2 = v_2$. Back to (118) and (117), we have $\|w_1\|_2 = \|w_2\|_2 = 1$, which implies $w_1 = v_1$ and $w_2 = v_2$.

Case 2: w_1 lies in Area 2 and w_2 lies in Area 1 in Figure 9

With $\theta_{12} = \theta_1 + \alpha/2$ and $\theta_{11} = \theta_1 - \alpha/2$,

$$\begin{aligned}
\cos\theta_{12} - \cos\alpha\cos\theta_{11} &= -\sin\alpha\sin\theta_{11},\\
\cos\alpha\cos\theta_{12} - \cos\theta_{11} &= -\sin\alpha\sin\theta_{12}.
\end{aligned}$$

With $\theta_{22} = \alpha/2 + \theta_2$ and $\theta_{21} = \alpha/2 - \theta_2$,

$$\begin{aligned}
\cos\theta_{22} - \cos\alpha\cos\theta_{21} &= \sin\alpha\sin\theta_{21},\\
\cos\alpha\cos\theta_{22} - \cos\theta_{21} &= -\sin\alpha\sin\theta_{22}.
\end{aligned}$$

In the interior of this case, $\theta_1 > \theta_2$, so the representative of the angle between w_1 and w_2 is $0 < \gamma = \theta_1 - \theta_2 < \pi$. We use this representative throughout the interior analysis below. And we have that

$$\begin{aligned}
&\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\\
&= \frac{1}{2}\Big[\cos(\theta_{11}-\theta_{22}) + \cos(\theta_{11}+\theta_{22})\Big] - \frac{1}{2}\Big[\cos(\theta_{12}-\theta_{21}) + \cos(\theta_{12}+\theta_{21})\Big]\\
&= \frac{1}{2}\Big[\cos\big((\theta_1-\alpha/2)-(\alpha/2+\theta_2)\big) + \cos\big((\theta_1-\alpha/2)+(\alpha/2+\theta_2)\big)\Big]\\
&\quad - \frac{1}{2}\Big[\cos\big((\theta_1+\alpha/2)-(\alpha/2-\theta_2)\big) + \cos\big((\theta_1+\alpha/2)+(\alpha/2-\theta_2)\big)\Big]\\
&= \frac{1}{2}\Big[\cos(\theta_1-\theta_2-\alpha) + \cos(\theta_1+\theta_2)\Big] - \frac{1}{2}\Big[\cos(\theta_1+\theta_2) + \cos(\theta_1-\theta_2+\alpha)\Big]\\
&= \frac{1}{2}\Big[\cos(\gamma-\alpha) + \cos(\theta_1+\theta_2)\Big] - \frac{1}{2}\Big[\cos(\theta_1+\theta_2) + \cos(\gamma+\alpha)\Big]\\
&= \frac{1}{2}\Big[\cos(\gamma-\alpha) - \cos(\gamma+\alpha)\Big]
\end{aligned}$$

$$= \sin\alpha \sin\gamma.$$

Recall that

$$\begin{aligned}
E_1(\theta_1, \theta_2) = &\Big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{11}\Big) \\
&\quad \cdot \Big(\sin\gamma\cos\theta_{12} + (\pi - \gamma)\cos\theta_{22}\Big) \\
&- \Big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12}) + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{12}\Big) \\
&\quad \cdot \Big(\sin\gamma\cos\theta_{11} + (\pi - \gamma)\cos\theta_{21}\Big) \\
&- \pi\Big(\big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\big)\cos\theta_{22} \\
&\quad - \big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\big)\cos\theta_{21}\Big) \\
= &\ 0.
\end{aligned}$$

Consider that

$$\begin{aligned}
&\Big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{11}\Big)\cos\theta_{12} \\
&\quad - \Big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12}) + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{12}\Big)\cos\theta_{11} \\
&= \big(\pi - \theta_{11}\big)\big(\cos\theta_{12} - \cos\alpha\cos\theta_{11}\big) + \big(\pi - \theta_{12}\big)\big(\cos\alpha\cos\theta_{12} - \cos\theta_{11}\big) \\
&= -\sin\alpha\Big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12}\Big).
\end{aligned}$$

Likewise, we have that

$$\begin{aligned}
&\Big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{11}\Big)\cos\theta_{22} \\
&\quad - \Big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12}) + (\sin\theta_{11} + \sin\theta_{12})\cos\theta_{12}\Big)\cos\theta_{21} \\
&= \big(\pi - \theta_{11}\big)\big(\cos\theta_{22} - \cos\alpha\cos\theta_{21}\big) + \big(\pi - \theta_{12}\big)\big(\cos\alpha\cos\theta_{22} - \cos\theta_{21}\big) \\
&\quad + (\sin\theta_{11} + \sin\theta_{12})\big(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21}\big) \\
&= \sin\alpha\Big((\pi - \theta_{11})\sin\theta_{21} - (\pi - \theta_{12})\sin\theta_{22} + (\sin\theta_{11} + \sin\theta_{12})\sin\gamma\Big).
\end{aligned}$$

Finally, we have that

$$\begin{aligned}
&\Big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\Big)\cos\theta_{22} - \Big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\Big)\cos\theta_{21} \\
&= (\pi - \theta_{21})\big(\cos\theta_{22} - \cos\alpha\cos\theta_{21}\big) + (\pi - \theta_{22})\big(\cos\alpha\cos\theta_{22} - \cos\theta_{21}\big) \\
&= \sin\alpha\Big((\pi - \theta_{21})\sin\theta_{21} - (\pi - \theta_{22})\sin\theta_{22}\Big).
\end{aligned}$$

And thus we have that

$$\begin{aligned}
E_1(\theta_1, \theta_2) = &-\sin\gamma\cdot\sin\alpha\Big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12}\Big) \\
&+ (\pi - \gamma)\cdot\sin\alpha\Big((\pi - \theta_{11})\sin\theta_{21} - (\pi - \theta_{12})\sin\theta_{22} + (\sin\theta_{11} + \sin\theta_{12})\sin\gamma\Big)
\end{aligned}$$

$$- \pi \cdot \sin\alpha\Big((\pi-\theta_{21})\sin\theta_{21} - (\pi-\theta_{22})\sin\theta_{22}\Big).$$

The same argument also works for $E_2(\theta_1,\theta_2)$. We have that

$$\begin{aligned}E_2(\theta_1,\theta_2) = &- \sin\gamma\cdot\sin\alpha\Big((\pi-\theta_{21})\sin\theta_{21} - (\pi-\theta_{22})\sin\theta_{22}\Big)\\ &+ (\pi-\gamma)\cdot\sin\alpha\Big((\pi-\theta_{21})\sin\theta_{11} + (\pi-\theta_{22})\sin\theta_{12} + (\sin\theta_{21}+\sin\theta_{22})\sin\gamma\Big)\\ &- \pi\cdot\sin\alpha\Big((\pi-\theta_{11})\sin\theta_{11} + (\pi-\theta_{12})\sin\theta_{12}\Big).\end{aligned}$$

And thus we have that

$$\begin{aligned}0 = &- \sin\gamma\cdot\Big((\pi-\theta_{11})\sin\theta_{11} + (\pi-\theta_{12})\sin\theta_{12}\Big)\\ &+ (\pi-\gamma)\cdot\Big((\pi-\theta_{11})\sin\theta_{21} - (\pi-\theta_{12})\sin\theta_{22} + (\sin\theta_{11}+\sin\theta_{12})\sin\gamma\Big)\\ &- \pi\cdot\Big((\pi-\theta_{21})\sin\theta_{21} - (\pi-\theta_{22})\sin\theta_{22}\Big),\\ 0 = &- \sin\gamma\cdot\Big((\pi-\theta_{21})\sin\theta_{21} - (\pi-\theta_{22})\sin\theta_{22}\Big)\\ &+ (\pi-\gamma)\cdot\Big((\pi-\theta_{21})\sin\theta_{11} + (\pi-\theta_{22})\sin\theta_{12} + (\sin\theta_{21}+\sin\theta_{22})\sin\gamma\Big)\\ &- \pi\cdot\Big((\pi-\theta_{11})\sin\theta_{11} + (\pi-\theta_{12})\sin\theta_{12}\Big),\end{aligned}$$

By $\gamma = \theta_{11}-\theta_{22}+\alpha$, $\theta_{21} = \alpha-\theta_{22}$ and $\theta_{12} = \theta_{11}+\alpha$, we have that

$$\begin{aligned}0 = &- \sin(\theta_{11}-\theta_{22}+\alpha)\cdot\Big((\pi-\theta_{11})\sin\theta_{11} + (\pi-\theta_{11}-\alpha)\sin(\theta_{11}+\alpha)\Big)\\ &+ (\pi-\theta_{11}+\theta_{22}-\alpha)\times\\ &\quad\Big((\pi-\theta_{11})\sin(\alpha-\theta_{22}) - (\pi-\theta_{11}-\alpha)\sin\theta_{22}\\ &\qquad + (\sin\theta_{11}+\sin(\theta_{11}+\alpha))\sin(\theta_{11}-\theta_{22}+\alpha)\Big)\\ &- \pi\cdot\Big((\pi-\alpha+\theta_{22})\sin(\alpha-\theta_{22}) - (\pi-\theta_{22})\sin\theta_{22}\Big),\end{aligned} \tag{129}$$

and that

$$\begin{aligned}0 = &- \sin(\theta_{11}-\theta_{22}+\alpha)\cdot\Big((\pi-\alpha+\theta_{22})\sin(\alpha-\theta_{22}) - (\pi-\theta_{22})\sin\theta_{22}\Big)\\ &+ (\pi-\theta_{11}+\theta_{22}-\alpha)\times\\ &\quad\Big((\pi-\alpha+\theta_{22})\sin\theta_{11} + (\pi-\theta_{22})\sin(\theta_{11}+\alpha)\\ &\qquad + (\sin(\alpha-\theta_{22})+\sin\theta_{22})\sin(\theta_{11}-\theta_{22}+\alpha)\Big)\\ &- \pi\cdot\Big((\pi-\theta_{11})\sin\theta_{11} + (\pi-\theta_{11}-\alpha)\sin(\theta_{11}+\alpha)\Big).\end{aligned} \tag{130}$$

For simplicity, we denote $x := \theta_{11}, y := \theta_{22}$. Grouping the terms that multiply $\sin(x-y+\alpha)$, and using

$$(\pi-x+y-\alpha)\big(\sin x+\sin(x+\alpha)\big) - \big((\pi-x)\sin x + (\pi-x-\alpha)\sin(x+\alpha)\big)$$

$$= (y-\alpha)\sin x + y\sin(x+\alpha),$$

and

$$\begin{aligned}
&(\pi - x + y - \alpha)\big(\sin(\alpha-y)+\sin y\big) - \big((\pi-\alpha+y)\sin(\alpha-y) - (\pi-y)\sin y\big) \\
&= -x\sin(\alpha-y) + (2\pi - x - \alpha)\sin y,
\end{aligned}$$

we have that

$$\begin{aligned}
0 &= (\pi - x + y - \alpha)\Big((\pi-x)\sin(\alpha-y) - (\pi-x-\alpha)\sin y\Big) \\
&\quad - \pi\Big((\pi-\alpha+y)\sin(\alpha-y) - (\pi-y)\sin y\Big) \\
&\quad + \sin(x-y+\alpha)\Big((y-\alpha)\sin x + y\sin(x+\alpha)\Big), \\
0 &= (\pi - x + y - \alpha)\Big((\pi-\alpha+y)\sin x + (\pi-y)\sin(x+\alpha)\Big) \\
&\quad - \pi\Big((\pi-x)\sin x + (\pi-x-\alpha)\sin(x+\alpha)\Big) \\
&\quad + \sin(x-y+\alpha)\Big(-x\sin(\alpha-y) + (2\pi-x-\alpha)\sin y\Big).
\end{aligned}$$

It is useful to first collect the scalar coefficients. Direct expansion gives

$$\begin{aligned}
(\pi - x + y - \alpha)(\pi - x) - \pi(\pi-\alpha+y) &= -\pi x - x(\pi-\alpha+y) + x^2 \\
&= x(-2\pi + \alpha - y + x), \\
\pi(\pi-y) - (\pi - x + y - \alpha)(\pi - x - \alpha) &= (x+\alpha-y)(\pi-x-\alpha) + \pi(x+\alpha-y) \\
&= (x+\alpha-y)(2\pi - x - \alpha), \\
(\pi - x + y - \alpha)(\pi-\alpha+y) - \pi(\pi-x) &= (2\pi - x)(y-\alpha) + (y-\alpha)^2 \\
&= (y-\alpha)(2\pi - x + y - \alpha), \\
(\pi - x + y - \alpha)(\pi-y) - \pi(\pi-x-\alpha) &= \pi y - (\pi-x-\alpha)y - y^2 \\
&= y(x+\alpha-y).
\end{aligned}$$

Using these identities in the previous display gives

$$\begin{aligned}
0 &= x(-2\pi+\alpha-y+x)\sin(\alpha-y) + (x+\alpha-y)(2\pi-x-\alpha)\sin y \\
&\quad + \sin(x-y+\alpha)\Big((y-\alpha)\sin x + y\sin(x+\alpha)\Big), \\
0 &= (y-\alpha)(2\pi-x+y-\alpha)\sin x + y(x+\alpha-y)\sin(x+\alpha) \\
&\quad + \sin(x-y+\alpha)\Big(-x\sin(\alpha-y) + (2\pi-x-\alpha)\sin y\Big).
\end{aligned}$$

This can be written as

$$\begin{aligned}
0 &= -\theta_{11}(2\pi-\gamma)\sin\theta_{21} + \gamma\cdot(2\pi-\theta_{12})\sin\theta_{22} + \sin\gamma\cdot\Big(\theta_{22}\sin\theta_{12} - \theta_{21}\sin\theta_{11}\Big), \\
0 &= -\theta_{21}(2\pi-\gamma)\sin\theta_{11} + \theta_{22}\gamma\sin\theta_{12} + \sin\gamma\cdot\Big((2\pi-\theta_{12})\sin\theta_{22} - \theta_{11}\sin\theta_{21}\Big).
\end{aligned}$$

We now give an analytic argument showing that the preceding displayed system has no solution in the interior of Case 2. Let

$$a = \theta_{11}, \qquad b = \theta_{21}, \qquad s = a + b.$$

Then

$$\theta_{12} = a + \alpha, \qquad \theta_{22} = \alpha - b, \qquad \gamma = s.$$

In the interior of this case,

$$0 < a < \pi - \alpha, \qquad 0 < b < \alpha, \qquad 0 < s < \pi.$$

Since the same four products occur several times in the two critical equations, write

$$U = a \sin b, \qquad V = b \sin a, \qquad M = (2\pi - a - \alpha)\sin(\alpha - b), \qquad N = (\alpha - b)\sin(a + \alpha).$$

Substituting these identities into the first equation in the preceding displayed system gives

$$\begin{aligned} 0 &= -a(2\pi - s)\sin b + s(2\pi - a - \alpha)\sin(\alpha - b) + \sin s\Big((\alpha - b)\sin(a + \alpha) - b\sin a\Big) \\ &= -(2\pi - s)U + sM + \sin s\, N - \sin s\, V. \end{aligned}$$

Thus

$$(2\pi - s)U + \sin s\, V = sM + \sin s\, N.$$

Similarly, substituting into the second equation in the preceding displayed system gives

$$\begin{aligned} 0 &= -b(2\pi - s)\sin a + s(\alpha - b)\sin(a + \alpha) + \sin s\Big((2\pi - a - \alpha)\sin(\alpha - b) - a\sin b\Big) \\ &= -(2\pi - s)V + sN + \sin s\, M - \sin s\, U, \end{aligned}$$

and therefore

$$\sin s\, U + (2\pi - s)V = \sin s\, M + sN.$$

Thus the two equations are equivalently

$$\begin{aligned} (2\pi - s)U + \sin s\, V &= sM + \sin s\, N, \\ \sin s\, U + (2\pi - s)V &= \sin s\, M + sN. \end{aligned} \tag{131}$$

Adding and subtracting the two equations in (131) gives

$$\begin{aligned} (2\pi - s + \sin s)(U + V) &= (s + \sin s)(M + N), \\ (2\pi - s - \sin s)(U - V) &= (s - \sin s)(M - N). \end{aligned} \tag{132}$$

We first record two consequences of (132). Let $r = \alpha - b$. Since $s = a + b$, we have $a = s - b = s - \alpha + r$. Consequently, $a + \alpha = s + r$ and $2\pi - a - \alpha = 2\pi - s - r$. Recalling the definitions of M and N, this gives $M = (2\pi - s - r)\sin r$ and $N = r\sin(s + r)$. Since $s + r = a + \alpha < \pi$, and hence $0 < r < s + r < \pi$, the function $t \mapsto \sin t/t$ is strictly decreasing on $(0, \pi)$. Moreover $2\pi - s - r > s + r$ because $s + r < \pi$. Hence $M = (2\pi - s - r)\sin r > r\sin(s + r) = N$. Since $0 < s < \pi$, both $2\pi - s - \sin s$ and $s - \sin s$ are positive. The second identity in (132) therefore implies $U > V$, namely $a \sin b > b \sin a$. Equivalently, $\sin b/b > \sin a/a$, and the strict monotonicity of $t \mapsto \sin t/t$ on $(0, \pi)$ gives

$$a > b. \tag{133}$$

Next we show that $b > \alpha/2$. Recall from (118) that

$$\|w_1\|_2 = \frac{(2\pi - \theta_{21} - \theta_{22})\cos(\alpha/2)\sin\theta_2 - (\theta_{22} - \theta_{21})\sin(\alpha/2)\cos\theta_2}{(\pi - \gamma)\sin(\theta_2 - \theta_1)}.$$

In Case 2, we have $\theta_{21} = b$, $\theta_{22} = \alpha - b$, and, from the angular parametrization of Area 1, $\theta_{21} = \alpha/2 - \theta_2$. Let $\xi = \alpha/2 - b$. Then $\xi = \theta_2$. Moreover,

$$s = a + b = \left(\theta_1 - \frac{\alpha}{2}\right) + \left(\frac{\alpha}{2} - \theta_2\right) = \theta_1 - \theta_2 = \gamma,$$

so $\theta_2 - \theta_1 = -s$. Hence

$$\begin{aligned} 2\pi - \theta_{21} - \theta_{22} &= 2\pi - \alpha, \\ \theta_{22} - \theta_{21} &= (\alpha - b) - b = 2\left(\frac{\alpha}{2} - b\right) = 2\xi, \\ \sin\theta_2 &= \sin\xi, \qquad \cos\theta_2 = \cos\xi, \\ (\pi - \gamma)\sin(\theta_2 - \theta_1) &= (\pi - s)\sin(-s) = -(\pi - s)\sin s. \end{aligned}$$

Substituting these identities into (118), we obtain

$$\|w_1\|_2 = \frac{(2\pi - \alpha)\cos(\alpha/2)\sin\xi - 2\xi\sin(\alpha/2)\cos\xi}{-(\pi - s)\sin s}.$$

The denominator $-(\pi - s)\sin s$ is negative. The numerator

$$(2\pi - \alpha)\cos(\alpha/2)\sin\xi - 2\xi\sin(\alpha/2)\cos\xi$$

has the same sign as ξ, and is zero when $\xi = 0$. Indeed, it factors as

$$\cos\xi\left((2\pi - \alpha)\cos(\alpha/2)\tan\xi - 2\xi\sin(\alpha/2)\right),$$

and for $0 < |\xi| \le \alpha/2 \le \pi/4$ we have $|\tan\xi| > |\xi|$, while $(2\pi - \alpha)\cos(\alpha/2) > 2\sin(\alpha/2)$ for $\alpha \in (0, \pi/2]$. If $\xi = 0$, then the displayed formula for $\|w_1\|_2$ would give $\|w_1\|_2 = 0$, impossible. Since $\|w_1\|_2 > 0$ and the denominator in the displayed formula for $\|w_1\|_2$ is negative, the numerator must be negative. Hence $\xi < 0$, and therefore

$$b > \frac{\alpha}{2}. \tag{134}$$

Combining (133) and (134), we have $0 < r = \alpha - b < b < a$. In particular, since $s = a + b$, the inequality $r < b < a$ gives $r < s/2$. Moreover, using $a = s - b = s - \alpha + r$, and the interior condition $a < \pi - \alpha$, we obtain $s - \alpha + r < \pi - \alpha$, which is equivalent to $r < \pi - s$. Hence

$$0 < r = \alpha - b < b < a, \qquad b < \frac{s}{2}, \qquad r < \min\left\{\frac{s}{2}, \pi - s\right\}. \tag{135}$$

For this fixed $s \in (0, \pi)$ and for $0 < x < s/2$, define

$$R_s(x) := \frac{(s-x)\sin x - x\sin(s-x)}{(s-x)\sin x + x\sin(s-x)}.$$

For $0 < x < \min\{s/2, \pi - s\}$, define

$$Q_s(x) := \frac{(2\pi - s - x)\sin x - x\sin(s+x)}{(2\pi - s - x)\sin x + x\sin(s+x)}.$$

Then $U = a\sin b = (s-b)\sin b$ and $V = b\sin a = b\sin(s-b)$, and hence

$$\frac{U-V}{U+V} = \frac{(s-b)\sin b - b\sin(s-b)}{(s-b)\sin b + b\sin(s-b)} = R_s(b).$$

Similarly, since $M = (2\pi - s - r)\sin r$ and $N = r\sin(s+r)$, we get

$$\frac{M-N}{M+N} = \frac{(2\pi - s - r)\sin r - r\sin(s+r)}{(2\pi - s - r)\sin r + r\sin(s+r)} = Q_s(r).$$

Therefore, $(U-V)/(U+V) = R_s(b)$ and $(M-N)/(M+N) = Q_s(r)$. Dividing the two identities in (132), we get

$$R_s(b) = \Lambda_s Q_s(r), \qquad \Lambda_s := \frac{s - \sin s}{s + \sin s} \cdot \frac{2\pi - s + \sin s}{2\pi - s - \sin s}. \tag{136}$$

By continuous extension at $x = 0$, $R_s(0) = (s - \sin s)/(s + \sin s)$ and $Q_s(0) = (2\pi - s - \sin s)/(2\pi - s + \sin s)$, we have $\Lambda_s = R_s(0)/Q_s(0)$.

We now prove two comparison estimates. First,

$$R_s(x) < \frac{s-x}{s} R_s(0), \qquad 0 < x < \frac{s}{2}. \tag{137}$$

For $t \in (0, \pi)$, let $A(t) = \log(t/\sin t)$. Fix $x \in (0, s/2)$. Then $0 < x < s - x < \pi$, so $A(x)$ and $A(s-x)$ are well-defined. From the definition of R_s and the identity $\tanh(\frac{1}{2}\log y) = (y-1)/(y+1)$ for $y > 0$,

$$\begin{aligned} R_s(x) &= \frac{(s-x)\sin x - x\sin(s-x)}{(s-x)\sin x + x\sin(s-x)} \\ &= \tanh\left(\frac{1}{2}\log\frac{(s-x)\sin x}{x\sin(s-x)}\right). \end{aligned}$$

The logarithm in this formula is

$$\begin{aligned} A(s-x) - A(x) &= \log\frac{s-x}{\sin(s-x)} - \log\frac{x}{\sin x} \\ &= \log\frac{(s-x)\sin x}{x\sin(s-x)}. \end{aligned}$$

Thus

$$R_s(x) = \tanh\left(\frac{A(s-x) - A(x)}{2}\right).$$

Since $A(x) > 0$ for $x \in (0, \pi)$, monotonicity of $\tanh$ gives

$$R_s(x) < \tanh\left(\frac{A(s-x)}{2}\right) = \frac{s-x-\sin(s-x)}{s-x+\sin(s-x)}.$$

Therefore, in order to prove (137), it is enough to prove

$$\frac{s-x-\sin(s-x)}{s-x+\sin(s-x)} \leq \frac{s-x}{s} \cdot \frac{s-\sin s}{s+\sin s}.$$

After dividing both sides by the positive factor $s-x$, this is exactly the monotonicity statement $F(s-x) \leq F(s)$, where $F(t) := (t-\sin t)/(t(t+\sin t))$ for $t \in (0, \pi)$. Thus it remains to show that F is increasing on $(0, \pi)$. By the quotient rule,

$$\begin{aligned} F'(t) &= \frac{(1-\cos t)t(t+\sin t) - (t-\sin t)(2t+\sin t + t\cos t)}{\big(t(t+\sin t)\big)^2} \\ &= \frac{\sin^2 t + 2t\sin t - t^2(1+2\cos t)}{\big(t(t+\sin t)\big)^2}. \end{aligned} \tag{138}$$

The expansion in the numerator of (138) is

$$\begin{aligned} &(1-\cos t)t(t+\sin t) - (t-\sin t)(2t+\sin t+t\cos t) \\ &= t^2 + t\sin t - t^2\cos t - t\sin t\cos t \\ &\quad - 2t^2 - t\sin t - t^2\cos t + 2t\sin t + \sin^2 t + t\sin t\cos t \\ &= \sin^2 t + 2t\sin t - t^2(1+2\cos t). \end{aligned}$$

It remains to show that the numerator in (138) is nonnegative. This numerator vanishes at $t = 0$, and its derivative is

$$\begin{aligned} \frac{d}{dt}\Big(\sin^2 t + 2t\sin t - t^2(1+2\cos t)\Big) &= 2\sin t\cos t + 2\sin t + 2t\cos t \\ &\quad - 2t(1+2\cos t) + 2t^2\sin t \\ &= 2t^2\sin t + 2\sin t\cos t + 2\sin t - 2t - 2t\cos t \\ &= 2\left(t^2\sin t - (1+\cos t)(t-\sin t)\right). \end{aligned}$$

Since $t - \sin t = \int_0^t (1-\cos u)\,du \leq t(1-\cos t)$, we obtain

$$(1+\cos t)(t-\sin t) \leq t(1-\cos^2 t) = t\sin^2 t \leq t^2 \sin t.$$

Thus the derivative of the numerator in (138) is nonnegative. Hence the numerator in (138) is nonnegative, and consequently $F'(t) \geq 0$. Therefore

$$R_s(x) < (s-x)F(s-x) \leq (s-x)F(s) = \frac{s-x}{s}R_s(0),$$

which proves (137).

Next we prove the second comparison estimate:

$$Q_s(x) \geq \frac{s-x}{s} Q_s(0), \qquad 0 < x < \min\left\{\frac{s}{2}, \pi - s\right\}. \tag{139}$$

For $t \in (0, \pi)$, let $B(t) = \log((2\pi - t)/\sin t)$. Fix $x \in (0, \min\{s/2, \pi - s\})$. Then $0 < x < \pi$ and $0 < s + x < \pi$, so $A(x)$ and $B(s+x)$ are well-defined. From the definition of Q_s and the identity $\tanh(\frac{1}{2}\log y) = (y-1)/(y+1)$ for $y > 0$,

$$Q_s(x) = \tanh\left(\frac{1}{2}\log\frac{(2\pi - s - x)\sin x}{x\sin(s+x)}\right).$$

The logarithm in this formula is

$$B(s+x) - A(x) = \log\frac{2\pi - s - x}{\sin(s+x)} - \log\frac{x}{\sin x} = \log\frac{(2\pi - s - x)\sin x}{x\sin(s+x)}.$$

Thus

$$Q_s(x) = \tanh\left(\frac{B(s+x) - A(x)}{2}\right).$$

We claim that

$$B(s+x) - A(x) \geq \frac{s-x}{s} B(s). \tag{140}$$

For this fixed s, let $H_s(x)$ be the left-hand side of (140) minus its right-hand side. Here we use the continuous extension

$$A(0) := \lim_{t \to 0^+} \log\frac{t}{\sin t} = 0.$$

With this convention, $H_s(0) = B(s) - A(0) - B(s) = 0$. For $x > 0$,

$$H_s(x) = B(s+x) - A(x) - \frac{s-x}{s} B(s).$$

Since $A'(x) = 1/x - \cot x$ and $B'(s+x) = \frac{d}{dx}\log\frac{2\pi - s - x}{\sin(s+x)} = -\frac{1}{2\pi - s - x} - \cot(s+x)$, we have

$$\begin{aligned} H_s'(x) &= B'(s+x) - A'(x) + \frac{B(s)}{s} \\ &= \frac{B(s)}{s} - \left(\frac{1}{x} - \cot x\right) - \left(\cot(s+x) + \frac{1}{2\pi - s - x}\right). \end{aligned}$$

On $(0, \pi)$,

$$\begin{aligned} \frac{d}{dt}\left(\frac{1}{t} - \cot t\right) &= \frac{1}{\sin^2 t} - \frac{1}{t^2} > 0, \\ \frac{d}{dt}\left(\cot t + \frac{1}{2\pi - t}\right) &= -\frac{1}{\sin^2 t} + \frac{1}{(2\pi - t)^2} < 0. \end{aligned}$$

Since $x < s/2$ and $s < s + x < \pi$, these two monotonicity facts imply

$$\frac{1}{x} - \cot x \leq \frac{2}{s} - \cot\frac{s}{2}, \qquad \cot(s+x) + \frac{1}{2\pi - s - x} \leq \cot s + \frac{1}{2\pi - s}.$$

Therefore

$$H_s'(x) \geq \frac{B(s)}{s} - \left(\frac{2}{s} - \cot\frac{s}{2}\right) - \left(\cot s + \frac{1}{2\pi - s}\right).$$

Using $\cot(s/2) = 1/\sin s + \cot s$, this lower bound is nonnegative provided that

$$B(s) \geq 2 - \frac{s}{\sin s} + \frac{s}{2\pi - s}. \tag{141}$$

We prove (141). Since $(2\pi - s)/\sin s > 1$, the elementary inequality $\log z \geq 2(z-1)/(z+1)$ for $z > 1$, gives

$$B(s) = \log\frac{2\pi - s}{\sin s} \geq \frac{2(2\pi - s - \sin s)}{2\pi - s + \sin s}.$$

Thus it is enough to show

$$\frac{2(2\pi - s - \sin s)}{2\pi - s + \sin s} \geq 2 - \frac{s}{\sin s} + \frac{s}{2\pi - s}.$$

After multiplying by the positive quantity $(2\pi - s)\sin s(2\pi - s + \sin s)$, this is equivalent to

$$s\big((2\pi - s)^2 - \sin^2 s\big) \geq 4(2\pi - s)\sin^2 s. \tag{142}$$

If $0 < s \leq 1$, then $\sin s \leq s$ and $(2\pi - s)/s \geq 2\pi - 1 > 2 + \sqrt{5}$. This implies

$$(2\pi - s)^2 - s^2 \geq 4(2\pi - s)s,$$

and therefore

$$s\big((2\pi - s)^2 - \sin^2 s\big) \geq s\big((2\pi - s)^2 - s^2\big) \geq 4(2\pi - s)s^2 \geq 4(2\pi - s)\sin^2 s.$$

If $1 \leq s < \pi$, then $\sin^2 s \leq 1$. Thus it suffices to prove

$$s\big((2\pi - s)^2 - 1\big) - 4(2\pi - s) > 0.$$

The function

$$t \longmapsto t\big((2\pi - t)^2 - 1\big) - 4(2\pi - t)$$

has second derivative $6t - 8\pi < 0$ on $[1, \pi]$. It is therefore concave on $[1, \pi]$, so its minimum on this interval is attained at an endpoint. The endpoint values are

$$4(\pi^2 - 3\pi + 1) > 0, \qquad \pi(\pi^2 - 5) > 0.$$

Therefore (142) holds in all cases. This proves (141), and hence $H_s'(x) \geq 0$. Since $H_s(0) = 0$, (140) follows.

Also, $0 < s < \pi$ gives $B(s) > 0$, $R_s(0) > 0$, $Q_s(0) > 0$, and $\Lambda_s > 0$. Finally, since $\tanh$ is increasing and concave on $\mathbb{R}_+$, and $\tanh(0) = 0$, (140) gives

$$Q_s(x) \geq \tanh\left(\frac{s - x}{s} \cdot \frac{B(s)}{2}\right) \geq \frac{s - x}{s}\tanh\left(\frac{B(s)}{2}\right) = \frac{s - x}{s}Q_s(0),$$

which proves (139).

We now apply (137) with $x = b$ and (139) with $x = r$. By (135), $0 < r < b$, and by the continuous-extension identity after (136), $\Lambda_s = R_s(0)/Q_s(0)$. Hence

$$R_s(b) < \frac{s-b}{s} R_s(0) < \frac{s-r}{s} R_s(0) = \Lambda_s \frac{s-r}{s} Q_s(0) \leq \Lambda_s Q_s(r).$$

This contradicts (136), which states that $R_s(b) = \Lambda_s Q_s(r)$. Therefore there is no critical point in the interior of Case 2. The argument above is the interior analysis of this open case; boundary configurations belong to the neighboring area cases, or to lemma 27 when $\gamma = 0$ or $\gamma = \pi$. The boundary point $a = 0$, $b = \alpha$ corresponds to $\bar{w}_1 = v_1$, $\bar{w}_2 = v_2$, which is one of the listed solutions.

Case 3: $\gamma \neq 0$, w_1 **lies at the Area** 1 **and** w_2 **lies at the Area** 3 **in Figure 9**

Without loss of generality, we assume $\theta_1 \in [0, \alpha/2]$. If the projection of w_2 lies inside the angle

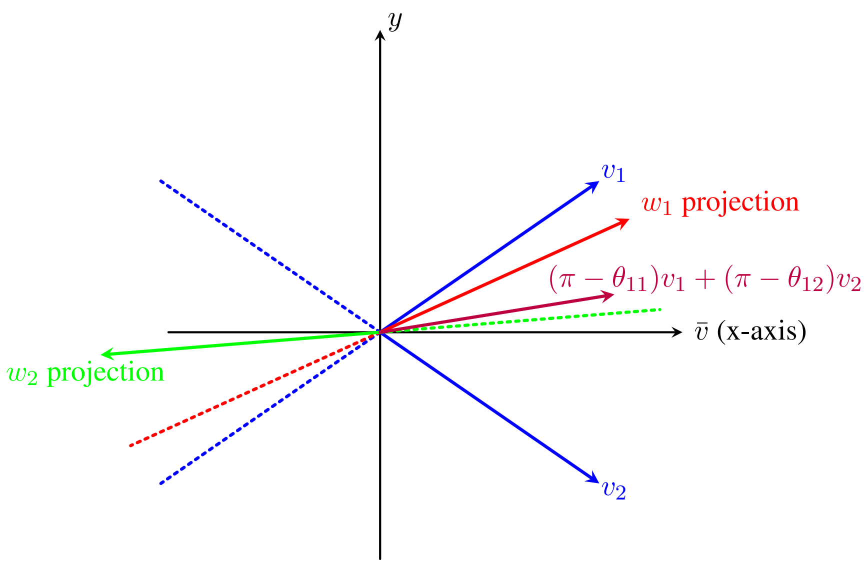

Figure 10: v_1, v_2, and the projections of w_1 and w_2 in the $x - y$ plane.

between $-w_1$ projection and the negative direction of the x-axis, as shown in Figure 10, then $-w_2$ lies strictly below $\bar{w}_1$. We now make this statement algebraic in coordinates. For any planar vector $\xi = (\xi_x, \xi_y)$, set

$$\mathcal{A}_{\bar{w}_1}(\xi) := (\bar{w}_1)_x \xi_y - (\bar{w}_1)_y \xi_x.$$

Since $\bar{w}_1 = (\cos\theta_1, \sin\theta_1)$, for a unit direction $u(\phi) = (\cos\phi, \sin\phi)$ we have

$$\mathcal{A}_{\bar{w}_1}(u(\phi)) = \cos\theta_1 \sin\phi - \sin\theta_1 \cos\phi = \sin(\phi - \theta_1).$$

Thus $\mathcal{A}_{\bar{w}_1}(\xi) < 0$ means that the direction ξ lies below $\bar{w}_1$ in the oriented plane. In the present subregion, $\theta_2 + \pi < \theta_1$, and hence

$$\mathcal{A}_{\bar{w}_1}(-\bar{w}_2) = \sin\big((\theta_2 + \pi) - \theta_1\big) < 0.$$

The teacher vector

$$A_1 := (\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2$$

also lies below $\bar{w}_1$, or is parallel to it on the boundary. Indeed, in Area 1 we have $\theta_1 \in [0, \alpha/2]$, so

$$\theta_{11} = \frac{\alpha}{2} - \theta_1, \qquad \theta_{12} = \frac{\alpha}{2} + \theta_1.$$

Let $h = \alpha/2$ and $p = \theta_1$. Then

$$\begin{aligned}\mathcal{A}_{\bar{w}_1}(A_1) &= (\pi - \theta_{11}) \sin\left(\frac{\alpha}{2} - \theta_1\right) + (\pi - \theta_{12}) \sin\left(-\frac{\alpha}{2} - \theta_1\right) \\ &= (\pi - h + p) \sin(h - p) - (\pi - h - p) \sin(h + p) \\ &= 2p \sin h \cos p - 2(\pi - h) \cos h \sin p.\end{aligned}$$

When $p = 0$, this signed area is zero. When $p > 0$,

$$\mathcal{A}_{\bar{w}_1}(A_1) = 2p \cos p \left(\sin h - (\pi - h) \cos h \frac{\tan p}{p}\right) < 0,$$

because $\tan p/p \geq 1$ and, for $h \in (0, \pi/4]$,

$$(\pi - h) \cos h \geq \frac{3\pi}{4} \cdot \frac{\sqrt{2}}{2} > 1 > \sin h.$$

Therefore $\mathcal{A}_{\bar{w}_1}(A_1) \leq 0$, while $\mathcal{A}_{\bar{w}_1}(-\bar{w}_2) < 0$. According to

$$\frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t} = \frac{1}{2\pi \|w_1\|_2} (I - \bar{w}_1 \bar{w}_1^\top) \Big((\pi - \theta_{11}) v_1 + (\pi - \theta_{12}) v_2 - (\pi - \gamma) w_2 \Big),$$

and since $\pi - \gamma > 0$, the vector $A_1 - (\pi - \gamma) w_2 = A_1 + (\pi - \gamma)(-w_2)$ has strictly negative signed area with $\bar{w}_1$. It cannot be parallel to $\bar{w}_1$, and therefore $(I - \bar{w}_1 \bar{w}_1^\top)(A_1 - (\pi - \gamma) w_2) \neq 0$. Hence $\mathrm{d}\bar{w}_1/\mathrm{d}t \neq 0$, contradicting the critical point condition. Thus the projection of w_2 lies outside the angle between $-w_1$ projection and the negative direction of the x-axis. Equivalently, the excluded subregion is precisely the one in which $\theta_1 - \theta_2 > \pi$: indeed, in the lower representative of Area 3, this subregion is

$$-\pi \leq \theta_2 < \theta_1 - \pi.$$

The boundary case $\theta_2 = \theta_1 - \pi$ corresponds to $\gamma = \pi$, which was already handled above. Thus, in Case 3, the same-side argument eliminates exactly the complementary-arc subcase $\theta_1 - \theta_2 > \pi$; the remaining analysis is restricted to $\theta_1 - \theta_2 \leq \pi$.

Let us spell out the resulting angular interval for θ_2. Since w_2 lies in Area 3, and we are using the representative of its angle on the lower side of the negative x-axis, we first have

$$-\pi \leq \theta_2 \leq \frac{\alpha}{2} - \pi.$$

The direction of $-w_1$ has angle $\theta_1 - \pi$, while the negative x-axis has angle $-\pi$. Therefore, the angle between $-w_1$ and the negative x-axis corresponds exactly to

$$-\pi \leq \theta_2 \leq \theta_1 - \pi.$$

The previous paragraph shows that w_2 cannot lie in this subregion at a critical point. Hence the only remaining part of Area 3 is

$$\theta_1 - \pi \leq \theta_2 \leq \frac{\alpha}{2} - \pi.$$

Thus $\theta_2 \in [\theta_1 - \pi, \alpha/2 - \pi]$. In particular,

$$\pi - \frac{\alpha}{2} \leq \theta_1 - \theta_2 \leq \pi.$$

Since the angle γ is the representative in $[0, \pi]$, the shorter angle is $\theta_1 - \theta_2$ in this remaining subcase. The complementary formula $2\pi - (\theta_1 - \theta_2)$ would apply only in the excluded subregion where $\theta_1 - \theta_2 > \pi$. Hence here

$$\gamma = \theta_1 - \theta_2.$$

Moreover, note that we have $\theta_{11} = \alpha/2 - \theta_1$, $\theta_{12} = \alpha/2 + \theta_1$, $\theta_{21} = 2\pi - \alpha/2 + \theta_2$, and $\theta_{22} = -\alpha/2 - \theta_2$ in this case, which yields the relations

$$\theta_{12} + \theta_{22} = \gamma, \qquad \theta_{11} + \theta_{21} = 2\pi - \gamma,$$

$$\theta_{12} = \alpha - \theta_{11}, \qquad \theta_{21} = 2\pi - \alpha - \theta_{22}.$$

$$\begin{aligned}
E_1(\theta_1, \theta_2) = \sin\gamma\Big(&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{12} \\
&- \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{11}\Big) \\
+ (\pi - \gamma)\Big(&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{22} \\
&- \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{21} \\
&+ (\sin\theta_{11} + \sin\theta_{12})(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21})\Big) \\
- \pi\Big(&\big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\big)\cos\theta_{22} \\
&- \big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\big)\cos\theta_{21}\Big).
\end{aligned}$$

Using $\theta_{12} = \alpha - \theta_{11}$,

$$\begin{aligned}
&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{12} - \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{11} \\
&= (\pi - \theta_{11})(\cos\theta_{12} - \cos\alpha\cos\theta_{11}) + (\pi - \theta_{12})(\cos\alpha\cos\theta_{12} - \cos\theta_{11}) \\
&= (\pi - \theta_{11})\big(\cos(\alpha - \theta_{11}) - \cos\alpha\cos\theta_{11}\big) + (\pi - \theta_{12})\big(\cos\alpha\cos\theta_{12} - \cos(\alpha - \theta_{12})\big) \\
&= (\pi - \theta_{11})(\sin\alpha\sin\theta_{11}) + (\pi - \theta_{12})(-\sin\alpha\sin\theta_{12}) \\
&= \sin\alpha\big((\pi - \theta_{11})\sin\theta_{11} - (\pi - \theta_{12})\sin\theta_{12}\big).
\end{aligned}$$

Using $\theta_{21} = 2\pi - \alpha - \theta_{22}$ (so $\cos\theta_{21} = \cos(\alpha + \theta_{22})$),

$$\begin{aligned}
&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{22} - \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{21} \\
&= (\pi - \theta_{11})(\cos\theta_{22} - \cos\alpha\cos\theta_{21}) + (\pi - \theta_{12})(\cos\alpha\cos\theta_{22} - \cos\theta_{21}) \\
&= (\pi - \theta_{11})\big(\cos\theta_{22} - \cos\alpha\cos(\alpha + \theta_{22})\big) + (\pi - \theta_{12})\big(\cos\alpha\cos\theta_{22} - \cos(\alpha + \theta_{22})\big) \\
&= (\pi - \theta_{11})(-\sin\alpha\sin\theta_{21}) + (\pi - \theta_{12})(\sin\alpha\sin\theta_{22}) \\
&= \sin\alpha\big(-(\pi - \theta_{11})\sin\theta_{21} + (\pi - \theta_{12})\sin\theta_{22}\big).
\end{aligned}$$

Using $\theta_{21} = 2\pi - \alpha - \theta_{22}$ (so $\cos\theta_{21} = \cos(\alpha + \theta_{22})$),

$$\begin{aligned}
&\big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\big)\cos\theta_{22} - \big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\big)\cos\theta_{21} \\
&= (\pi - \theta_{21})(\cos\theta_{22} - \cos\alpha\cos\theta_{21}) + (\pi - \theta_{22})(\cos\alpha\cos\theta_{22} - \cos\theta_{21}) \\
&= (\pi - \theta_{21})\big(\cos\theta_{22} - \cos\alpha\cos(\alpha + \theta_{22})\big) + (\pi - \theta_{22})\big(\cos\alpha\cos\theta_{22} - \cos(\alpha + \theta_{22})\big) \\
&= (\pi - \theta_{21})(-\sin\alpha\sin\theta_{21}) + (\pi - \theta_{22})(\sin\alpha\sin\theta_{22}) \\
&= \sin\alpha\big(-(\pi - \theta_{21})\sin\theta_{21} + (\pi - \theta_{22})\sin\theta_{22}\big).
\end{aligned}$$

Finally, using $\theta_{12} = \alpha - \theta_{11}$ and $\theta_{21} = 2\pi - \alpha - \theta_{22}$,

$$\begin{aligned}
&\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21} \\
&= \cos\theta_{11}\cos\theta_{22} - \cos(\alpha - \theta_{11})\cos(\alpha + \theta_{22}) \\
&= \cos\theta_{11}\cos\theta_{22} - (\cos\alpha\cos\theta_{11} + \sin\alpha\sin\theta_{11})(\cos\alpha\cos\theta_{22} - \sin\alpha\sin\theta_{22}) \\
&= \sin^2\alpha\big(\cos\theta_{11}\cos\theta_{22} + \sin\theta_{11}\sin\theta_{22}\big) + \sin\alpha\cos\alpha\big(\cos\theta_{11}\sin\theta_{22} - \sin\theta_{11}\cos\theta_{22}\big) \\
&= \sin^2\alpha\cos(\theta_{11} - \theta_{22}) - \sin\alpha\cos\alpha\sin(\theta_{11} - \theta_{22}) \\
&= \sin\alpha\sin\big(\alpha - (\theta_{11} - \theta_{22})\big) = \sin\alpha\sin\big((\alpha - \theta_{11}) + \theta_{22}\big) \\
&= \sin\alpha\sin(\theta_{12} + \theta_{22}) = \sin\alpha\sin\gamma.
\end{aligned}$$

And thus

$$\begin{aligned}
\frac{E_1(\theta_1, \theta_2)}{\sin\alpha} &= \sin\gamma\big((\pi - \theta_{11})\sin\theta_{11} - (\pi - \theta_{12})\sin\theta_{12}\big) \\
&\quad + (\pi - \gamma)\big(-(\pi - \theta_{11})\sin\theta_{21} + (\pi - \theta_{12})\sin\theta_{22} + (\sin\theta_{11} + \sin\theta_{12})\sin\gamma\big) \\
&\quad - \pi\big((\pi - \theta_{22})\sin\theta_{22} - (\pi - \theta_{21})\sin\theta_{21}\big).
\end{aligned}$$

We now show that this expression can vanish only at the angular bisectors. Put

$$h = \frac{\alpha}{2}, \qquad p = \theta_1, \qquad q = \theta_2 + \pi.$$

Since we assumed w_1 lies in Area 1 with $\theta_1 \in [0, \alpha/2]$, we immediately have

$$0 \le p \le h.$$

The preceding geometric reduction gave

$$\theta_2 \in [\theta_1 - \pi, \alpha/2 - \pi].$$

Adding π to this interval gives

$$q = \theta_2 + \pi \in [\theta_1, \alpha/2] = [p, h].$$

Hence, in the present subcase,

$$0 \le p \le q \le h.$$

Moreover,

$$\theta_{11} = h - p, \qquad \theta_{12} = h + p, \qquad \theta_{21} = \pi - h + q, \qquad \theta_{22} = \pi - h - q,$$

and

$$\gamma = \theta_1 - \theta_2 = \pi - (q - p).$$

Let $\delta = q - p$. Then $0 \le \delta \le q$, $\sin\gamma = \sin\delta$, and $\pi - \gamma = \delta$. Also,

$$\sin\theta_{21} = \sin(h - q), \qquad \sin\theta_{22} = \sin(h + q).$$

Define

$$\begin{aligned} A_p &\coloneqq (\pi - \theta_{11})\sin\theta_{11} - (\pi - \theta_{12})\sin\theta_{12}, \\ B_{p,q} &\coloneqq -(\pi - \theta_{11})\sin\theta_{21} + (\pi - \theta_{12})\sin\theta_{22} + (\sin\theta_{11} + \sin\theta_{12})\sin\delta, \\ C_q &\coloneqq (\pi - \theta_{22})\sin\theta_{22} - (\pi - \theta_{21})\sin\theta_{21}. \end{aligned}$$

Then

$$\frac{E_1(\theta_1, \theta_2)}{\sin\alpha} = \sin\delta\, A_p + \delta B_{p,q} - \pi C_q. \tag{143}$$

We first compute A_p. Since $\theta_{11} = h - p$ and $\theta_{12} = h + p$,

$$\begin{aligned} A_p &= (\pi - h + p)\sin(h - p) - (\pi - h - p)\sin(h + p) \\ &= 2p\sin h\cos p - 2(\pi - h)\cos h\sin p. \end{aligned}$$

Thus $A_0 = 0$, and for $p > 0$,

$$A_p = 2p\cos p\left(\sin h - (\pi - h)\cos h\frac{\tan p}{p}\right) < 0,$$

because $\tan p/p \ge 1$ and, for $h \in (0, \pi/4]$,

$$(\pi - h)\cos h \ge \frac{3\pi}{4}\cdot\frac{\sqrt{2}}{2} > 1 > \sin h.$$

Hence

$$A_p \le 0. \tag{144}$$

Next, using the same substitutions,

$$\begin{aligned} B_{p,q} &= -(\pi - h + p)\sin(h - q) + (\pi - h - p)\sin(h + q) + 2\sin h\cos p\sin\delta \\ &= 2(\pi - h)\cos h\sin q - 2p\sin h\cos q + 2\sin h\cos p\sin\delta, \end{aligned}$$

while

$$\begin{aligned} C_q &= (h + q)\sin(h + q) - (h - q)\sin(h - q) \\ &= 2q\sin h\cos q + 2h\cos h\sin q. \end{aligned}$$

Since $0 \le p \le q \le h \le \pi/4$, we have

$$\delta \le q, \qquad q(\pi - h) \le \pi h, \qquad \sin\delta \le \pi\cos q.$$

Therefore,

$$\begin{aligned}\delta B_{p,q} &\le 2\delta(\pi - h)\cos h \sin q + 2\delta \sin h \cos p \sin \delta \\ &\le 2q(\pi - h)\cos h \sin q + 2q \sin h \sin \delta \\ &\le 2\pi h \cos h \sin q + 2\pi q \sin h \cos q \\ &= \pi C_q.\end{aligned}$$

If $q > 0$, then $C_q > 0$, and the last comparison is strict. Combining this with (143) and (144), we obtain

$$\frac{E_1(\theta_1, \theta_2)}{\sin \alpha} < 0$$

for every $q > 0$. Hence $E_1 = 0$ forces $q = 0$. Since $0 \le p \le q$, this also gives $p = 0$. Consequently,

$$\theta_1 = 0, \qquad \theta_2 = -\pi,$$

i.e. $\bar{w}_1 = \bar{v}$ and $\bar{w}_2 = -\bar{v}$. Thus the only possible critical point in Case 3 lies on the two angle bisectors.

Case 4: $\gamma \neq 0$, w_1 **lies at the Area** 2 **and** w_2 **lies at the Area** 4 **in Figure 9**

We first rule out the complementary-arc subcase $\theta_1 - \theta_2 > \pi$. This exclusion is geometric and has to be checked with the present pair of regions. Since w_1 lies in Area 2 and w_2 lies in Area 4, we use the angular representatives

$$\frac{\alpha}{2} \le \theta_1 \le \pi - \frac{\alpha}{2}, \qquad -\pi + \frac{\alpha}{2} \le \theta_2 \le -\frac{\alpha}{2}.$$

If $\theta_1 - \theta_2 > \pi$, then $\theta_2 + \pi < \theta_1$. Hence the direction of $-w_2$, whose angle is $\theta_2 + \pi$, lies between the direction of v_1 and the direction of w_1:

$$\frac{\alpha}{2} \le \theta_2 + \pi < \theta_1.$$

Therefore $-w_2$ lies strictly on one side of $\bar{w}_1$. More precisely, for a planar vector $\xi = (\xi_x, \xi_y)$, define

$$\mathcal{A}_{\bar{w}_1}(\xi) \coloneqq (\bar{w}_1)_x \xi_y - (\bar{w}_1)_y \xi_x.$$

Since $\bar{w}_1 = (\cos\theta_1, \sin\theta_1)$, for a unit direction $u(\phi) = (\cos\phi, \sin\phi)$ we have

$$\mathcal{A}_{\bar{w}_1}(u(\phi)) = \cos\theta_1 \sin\phi - \sin\theta_1 \cos\phi = \sin(\phi - \theta_1).$$

Thus

$$\mathcal{A}_{\bar{w}_1}(-\bar{w}_2) = \sin\big((\theta_2 + \pi) - \theta_1\big) < 0.$$

The teacher part lies on the same side. Indeed, set

$$A_1 \coloneqq (\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2.$$

Since in Area 2 we have $\theta_{11} = \theta_1 - \alpha/2$ and $\theta_{12} = \theta_1 + \alpha/2$,

$$\mathcal{A}_{\bar{w}_1}(A_1) = (\pi - \theta_{11})\sin\left(\frac{\alpha}{2} - \theta_1\right) + (\pi - \theta_{12})\sin\left(-\frac{\alpha}{2} - \theta_1\right)$$

$$= -(\pi - \theta_{11})\sin\theta_{11} - (\pi - \theta_{12})\sin\theta_{12} < 0.$$

The last inequality is strict because

$$0 \le \theta_{11} \le \pi - \alpha, \qquad \alpha \le \theta_{12} \le \pi,$$

and the two nonnegative terms above cannot vanish simultaneously. Thus both A_1 and $-w_2$ have strictly negative signed area with $\bar{w}_1$. Since $\pi - \gamma > 0$ in the subcase $\theta_1 - \theta_2 > \pi$, the vector

$$A_1 - (\pi - \gamma)w_2 = A_1 + (\pi - \gamma)(-w_2)$$

also has strictly negative signed area with $\bar{w}_1$. It cannot be parallel to $\bar{w}_1$, and hence

$$(I - \bar{w}_1\bar{w}_1^\top)\big(A_1 - (\pi - \gamma)w_2\big) \neq 0.$$

This contradicts the critical point condition $\mathrm{d}\bar{w}_1/\mathrm{d}t = 0$. Therefore no critical point in Case 4 can satisfy $\theta_1 - \theta_2 > \pi$. We may now return to the original calculation and work only in the subcase $\theta_1 - \theta_2 \le \pi$. In this subcase $\gamma = \theta_1 - \theta_2$, and therefore

$$\theta_{12} = \theta_{11} + \alpha, \qquad \theta_{21} = \theta_{22} + \alpha, \qquad \gamma = \theta_{11} + \theta_{21} = \theta_{12} + \theta_{22}.$$

Write E_1 as the sum of the $\sin\gamma$ part, the $(\pi - \gamma)$ part, and the $-\pi$ part, and use the cancellations of the terms with $(\sin\theta_{11} + \sin\theta_{12})\cos\theta_{11}\cos\theta_{12}$. Then

$$\begin{aligned}
E_1(\theta_1, \theta_2) = \sin\gamma\Big(&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{12} \\
&- \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{11}\Big) \\
+ (\pi - \gamma)\Big(&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{22} \\
&- \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{21} \\
&+ (\sin\theta_{11} + \sin\theta_{12})(\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21})\Big) \\
- \pi\Big(&\big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\big)\cos\theta_{22} \\
&- \big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\big)\cos\theta_{21}\Big).
\end{aligned}$$

Using $\theta_{12} = \theta_{11} + \alpha$,

$$\begin{aligned}
&\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{12} - \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{11} \\
&= (\pi - \theta_{11})(\cos\theta_{12} - \cos\alpha\cos\theta_{11}) + (\pi - \theta_{12})(\cos\alpha\cos\theta_{12} - \cos\theta_{11}) \\
&= (\pi - \theta_{11})(\cos(\theta_{11} + \alpha) - \cos\alpha\cos\theta_{11}) + (\pi - \theta_{12})(\cos\alpha\cos\theta_{12} - \cos(\theta_{12} - \alpha)) \\
&= (\pi - \theta_{11})(-\sin\theta_{11}\sin\alpha) + (\pi - \theta_{12})(-\sin\alpha\sin\theta_{12}) \\
&= -\sin\alpha\big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12}\big).
\end{aligned}$$

Using $\theta_{22} = \theta_{21} - \alpha$,

$$\big((\pi - \theta_{11}) + (\pi - \theta_{12})\cos\alpha\big)\cos\theta_{22} - \big((\pi - \theta_{11})\cos\alpha + (\pi - \theta_{12})\big)\cos\theta_{21}$$

$$
\begin{aligned}
&= (\pi - \theta_{11})\big(\cos\theta_{22} - \cos\alpha\cos\theta_{21}\big) + (\pi - \theta_{12})\big(\cos\alpha\cos\theta_{22} - \cos\theta_{21}\big) \\
&= (\pi - \theta_{11})\big(\cos(\theta_{21} - \alpha) - \cos\alpha\cos\theta_{21}\big) + (\pi - \theta_{12})\big(\cos\alpha\cos\theta_{22} - \cos(\theta_{22} + \alpha)\big) \\
&= (\pi - \theta_{11})\big(\sin\theta_{21}\sin\alpha\big) + (\pi - \theta_{12})\big(\sin\alpha\sin\theta_{22}\big) \\
&= \sin\alpha\big((\pi - \theta_{11})\sin\theta_{21} + (\pi - \theta_{12})\sin\theta_{22}\big).
\end{aligned}
$$

Using $\theta_{22} = \theta_{21} - \alpha$,

$$
\begin{aligned}
&\big((\pi - \theta_{21}) + (\pi - \theta_{22})\cos\alpha\big)\cos\theta_{22} - \big((\pi - \theta_{21})\cos\alpha + (\pi - \theta_{22})\big)\cos\theta_{21} \\
&= (\pi - \theta_{21})(\cos\theta_{22} - \cos\alpha\cos\theta_{21}) + (\pi - \theta_{22})(\cos\alpha\cos\theta_{22} - \cos\theta_{21}) \\
&= (\pi - \theta_{21})(\cos(\theta_{21} - \alpha) - \cos\alpha\cos\theta_{21}) + (\pi - \theta_{22})(\cos\alpha\cos\theta_{22} - \cos(\theta_{22} + \alpha)) \\
&= (\pi - \theta_{21})(\sin\theta_{21}\sin\alpha) + (\pi - \theta_{22})(\sin\alpha\sin\theta_{22}) \\
&= \sin\alpha\big((\pi - \theta_{21})\sin\theta_{21} + (\pi - \theta_{22})\sin\theta_{22}\big).
\end{aligned}
$$

Finally, using $\theta_{12} = \theta_{11} + \alpha$ and $\theta_{21} = \theta_{22} + \alpha$,

$$
\begin{aligned}
&\cos\theta_{11}\cos\theta_{22} - \cos\theta_{12}\cos\theta_{21} \\
&= \cos\theta_{11}\cos\theta_{22} - \cos(\theta_{11} + \alpha)\cos(\theta_{22} + \alpha) \\
&= \cos\theta_{11}\cos\theta_{22} - \big(\cos\theta_{11}\cos\theta_{22}\cos^2\alpha - \cos\theta_{11}\sin\theta_{22}\sin\alpha\cos\alpha \\
&\qquad\qquad - \sin\theta_{11}\cos\theta_{22}\sin\alpha\cos\alpha + \sin\theta_{11}\sin\theta_{22}\sin^2\alpha\big) \\
&= \sin^2\alpha\big(\cos\theta_{11}\cos\theta_{22} - \sin\theta_{11}\sin\theta_{22}\big) + \sin\alpha\cos\alpha\big(\cos\theta_{11}\sin\theta_{22} + \sin\theta_{11}\cos\theta_{22}\big) \\
&= \sin\alpha\big(\sin\alpha\cos(\theta_{11} + \theta_{22}) + \cos\alpha\sin(\theta_{11} + \theta_{22})\big) \\
&= \sin\alpha\sin(\theta_{11} + \theta_{22} + \alpha) = \sin\alpha\sin\gamma.
\end{aligned}
$$

And thus

$$
\begin{aligned}
\frac{E_1(\theta_1, \theta_2)}{\sin\alpha} = &- \sin\gamma\big((\pi - \theta_{11})\sin\theta_{11} + (\pi - \theta_{12})\sin\theta_{12}\big) \\
&+ (\pi - \gamma)\big((\pi - \theta_{11})\sin\theta_{21} + (\pi - \theta_{12})\sin\theta_{22} + (\sin\theta_{11} + \sin\theta_{12})\sin\gamma\big) \\
&- \pi\big((\pi - \theta_{21})\sin\theta_{21} + (\pi - \theta_{22})\sin\theta_{22}\big).
\end{aligned}
$$

And thus we have that

$$
\begin{aligned}
\frac{E_1(\theta_1, \theta_2)}{\sin\alpha} = &\sin\gamma\Big(\theta_{11}\sin\theta_{11} + \theta_{12}\sin\theta_{12} - \gamma(\sin\theta_{11} + \sin\theta_{12})\Big) \\
&+ \Big((\pi - \gamma)(\pi - \theta_{11}) - \pi(\pi - \theta_{21})\Big)\sin\theta_{21} \\
&+ \Big((\pi - \gamma)(\pi - \theta_{12}) - \pi(\pi - \theta_{22})\Big)\sin\theta_{22}.
\end{aligned}
$$

Under this parametrization,

$$
\theta_{12} = \theta_{11} + \alpha, \qquad \theta_{21} = \theta_{22} + \alpha, \qquad \gamma = \theta_{11} + \theta_{21} = \theta_{12} + \theta_{22}.
$$

Equivalently,

$$
\theta_{21} = \gamma - \theta_{11}, \qquad \theta_{22} = \gamma - \theta_{12}.
$$

Therefore,

$$\begin{aligned}(\pi-\gamma)(\pi-\theta_{11})-\pi(\pi-\theta_{21}) &= (\pi-\gamma)(\pi-\theta_{11})-\pi(\pi-(\gamma-\theta_{11}))\\ &= -(2\pi-\gamma)\theta_{11},\\ (\pi-\gamma)(\pi-\theta_{12})-\pi(\pi-\theta_{22}) &= -(2\pi-\gamma)\theta_{12}.\end{aligned}$$

Hence

$$\begin{aligned}\frac{E_1(\theta_1,\theta_2)}{\sin\alpha} &= \sin\gamma\Big(\theta_{11}\sin\theta_{11}+\theta_{12}\sin\theta_{12}-\gamma(\sin\theta_{11}+\sin\theta_{12})\Big)\\ &\quad -(2\pi-\gamma)\Big(\theta_{11}\sin\theta_{21}+\theta_{12}\sin\theta_{22}\Big).\end{aligned}$$

Since $0<\gamma<\pi$, we have $\sin\gamma>0$ and $2\pi-\gamma>0$. Moreover,

$$0\le\theta_{11}\le\gamma,\qquad 0\le\theta_{12}\le\gamma,\qquad 0\le\theta_{21}\le\gamma,\qquad 0\le\theta_{22}\le\gamma.$$

Thus

$$\theta_{11}\sin\theta_{11}+\theta_{12}\sin\theta_{12}-\gamma(\sin\theta_{11}+\sin\theta_{12}) = -(\gamma-\theta_{11})\sin\theta_{11}-(\gamma-\theta_{12})\sin\theta_{12}\le 0,$$

and also

$$\theta_{11}\sin\theta_{21}+\theta_{12}\sin\theta_{22}\ge 0.$$

Consequently $E_1(\theta_1,\theta_2)\le 0$.

We now inspect the equality case. If $E_1(\theta_1,\theta_2)=0$, then both non-positive contributions above must vanish. From

$$\theta_{11}\sin\theta_{21}+\theta_{12}\sin\theta_{22}=0$$

and the non-negativity of all terms, we get $\theta_{11}\sin\theta_{21}=0$ and $\theta_{12}\sin\theta_{22}=0$. Here $\theta_{21}\ge\alpha>0$ and $\theta_{21}<\pi$, so $\sin\theta_{21}>0$. Also $\theta_{12}\ge\alpha>0$. Therefore

$$\theta_{11}=0,\qquad \sin\theta_{22}=0.$$

Since $0\le\theta_{22}\le\gamma<\pi$, this gives $\theta_{22}=0$. Therefore

$$\theta_{12}=\alpha,\qquad \theta_{21}=\alpha,\qquad \gamma=\alpha.$$

Hence $\bar{w}_1=v_1$ and $\bar{w}_2=v_2$. This is the only possible solution in Case 4, and substituting back into (118) and (117) gives $\|w_1\|_2=\|w_2\|_2=1$, so $w_1=v_1$ and $w_2=v_2$.

Case 5: w_1 and w_2 are both lie in Area 3 in Figure 9

Without loss of generality, we assume $\theta_1\le\theta_2$. Recall $\theta_2-\theta_1=\gamma$, Let $\theta_1+\theta_2=p$, we have

$$\begin{aligned}&\theta_1=\frac{p-\gamma}{2},\quad \theta_2=\frac{p+\gamma}{2},\quad \theta_1-\theta_2=-\gamma,\\ &\theta_{11}=\theta_1-\frac{\alpha}{2}=\frac{p-\gamma-\alpha}{2},\\ &\theta_{12}=2\pi-\theta_1-\frac{\alpha}{2}=2\pi+\frac{\gamma-p-\alpha}{2},\\ &\theta_{11}+\theta_{12}=2\pi-\alpha,\quad \theta_{12}-\theta_{11}=2\pi+\gamma-p,\end{aligned}$$

$$\theta_{21} = \theta_1 - \frac{\alpha}{2} + \gamma = \frac{p+\gamma-\alpha}{2},$$
$$\theta_{22} = 2\pi - \theta_1 - \frac{\alpha}{2} - \gamma = 2\pi - \frac{p+\gamma+\alpha}{2},$$
$$\theta_{21} + \theta_{22} = 2\pi - \alpha, \quad \theta_{22} - \theta_{21} = 2\pi - p - \gamma.$$

Let $\frac{\alpha}{2}\cos\frac{\alpha}{2} = A, \sin\frac{\alpha}{2} = B$. By (119) and (120), we have

$$\begin{aligned}
0 &= -2(\pi-\gamma)\sin\gamma \cdot A + 2\cos^2\frac{p-\gamma}{2}\cdot(\pi-\gamma)\sin\gamma\cdot B \\
&\quad + \pi\cos\frac{p-\gamma}{2}\left(2A\cdot\sin\frac{p+\gamma}{2} - (2\pi-p-\gamma)\cos\frac{p+\gamma}{2}\cdot B\right) \\
&\quad - \big(\sin\gamma\cos\frac{p-\gamma}{2} + (\pi-\gamma)\cos\frac{p+\gamma}{2}\big)\cdot\left(2A\cdot\sin\frac{p-\gamma}{2} - (2\pi+\gamma-p)\cos\frac{p-\gamma}{2}\cdot B\right) \\
&= \Big(-2(\pi-\gamma)\sin\gamma + \pi(\sin p+\sin\gamma) - \sin\gamma\sin(p-\gamma) - (\pi-\gamma)(\sin p - \sin\gamma)\Big)\cdot A \\
&\quad + \Big(\big(2(\pi-\gamma) + (2\pi+\gamma-p)\big)\sin\gamma\cos^2\frac{p-\gamma}{2} \\
&\qquad + \big(-\pi(2\pi-p-\gamma) + (\pi-\gamma)(2\pi+\gamma-p)\big)\cdot\cos\frac{p+\gamma}{2}\cdot\cos\frac{p-\gamma}{2}\Big)\cdot B \\
&= \Big(\gamma(\sin p+\sin\gamma) - \sin\gamma\sin(p-\gamma)\Big)\cdot A \\
&\quad + \Big((4\pi-\gamma-p)\sin\gamma\cdot\frac{1+\cos(p-\gamma)}{2} + (p\gamma-\gamma^2)\cdot\frac{\cos p+\cos\gamma}{2}\Big)\cdot B \\
&= \Big(\gamma(\sin p+\sin\gamma) - \sin\gamma\big(\sin p\cos\gamma - \cos p\sin\gamma\big)\Big)\cdot A \\
&\quad + \Big((4\pi-\gamma-p)\sin\gamma\cdot\big(\frac{1+\cos p\cos\gamma}{2} + \frac{\sin p\sin\gamma}{2}\big) + (p\gamma-\gamma^2)\cdot\frac{\cos p+\cos\gamma}{2}\Big)\cdot B
\end{aligned} \tag{145}$$

$$\begin{aligned}
0 &= -2A\cdot(\pi-\gamma)\sin\gamma + 2B\cdot\cos^2\frac{p+\gamma}{2}(\pi-\gamma)\sin\gamma \\
&\quad - \pi\cos\frac{p+\gamma}{2}\left(2A\cdot\sin\frac{p-\gamma}{2} - (2\pi+\gamma-p)\cos\frac{p-\gamma}{2}B\right) \\
&\quad + (\sin\gamma\cos\frac{p+\gamma}{2} + (\pi-\gamma)\cos\frac{p-\gamma}{2})\cdot\Big(2A\cdot\sin\frac{p+\gamma}{2} - (2\pi-p-\gamma)\cos\frac{p+\gamma}{2}B) \\
&= \Big(-2(\pi-\gamma)\sin\gamma - \pi(\sin p-\sin\gamma) + \sin\gamma\sin(p+\gamma) + (\pi-\gamma)(\sin p+\sin\gamma)\Big)\cdot A \\
&\quad + \Big(\big(2(\pi-\gamma) - (2\pi-p-\gamma)\big)\sin\gamma\cos^2\frac{p+\gamma}{2} \\
&\qquad + \big(\pi(2\pi+\gamma-p) - (\pi-\gamma)(2\pi-p-\gamma)\big)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\Big)\cdot B \\
&= \Big(-\gamma\sin p + \gamma\sin\gamma + \sin\gamma\sin(p+\gamma)\Big)\cdot A
\end{aligned}$$

$$
\begin{aligned}
&+\left((p-\gamma)\sin\gamma\frac{1+\cos(p+\gamma)}{2}+(4\pi\gamma-p\gamma-\gamma^2)\cos\frac{p+\gamma}{2}\cos\frac{p-\gamma}{2}\right)\cdot B\\
=&\left(-\gamma\sin p+\gamma\sin\gamma+\sin\gamma(\sin p\cos\gamma+\cos p\sin\gamma)\right)\cdot A\\
&+\left((p-\gamma)\sin\gamma(\frac{1+\cos p\cos\gamma}{2}-\frac{\sin p\sin\gamma}{2})\right.\\
&\qquad\left.+(4\pi\gamma-p\gamma-\gamma^2)\frac{\cos p+\cos\gamma}{2}\right)\cdot B. \qquad (146)
\end{aligned}
$$

(145)+(146) yields

$$
\begin{aligned}
0=&\underbrace{(2\gamma\sin\gamma+2\cos p\sin^2\gamma)}_{\text{I}}\cdot A\\
&+\underbrace{\begin{pmatrix}(2\pi-\gamma)\sin\gamma(1+\cos p\cos\gamma)+(2\pi-p)\sin p\sin^2\gamma\\ +(2\pi\gamma-\gamma^2)(\cos p+\cos\gamma)\end{pmatrix}}_{\text{II}}\cdot B, \qquad (147)
\end{aligned}
$$

and (145)−(146) yields

$$
\begin{aligned}
0=&\underbrace{(2\gamma\sin p-2\sin p\sin\gamma\cos\gamma)}_{\text{III}}\cdot A\\
&+\underbrace{\begin{pmatrix}(2\pi-p)\sin\gamma(1+\cos p\cos\gamma)+(2\pi-\gamma)\sin p\sin^2\gamma\\ +(p\gamma-2\pi\gamma)(\cos p+\cos\gamma)\end{pmatrix}}_{\text{IV}}\cdot B. \qquad (148)
\end{aligned}
$$

By, (147) and (148), we have $\text{I}\cdot\text{IV}=\text{II}\cdot\text{III}$. Substituting p with $2\pi-p$, we reduce to case 3. Then, we have $2\pi-p=0,\gamma=\alpha$. Revisiting equation (147), we see that the left-hand side is positive, which leads to a contradiction.

So far, we have identified all the critical points in the plane spanned by v_1 and v_2, which are as follows:

- $\bar{w}_1=\bar{w}_2=\bar{v}$, $\|w_1\|_2+\|w_2\|_2=\frac{2}{\pi}\left((\pi-\frac{\alpha}{2})\cos\frac{\alpha}{2}+\sin\frac{\alpha}{2}\right)$.
- $\bar{w}_1=\bar{w}_2=-\bar{v}$, $\|w_1\|_2+\|w_2\|_2=\frac{2}{\pi}\left(\sin\frac{\alpha}{2}-\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$.
- $\bar{w}_1=\bar{v}$, $\bar{w}_2=-\bar{v}$, $\|w_1\|_2=\frac{2}{\pi}\left((\pi-\frac{\alpha}{2})\cos\frac{\alpha}{2}+\sin\frac{\alpha}{2}\right)$, $\|w_2\|_2=\frac{2}{\pi}\left(\sin\frac{\alpha}{2}-\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$.
- $\bar{w}_1=-\bar{v}$, $\bar{w}_2=\bar{v}$, $\|w_1\|_2=\frac{2}{\pi}\left(\sin\frac{\alpha}{2}-\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$, $\|w_2\|_2=\frac{2}{\pi}\left((\pi-\frac{\alpha}{2})\cos\frac{\alpha}{2}+\sin\frac{\alpha}{2}\right)$.
- $\bar{w}_1=v_1$, $\bar{w}_2=v_2$, $\|w_1\|_2=\|w_2\|_2=1$.
- $\bar{w}_1=v_2$, $\bar{w}_2=v_1$, $\|w_1\|_2=\|w_2\|_2=1$.

In the following, we will show that all of these critical points are either saddle points or global minima, thereby completing the proof of the proposition.

Recall that

$$u_1 = \bar{v}, u_2 = \frac{v_1 - v_2}{\|v_1 - v_2\|_2}, u_3, \ldots, u_d \tag{149}$$

is a orthonormal basis of $\mathbb{R}^d$. Let $e_{2i-1} = (u_i, 0_d), e_{2i} = (0_d, u_i), i = 1, 2, \ldots, d$. Then $\{e_i\}_{i=1}^{2d}$ is a orthonormal basis of $\mathbb{R}^{2d}$. Let $H(w_1, w_2)$ denote the Hessian matrix expressed in this basis. It is sufficient to prove that H has a negative eigenvalue.

Case 1: $\bar{w}_1 = \bar{w}_2 = \bar{v}$, $\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right)$ or $\bar{w}_1 = \bar{w}_2 = -\bar{v}$, $\|w_1\|_2 + \|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$.

By lemma 14, we have

$$\begin{aligned}
\frac{\partial^2 L}{\partial w_1^2} &= \frac{1}{2}I_d + h_1(w_1, w_2) - h_1(w_1, v_1) - h_1(w_1, v_2)\\
&= \frac{1}{2}I_d - \frac{\sin(\alpha/2)}{2\pi\|w_1\|_2}\left(I_d - \bar{v}\bar{v}^\top + u_2u_2^\top\right) - \frac{\sin(\alpha/2)}{2\pi\|w_1\|_2}\left(I_d - \bar{v}\bar{v}^\top + u_2u_2^\top\right)\\
&= \left(\frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}\right)I_d + \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}\bar{v}\bar{v}^\top - \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}u_2u_2^\top,\\
\frac{\partial^2 L}{\partial w_2^2} &= \frac{1}{2}I_d + h_1(w_2, w_1) - h_1(w_2, v_1) - h_1(w_2, v_2)\\
&= \frac{1}{2}I_d - \frac{\sin(\alpha/2)}{2\pi\|w_2\|_2}\left(I_d - \bar{v}\bar{v}^\top + u_2u_2^\top\right) - \frac{\sin(\alpha/2)}{2\pi\|w_2\|_2}\left(I_d - \bar{v}\bar{v}^\top + u_2u_2^\top\right)\\
&= \left(\frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}\right)I_d + \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}\bar{v}\bar{v}^\top - \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}u_2u_2^\top,\\
\frac{\partial^2 L}{\partial w_1 \partial w_2} &= h_2(w_1, w_2) = \frac{1}{2}I_d,\\
\frac{\partial^2 L}{\partial w_2 \partial w_1} &= h_2(w_2, w_1) = \frac{1}{2}I_d.
\end{aligned} \tag{150}$$

Then we have

$$H = \begin{bmatrix}
a_1 & 1 & & & & & \\
1 & b_1 & & & & & \\
 & & a_2 & 1 & & & \\
 & & 1 & b_2 & & & \\
 & & & & \ddots & & \\
 & & & & & a_d & 1\\
 & & & & & 1 & b_d
\end{bmatrix},$$

where

$$
\begin{aligned}
&a_1 = b_1 = \frac{1}{2}, \quad a_2 = \frac{1}{2} - \frac{2\sin(\alpha/2)}{\pi\|w_1\|_2}, \quad b_2 = \frac{1}{2} - \frac{2\sin(\alpha/2)}{\pi\|w_2\|_2}, \\
&a_i = \frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}, \quad b_i = \frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}, \quad 3 \le i \le d.
\end{aligned} \tag{151}
$$

We will prove that

$$
\begin{bmatrix} a_2 & 1 \\ 1 & b_2 \end{bmatrix}
$$

has a negative eigenvalue, which implies H has a negative eigenvalue. It is suffice to prove that $a_2 b_2 < 1$. Notice that $a_2 \le \frac{1}{2}$ and $b_2 \le \frac{1}{2}$. If $a_2 b_2 \ge 1$, then $a_2 < 0$ and $b_2 < 0$. We have

$$
\begin{aligned}
a_2 > 0 &\Longrightarrow \|w_1\|_2 > \frac{\pi}{4\sin(\alpha/2)}, \\
b_2 > 0 &\Longrightarrow \|w_2\|_2 > \frac{\pi}{4\sin(\alpha/2)}.
\end{aligned}
$$

Notice that

$$
\begin{aligned}
\sin\frac{\alpha}{2}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right) &= \frac{1}{2}\left(\pi - \frac{\alpha}{2}\right)\sin\alpha + \sin^2\frac{\alpha}{2} \\
&\overset{(*)}{<} \frac{1}{2}\left(\pi - \frac{\alpha}{2}\right)\cdot\alpha + \left(\frac{\alpha}{2}\right)^2 \\
&= \frac{\pi}{2}\cdot\alpha \\
&\overset{(\star)}{\le} \frac{\pi^2}{4},
\end{aligned} \tag{152}
$$

where $(*)$ comes from $\sin\alpha < \alpha$ for all $\alpha > 0$ and $(\star)$ comes from $\alpha \le \frac{\pi}{2}$. Then, we can get

$$
\begin{aligned}
\|w_1\|_2 + \|w_2\|_2 &> \frac{\pi}{2\sin(\alpha/2)} \\
&> \frac{2}{\pi}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right), \\
&> \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - (\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2}\right).
\end{aligned}
$$

which contradicts with the assumption of sum of the weight norm.

Case 2: $\bar{w}_1 = \bar{v}$, $\bar{w}_2 = -\bar{v}$, $\|w_1\|_2 = \frac{2}{\pi}\left((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\right)$, $\|w_2\|_2 = \frac{2}{\pi}\left(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\right)$. By lemma 14, we have

$$
\begin{aligned}
\frac{\partial^2 L}{\partial w_1^2} &= \left(\frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}\right)I_d + \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}\bar{v}\bar{v}^\top - \frac{\sin(\alpha/2)}{\pi\|w_1\|_2}u_2u_2^\top, \\
\frac{\partial^2 L}{\partial w_2^2} &= \left(\frac{1}{2} - \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}\right)I_d + \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}\bar{v}\bar{v}^\top - \frac{\sin(\alpha/2)}{\pi\|w_2\|_2}u_2u_2^\top,
\end{aligned}
$$

$$\frac{\partial^2 L}{\partial w_1 \partial w_2} = 0,$$
$$\frac{\partial^2 L}{\partial w_1 \partial w_2} = 0. \tag{153}$$

Then Hessian matrix has negative eigenvalues $-\frac{\sin(\alpha/2)}{\pi\|w_1\|_2}$ and $-\frac{\sin(\alpha/2)}{\pi\|w_2\|_2}$.

Case 3: $\bar{w}_1 = -\bar{v}$, $\bar{w}_2 = -\bar{v}$, $\|w_1\|_2 = \frac{2}{\pi}\Big(\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}\Big)$, $\|w_2\|_2 = \frac{2}{\pi}\Big(\big(\pi - \frac{\alpha}{2}\big)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}\Big)$. The proof is similar to case 2. ■

Appendix C. Proofs for training dynamics

C.1. Phase 1: the students remain close to the bisector

We will show that with small initialization, two student weights will both keep close to the bisector of the v_1 and v_2. Recall that we use the Gaussian initialization such that $w_i(0) \sim \mathcal{N}(0, \sigma^2 I_d), \quad i = 1, 2$.

Lemma 29 *Suppose $d = \Omega\big(\log(n/\delta)\big)$. Then, with probability at least $1-\delta$, the initialization satisfies*

$$\frac{1}{2}\sigma\sqrt{d} \le \|w_i(0)\|_2 \le 2\sigma\sqrt{d}, \qquad \frac{\pi}{3} \le \angle(\bar{w}_i(0), \bar{v}) \le \frac{2\pi}{3}, \quad i = 1, 2.$$

The proof follows from Lemma 3 of [40].

Remark 30 *The only randomness comes from the initialization. Once the initialization is fixed, the rest of the proof is deterministic. For simplicity, we condition on this high-probability event throughout the proof.*

Let

$$C(\alpha) := 2l\left(\pi - \frac{\alpha}{2}\right) = \frac{\sin\frac{\alpha}{2} - \frac{\alpha}{2}\cos\frac{\alpha}{2}}{\pi}.$$

Lemma 31 *For $i = 1, 2$, we have*

$$C(\alpha) - \frac{1}{2}\big(\|w_1\|_2 + \|w_2\|_2\big) \le \frac{\mathrm{d}\|w_i\|_2}{\mathrm{d}t} \le 1 - \frac{1}{2}\|w_i\|_2 \le 1. \tag{154}$$

Proof It suffices to prove the claim for $i = 1$, since the case $i = 2$ is identical. By (9),

$$\frac{\mathrm{d}\|w_1\|_2}{\mathrm{d}t} = \frac{1}{2\pi}\big((\sin\theta_{11} + \sin\theta_{12}) + (\pi - \theta_{11})\cos\theta_{11} + (\pi - \theta_{12})\cos\theta_{12}\big) - \frac{1}{2}\|w_1\|_2 - \frac{\sin\gamma + (\pi-\gamma)\cos\gamma}{2\pi}\|w_2\|_2$$

$$\begin{aligned} &= l(\theta_{11}) + l(\theta_{12}) - \frac{1}{2}\|w_1\|_2 - l(\gamma)\|w_2\|_2 \\ &\leq \frac{1}{2} + \frac{1}{2} - \frac{1}{2}\|w_1\|_2 \\ &= 1 - \frac{1}{2}\|w_1\|_2. \end{aligned} \tag{155}$$

For the lower bound, we similarly have

$$\begin{aligned} \frac{\mathrm{d}\|w_1\|_2}{\mathrm{d}t} &= l(\theta_{11}) + l(\theta_{12}) - \frac{1}{2}\|w_1\|_2 - l(\gamma)\|w_2\|_2 \\ &\geq C(\alpha) - \frac{1}{2}\|w_1\|_2 - \frac{1}{2}\|w_2\|_2. \end{aligned} \tag{156}$$

Here the last inequality follows from lemma 10 and $l(\gamma) \leq 1/2$. ■

Corollary 32 *For all $t \geq 0$ and $i = 1, 2$,*

$$\|w_i(t)\|_2 \leq \|w_i(0)\|_2 + t \leq 2\sigma\sqrt{d} + t.$$

Corollary 33 *If $\|w_i(0)\|_2 \leq 2$ for $i = 1, 2$, then*

$$\|w_i(t)\|_2 \leq 2, \quad \forall t \geq 0, \quad i = 1, 2.$$

Proof By Equation (154),

$$\frac{\mathrm{d}\|w_i(t)\|_2}{\mathrm{d}t} \leq 1 - \frac{1}{2}\|w_i(t)\|_2 = -\frac{1}{2}\big(\|w_i(t)\|_2 - 2\big), \quad i = 1, 2.$$

Gronwall's inequality gives

$$\|w_i(t)\|_2 - 2 \leq \big(\|w_i(0)\|_2 - 2\big)e^{-t/2} \leq 0.$$

■

Corollary 34 *For $\alpha \in (0, \frac{\pi}{2}]$, we have*

$$0 < C(\alpha) < 1.$$

Proof Recall that

$$C(\alpha) = 2l\left(\pi - \frac{\alpha}{2}\right).$$

Since $\alpha \in (0, \frac{\pi}{2}]$, we have

$$\pi - \frac{\alpha}{2} \in \left[\frac{3\pi}{4}, \pi\right).$$

By lemma 9 and the strict monotonicity of l on $(0, \pi)$,

$$0 = l(\pi) < l\left(\pi - \frac{\alpha}{2}\right) < l(0) = \frac{1}{2}.$$

Therefore $0 < C(\alpha) < 1$. ■

Corollary 35 *If $\sigma \leq \frac{C(\alpha)}{200\sqrt{d}}$, then for all $0 \leq t \leq \frac{C(\alpha)}{100}$,*

$$\frac{C(\alpha)}{10} \leq \frac{\mathrm{d}\|w_i(t)\|_2}{\mathrm{d}t} \leq 1, \quad i = 1, 2.$$

Proof The upper bound in Equation (154) gives

$$\frac{\mathrm{d}\|w_i(t)\|_2}{\mathrm{d}t} \leq 1, \quad i = 1, 2.$$

Moreover, by corollary 32 and the assumptions on σ and t,

$$\|w_i(t)\|_2 \leq 2\sigma\sqrt{d} + t \leq \frac{C(\alpha)}{100} + t \leq \frac{C(\alpha)}{50}, \quad i = 1, 2.$$

Therefore, applying the lower bound in Equation (154), we obtain

$$\begin{aligned}\frac{\mathrm{d}\|w_i(t)\|_2}{\mathrm{d}t} &\geq C(\alpha) - \frac{1}{2}(\|w_1\|_2 + \|w_2\|_2) \\ &\geq C(\alpha) - \frac{1}{2}\left(\frac{C(\alpha)}{50} + \frac{C(\alpha)}{50}\right) \\ &= \frac{49}{50}C(\alpha) \\ &\geq \frac{C(\alpha)}{10}.\end{aligned}$$

■

Corollary 36 *If $\sigma \leq \frac{C(\alpha)}{200\sqrt{d}}$, then*

$$\frac{C(\alpha)^2}{1000} \leq \left\|w_i\left(\frac{C(\alpha)}{100}\right)\right\|_2 \leq \frac{C(\alpha)}{50}, \quad i = 1, 2.$$

Proof Let

$$t_\alpha := \frac{C(\alpha)}{100}.$$

By the previous corollary, for every $0 \leq t \leq t_\alpha$,

$$\frac{\mathrm{d}\|w_i(t)\|}{\mathrm{d}t} \geq \frac{C(\alpha)}{10}, \qquad i = 1, 2.$$

Integrating from 0 to t_α gives

$$\|w_i(t_\alpha)\| \geq \|w_i(0)\| + \frac{C(\alpha)}{10} t_\alpha \geq \frac{C(\alpha)^2}{1000}.$$

For the upper bound, corollary 32 and the assumption on σ give

$$\|w_i(t_\alpha)\| \leq 2\sigma\sqrt{d} + t_\alpha \leq \frac{C(\alpha)}{100} + \frac{C(\alpha)}{100} = \frac{C(\alpha)}{50}.$$

■

Corollary 37 *If $\sigma \leq \frac{C(\alpha)}{200\sqrt{d}}$, then for all $0 \leq t \leq \frac{C(\alpha)}{100}$,*

$$\frac{C(\alpha)}{10} \leq \frac{\|w_1(t)\|_2}{\|w_2(t)\|_2} \leq \frac{10}{C(\alpha)}.$$

Proof By the initialization bound and the previous corollary, for $i = 1, 2$,

$$\frac{1}{2}\sigma\sqrt{d} + \frac{C(\alpha)}{10} t \leq \|w_i(t)\|_2 \leq 2\sigma\sqrt{d} + t.$$

Hence, using $0 < C(\alpha) \leq 1$,

$$\frac{\|w_1(t)\|_2}{\|w_2(t)\|_2} \leq \frac{2\sigma\sqrt{d} + t}{\frac{1}{2}\sigma\sqrt{d} + \frac{C(\alpha)}{10} t} \leq \frac{10}{C(\alpha)}.$$

The lower bound follows by the same argument with w_1 and w_2 interchanged:

$$\frac{\|w_1(t)\|_2}{\|w_2(t)\|_2} \geq \frac{\frac{1}{2}\sigma\sqrt{d} + \frac{C(\alpha)}{10} t}{2\sigma\sqrt{d} + t} \geq \frac{C(\alpha)}{10}.$$

■

Lemma 38 *$\forall i \in \{1, 2\}$ and $\forall \bar{w}_i \in \mathbb{S}^{d-1}$, we have*

$$\left\langle \bar{v}, (I - \bar{w}_i \bar{w}_i^\top)\Big((\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2\Big) \right\rangle \geq 0.$$

Moreover,

$$\left\langle \bar{v}, (I - \bar{w}_i \bar{w}_i^\top)\Big((\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2\Big) \right\rangle = 0 \quad \Longleftrightarrow \quad \bar{w}_i = \bar{v} \text{ or } \bar{w}_i = -\bar{v}.$$

Proof Let

$$\begin{aligned}
g(\bar{w}_i) &:= \left\langle \bar{v}, (I - \bar{w}_i\bar{w}_i^\top)\big((\pi-\theta_{i1})v_1 + (\pi-\theta_{i2})v_2\big) \right\rangle \\
&= \left\langle \bar{v}, (\pi-\theta_{i1})v_1 + (\pi-\theta_{i2})v_2 \right\rangle \\
&\quad - \langle \bar{v}, \bar{w}_i\rangle \Big((\pi-\theta_{i1})\langle v_1, \bar{w}_i\rangle + (\pi-\theta_{i2})\langle v_2, \bar{w}_i\rangle \Big) \\
&= \cos\frac{\alpha}{2}\Big((\pi-\theta_{i1}) + (\pi-\theta_{i2}) \Big) \\
&\quad - \frac{\cos\theta_{i1} + \cos\theta_{i2}}{2\cos\frac{\alpha}{2}} \Big((\pi-\theta_{i1})\cos\theta_{i1} + (\pi-\theta_{i2})\cos\theta_{i2} \Big). \qquad (157)
\end{aligned}$$

Let $\theta_{i1}+\theta_{i2} = x, \theta_{i1}-\theta_{i2} = y$. Recall that $\theta_{i1} = \angle(\bar{w}_i, v_1), \theta_{i2} = \angle(\bar{w}_i, v_2)$. Without loss of generality, we assume $\theta_{i1} \geq \theta_{i2}$. By the triangular inequality, we have $\alpha \leq x \leq 2\pi - \alpha, 0 \leq y \leq \alpha \leq \frac{\pi}{2}$. If $\alpha \leq x \leq \pi$, we have

$$(2\pi - x)\cos^2\frac{\alpha}{2} - \sin x \geq \pi\cos^2\frac{\alpha}{2} - 1 \geq \frac{\pi}{2} - 1 > 0,$$

where the second inequality using the fact that $\alpha \in (0, \frac{\pi}{2}]$. If $\pi < x \leq 2\pi - \alpha$, we can also get

$$(2\pi - x)\cos^2\frac{\alpha}{2} - \sin x \geq (2\pi - x)\cos^2\frac{\alpha}{2} > 0.$$

Then, we have $(2\pi - x)\cos^2\frac{\alpha}{2} \geq \sin x$ for all $x \in [\alpha, 2\pi-\alpha]$. By (157), we have

$$\begin{aligned}
2\cos\frac{\alpha}{2}\cdot g(\bar{w}_i) &= 2\cos^2\frac{\alpha}{2}(2\pi - x) \\
&\quad - \left(\cos\frac{x+y}{2} + \cos\frac{x-y}{2} \right) \\
&\quad \cdot \left(\big(\pi - \frac{x+y}{2}\big)\cos\frac{x+y}{2} + \big(\pi - \frac{x-y}{2}\big)\cos\frac{x-y}{2} \right) \\
&= 2\cos^2\frac{\alpha}{2}(2\pi - x) - 2\cos\frac{x}{2}\cos\frac{y}{2} \\
&\quad \cdot \left((2\pi - x)\cos\frac{x}{2}\cos\frac{y}{2} + y\sin\frac{x}{2}\sin\frac{y}{2} \right). \\
&= 2(2\pi - x)\left(\cos^2\frac{\alpha}{2} - \cos^2\frac{x}{2}\cos^2\frac{y}{2} \right) \\
&\quad - 2y\sin\frac{x}{2}\cos\frac{x}{2}\sin\frac{y}{2}\cos\frac{y}{2} \\
&\overset{(\star)}{\geq} 2(2\pi - x)\left(\cos^2\frac{\alpha}{2} - \cos^2\frac{\alpha}{2}\cos^2\frac{y}{2} \right) \\
&\quad - 2y\sin\frac{x}{2}\cos\frac{x}{2}\sin\frac{y}{2}\cos\frac{y}{2} \\
&= 2(2\pi - x)\cos^2\frac{\alpha}{2}\sin^2\frac{y}{2}
\end{aligned}$$

$$
\begin{aligned}
&\quad - y\sin x \sin\frac{y}{2}\cos\frac{y}{2} \\
&= 2\sin\frac{y}{2}\bigg((2\pi - x)\cos^2\frac{\alpha}{2}\sin\frac{y}{2} \\
&\quad - \sin x\cdot\frac{y}{2}\cos\frac{y}{2}\bigg) \\
&\overset{(*)}{\geq} 2\sin\frac{y}{2}\bigg((2\pi - x)\cos^2\frac{\alpha}{2}\cdot\frac{y}{2}\cos\frac{y}{2} \\
&\quad - \sin x\cdot\frac{y}{2}\cos\frac{y}{2}\bigg) \\
&= 2\sin\frac{y}{2}\cdot\frac{y}{2}\cos\frac{y}{2} \\
&\quad \cdot\bigg((2\pi - x)\cos^2\frac{\alpha}{2} - \sin x\bigg) \\
&\geq 0,
\end{aligned} \tag{158}
$$

where $(*)$ comes from the fact that $\sin\frac{y}{2} \geq \frac{y}{2}\cos\frac{y}{2}, \forall y \in [0, \frac{\pi}{2}]$. It is obvious that $g(\bar{v}) = 0, g(-\bar{v}) = 0$. Vice versa, if $g(\bar{w}_i) = 0$, by (158), each inequality must hold. $(*)$ holds implies $y = 0$ and $(\star)$ holds implies $x = \alpha$ or $x = 2\pi - \alpha$, which is equivalent to $\bar{w} = \bar{v}$ or $\bar{w} = -\bar{v}$. ■

Lemma 39 $\forall\varepsilon > 0, \forall i \in \{1, 2\}$, *there exists* $\delta > 0$ *such that if* $\|\bar{w}_i - \bar{v}\| \geq \varepsilon$ *and* $\angle(\bar{w}_i, \bar{v}) \leq \frac{2\pi}{3}$, *then*

$$
\bigg\langle \bar{v}, (I - \bar{w}_i\bar{w}_i^\top)\Big((\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2\Big)\bigg\rangle \geq \delta,
$$

Proof Let

$$
g(\bar{w}_i) := \Big\langle \bar{v}, (I - \bar{w}_i\bar{w}_i^\top)\big((\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2\big)\Big\rangle. \tag{159}
$$

The constant obtained below is independent of the ambient dimension. Indeed, the formula (157) shows that g depends only on the two angles $(\theta_{i1}, \theta_{i2})$, or equivalently on

$$
x = \theta_{i1} + \theta_{i2}, \qquad y = \theta_{i1} - \theta_{i2}.
$$

The feasible angle pairs form a compact subset of the fixed two-dimensional region

$$
\alpha \leq x \leq 2\pi - \alpha, \qquad |y| \leq \alpha.
$$

Moreover, the additional conditions also depend only on the same angle pair. Indeed,

$$
\bar{v} = \frac{v_1 + v_2}{2\cos(\alpha/2)},
$$

and hence

$$
\langle \bar{w}_i, \bar{v}\rangle = \frac{\cos\theta_{i1} + \cos\theta_{i2}}{2\cos(\alpha/2)}.
$$

Therefore both

$$\|\bar{w}_i - \bar{v}\|^2 = 2 - 2\langle \bar{w}_i, \bar{v}\rangle$$

and

$$\angle(\bar{w}_i, \bar{v}) = \arccos\langle \bar{w}_i, \bar{v}\rangle$$

are continuous functions of $(\theta_{i1}, \theta_{i2})$. Hence the additional conditions

$$\|\bar{w}_i - \bar{v}\| \geq \varepsilon, \qquad \angle(\bar{w}_i, \bar{v}) \leq \frac{2\pi}{3}$$

cut out a closed subset of this same compact region.

If this closed subset is empty, then the implication in the statement is vacuous, and any positive δ works. Otherwise, by lemma 38, the only zeros of g are $\bar{v}$ and $-\bar{v}$. The first one is excluded by $\|\bar{w}_i - \bar{v}\| \geq \varepsilon$, and the second one is excluded by $\angle(\bar{w}_i, \bar{v}) \leq 2\pi/3$. Therefore g has a strictly positive minimum on this compact two-dimensional set. This minimum is the desired $\delta = \delta(\alpha, \varepsilon) > 0$. ■

Lemma 40 *For any $T_1 > 0$ and $\varepsilon > 0$, there exists $\sigma_0 = \sigma_0(\varepsilon, T_1) > 0$ such that, for every $\sigma \leq \sigma_0$, there exist times $t_1, t_2 \leq T_1$ satisfying*

$$\|\bar{w}_i(t_i) - \bar{v}\|_2 < \varepsilon, \quad i = 1, 2. \tag{160}$$

Proof We argue by contradiction. Suppose that the claim fails. Then there exist $T_1 > 0$ and $\varepsilon > 0$ such that, for every $\sigma_0 > 0$, one can choose $\sigma \leq \sigma_0$ for which at least one student remains outside the ε-neighborhood of $\bar{v}$ throughout $[0, T_1]$. By symmetry, we may assume that

$$\|\bar{w}_1(t) - \bar{v}\|_2 \geq \varepsilon, \quad 0 \leq t \leq T_1. \tag{161}$$

For each $i \in \{1, 2\}$, define

$$g(\bar{w}_i) := \left\langle \bar{v}, (I - \bar{w}_i\bar{w}_i^\top)\big((\pi - \theta_{i1})v_1 + (\pi - \theta_{i2})v_2\big) \right\rangle,$$

where $\theta_{ij} = \angle(\bar{w}_i, v_j)$ for $j = 1, 2$. By lemma 39, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$g(\bar{w}_1(t)) \geq \delta$$

whenever $\|\bar{w}_1(t) - \bar{v}\|_2 \geq \varepsilon$ and $\angle(\bar{w}_1(t), \bar{v}) \leq 2\pi/3$.

Set

$$B_\alpha := \frac{5}{C(\alpha)}, \qquad \eta_\alpha := \frac{\delta}{4\pi(B_\alpha + 1)}, \qquad T_* := \min\left\{T_1, \frac{C(\alpha)}{100}, \frac{\eta_\alpha}{2}\right\},$$

and choose

$$\sigma_0 := \min\left\{\frac{C(\alpha)}{200\sqrt{d}}, \frac{T_*}{2\sqrt{d}} \exp\left(-\frac{8\pi(B_\alpha + 1)}{\delta}\right)\right\}.$$

For the chosen $\sigma \le \sigma_0$, define

$$E_1(t) := \frac{1}{2}\|\bar{w}_1(t) - \bar{v}\|_2^2 = 1 - \cos\angle(\bar{w}_1(t), \bar{v}). \tag{162}$$

Whenever $\|w_2(t)\|_2/\|w_1(t)\|_2 \le 10/C(\alpha)$, we have

$$\begin{aligned}
\frac{\mathrm{d}E_1}{\mathrm{d}t} &= \left\langle \bar{w}_1 - \bar{v}, \frac{\mathrm{d}\bar{w}_1}{\mathrm{d}t} \right\rangle \\
&= \left\langle \bar{w}_1 - \bar{v}, \frac{1}{2\pi\|w_1\|_2}(I - \bar{w}_1\bar{w}_1^\top)\big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2 - (\pi - \gamma)w_2\big) \right\rangle \qquad (163)\\
&= -\frac{1}{2\pi\|w_1\|_2} g(\bar{w}_1) + \frac{\pi - \gamma}{2\pi}\frac{\|w_2\|_2}{\|w_1\|_2}\langle \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)\bar{w}_2\rangle \\
&\le -\frac{1}{2\pi\|w_1\|_2} g(\bar{w}_1) + B_\alpha. \qquad (164)
\end{aligned}$$

Let

$$\mathcal{T} := \left\{ s \ge 0 : \left.\frac{\mathrm{d}E_1}{\mathrm{d}t}\right|_{t=s} \ge 0 \right\}.$$

If $\mathcal{T} = \varnothing$, then $\mathrm{d}E_1/\mathrm{d}t < 0$ for all $t \ge 0$, so the integral estimate below applies directly on $[0, T_*]$. Thus, we assume $\mathcal{T} \ne \varnothing$ and set $\tau := \inf \mathcal{T}$. The initial ratio satisfies $\|w_1(0)\|_2/\|w_2(0)\|_2 \le 4 \le 10/C(\alpha)$. Since $\|w_1(0)\|_2 \le 2\sigma\sqrt{d} \le T_* \exp(-8\pi(B_\alpha + 1)/\delta)$, (164) yields

$$\begin{aligned}
\left.\frac{\mathrm{d}E_1}{\mathrm{d}t}\right|_{t=0} &\le -\frac{\delta}{2\pi T_*}\exp\left(\frac{8\pi(B_\alpha + 1)}{\delta}\right) + B_\alpha \\
&\le -\frac{\delta}{2\pi} \cdot \frac{8\pi(B_\alpha + 1)}{\delta} + B_\alpha \\
&= -4(B_\alpha + 1) + B_\alpha < 0.
\end{aligned}$$

Hence $\tau > 0$. We first show that $\tau > T_*$. Suppose instead that $\tau \le T_*$. Then $\frac{\mathrm{d}E_1}{\mathrm{d}t} < 0$ on $[0, \tau)$, and hence $E_1(s) \le E_1(0)$ for all $s < \tau$. By (162) and the initialization event, this implies

$$\angle(\bar{w}_1(s), \bar{v}) \le \angle(\bar{w}_1(0), \bar{v}) \le \frac{2\pi}{3}, \quad 0 \le s < \tau.$$

Together with (161), lemma 39 gives $g(\bar{w}_1(s)) \ge \delta$ for all $0 \le s < \tau$. Moreover, since $s < \tau \le T_*$,

$$\|w_i(s)\|_2 \le 2\sigma\sqrt{d} + s \le 2\sigma_0\sqrt{d} + T_* \le 2T_* \le \eta_\alpha, \quad i = 1, 2, \tag{165}$$

The bound above will be used to control the negative drift term. Separately, since $\sigma \le \sigma_0 \le \frac{C(\alpha)}{200\sqrt{d}}$ and $s \le T_* \le \frac{C(\alpha)}{100}$, the assumptions of corollary 37 are satisfied. Hence,

$$\frac{\|w_1(s)\|_2}{\|w_2(s)\|_2} \le \frac{10}{C(\alpha)}, \quad 0 \le s \le T_*. \tag{166}$$

Therefore, by (164),

$$\left.\frac{\mathrm{d}E_1}{\mathrm{d}t}\right|_{t=s} \le -\frac{\delta}{2\pi\|w_1(s)\|_2} + B_\alpha$$

$$\leq -\frac{\delta}{2\pi\eta_\alpha} + B_\alpha = -2(B_\alpha + 1) + B_\alpha < 0. \tag{167}$$

Letting $s \to \tau^-$ contradicts the definition of τ. Thus $\tau > T_*$.

Consequently, $\frac{\mathrm{d}E_1}{\mathrm{d}t} \leq 0$ on $[0, T_*]$. The same argument as above shows that, for all $0 \leq t \leq T_*$,

$$\angle(\bar{w}_1(t), \bar{v}) \leq \frac{2\pi}{3}, \qquad g(\bar{w}_1(t)) \geq \delta.$$

Combining (164) with corollary 32, we obtain

$$\begin{aligned}\frac{\mathrm{d}E_1}{\mathrm{d}t} &\leq -\frac{\delta}{2\pi\|w_1(t)\|_2} + B_\alpha \\ &\leq -\frac{\delta}{2\pi(2\sigma\sqrt{d} + t)} + B_\alpha.\end{aligned} \tag{168}$$

Hence,

$$\begin{aligned}E_1(0) - E_1(T_*) &\geq \int_0^{T_*} \left(\frac{\delta}{2\pi(2\sigma\sqrt{d} + t)} - B_\alpha\right) \mathrm{d}t \\ &= \frac{\delta}{2\pi} \log\left(1 + \frac{T_*}{2\sigma\sqrt{d}}\right) - B_\alpha T_* \\ &\geq \frac{\delta}{2\pi} \log\left(\frac{T_*}{2\sigma_0\sqrt{d}}\right) - B_\alpha T_* \\ &\geq 4(B_\alpha + 1) - B_\alpha T_* \\ &\geq 4(B_\alpha + 1) - B_\alpha > 2.\end{aligned} \tag{169}$$

This is impossible because $0 \leq E_1(t) \leq 2$ for all t. Therefore, the contradiction proves the desired claim. ■

Corollary 41 (Explicit initial alignment time) *On the initialization event in lemma 29, for every $\varepsilon > 0$, there exist constants*

$$K = K(\alpha, \varepsilon) > 0, \qquad c_\alpha = c_\alpha(\alpha, \varepsilon) > 0, \qquad C_\alpha = C_\alpha(\alpha, \varepsilon) > 0,$$

and $\sigma_0 = \sigma_0(\alpha, \varepsilon, d) > 0$ such that, for every $0 < \sigma \leq \sigma_0$, the time

$$\tau_{\text{init}} = K\sigma\sqrt{d}$$

satisfies

$$c_\alpha\sigma\sqrt{d} \leq \|w_i(t)\| \leq C_\alpha\sigma\sqrt{d}, \qquad 0 \leq t \leq \tau_{\text{init}}, \quad i = 1, 2,$$

and

$$\|\bar{w}_i(\tau_{\text{init}}) - \bar{v}\| < \varepsilon, \qquad i = 1, 2.$$

Proof Let $\delta_\varepsilon = \delta_\varepsilon(\alpha, \varepsilon) > 0$ be the constant from lemma 39 applied with tolerance $\varepsilon/2$. Thus

$$g(\bar{w}_i) \geq \delta_\varepsilon$$

whenever

$$\|\bar{w}_i - \bar{v}\| \geq \frac{\varepsilon}{2}, \qquad \angle(\bar{w}_i, \bar{v}) \leq \frac{2\pi}{3}.$$

Set

$$B_\alpha = \frac{5}{C(\alpha)}.$$

Choose $K = K(\alpha, \varepsilon) > 0$ so large that

$$\frac{\delta_\varepsilon}{2\pi} \log\left(1 + \frac{K}{4}\right) > 3. \tag{170}$$

We then choose σ_0 small enough so that

$$\sigma_0 \leq \frac{C(\alpha)}{200\sqrt{d}}, \qquad K\sigma_0\sqrt{d} \leq \frac{C(\alpha)}{100}, \tag{171}$$

and

$$B_\alpha K\sigma_0\sqrt{d} \leq \frac{\delta_\varepsilon}{2\pi}\left[\log\left(1 + \frac{K}{2}\right) - \log\left(1 + \frac{K}{4}\right)\right], \qquad B_\alpha K\sigma_0\sqrt{d} \leq \frac{\varepsilon^2}{8}. \tag{172}$$

For every $0 < \sigma \leq \sigma_0$, corollary 37 applies on $[0, K\sigma\sqrt{d}]$ by (171). In particular,

$$\frac{C(\alpha)}{10} \leq \frac{\|w_1(t)\|}{\|w_2(t)\|} \leq \frac{10}{C(\alpha)}, \qquad 0 \leq t \leq K\sigma\sqrt{d}. \tag{173}$$

Also, by the initialization event in lemma 29,

$$\frac{1}{2}\sigma\sqrt{d} \leq \|w_i(0)\| \leq 2\sigma\sqrt{d}, \qquad i = 1, 2.$$

Since the norm derivative is bounded by an absolute constant on this interval, for $0 \leq t \leq K\sigma\sqrt{d}$,

$$\|w_i(t)\| \leq \|w_i(0)\| + t \leq (2 + K)\sigma\sqrt{d}.$$

Reducing σ_0 further, we may assume

$$(2 + K)\sigma_0\sqrt{d} \leq \min\left\{\frac{C(\alpha)}{4}, \frac{\delta_\varepsilon}{4\pi B_\alpha}\right\}.$$

Hence, on $[0, K\sigma\sqrt{d}]$,

$$\|w_1(t)\| + \|w_2(t)\| \leq 2(2 + K)\sigma\sqrt{d} \leq \frac{C(\alpha)}{2}.$$

The lower bound in Equation (154) then gives

$$\frac{d}{dt}\|w_i(t)\| \geq C(\alpha) - \frac{1}{2}\big(\|w_1(t)\| + \|w_2(t)\|\big) \geq \frac{3C(\alpha)}{4} > 0.$$

Therefore

$$\|w_i(t)\| \geq \|w_i(0)\| \geq \frac{1}{2}\sigma\sqrt{d}, \qquad 0 \leq t \leq K\sigma\sqrt{d}.$$

Hence the stated norm estimate holds with

$$c_\alpha = \frac{1}{2}, \qquad C_\alpha = 2 + K.$$

It remains to prove the alignment at the endpoint. Suppose first that student i stays outside the $\varepsilon/2$-neighborhood of $\bar{v}$ on the entire interval:

$$\|\bar{w}_i(t) - \bar{v}\| \geq \frac{\varepsilon}{2}, \qquad 0 \leq t \leq K\sigma\sqrt{d}. \tag{174}$$

Define

$$E_i(t) = \frac{1}{2}\|\bar{w}_i(t) - \bar{v}\|^2.$$

The computation leading to (164) is symmetric in the two students. Hence, for $i = 1, 2$, whenever (173) holds,

$$\frac{d}{dt}E_i(t) \leq -\frac{1}{2\pi\|w_i(t)\|}g(\bar{w}_i(t)) + B_\alpha. \tag{175}$$

The initial event gives $\angle(\bar{w}_i(0), \bar{v}) \leq 2\pi/3$. We claim that, under (174),

$$\angle(\bar{w}_i(t), \bar{v}) \leq \frac{2\pi}{3}, \qquad 0 \leq t \leq K\sigma\sqrt{d}.$$

To see this, first note that (174), the initial angle bound, and lemma 39 give $g(\bar{w}_i(0)) \geq \delta_\varepsilon$. Moreover

$$\|w_i(0)\| \leq 2\sigma\sqrt{d} \leq (2 + K)\sigma_0\sqrt{d} \leq \frac{\delta_\varepsilon}{4\pi B_\alpha}.$$

Thus (175) gives $\frac{d}{dt}E_i(0) < 0$.

If there is no time in $[0, K\sigma\sqrt{d}]$ at which $dE_i/dt \geq 0$, then E_i is decreasing on the whole interval, and the claimed angle bound follows. Otherwise, let τ be the first such time. The previous paragraph shows that $\tau > 0$. For every $t < \tau$, the function E_i is decreasing, and therefore the angle to $\bar{v}$ cannot exceed its initial value. By continuity, the same angle bound holds at $t = \tau$. Together with (174) and lemma 39, this gives $g(\bar{w}_i(\tau)) \geq \delta_\varepsilon$. Since

$$\|w_i(\tau)\| \leq (2 + K)\sigma\sqrt{d} \leq \frac{\delta_\varepsilon}{4\pi B_\alpha},$$

(175) gives

$$\left.\frac{d}{dt}E_i(t)\right|_{t=\tau} \leq -\frac{\delta_\varepsilon}{2\pi\|w_i(\tau)\|} + B_\alpha \leq -B_\alpha < 0,$$

contradicting the definition of τ. Hence no such τ exists, E_i is decreasing on the whole interval, and the claimed angle bound follows. Therefore $g(\bar{w}_i(t)) \geq \delta_\varepsilon$ on this interval. Combining this with (175), (173), and $\|w_i(t)\| \leq 2\sigma\sqrt{d} + t$, we obtain

$$\frac{d}{dt}E_i(t) \leq -\frac{\delta_\varepsilon}{2\pi(2\sigma\sqrt{d} + t)} + B_\alpha.$$

Integrating from 0 to $K\sigma\sqrt{d}$ gives

$$\begin{aligned}E_i(0) - E_i(K\sigma\sqrt{d}) &\geq \int_0^{K\sigma\sqrt{d}} \left(\frac{\delta_\varepsilon}{2\pi(2\sigma\sqrt{d}+t)} - B_\alpha \right) dt \\ &= \frac{\delta_\varepsilon}{2\pi} \log\left(1 + \frac{K}{2}\right) - B_\alpha K\sigma\sqrt{d} \\ &\geq \frac{\delta_\varepsilon}{2\pi} \log\left(1 + \frac{K}{4}\right) > 3,\end{aligned}$$

where the penultimate inequality uses (172). This is impossible because $0 \leq E_i(t) \leq 2$. Hence for each i there exists $t_i \leq K\sigma\sqrt{d}$ such that

$$\|\bar{w}_i(t_i) - \bar{v}\| < \frac{\varepsilon}{2}.$$

Finally, (175) and $g(\bar{w}_i) \geq 0$ imply

$$\frac{d}{dt} E_i(t) \leq B_\alpha$$

whenever (173) holds. Therefore, using (172),

$$\begin{aligned}E_i(K\sigma\sqrt{d}) &\leq E_i(t_i) + B_\alpha(K\sigma\sqrt{d} - t_i) \\ &< \frac{\varepsilon^2}{8} + B_\alpha K\sigma\sqrt{d} \\ &\leq \frac{\varepsilon^2}{4}.\end{aligned}$$

Thus

$$\|\bar{w}_i(K\sigma\sqrt{d}) - \bar{v}\| = \sqrt{2E_i(K\sigma\sqrt{d})} < \varepsilon.$$

This proves the corollary. ■

Theorem 42 *For any $\varepsilon > 0$ and any $0 < t \leq \frac{C(\alpha)\min\{1,\varepsilon^2\}}{100}$, there exists $\sigma_0 = \sigma_0(\varepsilon, t) > 0$ such that, for every $\sigma \leq \sigma_0$,*

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon, \quad i = 1, 2.$$

Proof Set $B_\alpha := 5/C(\alpha)$. Apply corollary 41 with tolerance $\varepsilon/2$. This gives constants $K > 0$ and $\sigma_1 > 0$ such that, for every $0 < \sigma \leq \sigma_1$, the time

$$\tau_{\text{init}} = K\sigma\sqrt{d}$$

satisfies

$$\|\bar{w}_i(\tau_{\text{init}}) - \bar{v}\| < \frac{\varepsilon}{2}, \qquad i = 1, 2.$$

Choose

$$\sigma_0 := \min\left\{\sigma_1, \frac{t}{K\sqrt{d}}, \frac{C(\alpha)}{200\sqrt{d}}\right\}.$$

Then $\tau_{\text{init}} \le t$ for every $0 < \sigma \le \sigma_0$. Let

$$E_i(s) := \frac{1}{2}\|\bar{w}_i(s) - \bar{v}\|^2, \qquad i = 1, 2.$$

Then, for every $0 \le s \le t$, the assumptions of corollary 37 hold, since $s \le t \le C(\alpha)/100$ and $\sigma \le C(\alpha)/(200\sqrt{d})$. Hence (164) and its symmetric counterpart for $i = 2$ apply throughout $[0, t]$. In particular, by lemma 38,

$$\begin{aligned}\frac{\mathrm{d}E_i}{\mathrm{d}s} &\le -\frac{1}{2\pi\|w_i\|}g(\bar{w}_i) + B_\alpha \\ &\le B_\alpha, \qquad i = 1, 2.\end{aligned} \tag{176}$$

Therefore,

$$\begin{aligned}E_i(t) &\le E_i(\tau_{\text{init}}) + B_\alpha(t - \tau_{\text{init}}) \\ &< \frac{\varepsilon^2}{8} + B_\alpha t \\ &\le \frac{\varepsilon^2}{8} + \frac{5}{C(\alpha)} \cdot \frac{C(\alpha)\varepsilon^2}{100} \\ &= \frac{7}{40}\varepsilon^2 < \frac{1}{2}\varepsilon^2, \qquad i = 1, 2.\end{aligned} \tag{177}$$

Since $E_i(t) = \frac{1}{2}\|\bar{w}_i(t) - \bar{v}\|^2$, it follows that $\|\bar{w}_i(t) - \bar{v}\| < \varepsilon$ for $i = 1, 2$. ■

Recall that $u_1 = \bar{v}, u_2 = \frac{v_1 - v_2}{\|v_1 - v_2\|_2}$.

Lemma 43 *Let $r_1 = \bar{w}_1 - \bar{v}$. When $\|r_1\|_2 \to 0$, we have*

$$\begin{aligned}(I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big) &= 2\sin\frac{\alpha}{2}\langle r_1, e_2\rangle e_2 \\ &\quad - 2\Big(\pi - \frac{\alpha}{2}\Big)\cos\frac{\alpha}{2} \cdot r_1 + o(\|r_1\|_2).\end{aligned}$$

Let $r_i = \bar{w}_i - \bar{v}, i = 1, 2$ and $\|r\|_2 = \max\{\|r_1\|_2, \|r_2\|_2\}$.

Corollary 44 *When $\|r_1\|_2 \to 0$, we have*

$$\begin{aligned}\Big\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big)\Big\rangle &\le -2\Big((\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} - \sin\frac{\alpha}{2}\Big)\|r_1\|_2^2 \\ &\quad + o(\|r_1\|_2^2).\end{aligned}$$

Proof [Proof of lemma 43 and corollary 44] Let $r_1 = \bar{w}_1 - \bar{v}$, then $r_1 = o(1)$. We have

$$1 = \|\bar{w}_1\|_2^2 = \|r_1 + \bar{v}\|_2^2 = 1 + 2\langle r_1, \bar{v}\rangle + \|r_1\|_2^2, \tag{178}$$

which implies $\langle r_1, \bar{v}\rangle = o(\|r_1\|_2)$. By Taylor expansion, we have

$$\begin{aligned}\arccos(\langle \bar{w}_1, v_1\rangle) &= \arccos(\langle r_1, v_1\rangle + \langle \bar{v}, v_1\rangle)\\ &= \arccos(\langle \bar{v}, v_1\rangle) - \frac{1}{\sqrt{1-\langle \bar{v}, v_1\rangle^2}}\langle r_1, v_1\rangle + o(\langle r_1, v_1\rangle)\\ &= \frac{\alpha}{2} - \frac{1}{\sin\frac{\alpha}{2}}\langle r_1, v_1\rangle + o(\|r_1\|_2).\end{aligned} \tag{179}$$

Similarly, we have

$$\arccos(\langle \bar{w}_1, v_2\rangle) = \frac{\alpha}{2} - \frac{1}{\sin\frac{\alpha}{2}}\langle r_1, v_2\rangle + o(\|r_1\|_2). \tag{180}$$

Then, we have

$$\begin{aligned}(\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2 &= \big(\pi - \arccos(\langle \bar{w}_1, v_1\rangle)\big)v_1 + \big(\pi - \arccos(\langle \bar{w}_1, v_2\rangle)\big)v_2\\ &= \left(\pi - \frac{\alpha}{2} + \frac{1}{\sin\frac{\alpha}{2}}\langle r_1, v_1\rangle + o(\|r_1\|_2)\right)v_1\\ &\quad + \left(\pi - \frac{\alpha}{2} + \frac{1}{\sin\frac{\alpha}{2}}\langle r_1, v_2\rangle + o(\|r_1\|_2)\right)v_2\\ &= \underbrace{2\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\cdot\bar{v} + \frac{1}{\sin\frac{\alpha}{2}}\Big(\langle r_1, v_1\rangle v_1 + \langle r_1, v_2\rangle v_2\Big)}_{\mathrm{I}}\\ &\quad + o(\|r_1\|_2)\bar{v},\\ I - \bar{w}_1\bar{w}_1^\top &= I - (r_1 + \bar{v})(r_1 + \bar{v})^\top\\ &= I - \bar{v}\bar{v}^\top - r_1\bar{v}^\top - \bar{v}r_1^\top + o(\|r_1\|_2).\end{aligned} \tag{181}$$

Recall that we have

$$\begin{aligned}e_1 = \bar{v}, \quad e_2 = \frac{v_1 - v_2}{\|v_1 - v_2\|_2} = \frac{v_1 - v_2}{2\sin\frac{\alpha}{2}},\\ v_1 = \cos\frac{\alpha}{2}e_1 + \sin\frac{\alpha}{2}e_2, \quad v_2 = \cos\frac{\alpha}{2}e_1 - \sin\frac{\alpha}{2}e_2.\end{aligned}$$

Then, we have

$$\begin{aligned}(I - \bar{v}\bar{v}^\top)\cdot\mathrm{I} &= \frac{1}{\sin\frac{\alpha}{2}}(I - \bar{v}\bar{v}^\top)\cdot\bigg(\langle r_1, v_1\rangle\big(\cos\frac{\alpha}{2}e_1 + \sin\frac{\alpha}{2}e_2\big)\\ &\qquad + \langle r_1, v_2\rangle\big(\cos\frac{\alpha}{2}e_1 - \sin\frac{\alpha}{2}e_2\big)\bigg)\\ &= \frac{1}{\sin\frac{\alpha}{2}}(I - \bar{v}\bar{v}^\top)\cdot\sin\frac{\alpha}{2}\big(\langle r_1, v_1\rangle - \langle r_1, v_2\rangle\big)e_2\end{aligned}$$

$$
\begin{aligned}
&= 2\sin\frac{\alpha}{2}\langle r_1, e_2\rangle e_2,\\
r_1\bar{v}^\top \cdot \mathrm{I} &= 2\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} r_1 + o(\|r_1\|_2),\\
\bar{v}r_1^\top \cdot \mathrm{I} &= 2\left(\pi - \frac{\alpha}{2}\right)\langle r_1, \bar{v}\rangle\bar{v} + o(\|r_1\|_2)\\
&= o(\|r_1\|_2),
\end{aligned}
$$

where the last equality arises from the fact that $\langle r_1, \bar{v}\rangle = -\frac{1}{2}\|r_1\|_2^2 = o(\|r_1\|_2)$. Thus, we have

$$
\begin{aligned}
(I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big) &= 2\sin\frac{\alpha}{2}\langle r_1, e_2\rangle e_2\\
&\quad - 2\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\cdot r_1 + o(\|r_1\|_2).
\end{aligned}
$$

Moreover, we have

$$
\begin{aligned}
&\Big\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big)\Big\rangle\\
&\qquad = \Big\langle r_1, (I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big)\Big\rangle\\
&= 2\sin\frac{\alpha}{2}\langle r_1, e_2\rangle^2 - 2\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\|r_1\|_2^2 + o(\|r_1\|_2^2)\\
&\le 2\left(\sin\frac{\alpha}{2} - (\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2}\right)\|r_1\|_2^2 + o(\|r_1\|_2^2).
\end{aligned}
$$

■

Lemma 45 *When $\|r\|_2 \to 0$, we have*

$$
\Big\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)w_2\Big\rangle = -\langle r_1, r_1 - r_2\rangle\|w_2\|_2 + o(\|r\|_2^2).
$$

Proof Let $\|r\|_2 = \max\{\|r_1\|_2, \|r_2\|_2\} \to 0$. By Equation (178), we have $\langle r_i, \bar{v}\rangle = -\|r_i\|_2^2/2 = o(\|r_i\|_2) = o(\|r\|_2), i = 1, 2$. Then, we can get

$$
\begin{aligned}
&(I - \bar{v}\bar{v}^\top)w_2 = \|w_2\|_2(I - \bar{v}\bar{v}^\top)(\bar{v} + r_2) = \|w_2\|_2 r_2 + o(\|r\|_2),\\
&r_1\bar{v}^\top \cdot w_2 = \|w_2\|_2 \cdot \langle \bar{v}, \bar{v} + r_2\rangle \cdot r_1 = \|w_2\|_2 r_1 + o(\|r\|_2),\\
&\bar{v}r_1^\top \cdot w_2 = \|w_2\|_2\langle r_1, r_2 + \bar{v}\rangle\bar{v} = o(\|r\|_2).
\end{aligned}
$$

Thus, by (181), we have

$$
(I - \bar{w}_1\bar{w}_1^\top)w_2 = \|w_2\|_2(r_2 - r_1) + o(\|r\|_2).
$$

Moreover, we have

$$\left\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)w_2 \right\rangle = \left\langle r_1, \|w_2\|_2(r_2 - r_1) + o(\|r\|_2) \right\rangle$$
$$= -\langle r_1, r_1 - r_2\rangle \|w_2\|_2 + o(\|r\|_2^2).$$

■

Corollary 46 *There exists $\varepsilon_0 > 0$ such that for any $\varepsilon \le \varepsilon_0$, if $\|\bar{w}_1 - \bar{v}\|_2 \le \varepsilon$ and $\|\bar{w}_2 - \bar{v}\|_2 \le \varepsilon$, we have*

$$\left\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)\Big((\pi - \theta_{11})v_1 + (\pi - \theta_{12})v_2\Big) \right\rangle \le -\frac{3}{2}\Big(\big(\pi - \frac{\alpha}{2}\big)\cos\frac{\alpha}{2} - \sin\frac{\alpha}{2}\Big)\|r_1\|_2^2, \tag{182}$$

$$\left\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)w_2 \right\rangle \ge -\langle r_1, r_1 - r_2\rangle \|w_2\|_2. \tag{183}$$

Proof The first inequality is a direct result of corollary 44. For the second inequality, since $\bar{w}_i = \bar{v} + r_i$ are unit vectors,

$$\langle r_1, \bar{w}_1\rangle = \frac{1}{2}\|r_1\|_2^2, \qquad \langle \bar{w}_1, \bar{w}_2\rangle = 1 - \frac{1}{2}\|r_1 - r_2\|_2^2.$$

Therefore,

$$\left\langle \bar{w}_1 - \bar{v}, (I - \bar{w}_1\bar{w}_1^\top)w_2 \right\rangle$$
$$= \left(-\langle r_1, r_1 - r_2\rangle + \frac{1}{4}\|r_1\|_2^2\|r_1 - r_2\|_2^2\right)\|w_2\|_2 \ge -\langle r_1, r_1 - r_2\rangle\|w_2\|_2.$$

■

Theorem 47 *Let $T_\alpha = \frac{C(\alpha)}{100}$, and let ε_0 be the constant in corollary 46. For any $0 < \varepsilon \le \varepsilon_0$ and $0 < T_1 \le C(\alpha)\min\{1, \varepsilon^2\}/200$, there exists $\sigma_0 > 0$ such that for any $\sigma \le \sigma_0$,*

$$\|\bar{w}_i(t) - \bar{v}\|_2 < \varepsilon, \quad i = 1, 2, \quad \forall T_1 \le t \le T_\alpha.$$

Proof By theorem 42, there exists $\sigma_1 = \sigma_1(\varepsilon, T_1) > 0$ such that, for every $\sigma \le \sigma_1$,

$$\|\bar{w}_i(T_1) - \bar{v}\|_2 < \varepsilon, \quad i = 1, 2.$$

Choose

$$\sigma_0 := \min\left\{\sigma_1, \frac{C(\alpha)}{200\sqrt{d}}\right\}.$$

Let $r_i(t) \coloneqq \bar{w}_i(t) - \bar{v}$ and

$$E_i(t) \coloneqq \frac{1}{2}\|r_i(t)\|_2^2, \quad i = 1, 2.$$

Define the first exit time

$$\tau \coloneqq \inf\left\{t \geq T_1 : \exists i \in \{1,2\} \text{ such that } \|r_i(t)\|_2 = \varepsilon\right\}.$$

If $\tau > T_\alpha$, the claim follows. It remains to rule out the case $\tau \leq T_\alpha$.

Suppose, for contradiction, that $\tau \leq T_\alpha$. Without loss of generality,

$$\|r_1(\tau)\|_2 = \varepsilon \geq \|r_2(\tau)\|_2.$$

By corollary 32, for $i = 1, 2$,

$$\|w_i(\tau)\|_2 \leq 2\sigma_0\sqrt{d} + \tau \leq \frac{C(\alpha)}{100} + \frac{C(\alpha)}{100} = \frac{C(\alpha)}{50}. \tag{184}$$

Since τ is the first exit time and $\varepsilon \leq \varepsilon_0$, the bounds in corollary 46 apply at $t = \tau$. Hence

$$\left\langle r_1(\tau), (I - \bar{w}_1(\tau)\bar{w}_1(\tau)^\top)\Big(\big(\pi - \theta_{11}(\tau)\big)v_1 + \big(\pi - \theta_{12}(\tau)\big)v_2\Big)\right\rangle \leq -\frac{3}{2}A_\alpha\|r_1(\tau)\|_2^2,$$

$$\begin{aligned}\left\langle r_1(\tau), (I - \bar{w}_1(\tau)\bar{w}_1(\tau)^\top)w_2(\tau)\right\rangle &\geq -\langle r_1(\tau), r_1(\tau) - r_2(\tau)\rangle\|w_2(\tau)\|_2 \\ &\geq -2\|r_1(\tau)\|_2^2\|w_2(\tau)\|_2,\end{aligned}$$

where

$$A_\alpha \coloneqq \left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} - \sin\frac{\alpha}{2}.$$

The last inequality uses $\|r_1(\tau)\|_2 \geq \|r_2(\tau)\|_2$. By (163), at time $t = \tau$ we have

$$\begin{aligned}\left.\frac{\mathrm{d}E_1}{\mathrm{d}t}\right|_{t=\tau} &\leq \frac{\|r_1(\tau)\|_2^2}{2\pi\|w_1(\tau)\|_2}\left(-\frac{3}{2}A_\alpha + 2\pi\|w_2(\tau)\|_2\right) \\ &\leq \frac{\|r_1(\tau)\|_2^2}{2\pi\|w_1(\tau)\|_2}\left(-\frac{3}{2}A_\alpha + \frac{2\pi C(\alpha)}{50}\right).\end{aligned} \tag{185}$$

The function

$$\left(\pi - \frac{x}{2}\right)\cos\frac{x}{2} - \sin\frac{x}{2}$$

is decreasing on $(0, \pi/2]$. Therefore $A_\alpha \geq A_{\pi/2}$, and since $0 < C(\alpha) < 1$,

$$-\frac{3}{2}A_\alpha + \frac{2\pi C(\alpha)}{50} \leq -\frac{3}{2}\left(\frac{3\pi}{4}\cos\frac{\pi}{4} - \sin\frac{\pi}{4}\right) + \frac{2\pi}{50} < 0. \tag{186}$$

Hence $\left.\frac{\mathrm{d}E_1}{\mathrm{d}t}\right|_{t=\tau} < 0$. Since $E_1(\tau) = \varepsilon^2/2$, there exists $h > 0$ such that

$$E_1(\tau - h) > E_1(\tau) = \frac{\varepsilon^2}{2}.$$

Thus $\|r_1(\tau - h)\|_2 > \varepsilon$, contradicting the definition of τ as the first exit time. Therefore $\tau > T_\alpha$, which proves the theorem. ■

Lemma 48 (Finite-time radial lower bound under angular confinement) *Let $T_\alpha = C(\alpha)/100$ and*

$$m_\alpha := \frac{C(\alpha)^2}{1000}.$$

For every finite $T_2 > T_\alpha$, there exists $\varepsilon_{\rm rad} = \varepsilon_{\rm rad}(\alpha, T_2) > 0$ such that the following holds. Suppose that, on an interval $[T_\alpha, T] \subset [T_\alpha, T_2]$,

$$\|\bar{w}_i(t) - \bar{v}\| \le \varepsilon_{\rm rad}, \qquad i = 1, 2,$$

and that

$$m_\alpha \le \|w_i(T_\alpha)\| \le \frac{C(\alpha)}{50}, \qquad i = 1, 2.$$

Then

$$\|w_i(t)\| \ge \frac{m_\alpha}{2}, \qquad T_\alpha \le t \le T, \quad i = 1, 2.$$

Proof Let

$$\Delta(t) := \big|\|w_1(t)\| - \|w_2(t)\|\big|.$$

Also set

$$r_s := 2l\left(\frac{\alpha}{2}\right) = \frac{\left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}}{\pi}.$$

Since

$$\frac{d}{d\alpha} r_s = -\frac{(\pi - \alpha/2)\sin(\alpha/2)}{2\pi} < 0,$$

we have

$$r_s \ge \frac{\frac{3\pi}{4}\cos\frac{\pi}{4} + \sin\frac{\pi}{4}}{\pi} > \frac{1}{2}.$$

Together with $0 < C(\alpha) < 1$, this implies

$$\eta_\alpha := \frac{1}{2}\left(r_s - \frac{C(\alpha)}{50} - m_\alpha\right) > 0.$$

By continuity of the map

$$\bar{w} \mapsto l(\angle(\bar{w}, v_1)) + l(\angle(\bar{w}, v_2))$$

at $\bar{v}$, after decreasing $\varepsilon_{\rm rad}$ if necessary, we have

$$l(\theta_{i1}(t)) + l(\theta_{i2}(t)) \ge \frac{r_s}{2}, \qquad i = 1, 2, \tag{187}$$

whenever $\|\bar{w}_i(t) - \bar{v}\| \le \varepsilon_{\rm rad}$. Moreover, in the same neighborhood this map is Lipschitz, so for some constant $L_\alpha > 0$,

$$|l(\theta_{11}) + l(\theta_{12}) - l(\theta_{21}) - l(\theta_{22})| \le L_\alpha\big(\|\bar{w}_1 - \bar{v}\| + \|\bar{w}_2 - \bar{v}\|\big).$$

We decrease $\varepsilon_{\rm rad}$ once more so that

$$2L_\alpha T_2 \varepsilon_{\rm rad} \le \eta_\alpha.$$

Subtracting the two norm equations (9)–(10) gives

$$\frac{d}{dt}\big(\|w_1\| - \|w_2\|\big) = l(\theta_{11}) + l(\theta_{12}) - l(\theta_{21}) - l(\theta_{22}) + \left(l(\gamma) - \frac{1}{2}\right)\big(\|w_1\| - \|w_2\|\big).$$

Since $l(\gamma) \leq 1/2$, the upper-Dini derivative of Δ satisfies

$$D^+\Delta(t) \leq L_\alpha\big(\|\bar{w}_1(t) - \bar{v}\| + \|\bar{w}_2(t) - \bar{v}\|\big) \leq 2L_\alpha\varepsilon_{\text{rad}}.$$

Thus, for every $T_\alpha \leq t \leq T$,

$$\Delta(t) \leq \Delta(T_\alpha) + 2L_\alpha T_2\varepsilon_{\text{rad}} \leq \frac{C(\alpha)}{50} + \eta_\alpha.$$

We now prove the lower bound by a barrier argument. Suppose that the claim fails, and let $\tau \in [T_\alpha, T]$ be the first time at which $\|w_i(\tau)\| = m_\alpha/2$ for some i. Then

$$\|w_i(\tau)\| = \frac{m_\alpha}{2}, \qquad \|w_{3-i}(\tau)\| \leq \frac{m_\alpha}{2} + \frac{C(\alpha)}{50} + \eta_\alpha$$

by (C.1). Using the norm equation for $\|w_i\|$, (187), and $l(\gamma) \leq 1/2$, we obtain at $t = \tau$

$$\begin{aligned}
\frac{d}{dt}\|w_i(t)\|\bigg|_{t=\tau} &= l(\theta_{i1}) + l(\theta_{i2}) - \frac{1}{2}\|w_i(\tau)\| - l(\gamma)\|w_{3-i}(\tau)\| \\
&\geq \frac{r_s}{2} - \frac{m_\alpha}{4} - \frac{1}{2}\left(\frac{m_\alpha}{2} + \frac{C(\alpha)}{50} + \eta_\alpha\right) \\
&= \frac{r_s - m_\alpha}{4} - \frac{C(\alpha)}{200} > 0.
\end{aligned}$$

Here the last inequality follows from $r_s - m_\alpha - C(\alpha)/50 > 0$, which is exactly the positivity of η_α. This contradicts the definition of τ as the first downward hitting time of $m_\alpha/2$. Hence no such τ exists, which proves the claim. ■

Theorem 49 *For any $\varepsilon > 0$ and any T_1, T_2 satisfying $0 < T_1 \leq C(\alpha)\min\{1, \varepsilon^2\}/200 < T_2$, there exists $\sigma_0 > 0$ such that for any $\sigma \leq \sigma_0$,*

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon, \quad i = 1, 2, \quad \forall T_1 \leq t \leq T_2.$$

Proof Let $T_\alpha := C(\alpha)/100$ and $K_\alpha := 8000/C(\alpha)^2$. We also write

$$A_\alpha := \left(\pi - \frac{\alpha}{2}\right)\cos\frac{\alpha}{2} - \sin\frac{\alpha}{2}.$$

Let ε_0 be the constant in corollary 46. If $T_2 > T_\alpha$, let $\varepsilon_{\text{rad}} = \varepsilon_{\text{rad}}(\alpha, T_2)$ be the constant in lemma 48; otherwise set $\varepsilon_{\text{rad}} = 1$. Put

$$\varepsilon_* := \min\{\varepsilon, \varepsilon_0, \varepsilon_{\text{rad}}\}.$$

Choose

$$0 < \varepsilon_1 < \varepsilon_* \exp\left(-\frac{K_\alpha T_2}{2}\right),$$

and define

$$T_0 := \min\left\{T_1, \frac{C(\alpha)\min\{1, \varepsilon_1^2\}}{200}\right\}.$$

By theorem 47, there exists $\sigma_1 > 0$ such that, for every $\sigma \le \sigma_1$,

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_1, \quad i = 1, 2, \quad \forall T_0 \le t \le T_\alpha. \tag{188}$$

If $T_2 \le T_\alpha$, then (188) proves the claim because $T_0 \le T_1$ and $\varepsilon_1 < \varepsilon$. Hence we may assume $T_2 > T_\alpha$.

Choose $\sigma_0 \le \sigma_1$ small enough so that $\sigma_0 \le C(\alpha)/(200\sqrt{d})$ and $\|w_i(0)\| \le 2$ for $i = 1, 2$. Then corollary 33 gives

$$\|w_i(t)\| \le 2, \quad t \ge 0, \quad i = 1, 2.$$

Let

$$r_i(t) := \bar{w}_i(t) - \bar{v}, \qquad E_i(t) := \frac{1}{2}\|r_i(t)\|^2, \quad i = 1, 2.$$

Let

$$\mathcal{T}_* := \left\{t \ge T_\alpha : \max_{i=1,2}\|\bar{w}_i(t) - \bar{v}\| = \varepsilon_*\right\}.$$

If $\mathcal{T}_* = \varnothing$, then the conclusion is immediate. Otherwise, let $\tau := \inf \mathcal{T}_*$. We prove that $\tau > T_2$.

Suppose, for contradiction, that $\tau \le T_2$. For $t \in [T_\alpha, \tau]$, the bounds in corollary 46 apply. Also, corollary 36 and lemma 48 give, for $T_\alpha \le t \le \tau$,

$$\|w_i(t)\| \ge \frac{C(\alpha)^2}{2000}, \quad i = 1, 2. \tag{189}$$

If $E_1(t) \ge E_2(t)$, then the same estimate as in (185), together with (189) and $\|w_2(t)\| \le 2$, gives

$$\begin{aligned}
\frac{\mathrm{d}E_1}{\mathrm{d}t} &\le \frac{1}{2\pi\|w_1(t)\|}\bigg(-\frac{3}{2}A_\alpha\|r_1(t)\|^2 + 2\pi\|r_1(t)\|^2\|w_2(t)\|\bigg) \\
&\le \frac{2000}{2\pi C(\alpha)^2}\cdot 4\pi\|r_1(t)\|^2 \\
&\le K_\alpha E_1(t).
\end{aligned} \tag{190}$$

Similarly, if $E_2(t) \ge E_1(t)$, then $\frac{\mathrm{d}E_2}{\mathrm{d}t} \le K_\alpha E_2(t)$. Thus, for

$$M(t) := \max\{E_1(t), E_2(t)\},$$

we can compare M as follows. On any subinterval where $M(t) = E_1(t)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-K_\alpha t}M(t)\right) = e^{-K_\alpha t}\left(\frac{\mathrm{d}E_1}{\mathrm{d}t} - K_\alpha E_1(t)\right) \le 0.$$

On any subinterval where $M(t) = E_2(t)$, the same argument gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-K_\alpha t}M(t)\right) = e^{-K_\alpha t}\left(\frac{\mathrm{d}E_2}{\mathrm{d}t} - K_\alpha E_2(t)\right) \le 0.$$

At the switching times, where the maximizer changes between E_1 and E_2, the function M is continuous. Therefore the above inequalities can be concatenated over $[T_\alpha, t]$, and for every $T_\alpha \le t \le \tau$,

$$\begin{aligned} M(t) &\le M(T_\alpha)\exp\left(K_\alpha(t - T_\alpha)\right) \\ &< \frac{\varepsilon_1^2}{2}\exp(K_\alpha T_2) \\ &< \frac{\varepsilon_*^2}{2}. \end{aligned} \tag{191}$$

This contradicts the definition of τ, since $M(\tau) = \varepsilon_*^2/2$. Hence $\tau > T_2$, and therefore

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_* \le \varepsilon, \quad i = 1, 2, \quad \forall T_1 \le t \le T_2.$$

■

Corollary 50 *For any* $\varepsilon > 0$ *and* $0 < T_1 \le \frac{\min\{1,\varepsilon^2\}}{200} < T_2$, *there exists* $\sigma_0 > 0$ *such that for any* $\sigma \le \sigma_0$,

$$|\theta_{ij}(t) - \frac{\alpha}{2}| < \varepsilon, |\sin\theta_{ij}(t) - \sin\frac{\alpha}{2}| < \varepsilon, \quad i = 1, 2, j = 1, 2, \forall T_1 \le t \le T_2.$$

Proof Fix $\varepsilon > 0$ and let $c_j := \langle \bar{v}, v_j \rangle = \cos\frac{\alpha}{2}$ for $j = 1, 2$. Since $\alpha \in (0, \frac{\pi}{2}]$, we have $c_j \in (-1, 1)$, and hence $x \mapsto \arccos x$ is Lipschitz in a neighborhood of c_j. More precisely, there exist $\rho > 0$ and $L > 0$ such that

$$|\arccos x - \arccos y| \le L|x - y|$$

whenever $|x - c_j| \le \rho$ and $|y - c_j| \le \rho$. Choose

$$0 < \eta \le \min\left\{1, \varepsilon, \frac{\rho}{2}, \frac{\varepsilon}{2L}\right\}.$$

Let

$$T_1' := \min\left\{T_1, \frac{C(\alpha)\min\{1, \eta^2\}}{200}\right\}.$$

Since $\eta \le \varepsilon$ and $0 < C(\alpha) < 1$, we have

$$\frac{C(\alpha)\min\{1, \eta^2\}}{200} \le \frac{\min\{1, \varepsilon^2\}}{200} < T_2.$$

Therefore, by theorem 49, there exists $\sigma_0 > 0$ such that for any $\sigma \le \sigma_0$,

$$\|\bar{w}_i(t) - \bar{v}\| < \eta, \quad i = 1, 2, \quad \forall T_1' \le t \le T_2.$$

Thus, for all $T_1 \le t \le T_2$ and $i, j = 1, 2$,

$$\left|\langle \bar{w}_i(t), v_j\rangle - \langle \bar{v}, v_j\rangle\right| \le \|\bar{w}_i(t) - \bar{v}\| < \eta \le \rho.$$

Since $\theta_{ij}(t) = \arccos(\langle \bar{w}_i(t), v_j\rangle)$ and $\arccos(\langle \bar{v}, v_j\rangle) = \frac{\alpha}{2}$, the Lipschitz bound gives

$$\begin{aligned}\left|\theta_{ij}(t) - \frac{\alpha}{2}\right| &= \left|\arccos(\langle \bar{w}_i(t), v_j\rangle) - \arccos(\langle \bar{v}, v_j\rangle)\right| \\ &\le L\left|\langle \bar{w}_i(t) - \bar{v}, v_j\rangle\right| \\ &\le L\|\bar{w}_i(t) - \bar{v}\| < \varepsilon.\end{aligned}$$

Finally,

$$\begin{aligned}\left|\sin\theta_{ij}(t) - \sin\frac{\alpha}{2}\right| &= 2\left|\sin\frac{\theta_{ij}(t) - \frac{\alpha}{2}}{2}\cos\frac{\theta_{ij}(t) + \frac{\alpha}{2}}{2}\right| \\ &\le \left|\theta_{ij}(t) - \frac{\alpha}{2}\right| < \varepsilon.\end{aligned}$$

■

Lemma 51 *For any $\varepsilon > 0$ and any finite time $T > 0$, there exists $\sigma_0 > 0$ such that for any $0 < \sigma \le \sigma_0$,*

$$\left|\|w_1(t)\| - \|w_2(t)\|\right| \le \varepsilon, \quad \forall 0 \le t \le T. \tag{192}$$

Proof Choose $0 < \varepsilon_1 \le 1$ small enough so that

$$9\varepsilon_1 e^{T/2} < \varepsilon \quad \text{and} \quad t_0 := \frac{C(\alpha)\min\{1, \varepsilon_1^2\}}{200} < T.$$

By theorem 49, there exists $\sigma_1 > 0$ such that for all $\sigma \le \sigma_1$,

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_1, \quad i = 1, 2, \quad \forall t_0 \le t \le T. \tag{193}$$

Choose $\sigma_0 \le \min\{\sigma_1, \varepsilon_1/(100\sqrt{d})\}$. Let

$$\Delta(t) := \left|\|w_1(t)\| - \|w_2(t)\|\right|.$$

For $0 \le t \le t_0$, corollary 32 gives

$$\begin{aligned}\Delta(t) &\le \|w_1(t)\| + \|w_2(t)\| \\ &\le 4\sigma\sqrt{d} + 2t_0\end{aligned}$$

$$\leq \frac{\varepsilon_1}{25} + \frac{C(\alpha)\varepsilon_1^2}{100} \leq \varepsilon_1 < \varepsilon. \tag{194}$$

It remains to consider $t \in [t_0, T]$. By (193), for $i = 1, 2$,

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_1.$$

Hence, by the triangle inequality,

$$\|\bar{w}_1(t) - \bar{w}_2(t)\| \leq \|\bar{w}_1(t) - \bar{v}\| + \|\bar{w}_2(t) - \bar{v}\| < 2\varepsilon_1.$$

Since $\bar{w}_1(t)$ and $\bar{w}_2(t)$ are unit vectors,

$$\|\bar{w}_1(t) - \bar{w}_2(t)\| = 2\sin\frac{\gamma(t)}{2}.$$

For $0 \leq x \leq \pi/2$, we use $\sin x \geq x/2$. Therefore,

$$\gamma(t) \leq 2\|\bar{w}_1(t) - \bar{w}_2(t)\| < 4\varepsilon_1, \quad t_0 \leq t \leq T.$$

We also record a Lipschitz estimate for the teacher forcing. Since

$$l'(\theta) = -\frac{(\pi - \theta)\sin\theta}{2\pi},$$

we have $|l'(\theta)| \leq 1/2$ for all $\theta \in [0, \pi]$. Moreover, the angle map is 1-Lipschitz:

$$|\theta_{1j} - \theta_{2j}| = \big|\angle(\bar{w}_1, v_j) - \angle(\bar{w}_2, v_j)\big| \leq \angle(\bar{w}_1, \bar{w}_2) = \gamma.$$

Hence

$$\left|\sum_{j=1}^{2} \big(l(\theta_{1j}) - l(\theta_{2j})\big)\right| \leq \frac{1}{2}\sum_{j=1}^{2} |\theta_{1j} - \theta_{2j}| \leq \gamma.$$

Using (9) and (10), we obtain, for $t_0 \leq t \leq T$,

$$\begin{aligned} D^+\Delta(t) &\leq \left|\sum_{i=1}^{2} \big(l(\theta_{1i}) - l(\theta_{2i})\big) + \left(l(\gamma) - \frac{1}{2}\right)\big(\|w_1\| - \|w_2\|\big)\right| \\ &\leq \gamma + \left|l(\gamma) - \frac{1}{2}\right|\Delta(t) \\ &\leq 4\varepsilon_1 + \frac{1}{2}\Delta(t). \end{aligned} \tag{195}$$

Gronwall's inequality for upper-Dini derivatives then yields, for $t_0 \leq t \leq T$,

$$\begin{aligned} \Delta(t) &\leq e^{(t-t_0)/2}\Delta(t_0) + 8\varepsilon_1\big(e^{(t-t_0)/2} - 1\big) \\ &\leq e^{T/2}\big(\Delta(t_0) + 8\varepsilon_1\big) \\ &\leq 9\varepsilon_1 e^{T/2} < \varepsilon. \end{aligned} \tag{196}$$

■

Let $r_s = \frac{(\pi - \frac{\alpha}{2})\cos\frac{\alpha}{2} + \sin\frac{\alpha}{2}}{\pi}$, $w(t) = (w_1(t), w_2(t)) \in \mathbb{R}^{2d}$, $w_{\text{saddle}} = (w_s, w_s) \in \mathbb{R}^{2d}$, where $w_s = r_s\bar{v}$. We have $\frac{\mathrm{d}w}{\mathrm{d}t}\big|_{w=w_s} = 0$, which implies $\nabla L(w_{\text{saddle}}) = 0$. Let $\{u_1 = \bar{v}, u_2 = \frac{v_1 - v_2}{\|v_1 - v_2\|_2}, u_3, u_4, \ldots, u_d\}$ be an orthogonal basis of $\mathbb{R}^d$. Under this basis, we assume $w_i = \sum_{j=1}^{d} z_{i,j}u_j$.

Lemma 52 *For any $\varepsilon > 0$ and any finite time $T > 0$, there exists $\sigma_0 > 0$ such that for any $0 < \sigma \leq \sigma_0$,*

$$|z_{1,1}(t) - z_{2,1}(t)| < \varepsilon, \quad \forall 0 \leq t \leq T.$$

Proof Choose $0 < \varepsilon_1 \leq 1$ small enough so that

$$\varepsilon_1 + \varepsilon_1^2 < \varepsilon \quad \text{and} \quad t_0 := \frac{C(\alpha)\min\{1, \varepsilon_1^2\}}{200} < T.$$

Choose $\sigma_0 \leq \varepsilon_1/(100\sqrt{d})$. For $0 \leq t \leq t_0$, corollary 32 gives

$$\begin{aligned} |z_{1,1}(t) - z_{2,1}(t)| &\leq |z_{1,1}(t)| + |z_{2,1}(t)| \\ &\leq \|w_1(t)\| + \|w_2(t)\| \\ &\leq 4\sigma\sqrt{d} + 2t_0 \\ &\leq \frac{\varepsilon_1}{25} + \frac{C(\alpha)\varepsilon_1^2}{100} \leq \varepsilon_1 + \varepsilon_1^2 < \varepsilon. \end{aligned} \tag{197}$$

We now consider $t_0 \leq t \leq T$. By theorem 49 and lemma 51, after decreasing σ_0 if necessary, for every $0 < \sigma \leq \sigma_0$ we have

$$\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_1, \quad i = 1, 2, \qquad \big|\|w_1(t)\| - \|w_2(t)\|\big| \leq \varepsilon_1, \quad \forall t_0 \leq t \leq T.$$

Since $u_1 = \bar{v}$ and $z_{i,1}(t) = \langle w_i(t), \bar{v}\rangle$, we have

$$\frac{z_{i,1}(t)}{\|w_i(t)\|} = \langle \bar{w}_i(t), \bar{v}\rangle, \quad i = 1, 2.$$

The directional bound gives

$$2 - 2\langle \bar{w}_i(t), \bar{v}\rangle = \|\bar{w}_i(t) - \bar{v}\|^2 < \varepsilon_1^2,$$

and hence

$$1 - \frac{\varepsilon_1^2}{2} < \frac{z_{i,1}(t)}{\|w_i(t)\|} \leq 1, \quad i = 1, 2.$$

Therefore,

$$\left| \frac{z_{1,1}(t)}{\|w_1(t)\|} - \frac{z_{2,1}(t)}{\|w_2(t)\|} \right| \leq \frac{\varepsilon_1^2}{2}.$$

Since $\sigma_0 \leq \varepsilon_1/(100\sqrt{d})$ and $\varepsilon_1 \leq 1$, the initialization bound gives $\|w_i(0)\| \leq 2\sigma\sqrt{d} \leq \varepsilon_1/50 < 2$. Hence corollary 33 implies $\|w_i(t)\| \leq 2$ for all $t \geq 0$. Also,

$$\left| \frac{z_{2,1}(t)}{\|w_2(t)\|} \right| = |\langle \bar{w}_2(t), \bar{v}\rangle| \leq 1.$$

We obtain

$$\begin{aligned}
|z_{1,1}(t) - z_{2,1}(t)| &= \left| \|w_1(t)\| \left(\frac{z_{1,1}(t)}{\|w_1(t)\|} - \frac{z_{2,1}(t)}{\|w_2(t)\|} \right) + (\|w_1(t)\| - \|w_2(t)\|) \frac{z_{2,1}(t)}{\|w_2(t)\|} \right| \\
&\leq \|w_1(t)\| \left| \frac{z_{1,1}(t)}{\|w_1(t)\|} - \frac{z_{2,1}(t)}{\|w_2(t)\|} \right| + |\|w_1(t)\| - \|w_2(t)\|| \left| \frac{z_{2,1}(t)}{\|w_2(t)\|} \right| \\
&\leq 2 \cdot \frac{\varepsilon_1^2}{2} + \varepsilon_1 \\
&= \varepsilon_1^2 + \varepsilon_1 < \varepsilon.
\end{aligned}$$

■

Theorem 53 *For every $\rho > 0$, there exist $t_0 = t_0(\alpha, \rho) > 0$ and $\sigma_0 = \sigma_0(\alpha, \rho, d) > 0$ such that, for every $0 < \sigma \leq \sigma_0$, if*

$$t_3 \coloneqq \inf\{t > 0 : \|w(t) - w_{\text{saddle}}\| \leq \rho\},$$

then $t_3 \leq t_0$.

Proof Recall that

$$r_s = \frac{\left(\pi - \frac{\alpha}{2}\right) \cos \frac{\alpha}{2} + \sin \frac{\alpha}{2}}{\pi}.$$

Choose $0 < \varepsilon_1 < 1$ small enough so that

$$9\varepsilon_1 + \frac{C(\alpha)\varepsilon_1^2}{200} < \frac{r_s}{2}, \qquad 2\left((10\varepsilon_1)^2 + \frac{4\varepsilon_1^2}{1 - \varepsilon_1^2} \right) < \rho^2. \tag{198}$$

Set

$$t_* \coloneqq \frac{C(\alpha)\varepsilon_1^2}{200}.$$

Choose T such that

$$T > \max\left\{ t_*, \frac{\varepsilon_1^2}{200}, \frac{r_s + 1}{\varepsilon_1} \right\}.$$

By theorem 49, corollary 50, lemma 51, and lemma 52, after choosing $\sigma_0 > 0$ sufficiently small, for every $0 < \sigma \leq \sigma_0$ we have, for all $t_* \leq t \leq T$,

$$\begin{aligned}
&\|\bar{w}_i(t) - \bar{v}\| < \varepsilon_1, \quad \left|\theta_{ij}(t) - \frac{\alpha}{2}\right| < \varepsilon_1, \quad \left|\sin \theta_{ij}(t) - \sin \frac{\alpha}{2}\right| < \varepsilon_1, \quad i, j = 1, 2, \\
&|z_{1,1}(t) - z_{2,1}(t)| < \varepsilon_1, \quad \left|\|w_1(t)\| - \|w_2(t)\|\right| < \varepsilon_1.
\end{aligned} \tag{199}$$

We also assume $\sigma_0 \leq \min\{1/(1000\sqrt{d}), r_s/(8\sqrt{d})\}$. Then the initialization bound and corollary 33 imply

$$\|w_i(t)\| \leq 2, \quad i = 1, 2, \quad \forall t \geq 0.$$

From corollary 32,

$$|z_{1,1}(t_*)| \leq \|w_1(t_*)\| \leq 2\sigma\sqrt{d} + t_* \leq \frac{r_s}{4} + t_* < r_s - 9\varepsilon_1,$$

where the last inequality follows from (198). Define

$$\mathcal{T} := \{t \in [t_*, T] : |z_{1,1}(t) - r_s| \leq 9\varepsilon_1\}.$$

We first show that $\mathcal{T}$ is nonempty. Suppose otherwise. Since $z_{1,1}(t_*) < r_s - 9\varepsilon_1$, continuity implies

$$r_s - z_{1,1}(t) > 9\varepsilon_1, \quad \forall t_* \leq t \leq T.$$

On this interval, (199) gives

$$\langle \bar{w}_i(t), \bar{v} \rangle > 1 - \frac{\varepsilon_1^2}{2}, \quad i = 1, 2.$$

Hence $z_{2,1}(t) \geq 0$, and

$$\begin{aligned}\left| z_{2,1}(t) - \frac{\|w_2(t)\|}{\|w_1(t)\|} z_{1,1}(t) \right| &= \|w_2(t)\| \left| \frac{z_{2,1}(t)}{\|w_2(t)\|} - \frac{z_{1,1}(t)}{\|w_1(t)\|} \right| \\ &\leq 2 \cdot \frac{\varepsilon_1^2}{2} = \varepsilon_1^2.\end{aligned}$$

The teacher part in the $z_{1,1}$-equation satisfies

$$\begin{aligned}&\frac{1}{2\pi}\left((2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} + (\sin \theta_{11} + \sin \theta_{12}) \langle \bar{w}_1, \bar{v} \rangle \right) \\ &\qquad \geq r_s - 2\varepsilon_1 - \varepsilon_1^2.\end{aligned}$$

Indeed, this follows from

$$\left| \theta_{1j} - \frac{\alpha}{2} \right| < \varepsilon_1, \qquad \left| \sin \theta_{1j} - \sin \frac{\alpha}{2} \right| < \varepsilon_1, \qquad \langle \bar{w}_1, \bar{v} \rangle > 1 - \frac{\varepsilon_1^2}{2},$$

and the identity

$$r_s = \frac{1}{2\pi} \left((2\pi - \alpha) \cos \frac{\alpha}{2} + 2 \sin \frac{\alpha}{2} \right).$$

Also, $\gamma - \sin\gamma \geq 0$ and $z_{2,1}(t) \geq 0$, so the first coupling term below is nonnegative. The last coupling term is bounded below by $-\varepsilon_1^2$, because $\sin\gamma/(2\pi) \leq 1$ and the preceding display gives the needed absolute bound. Using Equation (5), we then have

$$\begin{aligned}\frac{\mathrm{d}z_{1,1}(t)}{\mathrm{d}t} &= \frac{1}{2\pi}\left((2\pi - \theta_{11} - \theta_{12}) \cos \frac{\alpha}{2} + (\sin \theta_{11} + \sin \theta_{12}) \langle \bar{w}_1, \bar{v} \rangle \right) \\ &\quad - z_{1,1}(t) - \frac{1}{2}\big(z_{2,1}(t) - z_{1,1}(t) \big) + \frac{\gamma - \sin\gamma}{2\pi} z_{2,1}(t) \\ &\quad + \frac{\sin\gamma}{2\pi} \left(z_{2,1}(t) - \frac{\|w_2(t)\|}{\|w_1(t)\|} z_{1,1}(t) \right) \\ &\geq r_s - 2\varepsilon_1 - \varepsilon_1^2 - z_{1,1}(t) - \frac{\varepsilon_1}{2} - \varepsilon_1^2 \\ &= \big(r_s - z_{1,1}(t) \big) - \frac{5}{2}\varepsilon_1 - 2\varepsilon_1^2 \\ &\geq \varepsilon_1.\end{aligned}$$

Therefore,

$$\begin{aligned} z_{1,1}(T) &\geq z_{1,1}(t_*) + \varepsilon_1(T - t_*) \\ &\geq \varepsilon_1 T - \varepsilon_1 t_* - |z_{1,1}(t_*)| \\ &\geq r_s + 1 - \frac{1}{200} - \left(2\sigma\sqrt{d} + t_*\right) \\ &\geq r_s + 1 - \frac{1}{200} - \frac{1}{500} - \frac{1}{200} \\ &> r_s, \end{aligned}$$

contradicting the assumption that $\mathcal{T}$ is empty. Thus $\mathcal{T} \neq \varnothing$. Since $z_{1,1}$ is continuous, $\mathcal{T}$ is a closed subset of the compact interval $[t_*, T]$, and hence it has a minimum. Let $\tau \coloneqq \min \mathcal{T}$. Then $\tau \leq T$ and

$$|z_{1,1}(\tau) - r_s| \leq 9\varepsilon_1.$$

Together with (199), this gives

$$\begin{aligned} |z_{2,1}(\tau) - r_s| &\leq |z_{2,1}(\tau) - z_{1,1}(\tau)| + |z_{1,1}(\tau) - r_s| \\ &\leq 10\varepsilon_1. \end{aligned}$$

Hence $|z_{i,1}(\tau) - r_s| \leq 10\varepsilon_1$ for $i = 1, 2$.

By (199),

$$\langle \bar{w}_i(\tau), \bar{v} \rangle = \frac{z_{i,1}(\tau)}{\sqrt{\sum_{j=1}^{d} z_{i,j}(\tau)^2}} > 1 - \frac{\varepsilon_1^2}{2}, \quad i = 1, 2.$$

Therefore,

$$\frac{\sum_{j=1}^{d} z_{i,j}(\tau)^2}{z_{i,1}(\tau)^2} < \frac{1}{\left(1 - \frac{\varepsilon_1^2}{2}\right)^2} < \frac{1}{1 - \varepsilon_1^2},$$

and hence

$$\sum_{j=2}^{d} z_{i,j}(\tau)^2 < \frac{\varepsilon_1^2}{1 - \varepsilon_1^2} z_{i,1}(\tau)^2, \quad i = 1, 2.$$

Since $z_{i,1}(\tau) \leq \|w_i(\tau)\| \leq 2$, we obtain

$$\begin{aligned} \|w(\tau) - w_{\text{saddle}}\|^2 &= \sum_{i=1}^{2} (z_{i,1}(\tau) - r_s)^2 + \sum_{i=1}^{2} \sum_{j=2}^{d} z_{i,j}(\tau)^2 \\ &\leq \sum_{i=1}^{2} \left((z_{i,1}(\tau) - r_s)^2 + \frac{\varepsilon_1^2}{1 - \varepsilon_1^2} z_{i,1}(\tau)^2 \right) \\ &\leq 2 \left((10\varepsilon_1)^2 + \frac{4\varepsilon_1^2}{1 - \varepsilon_1^2} \right) \\ &< \rho^2. \end{aligned}$$

Thus $\|w(\tau) - w_{\text{saddle}}\| < \rho$, so $t_3 \leq \tau \leq T$. Taking $t_0 = T$ proves the theorem. ■

C.2. Phase 2: saddle avoidance

Theorem 54 (Generic saddle avoidance) *For Lebesgue-almost every initialization, the gradient-flow trajectory (3) is defined for all finite time, does not reach a zero student weight, and does not converge to a saddle point. Consequently, by theorem 3, it converges to a global minimum.*

C.2.1. PROOF OF THEOREM 54

Proof We verify the hypotheses of the standard stable-manifold argument for gradient flows. It suffices to consider the open set

$$\Omega \coloneqq \{(w_1, w_2) \in (\mathbb{R}^d)^2 : w_1 \neq 0,\ w_2 \neq 0\}.$$

The complement of Ω has Lebesgue measure zero. On Ω, the closed-form ReLU kernel

$$\mathbb{E}[\phi(\langle u, x\rangle)\phi(\langle v, x\rangle)] = \frac{\|u\|_2\|v\|_2}{2\pi}\big(\sin\theta_{u,v} + (\pi - \theta_{u,v})\cos\theta_{u,v}\big),$$

shows that L is C^2 in a neighborhood of every nonzero stationary point. The apparent singularities at collinear pairs are removable at the C^2 level: the bracket equals $\pi - \frac{\pi}{2}\theta^2 + O(\theta^3)$ as $\theta \to 0$ and equals $O((\pi-\theta)^3)$ as $\theta \to \pi$. Thus the gradient-flow vector field is C^1 near each saddle listed in theorem 3.

We also need to exclude finite-time exits from Ω. We claim that the set of initializations in Ω whose maximal classical trajectory reaches $\partial\Omega$ in finite time has Lebesgue measure zero. It is enough to treat the face $\{w_1 = 0, w_2 \neq 0\}$, the other face being identical. Write $w_1 = ru$ with $r > 0$, $u \in \mathbb{S}^{d-1}$, and write $w_2 = Ra$ with $R > 0$ and $a \in \mathbb{S}^{d-1}$. With $l(\theta) = ((\pi - \theta)\cos\theta + \sin\theta)/(2\pi)$, define the residual correlation

$$\Psi_{w_2}(u) = l(\angle(u, v_1)) + l(\angle(u, v_2)) - R\, l(\angle(u, a)).$$

The radial and angular equations near $r = 0$ are

$$\dot r = \Psi_{w_2}(u) - \frac{1}{2}r, \qquad \dot u = \frac{1}{r}\nabla_{\mathbb{S}^{d-1}}\Psi_{w_2}(u),$$

where the second identity is the spherical-gradient form of (7). After the time change $d\tau = dt/r$, the equations extend to the blown-up boundary $r = 0$ as the definable C^1 system

$$r' = r\Big(\Psi_{w_2}(u) - \frac{1}{2}r\Big), \qquad u' = \nabla_{\mathbb{S}^{d-1}}\Psi_{w_2}(u), \qquad w_2' = r\dot w_2.$$

If a trajectory reaches $r = 0$ in finite original time, then in the rescaled time it converges to the boundary critical set

$$\mathcal{C}_- = \{r = 0 : \nabla_{\mathbb{S}^{d-1}}\Psi_{w_2}(u) = 0, \Psi_{w_2}(u) < 0\}.$$

The case $\Psi_{w_2}(u) = 0$ gives only $r' = -r^2/2 + o(r^2)$ and therefore $\int r(\tau)\, d\tau = \infty$, so it cannot give a finite-time collision. At each point of $\mathcal{C}_-$, the angular critical point is not a local maximum of Ψ_{w_2}. Indeed, all relevant directions lie in $\operatorname{span}\{v_1, v_2, a\}$. In this at-most three-dimensional span, a scalar second-variation check on great circles through the critical point gives a tangent direction

e for which, with $u(\varphi) = \cos\varphi\, u + \sin\varphi\, e$, the function $\psi(\varphi) = \Psi_{w_2}(u(\varphi))$ satisfies $\psi''(0) > 0$ whenever $\psi'(0) = 0$ and $\psi(0) < 0$. The check uses only

$$l'(\theta) = -\frac{(\pi - \theta)\sin\theta}{2\pi}, \qquad l''(\theta) = \frac{\sin\theta - (\pi - \theta)\cos\theta}{2\pi};$$

after ordering the signed angles to v_1, v_2, a, the equation $\psi'(0) = 0$ eliminates the coefficient R. This gives

$$\psi''(0) = -\psi(0) + \frac{1}{\pi}\big(\sin\theta_1 + \sin\theta_2 - R\sin\theta_a\big),$$

and the ordered critical equation makes the sine term nonnegative whenever $\psi(0) < 0$; degenerate collinear cases follow by taking limits. Thus the desingularized flow has an expanding tangent direction along $\mathcal{C}_-$. The local center-stable manifold theorem, applied on the finite smooth stratification of this definable boundary critical set, shows that its stable set has codimension at least one in the blown-up coordinates. Blowing down $(r, u, w_2) \mapsto (ru, w_2)$ and taking a countable chart cover preserves Lebesgue null sets. Hence finite-time collisions with $\partial\Omega$ occur only for a null set of initializations.

The trajectory is also precompact. Indeed, writing $f_w(x) = \phi(\langle w_1, x\rangle) + \phi(\langle w_2, x\rangle)$, we have

$$\frac{1}{2}\big(\|w_1\|_2^2 + \|w_2\|_2^2\big) \le \mathbb{E} f_w(x)^2 \le 2\mathbb{E} f^*(x)^2 + 4L(w_1, w_2),$$

where the first inequality uses the nonnegativity of the student cross term, and the second is the triangle inequality in L^2. Since L decreases along gradient flow, every trajectory in Ω stays in a compact sublevel set. The same closed-form expression makes L definable on this compact set, so the Kurdyka–Lojasiewicz convergence theorem for bounded definable gradient flows [11] implies that the trajectory converges to a stationary point.

It remains to exclude convergence to nonglobal stationary points for almost every initialization. The closed-form stationary equations are definable, so the nonglobal stationary set admits a finite smooth stratification. By theorem 3, these strata are exactly the in-plane saddle families and the scalar-parameterized out-of-plane saddle family. Along each stratum the gradient is identically zero, so tangent directions lie in the kernel of the Hessian. The proofs of lemmas 25 and 26 show that every nonglobal stationary point has a strictly negative Hessian direction transverse to its critical stratum. Equivalently, the linearized gradient flow has a strictly expanding normal direction.

The local center-stable manifold theorem therefore gives, near each saddle stratum, a center-stable set of codimension at least one containing all trajectories that converge to that stratum. These local sets have Lebesgue measure zero, and a countable cover of the saddle strata gives a countable union. Pulling these sets back by the time-n flow maps, $n \in \mathbb{N}$, still gives measure-zero sets because the flow maps are local C^1 diffeomorphisms on Ω. Hence the set of initializations whose trajectories converge to a saddle has Lebesgue measure zero.

For every remaining initialization, the limiting stationary point is not a saddle. The landscape classification then leaves only the two global minima, (v_1, v_2) and (v_2, v_1). ■

C.3. Phase 2: escape directions near the bisector saddle

Lemma 55 (Hessian eigenspaces at the bisector saddle) *At $w_{\text{saddle}} = (r_s\bar{v}, r_s\bar{v})$, the negative eigenspace of $\nabla^2 L(w_{\text{saddle}})$ is*

$$\operatorname{span}\{(v_\Delta, -v_\Delta), (u_j, -u_j) : 3 \le j \le d\}.$$

The planar antisymmetric eigenvalue is $-2\sin(\alpha/2)/(\pi r_s)$, while each out-of-plane antisymmetric eigenvalue is $-\sin(\alpha/2)/(\pi r_s)$. The positive eigenspace is

$$\operatorname{span}\{(\bar{v}, \bar{v}), (v_\Delta, v_\Delta), (u_j, u_j) : 3 \le j \le d\},$$

and the kernel is $\operatorname{span}\{(\bar{v}, -\bar{v})\}$.

C.3.1. PROOF OF LEMMA 55

Let w_s be the vector lies on the angular bisector of the two teacher vectors v_1 and v_2, and with norm $\|w_s\|_2 = \frac{(\pi - \frac{\alpha}{2})\cos(\frac{\alpha}{2}) + \sin(\frac{\alpha}{2})}{\pi}$. Let the saddle point be $w_{\text{saddle}} = (w_s, w_s)$, and let $w(t) = (w_1(t), w_2(t))$ be the student vectors.

We start with the following equation:

$$\begin{aligned}
\frac{1}{2}\frac{d\|w - w_{\text{saddle}}\|_2^2}{dt} &= \langle w - w_{\text{saddle}}, \frac{dw}{dt}\rangle \\
&= \langle w - w_{\text{saddle}}, -\nabla_w L(w)\rangle \\
&= \langle w - w_{\text{saddle}}, -H(w - w_{\text{saddle}}) + o(\|w - w_{\text{saddle}}\|_2)\rangle \\
&= -\langle w - w_{\text{saddle}}, H(w - w_{\text{saddle}})\rangle + o(\|w - w_{\text{saddle}}\|_2^2),
\end{aligned}$$

where H is the Hessian matrix of the loss function with respect to the student vectors. And we need to calculate the Hessian matrix H of the loss function with respect to the student vectors. By Theorem 5 in [33], we have that the Hessian matrix is given by

$$H = \nabla_w^2 L(w_s, w_s) = \begin{pmatrix} A & B \\ B & A \end{pmatrix},$$

where A and B are given by

$$\begin{aligned}
A &= \frac{1}{2}I - \frac{\sin(\alpha/2)}{\pi r_\star}\left(I - \bar{v}\bar{v}^\top + v_\Delta v_\Delta^\top\right), \\
B &= \frac{1}{2}I,
\end{aligned}$$

where $\bar{v} = \frac{v_1 + v_2}{\|v_1 + v_2\|_2}$ is the normalized mean of the teacher vectors, $v_\Delta = \frac{v_1 - v_2}{\|v_1 - v_2\|_2}$ is the normalized vector orthogonal to $\bar{v}$, and $r_s = \frac{(\pi - \frac{\alpha}{2})\cos(\frac{\alpha}{2}) + \sin(\frac{\alpha}{2})}{\pi}$.

By linear algebra, the eigenvalues of the matrix $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ are given by the eigenvalues of the matrix $A + B$ and $A - B$. In addition, the eigenvectors of the matrix $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$ are given by

$\begin{pmatrix}\mathbf{x}\\\mathbf{x}\end{pmatrix}$ where $\mathbf{x}$ is the eigenvector of $A+B$, and $\begin{pmatrix}\mathbf{x}\\-\mathbf{x}\end{pmatrix}$ where $\mathbf{x}$ is the eigenvector of $A-B$. The eigenvalues of the matrix $A+B$ are given by

$$\begin{aligned}
\lambda_{1,2,\ldots,d-2} &= 1-\frac{\sin(\alpha/2)}{\pi r_s},\ \text{with } d-2 \text{ eigenvectors } \mathbf{x}\perp \text{span}\{\bar{v}, v_\Delta\},\\
\lambda_{d-1} &= 1-\frac{2\sin(\alpha/2)}{\pi r_s},\ \text{with 1 eigenvector } v_\Delta,\\
\lambda_d &= 1,\ \text{with 1 eigenvector } \bar{v}.
\end{aligned}$$

The eigenvalues of the matrix $A-B$ are given by

$$\begin{aligned}
\lambda_{d+1,d+2,\ldots,2d-2} &= -\frac{\sin(\alpha/2)}{\pi r_s},\ \text{with } d-2 \text{ eigenvectors } \mathbf{x}\perp \text{span}\{\bar{v}, v_\Delta\},\\
\lambda_{2d-1} &= -\frac{2\sin(\alpha/2)}{\pi r_s},\ \text{with 1 eigenvector } v_\Delta,\\
\lambda_{2d} &= 0,\ \text{with 1 eigenvector } \bar{v}.
\end{aligned}$$

In the notation of the main text, $\hat{v}=\bar{v}=u_1$ and $\hat{u}=v_\Delta=u_2$. We order the orthonormal Hessian eigenbasis as follows:

$$\begin{gathered}
\mathbf{e}_j=\frac{1}{\sqrt{2}}(u_{j+2},u_{j+2}),\quad 1\le j\le d-2,\qquad \mathbf{e}_{d-1}=\frac{1}{\sqrt{2}}(u_2,u_2),\qquad \mathbf{e}_d=\frac{1}{\sqrt{2}}(\bar{v},\bar{v}),\\
\mathbf{e}_{d+j}=\frac{1}{\sqrt{2}}(u_{j+2},-u_{j+2}),\quad 1\le j\le d-2,\\
\mathbf{e}_{2d-1}=\frac{1}{\sqrt{2}}(u_2,-u_2),\qquad \mathbf{e}_{2d}=\frac{1}{\sqrt{2}}(\bar{v},-\bar{v}).
\end{gathered}$$

Let $\{\mathbf{e}_i\}_{i=1}^{2d}$ be the orthonormal Hessian eigenbasis at w_{saddle}, with eigenvalues λ_i, and write

$$w(t)-w_{\text{saddle}}=\sum_{i=1}^{2d}x_i(t)\mathbf{e}_i.$$

We define the nonnegative- and negative-eigenspace components by

$$V_+(t)=\sum_{i:\lambda_i\ge 0}x_i(t)\mathbf{e}_i,\qquad V_-(t)=\sum_{i:\lambda_i<0}x_i(t)\mathbf{e}_i.$$

Let t_{in} be the first time at which $\|w(t_{\text{in}})-w_{\text{saddle}}\|_2=\varepsilon$. We use two stopping times measured from t_{in}: Δt_1 is the rebalancing time, when the nonnegative-eigenspace component has become comparable to the negative-eigenspace component, and Δt_2 is the escape time, when the trajectory first reaches radius ε_1 from the saddle. More precisely,

$$\begin{aligned}
\Delta t_1 &= \inf\{\tau\ge 0:\|V_+(t_{\text{in}}+\tau)\|_2\le C\|V_-(t_{\text{in}}+\tau)\|_2\},\\
\Delta t_2 &= \inf\{\tau\ge 0:\|w(t_{\text{in}}+\tau)-w_{\text{saddle}}\|_2=\varepsilon_1\}.
\end{aligned}$$

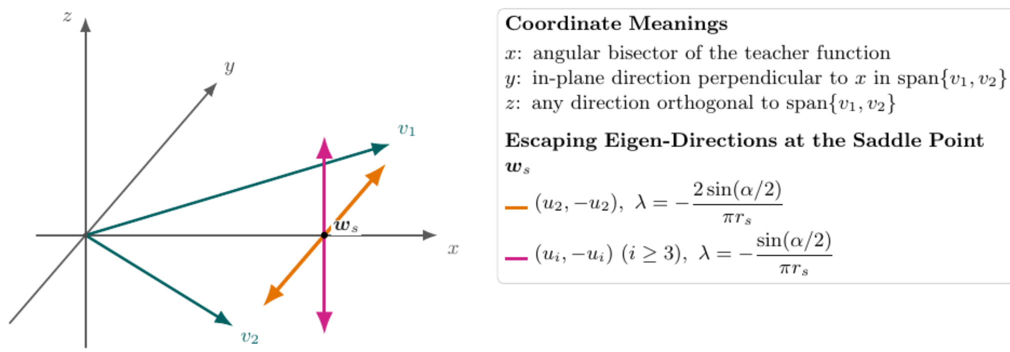

Figure 11: Hessian eigendirections at the positive-bisector saddle. The unstable directions split the two weights antisymmetrically.

Proposition 56 *Let $\Delta t_0(\varepsilon, \varepsilon_1, \sigma_0) = \min\{\Delta t_1, \Delta t_2\}$. Then $\mathbb{P}(\Delta t_0 < +\infty) = 1$.*

Theorem 57 (Escape time near the bisector saddle) *Let $V_-(t)$ and $V_+(t)$ be the projections of $w(t) - w_{\text{saddle}}$ onto the negative and nonnegative Hessian eigenspaces at w_{saddle}, respectively. There exist constants $C > 1$ and $\varepsilon_0 > 0$, depending only on α, such that the following holds. Let $0 < \varepsilon, \varepsilon_1 \leq \varepsilon_0$ satisfy $C\varepsilon \leq \varepsilon_1$. Let t_{in} be the first time such that*

$$\|w(t_{\text{in}}) - w_{\text{saddle}}\|_2 = \varepsilon,$$

and define

$$\Delta t_2 = \inf\{\tau \geq 0 : \|w(t_{\text{in}} + \tau) - w_{\text{saddle}}\|_2 = \varepsilon_1\}.$$

If $\|V_+(t_{\text{in}})\|_2 \leq C\|V_-(t_{\text{in}})\|_2$, then

$$\Delta t_2 \asymp_\alpha \log \frac{\varepsilon_1}{\varepsilon}.$$

If $\|V_+(t_{\text{in}})\|_2 > C\|V_-(t_{\text{in}})\|_2$, let

$$\Delta t_1 = \inf\{\tau \geq 0 : \|V_+(t_{\text{in}} + \tau)\|_2 \leq C\|V_-(t_{\text{in}} + \tau)\|_2\}.$$

Then $\Delta t_1 \leq \Delta t_2$ and

$$\Delta t_2 - \Delta t_1 \asymp_\alpha \log \frac{\varepsilon_1}{\|w(t_{\text{in}} + \Delta t_1) - w_{\text{saddle}}\|_2}.$$

Here the implicit constants depend only on α.

Proof [Proof of theorem 57] We use the Hessian eigenbasis from appendix C.3.1. The zero eigendirection is tangent to the saddle family, so the transverse estimates below only use the strictly positive and negative eigenspaces. Thus, within this proof, V_+ denotes the strictly positive component. Write

$$\mathbf{x}(t) = w(t) - w_{\text{saddle}} = \sum_{i=1}^{2d} x_i(t)\mathbf{e}_i, \qquad \|V_+(t)\|_2^2 = \sum_{i:\lambda_i>0} x_i(t)^2, \qquad \|V_-(t)\|_2^2 = \sum_{i:\lambda_i<0} x_i(t)^2. \tag{200}$$

By Taylor expansion of the gradient at w_{saddle},

$$\begin{aligned} \frac{\mathrm{d}w(t)}{\mathrm{d}t} &= -\nabla L(w(t)) = -\nabla^2 L(w_{\text{saddle}})\big(w(t) - w_{\text{saddle}}\big) + o(\|w(t) - w_{\text{saddle}}\|_2) \\ &= -\sum_{i=1}^{2d}\big(\lambda_i x_i(t) + r_i(\mathbf{x}(t))\big)\mathbf{e}_i. \end{aligned} \tag{201}$$

Set

$$t_3 = t_{\text{in}}, \qquad \|w(t_3) - w_{\text{saddle}}\|_2 = \varepsilon.$$

For the spectral constants, define

$$\begin{aligned} \lambda^+_{\text{max}} &= \max_{i:\lambda_i>0} \lambda_i, & \lambda^+_{\text{min}} &= \min_{i:\lambda_i>0} \lambda_i, & \lambda^-_{\text{max}} &= \max_{i:\lambda_i<0} \lambda_i, \\ \lambda^-_{\text{min}} &= \min_{i:\lambda_i<0} \lambda_i, & \lambda_{\text{min}} &= \min_{1\le i\le 2d-1} |\lambda_i|, & \lambda_{\text{max}} &= \max_{1\le i\le 2d} |\lambda_i|. \end{aligned}$$

The constant C in the theorem will be chosen large. In the proof it is useful to use an auxiliary balancing threshold $C_{\text{bal}} < C$. Choose C large enough that the denominator below is positive. Define

$$\widetilde{C} = \sqrt{\frac{-\lambda^-_{\text{min}} + \frac{2\lambda^+_{\text{max}}}{C^2}}{\frac{1}{4}\lambda^+_{\text{min}} - \frac{2\lambda^+_{\text{max}}}{C^2}}}.$$

We then choose C_{bal} so that

$$\max\{1, \widetilde{C}\} < C_{\text{bal}} < C, \qquad C > 1 + C_{\text{bal}}^2.$$

This is possible after increasing C, since $\widetilde{C}$ stays bounded as $C \to \infty$. Finally set

$$C'' = \frac{\lambda_{\text{min}}}{2(C_{\text{bal}} + 1)d},$$

and decrease ε_0 to make the Taylor remainder small in the energy estimates. More explicitly, if

$$S_+ = \{i : \lambda_i > 0\}, \qquad S_- = \{i : \lambda_i < 0\},$$

then, whenever $\|\mathbf{x}(t)\|_2 \le \varepsilon_0$,

$$2\left|\sum_{i\in S_\pm} x_i(t) r_i(\mathbf{x}(t))\right| \le C''\|V_\pm(t)\|_2\|\mathbf{x}(t)\|_2.$$

These are the only Taylor remainders used below when differentiating $\|V_+(t)\|_2^2$ and $\|V_-(t)\|_2^2$.

Case 1. Suppose that

$$\|V_+(t_3)\|_2 \le C_{\text{bal}} \|V_-(t_3)\|_2.$$

This comparability persists until the escape time. To see this, write

$$F(t) = \|V_+(t)\|_2^2 - C_{\text{bal}}^2 \|V_-(t)\|_2^2.$$

We also use the elementary bound

$$\|\mathbf{x}(t)\|_2 \le \|V_+(t)\|_2 + \|V_-(t)\|_2,$$

which follows directly from the definitions in (200). We show that F cannot cross from nonpositive to positive. Let $s \le t_3 + \Delta t_2$ be any boundary time with $F(s) = 0$. Then

$$\|V_+(s)\|_2 = C_{\text{bal}} \|V_-(s)\|_2.$$

Using (201), we get

$$\begin{aligned} F'(s) = &-2 \sum_{i \in S_+} \lambda_i x_i(s)^2 + 2C_{\text{bal}}^2 \sum_{i \in S_-} \lambda_i x_i(s)^2 \\ &- 2 \sum_{i \in S_+} x_i(s) r_i(\mathbf{x}(s)) + 2C_{\text{bal}}^2 \sum_{i \in S_-} x_i(s) r_i(\mathbf{x}(s)). \end{aligned}$$

The first line is the linearized Hessian contribution. Since the positive eigenvalues are at least $\lambda_{\min}$ and the negative eigenvalues are at most $-\lambda_{\min}$, the crossing relation gives

$$-2 \sum_{i \in S_+} \lambda_i x_i(s)^2 + 2C_{\text{bal}}^2 \sum_{i \in S_-} \lambda_i x_i(s)^2 \le -4\lambda_{\min} \|V_+(s)\|_2^2.$$

The second line is the Taylor remainder. By the preceding remainder bound and the equality $\|V_+(s)\|_2 = C_{\text{bal}} \|V_-(s)\|_2$, its absolute value is at most

$$C'' \|V_+(s)\|_2 \|\mathbf{x}(s)\|_2 + C_{\text{bal}}^2 C'' \|V_-(s)\|_2 \|\mathbf{x}(s)\|_2 \le \lambda_{\min} \|V_+(s)\|_2^2.$$

Hence

$$F'(s) \le -3\lambda_{\min} \|V_+(s)\|_2^2 < 0,$$

so every boundary point points back into the region $F < 0$. Since $F(t_3) \le 0$, this rules out a crossing from $F \le 0$ to $F > 0$, including the case $F(t_3) = 0$. Hence, throughout $t_3 \le t \le t_3 + \Delta t_2$,

$$\|V_+(t)\|_2 \le C_{\text{bal}} \|V_-(t)\|_2, \qquad \|V_-(t)\|_2^2 \le \|\mathbf{x}(t)\|_2^2 \le (1 + C_{\text{bal}}^2) \|V_-(t)\|_2^2.$$

For the negative component, (201) gives, for $t \in [t_3, t_3 + \Delta t_2]$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \|V_-(t)\|_2^2 = -2 \sum_{i \in S_-} \lambda_i x_i(t)^2 - 2 \sum_{i \in S_-} x_i(t) r_i(\mathbf{x}(t)).$$

On the same interval, the comparability above and the Taylor remainder bound give

$$2\left|\sum_{i\in S_-} x_i(t)r_i(\mathbf{x}(t))\right| \le C''(1+C_{\mathrm{bal}})\|V_-(t)\|_2^2.$$

Also, for every $t \in [t_3, t_3+\Delta t_2]$, since $\lambda_{\min}^- \le \lambda_i \le \lambda_{\max}^- < 0$ for all $i \in S_-$,

$$2(-\lambda_{\max}^-)\|V_-(t)\|_2^2 \le -2\sum_{i\in S_-}\lambda_i x_i(t)^2 \le 2(-\lambda_{\min}^-)\|V_-(t)\|_2^2.$$

Thus we may take

$$a_- = 2(-\lambda_{\max}^-) - C''(1+C_{\mathrm{bal}}), \qquad b_- = 2(-\lambda_{\min}^-) + C''(1+C_{\mathrm{bal}}).$$

Combining the last two estimates with the decomposition of $\frac{\mathrm{d}}{\mathrm{d}t}\|V_-(t)\|_2^2$, we get, for every $t \in [t_3, t_3+\Delta t_2]$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|V_-(t)\|_2^2 \ge 2(-\lambda_{\max}^-)\|V_-(t)\|_2^2 - C''(1+C_{\mathrm{bal}})\|V_-(t)\|_2^2 = a_-\|V_-(t)\|_2^2,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\|V_-(t)\|_2^2 \le 2(-\lambda_{\min}^-)\|V_-(t)\|_2^2 + C''(1+C_{\mathrm{bal}})\|V_-(t)\|_2^2 = b_-\|V_-(t)\|_2^2.$$

These constants satisfy $0 < a_- < b_-$. Indeed, since $-\lambda_{\max}^- \ge \lambda_{\min}$ and $C''(1+C_{\mathrm{bal}}) = \lambda_{\min}/(2d)$,

$$a_- \ge 2\lambda_{\min} - \frac{\lambda_{\min}}{2d} > 0.$$

Moreover,

$$b_- - a_- = 2(\lambda_{\max}^- - \lambda_{\min}^-) + 2C''(1+C_{\mathrm{bal}}) > 0.$$

All constants used here depend only on α: the eigenvalue bounds $\lambda_{\max}^{\pm}, \lambda_{\min}^{\pm}, \lambda_{\min}$ are fixed by the Hessian at the saddle; C_{bal} is chosen from these spectral constants; C'' is then defined from $\lambda_{\min}$ and C_{bal}. Thus a_- and b_- also depend only on α. We have proved that, for every $t \in [t_3, t_3+\Delta t_2]$,

$$a_-\|V_-(t)\|_2^2 \le \frac{\mathrm{d}}{\mathrm{d}t}\|V_-(t)\|_2^2 \le b_-\|V_-(t)\|_2^2. \tag{202}$$

Integrating gives

$$\|V_-(t_3)\|_2^2 e^{a_-(t-t_3)} \le \|V_-(t)\|_2^2 \le \|V_-(t_3)\|_2^2 e^{b_-(t-t_3)}.$$

Applying the norm comparison at $t = t_3$ and at $t = t_3 + \Delta t_2$, we obtain

$$\frac{1}{b_-}\log\frac{\varepsilon_1^2}{(1+C_{\mathrm{bal}}^2)\varepsilon^2} \le \Delta t_2 \le \frac{1}{a_-}\log\frac{(1+C_{\mathrm{bal}}^2)\varepsilon_1^2}{\varepsilon^2}. \tag{203}$$

The only radius gap used in the last display is

$$\varepsilon_1 \ge (1+C_{\mathrm{bal}}^2)\varepsilon. \tag{204}$$

Since $1 + C_{\text{bal}}^2$ depends only on α, $C > 1 + C_{\text{bal}}^2$, and $\varepsilon_1 \geq C\varepsilon$, (204) holds. Therefore

$$\Delta t_2 \asymp_\alpha \log \frac{\varepsilon_1}{\varepsilon}$$

in the comparable case.

For the remaining two cases we use one common observation. Suppose that

$$\|V_+(t_3)\|_2 > C_{\text{bal}} \|V_-(t_3)\|_2.$$

Introduce the auxiliary balancing time

$$\begin{aligned} \Delta t_{\text{bal}} &= \inf\{\tau \geq 0 \colon \|V_+(t_3 + \tau)\|_2 \leq C_{\text{bal}} \|V_-(t_3 + \tau)\|_2\}, \\ \Delta t_2 &= \inf\{\tau \geq 0 \colon \|w(t_3 + \tau) - w_{\text{saddle}}\|_2 = \varepsilon_1\}. \end{aligned}$$

We first show that $\Delta t_{\text{bal}} \leq \Delta t_2$. Suppose, to the contrary, that $\Delta t_2 < \Delta t_{\text{bal}}$. Then

$$\|V_+(t)\|_2 > C_{\text{bal}} \|V_-(t)\|_2, \qquad t_3 \leq t \leq t_3 + \Delta t_2.$$

We will show that the endpoint instead satisfies $\|V_+(t_3 + \Delta t_2)\|_2 < \widetilde{C} \|V_-(t_3 + \Delta t_2)\|_2$, which contradicts $\widetilde{C} < C_{\text{bal}}$. For any $t \in [t_3, t_3 + \Delta t_2]$, the decomposition (201) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \|V_+(t)\|_2^2 = -2 \sum_{i \in S_+} \lambda_i x_i(t)^2 - 2 \sum_{i \in S_+} x_i(t) r_i(\mathbf{x}(t)).$$

The Taylor remainder bound gives

$$2 \left| \sum_{i \in S_+} x_i(t) r_i(\mathbf{x}(t)) \right| \leq C'' \|V_+(t)\|_2 \|\mathbf{x}(t)\|_2.$$

On this interval, the contradiction assumption gives $\|V_-(t)\|_2 < C_{\text{bal}}^{-1} \|V_+(t)\|_2$. Hence

$$\|\mathbf{x}(t)\|_2 \leq \|V_+(t)\|_2 + \|V_-(t)\|_2 \leq (1 + C_{\text{bal}}^{-1}) \|V_+(t)\|_2 \leq 2 \|V_+(t)\|_2.$$

Combining these estimates, for every $t \in [t_3, t_3 + \Delta t_2]$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|V_+(t)\|_2^2 &\leq -2\lambda_{\min}^+ \|V_+(t)\|_2^2 + 2C'' \|V_+(t)\|_2^2 \\ &\leq -\lambda_{\min}^+ \|V_+(t)\|_2^2 < 0, \end{aligned} \tag{205}$$

where the last inequality uses

$$2C'' = \frac{\lambda_{\min}}{(C_{\text{bal}} + 1)d} \leq \lambda_{\min} \leq \lambda_{\min}^+.$$

Here the first inequality follows from $(C_{\text{bal}} + 1)d \geq 1$, and the second follows from the definition $\lambda_{\min} = \min_{1 \leq i \leq 2d-1} |\lambda_i|$. It follows that

$$\|V_+(t)\|_2 \leq \|V_+(t_3)\|_2, \qquad t \in [t_3, t_3 + \Delta t_2].$$

Next we compare the loss. Using the coordinate definition (200), Taylor's theorem at w_{saddle} gives

$$L(w(t)) - L(w_{\text{saddle}}) = \langle \nabla L(w_{\text{saddle}}), \mathbf{x}(t)\rangle + \frac{1}{2}\langle \mathbf{x}(t), \nabla^2 L(w_{\text{saddle}})\mathbf{x}(t)\rangle + R_L(t),$$

where $R_L(t) = o(\|\mathbf{x}(t)\|_2^2)$. The linear term is zero because $\nabla L(w_{\text{saddle}}) = 0$. Also, since $\{\mathbf{e}_i\}_{i=1}^{2d}$ is an orthonormal Hessian eigenbasis with $\nabla^2 L(w_{\text{saddle}})\mathbf{e}_i = \lambda_i \mathbf{e}_i$,

$$\langle \mathbf{x}(t), \nabla^2 L(w_{\text{saddle}})\mathbf{x}(t)\rangle = \sum_{i=1}^{2d} \lambda_i x_i(t)^2.$$

Thus

$$L(w(t)) - L(w_{\text{saddle}}) = \frac{1}{2}\sum_{i=1}^{2d} \lambda_i x_i(t)^2 + R_L(t).$$

After decreasing ε_0, we may assume that, whenever $\|\mathbf{x}(t)\|_2 \le \varepsilon_0$,

$$|R_L(t)| \le \min\left\{\frac{1}{2}\lambda_{\max}^+, \frac{1}{4}\lambda_{\min}^+, -\frac{1}{2}\lambda_{\min}^-\right\}\left(\|V_+(t)\|_2^2 + \|V_-(t)\|_2^2\right).$$

On the interval $t \in [t_3, t_3 + \Delta t_2]$, we are still under the temporary assumption $\Delta t_2 < \Delta t_{\text{bal}}$. By the definition of Δt_{bal}, the trajectory has not yet entered the balanced region, so

$$\|V_+(t)\|_2 > C_{\text{bal}}\|V_-(t)\|_2.$$

Since $C_{\text{bal}} > 1$, this gives $\|V_-(t)\|_2 \le \|V_+(t)\|_2$. Therefore

$$\begin{aligned} L(w(t)) - L(w_{\text{saddle}}) &\le \frac{1}{2}\lambda_{\max}^+\|V_+(t)\|_2^2 + \lambda_{\max}^+\|V_+(t)\|_2^2 \\ &< 2\lambda_{\max}^+\|V_+(t)\|_2^2 \\ &\le 2\lambda_{\max}^+\|V_+(t_3)\|_2^2 \\ &\le 2\lambda_{\max}^+\|\mathbf{x}(t_3)\|_2^2 = 2\lambda_{\max}^+\varepsilon^2. \end{aligned}$$

On the other hand, for every $t \in [t_3, t_3 + \Delta t_2]$,

$$\begin{aligned} L(w(t)) - L(w_{\text{saddle}}) &\ge \frac{1}{2}\lambda_{\min}^+\|V_+(t)\|_2^2 + \frac{1}{2}\lambda_{\min}^-\|V_-(t)\|_2^2 \\ &\quad - \frac{1}{4}\lambda_{\min}^+\|V_+(t)\|_2^2 + \frac{1}{2}\lambda_{\min}^-\|V_-(t)\|_2^2 \\ &= \frac{1}{4}\lambda_{\min}^+\|V_+(t)\|_2^2 + \lambda_{\min}^-\|V_-(t)\|_2^2. \end{aligned}$$

At the escape time $t_3 + \Delta t_2$, the definition of Δt_2 and $\varepsilon_1 \ge C\varepsilon$ give the first relation below. The loss comparison above gives the second:

$$\|V_+(t_3 + \Delta t_2)\|_2^2 + \|V_-(t_3 + \Delta t_2)\|_2^2 = \varepsilon_1^2 \ge C^2\varepsilon^2, \tag{206}$$

$$\frac{1}{4}\lambda_{\min}^+\|V_+(t_3 + \Delta t_2)\|_2^2 + \lambda_{\min}^-\|V_-(t_3 + \Delta t_2)\|_2^2 < 2\lambda_{\max}^+\varepsilon^2. \tag{207}$$

By (206),

$$\varepsilon^2 \leq \frac{1}{C^2}\left(\|V_+(t_3+\Delta t_2)\|_2^2 + \|V_-(t_3+\Delta t_2)\|_2^2\right).$$

Substituting this upper bound for ε^2 into (207) gives

$$\begin{aligned}\frac{1}{4}\lambda_{\min}^+\|V_+(t_3+\Delta t_2)\|_2^2 + \lambda_{\min}^-\|V_-(t_3+\Delta t_2)\|_2^2 \\ < \frac{2\lambda_{\max}^+}{C^2}\left(\|V_+(t_3+\Delta t_2)\|_2^2 + \|V_-(t_3+\Delta t_2)\|_2^2\right).\end{aligned}$$

Rearranging,

$$\left(\frac{1}{4}\lambda_{\min}^+ - \frac{2\lambda_{\max}^+}{C^2}\right)\|V_+(t_3+\Delta t_2)\|_2^2 < \left(-\lambda_{\min}^- + \frac{2\lambda_{\max}^+}{C^2}\right)\|V_-(t_3+\Delta t_2)\|_2^2.$$

By the definition of $\widetilde{C}$, this is

$$\|V_+(t_3+\Delta t_2)\|_2^2 < \widetilde{C}^2\|V_-(t_3+\Delta t_2)\|_2^2.$$

Hence

$$\begin{aligned}\|V_+(t_3+\Delta t_2)\|_2 &< \widetilde{C}\|V_-(t_3+\Delta t_2)\|_2 \\ &< C_{\text{bal}}\|V_-(t_3+\Delta t_2)\|_2,\end{aligned}$$

contradicting the assumption $\Delta t_2 < \Delta t_{\text{bal}}$. Therefore

$$\Delta t_{\text{bal}} \leq \Delta t_2.$$

Case 2. Suppose that the initial ratio lies between the auxiliary threshold and the theorem threshold:

$$C_{\text{bal}}\|V_-(t_3)\|_2 < \|V_+(t_3)\|_2 \leq C\|V_-(t_3)\|_2,$$

Then the ratio is already bounded by a constant depending only on α. On any time interval where

$$C_{\text{bal}}\|V_-(t)\|_2 < \|V_+(t)\|_2 \leq C\|V_-(t)\|_2,$$

The V_+-estimate is the same calculation as (205). The V_--estimate is the same calculation as (202), with the factor $1 + C_{\text{bal}}$ replaced by $1 + C$. Hence there are constants $c_+, a_{\text{mid}}, b_{\text{mid}} > 0$, depending only on α, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|V_+(t)\|_2^2 \leq -c_+\|V_+(t)\|_2^2, \qquad a_{\text{mid}}\|V_-(t)\|_2^2 \leq \frac{\mathrm{d}}{\mathrm{d}t}\|V_-(t)\|_2^2 \leq b_{\text{mid}}\|V_-(t)\|_2^2.$$

Thus the ratio

$$R(t) = \frac{\|V_+(t)\|_2^2}{\|V_-(t)\|_2^2}$$

satisfies

$$R'(t) \leq -(c_+ + a_{\text{mid}})R(t)$$

as long as $C_{\mathrm{bal}}^2 < R(t) \leq C^2$. Since $R(t_3) \leq C^2$, the ratio cannot cross above C^2 before reaching C_{bal}^2. Hence the estimate applies up to Δt_{bal}, and

$$\Delta t_{\mathrm{bal}} \leq \frac{1}{c_+ + a_{\mathrm{mid}}} \log \frac{C^2}{C_{\mathrm{bal}}^2} = O_\alpha(1).$$

The same exponential bounds show that $\|V_-(t_3 + \Delta t_{\mathrm{bal}})\|_2$ is comparable to $\|V_-(t_3)\|_2$, with constants depending only on α. Since $\|\mathbf{x}(t_3)\|_2 = \varepsilon$ and the initial ratio lies between C_{bal} and C, we have $\|V_-(t_3)\|_2 \asymp_\alpha \varepsilon$. At $t_3 + \Delta t_{\mathrm{bal}}$, the defining relation $\|V_+(t_3 + \Delta t_{\mathrm{bal}})\|_2 = C_{\mathrm{bal}} \|V_-(t_3 + \Delta t_{\mathrm{bal}})\|_2$ then gives

$$\varepsilon_{\mathrm{bal}} := \|\mathbf{x}(t_3 + \Delta t_{\mathrm{bal}})\|_2 \asymp_\alpha \varepsilon.$$

It remains to apply the comparable case from this new starting time. To avoid changing the meaning of the original Δt_2, write the remaining escape time as

$$\widehat{\Delta t}_2 = \inf\{\tau \geq 0 : \|w(t_3 + \Delta t_{\mathrm{bal}} + \tau) - w_{\mathrm{saddle}}\|_2 = \varepsilon_1\}.$$

Since $\Delta t_{\mathrm{bal}} \leq \Delta t_2$, the original escape time decomposes as

$$\Delta t_2 = \Delta t_{\mathrm{bal}} + \widehat{\Delta t}_2.$$

There are two subcases. First suppose that

$$\varepsilon_1 \geq (1 + C_{\mathrm{bal}}^2)\varepsilon_{\mathrm{bal}},$$

which is (204) with ε replaced by $\varepsilon_{\mathrm{bal}}$. The comparable case, applied with initial radius $\varepsilon_{\mathrm{bal}}$, gives

$$\widehat{\Delta t}_2 \asymp_\alpha \log \frac{\varepsilon_1}{\varepsilon_{\mathrm{bal}}}.$$

Since C has already been fixed as a constant depending only on α, the comparison $\varepsilon_{\mathrm{bal}} \asymp_\alpha \varepsilon$ gives

$$\log \frac{\varepsilon_1}{\varepsilon_{\mathrm{bal}}} = \log \frac{\varepsilon_1}{\varepsilon} + O_\alpha(1).$$

Together with $\Delta t_{\mathrm{bal}} = O_\alpha(1)$, this gives

$$\Delta t_2 \asymp_\alpha \log \frac{\varepsilon_1}{\varepsilon}.$$

It remains to consider the complementary subcase

$$\varepsilon_1 < (1 + C_{\mathrm{bal}}^2)\varepsilon_{\mathrm{bal}}.$$

The comparison $\varepsilon_{\mathrm{bal}} \asymp_\alpha \varepsilon$ then gives

$$C \leq \frac{\varepsilon_1}{\varepsilon} \leq O_\alpha(1),$$

so $\log(\varepsilon_1/\varepsilon) \asymp_\alpha 1$. We next bound Δt_2 from above. At time $t_3 + \Delta t_{\mathrm{bal}}$, the ratio is exactly C_{bal}, so the comparable-case argument applies from this new starting time. The upper bound in (203), with initial radius $\varepsilon_{\mathrm{bal}}$, gives

$$\widehat{\Delta t}_2 \leq \frac{1}{a_-} \log \frac{(1 + C_{\mathrm{bal}}^2)\varepsilon_1^2}{\varepsilon_{\mathrm{bal}}^2}.$$

In the present subcase $\varepsilon_1 < (1 + C_{\rm bal}^2)\varepsilon_{\rm bal}$, and $\varepsilon_{\rm bal} \le \varepsilon_1$ because $t_3 + \Delta t_{\rm bal} \le t_3 + \Delta t_2$ and $t_3 + \Delta t_2$ is the first time the radius reaches ε_1. Hence the logarithm above is bounded by a constant depending only on α. Thus $\widehat{\Delta t}_2 = O_\alpha(1)$, and therefore

$$\Delta t_2 = \Delta t_{\rm bal} + \widehat{\Delta t}_2 = O_\alpha(1).$$

For the lower bound, the spectral constant $\lambda_{\max}$ defined above controls the linear part in (201). The Taylor expansion there has remainder vector $r(\mathbf{x}(t)) = (r_i(\mathbf{x}(t)))_{i=1}^{2d} = o(\|\mathbf{x}(t)\|_2)$. After decreasing ε_0, we may assume

$$\|r(\mathbf{x}(t))\|_2 \le \|\mathbf{x}(t)\|_2 \qquad \text{whenever } \|\mathbf{x}(t)\|_2 \le \varepsilon_0.$$

Thus, as long as $\|\mathbf{x}(t)\|_2 \le \varepsilon_0$,

$$\|\dot{\mathbf{x}}(t)\|_2 \le \lambda_{\max}\|\mathbf{x}(t)\|_2 + \|r(\mathbf{x}(t))\|_2 \le (\lambda_{\max} + 1)\|\mathbf{x}(t)\|_2.$$

By the coordinate definition in (200), $\mathbf{x}(t)$ is the displacement from $w_{\rm saddle}$ in the Hessian eigenbasis. Thus, at any time with $\mathbf{x}(t) \neq 0$, Cauchy–Schwarz gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2 = \frac{\langle \mathbf{x}(t), \dot{\mathbf{x}}(t)\rangle}{\|\mathbf{x}(t)\|_2} \le \|\dot{\mathbf{x}}(t)\|_2.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2 \le (\lambda_{\max} + 1)\|\mathbf{x}(t)\|_2.$$

We can take $M_\alpha = \lambda_{\max} + 1$ and obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2 \le M_\alpha\|\mathbf{x}(t)\|_2.$$

Integrating this differential inequality on $[t_3, t_3 + \Delta t_2]$ gives

$$\varepsilon_1 = \|\mathbf{x}(t_3 + \Delta t_2)\|_2 \le \|\mathbf{x}(t_3)\|_2 e^{M_\alpha \Delta t_2} = \varepsilon e^{M_\alpha \Delta t_2}.$$

Hence

$$\Delta t_2 \ge \frac{1}{M_\alpha}\log\frac{\varepsilon_1}{\varepsilon}.$$

Thus the same conclusion holds in the complementary subcase. This proves the first assertion of the theorem for the whole range $\|V_+(t_3)\|_2 \le C\|V_-(t_3)\|_2$.

Case 3. Finally suppose that $\|V_+(t_3)\|_2 > C\|V_-(t_3)\|_2$, and let Δt_1 be the time defined in the theorem, with threshold C. Since $C_{\rm bal} < C$ and $\Delta t_{\rm bal} \le \Delta t_2$, continuity gives

$$\Delta t_1 \le \Delta t_{\rm bal} \le \Delta t_2.$$

On the interval from $t_3 + \Delta t_1$ to $t_3 + \Delta t_{\rm bal}$, the ratio lies between $C_{\rm bal}$ and C. Hence the same bounded-ratio estimate gives

$$\begin{aligned} \Delta t_{\rm bal} - \Delta t_1 &= O_\alpha(1), \\ \|\mathbf{x}(t_3 + \Delta t_{\rm bal})\|_2 &\asymp_\alpha \|\mathbf{x}(t_3 + \Delta t_1)\|_2. \end{aligned} \tag{208}$$

Write

$$\widehat{\Delta t}_2 = \inf\{\tau \geq 0 : \|w(t_3 + \Delta t_{\mathrm{bal}} + \tau) - w_{\mathrm{saddle}}\|_2 = \varepsilon_1\}.$$

Then

$$\Delta t_2 - \Delta t_1 = (\Delta t_{\mathrm{bal}} - \Delta t_1) + \widehat{\Delta t}_2. \tag{209}$$

The comparable case, applied from $t_3 + \Delta t_{\mathrm{bal}}$, gives

$$\widehat{\Delta t}_2 \asymp_\alpha \log \frac{\varepsilon_1}{\|\mathbf{x}(t_3 + \Delta t_{\mathrm{bal}})\|_2}. \tag{210}$$

We now combine (208), (209), and (210). The norm comparison in (208) implies

$$\log \frac{\varepsilon_1}{\|\mathbf{x}(t_3 + \Delta t_{\mathrm{bal}})\|_2} = \log \frac{\varepsilon_1}{\|\mathbf{x}(t_3 + \Delta t_1)\|_2} + O_\alpha(1).$$

The time comparison in (208) contributes only the additive term $\Delta t_{\mathrm{bal}} - \Delta t_1 = O_\alpha(1)$ in (209). Therefore

$$\Delta t_2 - \Delta t_1 \asymp_\alpha \log \frac{\varepsilon_1}{\|w(t_3 + \Delta t_1) - w_{\mathrm{saddle}}\|_2}.$$

This completes the proof. ■

C.4. Phase 2: homogeneous exclusion of out-of-plane saddles

By the definition of the gradient flow and $z_{i,j}$ as the u_j-coordinate of the i-th student weight, we have

$$\begin{aligned}
\frac{\mathrm{d}z_{1,1}}{\mathrm{d}t} &= \frac{1}{2\pi}\left((2\pi - \theta_{11} - \theta_{12})\cos\frac{\alpha}{2} + \Big(\frac{\sin\theta_{11} + \sin\theta_{12}}{\|w_1\|_2} - \pi - \sin\gamma\frac{\|w_2\|_2}{\|w_1\|_2}\Big)z_{1,1} - (\pi - \gamma)z_{2,1}\right),\\
\frac{\mathrm{d}z_{2,1}}{\mathrm{d}t} &= \frac{1}{2\pi}\left((2\pi - \theta_{21} - \theta_{22})\cos\frac{\alpha}{2} + \Big(\frac{\sin\theta_{21} + \sin\theta_{22}}{\|w_2\|_2} - \pi - \sin\gamma\frac{\|w_1\|_2}{\|w_2\|_2}\Big)z_{2,1} - (\pi - \gamma)z_{1,1}\right),\\
\frac{\mathrm{d}z_{1,2}}{\mathrm{d}t} &= \frac{1}{2\pi}\left((\theta_{12} - \theta_{11})\sin\frac{\alpha}{2} + \Big(\frac{\sin\theta_{11} + \sin\theta_{12}}{\|w_1\|_2} - \pi - \sin\gamma\frac{\|w_2\|_2}{\|w_1\|_2}\Big)z_{1,2} - (\pi - \gamma)z_{2,2}\right),\\
\frac{\mathrm{d}z_{2,2}}{\mathrm{d}t} &= \frac{1}{2\pi}\left((\theta_{22} - \theta_{21})\sin\frac{\alpha}{2} + \Big(\frac{\sin\theta_{21} + \sin\theta_{22}}{\|w_2\|_2} - \pi - \sin\gamma\frac{\|w_1\|_2}{\|w_2\|_2}\Big)z_{2,2} - (\pi - \gamma)z_{1,2}\right),\\
\frac{\mathrm{d}z_{1,j}}{\mathrm{d}t} &= \frac{1}{2\pi}\left(\Big(\frac{\sin\theta_{11} + \sin\theta_{12}}{\|w_1\|_2} - \pi - \sin\gamma\frac{\|w_2\|_2}{\|w_1\|_2}\Big)z_{1,j} - (\pi - \gamma)z_{2,j}\right),\\
\frac{\mathrm{d}z_{2,j}}{\mathrm{d}t} &= \frac{1}{2\pi}\left(\Big(\frac{\sin\theta_{21} + \sin\theta_{22}}{\|w_2\|_2} - \pi - \sin\gamma\frac{\|w_1\|_2}{\|w_2\|_2}\Big)z_{2,j} - (\pi - \gamma)z_{1,j}\right).
\end{aligned} \tag{211}$$

Lemma 58 (Planar split sign preservation) *If $z_{1,2}(0) > 0 > z_{2,2}(0)$, then*

$$z_{1,2}(t) > 0 > z_{2,2}(t) \qquad \textit{for all } t \geq 0.$$

Proof Let τ be the first time at which either $z_{1,2}$ or $z_{2,2}$ vanishes. If $\tau < \infty$, uniqueness for the ODE rules out $z_{1,2}(\tau) = z_{2,2}(\tau) = 0$, since this would force $z_{1,2}(0) = z_{2,2}(0) = 0$. If $z_{1,2}(\tau) = 0$ and $z_{2,2}(\tau) < 0$, then (211) gives $\dot{z}_{1,2}(\tau) = -(\pi - \gamma)z_{2,2}(\tau)/(2\pi) > 0$, so $z_{1,2}$ could not have reached zero from the positive side. The case $z_{2,2}(\tau) = 0 < z_{1,2}(\tau)$ is identical. Hence no such first time exists. ■

Lemma 59 (Homogeneous coordinate domination) *Under assumption 60, let C_{hom} be a constant for which*

$$|z_{i,j}(0)| \leq C_{\text{hom}}|z_{i,2}(0)| \qquad (i \in \{1,2\},\ 3 \leq j \leq d)$$

holds. Then, for every $j \geq 3$ and every $t \geq 0$,

$$|z_{1,j}(t) - z_{2,j}(t)| \leq C_{\text{hom}}|z_{1,2}(t) - z_{2,2}(t)|.$$

Proof We first fix the sign convention used in the rest of the proof. The two student neurons are exchangeable, so after relabeling the neurons we may assume

$$z_{1,2}(0) > 0 > z_{2,2}(0).$$

Then lemma 58 gives

$$z_{1,2}(t) > 0 > z_{2,2}(t) \qquad \text{for all } t \geq 0.$$

Next fix $j \geq 3$. The orientation of the basis vector u_j is arbitrary. Replacing u_j by $-u_j$ sends $(z_{1,j}, z_{2,j})$ to $(-z_{1,j}, -z_{2,j})$, leaves (211) unchanged, and leaves the desired absolute-value estimate unchanged. Hence the sign pattern $z_{1,j}(0) > 0 > z_{2,j}(0)$ and its reflected pattern are the same case. In the rest of the proof for this fixed j, we assume

$$z_{1,j}(0) > 0 > z_{2,j}(0).$$

Under this convention, the sign of the transverse pair is preserved. Indeed, if $z_{1,j}$ were the first of the two coordinates to vanish while $z_{2,j} < 0$, then the last two lines of (211) would give $\dot{z}_{1,j} = -(\pi - \gamma)z_{2,j}/(2\pi) > 0$, so $z_{1,j}$ could not cross from positive to nonpositive. The same argument gives $z_{2,j}(t) < 0$ for all $t \geq 0$.

By lemma 58, $z_{1,2}(t) > 0 > z_{2,2}(t)$ for all $t \geq 0$. Set

$$a_j(t) = C_{\text{hom}}z_{1,2}(t) - z_{1,j}(t), \qquad b_j(t) = C_{\text{hom}}z_{2,2}(t) - z_{2,j}(t).$$

The initialization gives $a_j(0) > 0$ and $b_j(0) < 0$. We claim that these signs are preserved. Suppose first that a_j reaches zero before either sign condition fails. Since $z_{1,2} > 0 > z_{2,2}$, the teacher geometry gives $\theta_{12} > \theta_{11}$ and $\theta_{22} < \theta_{21}$. Using (211), at the first time with $a_j = 0$ and $b_j \leq 0$,

$$\dot{a}_j = \frac{1}{2\pi}\left[C_{\text{hom}}(\theta_{12} - \theta_{11})\sin\frac{\alpha}{2} - (\pi - \gamma)b_j\right] > 0.$$

Thus a_j cannot cross from positive to nonpositive. Similarly, at the first time with $b_j = 0$ and $a_j \geq 0$,

$$\dot{b}_j = \frac{1}{2\pi}\left[C_{\text{hom}}(\theta_{22} - \theta_{21})\sin\frac{\alpha}{2} - (\pi - \gamma)a_j\right] < 0,$$

so b_j cannot cross from negative to nonnegative.

Therefore $0 < z_{1,j}(t) < C_{\text{hom}}z_{1,2}(t)$ and $C_{\text{hom}}z_{2,2}(t) < z_{2,j}(t) < 0$. Since both coordinate pairs have opposite signs,

$$|z_{1,j}(t) - z_{2,j}(t)| = z_{1,j}(t) - z_{2,j}(t) \leq C_{\text{hom}}\big(z_{1,2}(t) - z_{2,2}(t)\big) = C_{\text{hom}}|z_{1,2}(t) - z_{2,2}(t)|.$$

■

Assumption 60 (Homogeneous initialization) *In the basis extending $u_1 = \bar{v}$ and $u_2 = v_\Delta$, write $w_i(0) = \sum_j z_{i,j}(0)u_j$. The planar splitting coordinates have opposite signs, $z_{1,2}(0)z_{2,2}(0) < 0$, and every out-of-plane coordinate is dominated by the corresponding planar splitting coordinate: for some absolute constant $C > 0$,*

$$|z_{i,j}(0)/z_{i,2}(0)| \leq C, \qquad i \in \{1,2\}, \quad 3 \leq j \leq d.$$

Theorem 61 (Homogeneous initialization excludes out-of-plane saddles) *Under assumption 60, the gradient flow satisfies*

$$|z_{1,j}(t) - z_{2,j}(t)| \leq C|z_{1,2}(t) - z_{2,2}(t)|$$

for every $j \geq 3$ and $t \geq 0$. Consequently, the trajectory remains a positive distance away from the out-of-plane saddle family.

C.4.1. PROOF OF THEOREM 61

Under assumption 60, let C_{hom} be a constant for which

$$|z_{i,j}(0)| \leq C_{\text{hom}}|z_{i,2}(0)| \qquad (i \in \{1,2\},\ 3 \leq j \leq d)$$

holds. The coordinate estimate is lemma 59. It remains to separate the resulting cone from the out-of-plane saddle family. If $d = 2$, this family is empty. Assume $d \geq 3$.

Let

$$w(t) = \big(w_1(t), w_2(t)\big), \qquad p(w(t)) = |z_{1,2} - z_{2,2}|, \qquad q(w(t)) = \left(\sum_{j=3}^{d} |z_{1,j} - z_{2,j}|^2\right)^{1/2}.$$

By lemma 59, the trajectory lies in the cone

$$q(w(t)) \leq C_{\text{hom}}\sqrt{d-2}\, p(w(t)).$$

On the other hand, every out-of-plane saddle has the form

$$W_1 = r(\cos\kappa\, u_1 + \sin\kappa\, \xi), \qquad W_2 = r(\cos\kappa\, u_1 - \sin\kappa\, \xi), \qquad \xi \perp \operatorname{span}\{v_1, v_2\},$$

as in theorem 3 and lemma 25. Hence $p(W) = 0$ and $q(W) = 2r\sin\kappa$. For fixed α, the scalar characterization of the out-of-plane saddles gives

$$m_\alpha := \inf_{W \in \mathcal{S}_\perp} q(W) > 0.$$

The closed cone $q \leq C_{\text{hom}}\sqrt{d-2}\, p$ is therefore a positive Euclidean distance from the closed set $\{p = 0,\ q \geq m_\alpha\}$. This distance is a constant $c_{\alpha, C_{\text{hom}}, d} > 0$. Since the trajectory stays in the cone,

$$\operatorname{dist}\big(w(t), \mathcal{S}_\perp\big) \geq c_{\alpha, C_{\text{hom}}, d} \qquad \text{for all } t \geq 0.$$

Theorem 62 *Let $\mathcal{G} = \{(v_1, v_2), (v_2, v_1)\}$. Fix $\epsilon > 0$, and let $t_{\text{reach global}}$ be the first time such that $\operatorname{dist}(w(t), \mathcal{G}) \le \epsilon$. Suppose there is a finite time $t_{\text{esc}} < t_{\text{reach global}}$ such that, on $[t_{\text{esc}}, t_{\text{reach global}})$, the trajectory stays a fixed positive distance from every nonglobal stationary point. Then there exists $c_{\text{grad}}(\epsilon) > 0$ such that $\|\nabla_w L(w(t))\|_2 \ge c_{\text{grad}}(\epsilon)$ for all $t \in [t_{\text{esc}}, t_{\text{reach global}})$. Moreover,*

$$t_{\text{reach global}} - t_{\text{esc}} \le \frac{L(w(t_{\text{esc}}))}{c_{\text{grad}}(\epsilon)^2}.$$

Proof [Proof of theorem 62]

By the landscape classification in theorem 3, the only stationary points outside $\mathcal{G}$ are saddles. The assumed positive distance from those saddles, together with the condition $\operatorname{dist}(w(t), \mathcal{G}) \ge \epsilon$ for $t < t_{\text{reach global}}$, gives a compact region on which $\|\nabla_w L\|_2$ has a positive minimum. Denote this minimum by $c_{\text{grad}}(\epsilon)$.

Along gradient flow, $\frac{\mathrm{d}}{\mathrm{d}t} L(w(t)) = -\|\nabla_w L(w(t))\|_2^2$. Integrating from t_{esc} to $t_{\text{reach global}}$ gives

$$\begin{aligned} L(w(t_{\text{esc}})) &\ge L(w(t_{\text{esc}})) - L(w(t_{\text{reach global}})) \\ &= \int_{t_{\text{esc}}}^{t_{\text{reach global}}} \|\nabla_w L(w(t))\|_2^2 \, dt \\ &\ge c_{\text{grad}}(\epsilon)^2 \big(t_{\text{reach global}} - t_{\text{esc}}\big). \end{aligned}$$

This proves the stated bound. ∎

C.5. Phase 2: a quantitative escape theorem under structured initialization

C.5.1. PROOF OF THEOREM 6

Theorem 63 (Structured logarithmic escape) *Fix $\alpha \in (0, \pi/2]$ and $0 < m < M < \infty$. Set*

$$\mu_\alpha = \frac{2\sin(\alpha/2)}{\pi r_s},$$

where r_s is the positive-bisector saddle radius. Use the teacher-adapted basis from (46), extended to an orthonormal basis of $\mathbb{R}^d$. There are constants $\rho_0 > 0$, $\sigma_0 > 0$, $D_0 < \infty$, and $\varepsilon_0 = \varepsilon_0(\alpha) > 0$ such that the following holds. Let $d \ge D_0$ and let σ satisfy

$$\sigma \le \sigma_0, \qquad \sigma\sqrt{d} \le \rho_0.$$

For each time write

$$w_i(t) = z_{i,1}(t) u_1 + z_{i,2}(t) u_2 + o_i(t), \qquad o_i(t) \perp \operatorname{span}\{u_1, u_2\}.$$

Assume the initial data satisfy

$$m\sigma\sqrt{d} \le \frac{z_{1,1}(0) + z_{2,1}(0)}{2} \le M\sigma\sqrt{d}, \qquad m\sigma \le z_{1,2}(0) - z_{2,2}(0) \le M\sigma,$$
$$z_{1,2}(0) > 0 > z_{2,2}(0), \qquad \frac{|z_{1,1}(0) - z_{2,1}(0)|}{2} + \big(\|o_1(0)\|_2^2 + \|o_2(0)\|_2^2\big)^{1/2} \le M\sigma.$$

For every $0 < \varepsilon \le \varepsilon_0$*, define*

$$T_\varepsilon = \inf\left\{t \ge 0 : z_{1,2}(t)^2 + z_{2,2}(t)^2 + \|o_1(t)\|_2^2 + \|o_2(t)\|_2^2 \ge \varepsilon^2\right\}.$$

Then T_ε *is finite and, for a constant* $C = C_{\alpha,m,M,\varepsilon}$ *independent of* σ *and* d*,*

$$\left|T_\varepsilon - \left[\left(\frac{1}{\mu_\alpha} - 1\right)\log\frac{1}{\sigma} + \frac{1}{2}\log d\right]\right| \le C.$$

Proof [Proof of theorems 6 and 63] All constants in this proof are independent of σ and d. We use the following convention. A constant written as C_α may depend only on α. A constant written as $C_{\alpha,m,M}$ may depend only on α, m, M. A constant written as $C_{\alpha,m,M,\varepsilon_{\text{in}},\varepsilon}$ may also depend on the two fixed saddle-tube radii. Constants with bootstrap parameters in the subscript, such as $C_{\alpha,\varepsilon,K_R}$, may depend on those displayed parameters in addition to α. The same symbol may change from line to line, but its allowed dependencies are always those shown in the subscript. Inside a lemma proof, an unsubscripted C has the dependency class stated in that lemma, unless the line displays a smaller dependency such as C_α. When a numerical constant is needed, we write it explicitly, as in $1/8$, $7/8$, or 2. The constants ρ_0, σ_0, κ, and D_0 in theorem 6 are chosen so that $\sigma\sqrt{d}$ and σ are small, d is large, and σ is bounded below by $e^{-\kappa d}$. Since $\alpha \in (0, \pi/2]$, the positivity of μ_α is immediate. For the upper bound,

$$\begin{aligned}\mu_\alpha < 1 &\iff 2\sin(\alpha/2) < (\pi - \alpha/2)\cos(\alpha/2) + \sin(\alpha/2)\\ &\iff \tan(\alpha/2) < \pi - \alpha/2.\end{aligned}$$

The inequality $\tan(\alpha/2) < \pi - \alpha/2$ holds because $\alpha/2 \in (0, \pi/4]$ and the function $x \mapsto \pi - x - \tan x$ is positive on this interval.

For the rest of the proof write, on every stopped interval under discussion,

$$w_i(t) = z_{i,1}(t)u_1 + z_{i,2}(t)u_2 + o_i(t), \qquad o_i(t) \perp \operatorname{span}\{u_1, u_2\},$$

and set

$$\begin{aligned}R(t) &= \frac{z_{1,1}(t) + z_{2,1}(t)}{2},\\ \delta(t) &= \frac{z_{1,1}(t) - z_{2,1}(t)}{2},\\ a(t) &= z_{1,2}(t) - z_{2,2}(t),\\ Y(t)^2 &= \|o_1(t)\|_2^2 + \|o_2(t)\|_2^2,\\ N_\perp(t)^2 &= z_{1,2}(t)^2 + z_{2,2}(t)^2 + Y(t)^2.\end{aligned}$$

Write $R_0 = R(0)$, $\delta_0 = \delta(0)$, $a_0 = a(0)$, and $Y_0 = Y(0)$. The component equations are those in (211). When (215) applies, it gives $z_{i,1} \ge 7R/8 > 0$. When (247) or (269) applies, it gives $z_{i,1} \ge r_s/8 > 0$. Thus, in every part of the proof where (211) is used, the two student directions remain in the positive radial sector, and hence $\gamma < \pi$. If $z_{1,2} = 0$ and $z_{2,2} < 0$, then

$$\dot{z}_{1,2} = -\frac{\pi - \gamma}{2\pi} z_{2,2} > 0,$$

and if $z_{2,2} = 0$ and $z_{1,2} > 0$, then

$$\dot{z}_{2,2} = -\frac{\pi - \gamma}{2\pi} z_{1,2} < 0.$$

Hence a first time with $z_{1,2} = 0$ and $z_{2,2} < 0$ cannot occur, and a first time with $z_{2,2} = 0$ and $z_{1,2} > 0$ cannot occur. The corner $z_{1,2} = z_{2,2} = 0$ also cannot be reached at a first finite time. Indeed, at that corner the two angle differences $\theta_{12} - \theta_{11}$ and $\theta_{22} - \theta_{21}$ vanish, and the third and fourth equations in (211) have zero right-hand side. By uniqueness for the locally Lipschitz ODE in the positive radial sector, a trajectory that reaches this corner at a finite time must have had $z_{1,2} = z_{2,2} = 0$ on the preceding interval, contradicting the strict initial signs. Since $z_{1,2}(0) > 0 > z_{2,2}(0)$, for every time considered in the proof

$$z_{1,2}(t) > 0 > z_{2,2}(t), \qquad a(t) = |z_{1,2}(t)| + |z_{2,2}(t)|, \qquad |z_{1,2}(t) + z_{2,2}(t)| \le a(t).$$

We first prove the estimates used before the trajectory enters the saddle chart as a separate lemma. This also isolates the Taylor calculations for the teacher angles.

Lemma 64 (Thin-cone Taylor and transport estimates) *There are constants $\eta_0 = \eta_0(\alpha) > 0$ and $C_\alpha > 0$ such that, whenever*

$$\frac{a + Y + |\delta|}{R} \le \eta_0, \qquad 0 < R \le \frac{r_s}{2},$$

one has

$$\left| \dot{R} - (r_s - R) \right| + |\dot{\delta}| \le C_\alpha \frac{a^2 + Y^2}{R^2}. \tag{212}$$

$$\left| \frac{d}{dt} \log a - \frac{2\sin(\alpha/2)}{\pi R} \right| \le C_\alpha \left(1 + \frac{|\delta|}{R^2} + \frac{a^2 + Y^2}{R^3} \right), \tag{213}$$

and

$$\frac{d}{dt}(Y^2 + \xi)^{1/2} \le \left[\frac{\sin(\alpha/2)}{\pi R} + C_\alpha \left(1 + \frac{|\delta|}{R^2} + \frac{a^2 + Y^2}{R^3} \right) \right] (Y^2 + \xi)^{1/2} \qquad \textit{for every } \xi > 0. \tag{214}$$

Proof We justify (212)–(214) now. In the cone, the comparison between $z_{i,1}$, $\|w_i\|_2$, and R is as follows. Since $z_{i,1} = R \pm \delta$, $|z_{i,2}| \le a$, and $\|o_i\|_2 \le Y$, if $\eta_0 \le 1/8$, then

$$\begin{aligned} &\frac{7}{8} R \le R - |\delta| \le z_{i,1} \le R + |\delta| \le \frac{9}{8} R, \\ &\frac{7}{8} R \le z_{i,1} \le \|w_i\|_2 = (z_{i,1}^2 + z_{i,2}^2 + \|o_i\|_2^2)^{1/2} \\ &\quad \le z_{i,1} + |z_{i,2}| + \|o_i\|_2 \le R + |\delta| + a + Y \le \frac{9}{8} R + \eta_0 R \le \frac{5}{4} R. \end{aligned} \tag{215}$$

The comparison bounds in (215) give the cone comparability used in the Taylor estimates. For each i,

$$\cos\theta_{i1} = \frac{z_{i,1}\cos(\alpha/2) + z_{i,2}\sin(\alpha/2)}{(z_{i,1}^2 + z_{i,2}^2 + \|o_i\|_2^2)^{1/2}},$$

$$\cos\theta_{i2} = \frac{z_{i,1}\cos(\alpha/2) - z_{i,2}\sin(\alpha/2)}{(z_{i,1}^2 + z_{i,2}^2 + \|o_i\|_2^2)^{1/2}}.$$

We now expand these quantities. By (215),

$$0 \le \frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2} \le C\eta_0^2.$$

Thus, after decreasing η_0 once, Taylor's theorem applied to $x \mapsto (1+x)^{-1/2}$ gives a constant C_α such that, for $0 \le x \le C_\alpha \eta_0^2$,

$$\left|(1+x)^{-1/2} - 1 + \frac{x}{2}\right| \le C_\alpha x^2.$$

Applying this with $x = z_{i,2}^2/z_{i,1}^2 + \|o_i\|_2^2/z_{i,1}^2$ gives

$$\left|(1 + z_{i,2}^2/z_{i,1}^2 + \|o_i\|_2^2/z_{i,1}^2)^{-1/2} - 1 + \frac{1}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right| \le C\frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}.$$

Here we used (215) to replace powers of $z_{i,1}$ in the error by powers of R. Hence there is a constant C_α such that

$$\begin{aligned}
&\left|\cos\theta_{i1} - \cos(\alpha/2) - \frac{z_{i,2}}{z_{i,1}}\sin(\alpha/2) + \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right| \\
&\le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right), \\
&\left|\cos\theta_{i2} - \cos(\alpha/2) + \frac{z_{i,2}}{z_{i,1}}\sin(\alpha/2) + \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right| \\
&\le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right).
\end{aligned} \tag{216}$$

We shall use the following Taylor bound, valid for x close to $\cos(\alpha/2)$:

$$\left|\arccos x - \frac{\alpha}{2} + \frac{x - \cos(\alpha/2)}{\sin(\alpha/2)} + \frac{\cos(\alpha/2)}{2\sin^3(\alpha/2)}(x - \cos(\alpha/2))^2\right| \le C_\alpha |x - \cos(\alpha/2)|^3. \tag{217}$$

Substituting the first-order part of the two formulas for $\cos\theta_{i1}$ and $\cos\theta_{i2}$ from (216) into (217) gives the rough angle estimates

$$\begin{aligned}
\left|\theta_{i1} - \frac{\alpha}{2} + \frac{z_{i,2}}{z_{i,1}}\right| &\le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2}, \\
\left|\theta_{i2} - \frac{\alpha}{2} - \frac{z_{i,2}}{z_{i,1}}\right| &\le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2}.
\end{aligned} \tag{218}$$

To prove the estimate for $\theta_{i2} - \theta_{i1} - 2z_{i,2}/z_{i,1}$, write the angle difference directly as

$$\theta_{i2} - \theta_{i1} = \int_{-z_{i,2}/z_{i,1}}^{z_{i,2}/z_{i,1}} \frac{d}{dx}\arccos\frac{\cos(\alpha/2) - x\sin(\alpha/2)}{\left(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2\right)^{1/2}}\,dx.$$

First compute the inner derivative:

$$
\begin{aligned}
&\frac{d}{dx}\left[\big(\cos(\alpha/2) - x\sin(\alpha/2)\big)\left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{-1/2}\right] \\
&= -\sin(\alpha/2)\left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{-1/2} \\
&\quad - x\big(\cos(\alpha/2) - x\sin(\alpha/2)\big)\left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{-3/2} \\
&= -\frac{\sin(\alpha/2)(1 + \|o_i\|_2^2/z_{i,1}^2) + x\cos(\alpha/2)}{\Big(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2\Big)^{3/2}}.
\end{aligned}
$$

The numerator $\sin(\alpha/2)(1 + \|o_i\|_2^2/z_{i,1}^2) + x\cos(\alpha/2)$ follows from the explicit cancellation

$$
\begin{aligned}
&\sin(\alpha/2)\left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) + x(\cos(\alpha/2) - x\sin(\alpha/2)) \\
&= \sin(\alpha/2)\left(1 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) + x\cos(\alpha/2).
\end{aligned}
$$

The square-root factor from differentiating $\arccos$ is

$$
\begin{aligned}
1 - \frac{(\cos(\alpha/2) - x\sin(\alpha/2))^2}{1 + x^2 + \|o_i\|_2^2/z_{i,1}^2} &= \frac{\begin{array}{c}1 + x^2 + \|o_i\|_2^2/z_{i,1}^2 \\ -(\cos(\alpha/2) - x\sin(\alpha/2))^2\end{array}}{1 + x^2 + \|o_i\|_2^2/z_{i,1}^2} \\
&= \frac{\sin^2(\alpha/2) + 2x\sin(\alpha/2)\cos(\alpha/2) + x^2\cos^2(\alpha/2) + \|o_i\|_2^2/z_{i,1}^2}{1 + x^2 + \|o_i\|_2^2/z_{i,1}^2} \\
&= \frac{\big(\sin(\alpha/2) + x\cos(\alpha/2)\big)^2}{1 + x^2 + \|o_i\|_2^2/z_{i,1}^2} \\
&\quad + \frac{\|o_i\|_2^2/z_{i,1}^2}{1 + x^2 + \|o_i\|_2^2/z_{i,1}^2}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\frac{d}{dx}\arccos\frac{\cos(\alpha/2) - x\sin(\alpha/2)}{\Big(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2\Big)^{1/2}} \\
&= \frac{\sin(\alpha/2)(1 + \|o_i\|_2^2/z_{i,1}^2) + x\cos(\alpha/2)}{\Big(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2\Big)\Big((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \|o_i\|_2^2/z_{i,1}^2\Big)^{1/2}}.
\end{aligned}
$$

For $|x| \le 2|z_{i,2}|/z_{i,1}$, the cone condition gives $|x| + \|o_i\|_2/z_{i,1} \le C_\alpha\eta_0$. After decreasing η_0 if needed,

$$
\sin(\alpha/2) + x\cos(\alpha/2) \ge \frac{1}{2}\sin(\alpha/2),
$$

$$\left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{1/2} \ge \frac{1}{2}\sin(\alpha/2),$$

$$\left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{1/2} \le C_\alpha.$$

Therefore

$$\begin{aligned}
&\left|\frac{d}{dx}\arccos\frac{\cos(\alpha/2) - x\sin(\alpha/2)}{\left(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2\right)^{1/2}} - 1\right| \\
&\le C_\alpha \left|\sin(\alpha/2) + x\cos(\alpha/2) + \sin(\alpha/2)\frac{\|o_i\|_2^2}{z_{i,1}^2}\right. \\
&\quad \left. - \left(1 + x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{1/2}\right| \\
&\le C_\alpha \frac{\|o_i\|_2^2}{z_{i,1}^2} \\
&\quad + C_\alpha \left|\sin(\alpha/2) + x\cos(\alpha/2) - \left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{1/2}\right| \\
&\quad + C_\alpha \left(x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)^{1/2} \\
&\le C_\alpha \frac{\|o_i\|_2^2}{z_{i,1}^2} \\
&\quad + C_\alpha \frac{\|o_i\|_2^2/z_{i,1}^2}{\left((\sin(\alpha/2) + x\cos(\alpha/2))^2 + \|o_i\|_2^2/z_{i,1}^2\right)^{1/2} + \sin(\alpha/2) + x\cos(\alpha/2)} \\
&\quad + C_\alpha \left(x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) \\
&\le C_\alpha \left(x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\left|\theta_{i2} - \theta_{i1} - \frac{2z_{i,2}}{z_{i,1}}\right| &= \left|\int_{-z_{i,2}/z_{i,1}}^{z_{i,2}/z_{i,1}} \left[\frac{d}{dx}\arccos\frac{\cos(\alpha/2) - x\sin(\alpha/2)}{(1 + x^2 + \|o_i\|_2^2/z_{i,1}^2)^{1/2}} - 1\right] dx\right| \\
&\le \int_{-|z_{i,2}|/z_{i,1}}^{|z_{i,2}|/z_{i,1}} C_\alpha \left(x^2 + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) dx \\
&\le C_\alpha \left(\frac{|z_{i,2}|^3}{R^3} + \frac{|z_{i,2}|\|o_i\|_2^2}{R^3}\right) \le C_\alpha \frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3}.
\end{aligned} \tag{219}$$

For the sum $\theta_{i1} + \theta_{i2}$, the first-order terms in the two cosine expansions cancel, so the second-order terms must be kept. To make this explicit, write (216) as

$$\cos\theta_{i1} - \cos(\alpha/2) = \sin(\alpha/2)\frac{z_{i,2}}{z_{i,1}} - \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) + E_{i1},$$
$$\cos\theta_{i2} - \cos(\alpha/2) = -\sin(\alpha/2)\frac{z_{i,2}}{z_{i,1}} - \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right) + E_{i2},$$

where the two remainders satisfy

$$|E_{i1}| + |E_{i2}| \le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right). \tag{220}$$

Adding the two displayed identities cancels $\pm\sin(\alpha/2)z_{i,2}/z_{i,1}$ and gives

$$\begin{aligned}&\left|(\cos\theta_{i1} - \cos(\alpha/2)) + (\cos\theta_{i2} - \cos(\alpha/2)) + \cos(\alpha/2)\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right| \\ &\le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right).\end{aligned} \tag{221}$$

The squared estimate follows from the same decomposition. Squaring the two lines and adding gives

$$\begin{aligned}&(\cos\theta_{i1} - \cos(\alpha/2))^2 + (\cos\theta_{i2} - \cos(\alpha/2))^2 - 2\sin^2(\alpha/2)\frac{z_{i,2}^2}{z_{i,1}^2} \\ &= 2\left[\frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right]^2 \\ &\quad + 2\left[\sin(\alpha/2)\frac{z_{i,2}}{z_{i,1}} - \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right]E_{i1} \\ &\quad + 2\left[-\sin(\alpha/2)\frac{z_{i,2}}{z_{i,1}} - \frac{\cos(\alpha/2)}{2}\left(\frac{z_{i,2}^2}{z_{i,1}^2} + \frac{\|o_i\|_2^2}{z_{i,1}^2}\right)\right]E_{i2} + E_{i1}^2 + E_{i2}^2.\end{aligned}$$

The first term on the right is bounded by $C_\alpha(z_{i,2}^2 + \|o_i\|_2^2)^2/R^4$. The terms involving E_{i1}, E_{i2} are bounded by (220), using $z_{i,1} \simeq R$ from (215) and the cone smallness. Therefore

$$\begin{aligned}&\left|(\cos\theta_{i1} - \cos(\alpha/2))^2 + (\cos\theta_{i2} - \cos(\alpha/2))^2 - 2\sin^2(\alpha/2)\frac{z_{i,2}^2}{z_{i,1}^2}\right| \\ &\le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right).\end{aligned} \tag{222}$$

Substituting (221) and (222) into (217) gives

$$\begin{aligned}\left|\theta_{i1} + \theta_{i2} - \alpha - \frac{\cos(\alpha/2)}{\sin(\alpha/2)}\frac{\|o_i\|_2^2}{z_{i,1}^2}\right| &\le C_\alpha\left(\frac{|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2)}{R^3} + \frac{(z_{i,2}^2 + \|o_i\|_2^2)^2}{R^4}\right), \\ |\theta_{i1} + \theta_{i2} - \alpha| &\le C_\alpha\frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2}.\end{aligned} \tag{223}$$

For the sine term, use (218). Since Taylor's theorem gives $|\sin(\alpha/2+y)-\sin(\alpha/2)-y\cos(\alpha/2)| \le C_\alpha y^2$,

$$\left|\sin\theta_{i1} - \sin(\alpha/2) + \frac{z_{i,2}}{z_{i,1}}\cos(\alpha/2)\right| \le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2},$$
$$\left|\sin\theta_{i2} - \sin(\alpha/2) - \frac{z_{i,2}}{z_{i,1}}\cos(\alpha/2)\right| \le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2}.$$

Therefore

$$|\sin\theta_{i1} + \sin\theta_{i2} - 2\sin(\alpha/2)| \le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2},$$
$$|\|w_i\|_2 - z_{i,1}| = \frac{z_{i,2}^2 + \|o_i\|_2^2}{\|w_i\|_2 + z_{i,1}} \le C\frac{z_{i,2}^2 + \|o_i\|_2^2}{R},$$

and hence

$$\left|\frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} - \frac{2\sin(\alpha/2)}{z_{i,1}}\right| \le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^3}. \tag{224}$$

The dependence on o_i begins through $\|o_i\|_2^2/z_{i,1}^2$, so these constants do not depend on d.

On the positive bisector sheet $z_{1,2} = z_{2,2} = 0$ and $o_1 = o_2 = 0$, the first two lines of (211), in the first two component coordinates, give

$$\dot z_{1,1} = \dot z_{2,1} = r_s - R, \qquad \dot R = r_s - R, \qquad \dot\delta = 0$$

for all small δ. Hence the equations for $\dot R - (r_s - R)$ and $\dot\delta$ have no pure δ term. In the first two lines of (211), the terms containing γ are $-\sin\gamma(\|w_2\|_2/\|w_1\|_2)z_{1,1} - (\pi-\gamma)z_{2,1}$ in the equation for $\dot z_{1,1}$, and $-\sin\gamma(\|w_1\|_2/\|w_2\|_2)z_{2,1} - (\pi-\gamma)z_{1,1}$ in the equation for $\dot z_{2,1}$. After expanding $-(\pi-\gamma)z_{j,1} = -\pi z_{j,1} + \gamma z_{j,1}$ and putting $-\pi z_{i,1} - \pi z_{j,1}$ with the terms that do not contain γ, the parts that still contain γ are

$$\gamma z_{2,1} - \sin\gamma\frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}, \qquad \gamma z_{1,1} - \sin\gamma\frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}.$$

To bound γ, first apply (215). For each $i = 1, 2$,

$$\begin{aligned}\left\|\frac{w_i}{\|w_i\|_2} - u_1\right\|_2 &\le \left|1 - \frac{z_{i,1}}{\|w_i\|_2}\right| + \frac{|z_{i,2}| + \|o_i\|_2}{\|w_i\|_2}\\ &= \frac{\|w_i\|_2 - z_{i,1}}{\|w_i\|_2} + \frac{|z_{i,2}| + \|o_i\|_2}{\|w_i\|_2}\\ &\le C\frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2} + C\frac{a+Y}{R} \le C\frac{a+Y}{R}.\end{aligned}$$

Hence

$$\begin{aligned}\left\|\frac{w_1}{\|w_1\|_2} - \frac{w_2}{\|w_2\|_2}\right\|_2 &\le \left\|\frac{w_1}{\|w_1\|_2} - u_1\right\|_2 + \left\|\frac{w_2}{\|w_2\|_2} - u_1\right\|_2 \le C\frac{a+Y}{R},\\ \gamma &\le C\left\|\frac{w_1}{\|w_1\|_2} - \frac{w_2}{\|w_2\|_2}\right\|_2 \le C\frac{a+Y}{R}.\end{aligned} \tag{225}$$

The inequality $\gamma \le C\|w_1/\|w_1\|_2 - w_2/\|w_2\|_2\|_2$ in (225) holds after decreasing η_0, because $\|w_1/\|w_1\|_2 - w_2/\|w_2\|_2\|_2 \le C\eta_0$ gives $\gamma \le \pi/2$, and on $[0, \pi/2]$ one has $\gamma \le (\pi/2)\|w_1/\|w_1\|_2 - w_2/\|w_2\|_2\|_2$. Also

$$\left|\|w_i\|_2 - z_{i,1}\right| = \frac{z_{i,2}^2 + \|o_i\|_2^2}{\|w_i\|_2 + z_{i,1}} \le C\frac{z_{i,2}^2 + \|o_i\|_2^2}{R},$$

$$\begin{aligned}\left|z_{2,1} - \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| &= \left|\frac{z_{2,1}(\|w_1\|_2 - z_{1,1}) - z_{1,1}(\|w_2\|_2 - z_{2,1})}{\|w_1\|_2}\right| \\ &\le C\frac{z_{1,2}^2 + z_{2,2}^2 + \|o_1\|_2^2 + \|o_2\|_2^2}{R} \le C\frac{a^2 + Y^2}{R}.\end{aligned}$$

Interchanging the two indices in the displayed fraction gives

$$\left|z_{1,1} - \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| \le C\frac{a^2 + Y^2}{R}.$$

For the part of the first two lines of (211) that does not contain γ, namely

$$(2\pi - \theta_{i1} - \theta_{i2})\cos(\alpha/2) + \frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} z_{i,1} - \pi z_{i,1} - \pi z_{j,1},$$

(223) and (224) give, for $i = 1, 2$ and $\{j\} = \{1, 2\} \setminus \{i\}$,

$$\begin{aligned}&\left|(2\pi - \theta_{i1} - \theta_{i2})\cos(\alpha/2) + \frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} z_{i,1} - \pi z_{i,1} - \pi z_{j,1} - 2\pi(r_s - R)\right| \\ &\le C_\alpha \frac{z_{i,2}^2 + \|o_i\|_2^2}{R^2},\end{aligned} \tag{226}$$

because $z_{1,1} + z_{2,1} = 2R$ and

$$(2\pi - \alpha)\cos(\alpha/2) + 2\sin(\alpha/2) - 2\pi R = 2\pi(r_s - R).$$

This calculation shows again that no term depending only on δ appears in the first two lines of (211). For the expression $\gamma z_{2,1} - \sin\gamma \|w_2\|_2 z_{1,1}/\|w_1\|_2$,

$$\begin{aligned}\left|\gamma z_{2,1} - \sin\gamma \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| &\le |\gamma| \left|z_{2,1} - \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| + |\gamma - \sin\gamma| \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2} \\ &\le C\frac{a+Y}{R}\frac{a^2 + Y^2}{R} + C\left(\frac{a+Y}{R}\right)^3 R \\ &\le C(a+Y)\frac{a^2 + Y^2}{R^2} + C(a+Y)\frac{(a+Y)^2}{R^2} \\ &\le C\frac{a^2 + Y^2}{R^2},\end{aligned} \tag{227}$$

because $a + Y \le \eta_0 R \le \eta_0 r_s/2$ by the hypotheses and $(a+Y)^2 \le 2(a^2 + Y^2)$. Interchanging the indices gives the same type of bound for $\gamma z_{1,1} - \sin\gamma \|w_1\|_2 z_{2,1}/\|w_2\|_2$:

$$\begin{aligned}\left|\gamma z_{1,1} - \sin\gamma \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| &\le |\gamma| \left|z_{1,1} - \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| + |\gamma - \sin\gamma| \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2} \\ &\le C\frac{a^2 + Y^2}{R^2},\end{aligned}$$

using (225) and the exchanged estimate after (227). Therefore (226) and (227) give

$$|\dot z_{1,1}-(r_s-R)|\le C_\alpha\frac{a^2+Y^2}{R^2},\qquad |\dot z_{2,1}-(r_s-R)|\le C_\alpha\frac{a^2+Y^2}{R^2}.$$

Here there is no term proportional only to δ, because the identities $\dot R=r_s-R$ and $\dot\delta=0$ hold on the whole positive bisector sheet. Thus

$$\begin{aligned}\left|\dot R-(r_s-R)\right|&=\left|\frac{\dot z_{1,1}+\dot z_{2,1}}{2}-(r_s-R)\right|\le C_\alpha\frac{a^2+Y^2}{R^2},\\ |\dot\delta|&=\left|\frac{\dot z_{1,1}-\dot z_{2,1}}{2}\right|\le C_\alpha\frac{a^2+Y^2}{R^2}.\end{aligned}\tag{228}$$

For $a=z_{1,2}-z_{2,2}$, subtracting the third and fourth lines of (211) gives

$$\begin{aligned}\dot a=\frac{1}{2\pi}\Bigg[&\\ &\big((\theta_{12}-\theta_{11})-(\theta_{22}-\theta_{21})\big)\sin(\alpha/2)\\ &+\left(\frac{B_1+B_2}{2}+\pi-\gamma\right)a\\ &+\frac{B_1-B_2}{2}(z_{1,2}+z_{2,2})\Bigg],\end{aligned}\tag{229}$$

where

$$\begin{aligned}B_1&=\frac{\sin\theta_{11}+\sin\theta_{12}}{\|w_1\|_2}-\pi-\sin\gamma\frac{\|w_2\|_2}{\|w_1\|_2},\\ B_2&=\frac{\sin\theta_{21}+\sin\theta_{22}}{\|w_2\|_2}-\pi-\sin\gamma\frac{\|w_1\|_2}{\|w_2\|_2}.\end{aligned}$$

We now compute the leading part of the right side of (229). The difference estimate (219) gives

$$\begin{aligned}&\left|\big((\theta_{12}-\theta_{11})-(\theta_{22}-\theta_{21})\big)\sin(\alpha/2)-2\sin(\alpha/2)\left(\frac{z_{1,2}}{z_{1,1}}-\frac{z_{2,2}}{z_{2,1}}\right)\right|\le C_\alpha\frac{a(a^2+Y^2)}{R^3},\\ &\left|\big((\theta_{12}-\theta_{11})-(\theta_{22}-\theta_{21})\big)\sin(\alpha/2)-\frac{2\sin(\alpha/2)}{R}a\right|\le C_\alpha\left(\frac{a|\delta|}{R^2}+\frac{a(a^2+Y^2)}{R^3}\right),\end{aligned}\tag{230}$$

where we used

$$\begin{aligned}\frac{z_{1,2}}{z_{1,1}}-\frac{z_{2,2}}{z_{2,1}}&=\frac{R(z_{1,2}-z_{2,2})-\delta(z_{1,2}+z_{2,2})}{R^2-\delta^2},\\ \left|\frac{z_{1,2}}{z_{1,1}}-\frac{z_{2,2}}{z_{2,1}}-\frac{a}{R}\right|&\le C\frac{a|\delta|}{R^2}.\end{aligned}\tag{231}$$

For the coefficient involving B_1 and B_2, the sine-ratio estimate (224) gives

$$\left|B_1 - \left(\frac{2\sin(\alpha/2)}{z_{1,1}} - \pi - \sin\gamma\frac{\|w_2\|_2}{\|w_1\|_2}\right)\right| \le C_\alpha \frac{z_{1,2}^2 + \|o_1\|_2^2}{R^3},$$
$$\left|B_2 - \left(\frac{2\sin(\alpha/2)}{z_{2,1}} - \pi - \sin\gamma\frac{\|w_1\|_2}{\|w_2\|_2}\right)\right| \le C_\alpha \frac{z_{2,2}^2 + \|o_2\|_2^2}{R^3}.$$

Also,

$$\begin{gathered}\left|\frac{1}{z_{1,1}} + \frac{1}{z_{2,1}} - \frac{2}{R}\right| = \left|\frac{2R}{R^2-\delta^2} - \frac{2}{R}\right| \le C\frac{|\delta|}{R^2}, \quad \left|\frac{1}{z_{1,1}} - \frac{1}{z_{2,1}}\right| = \frac{2|\delta|}{R^2-\delta^2} \le C\frac{|\delta|}{R^2}, \\ |\sin\gamma| \le C\frac{a+Y}{R} \le C, \quad \frac{\|w_1\|_2}{\|w_2\|_2} + \frac{\|w_2\|_2}{\|w_1\|_2} \le C.\end{gathered} \tag{232}$$

Therefore

$$\begin{aligned}\left|\frac{B_1+B_2}{2} + \pi - \gamma - \frac{2\sin(\alpha/2)}{R}\right| &\le C_\alpha\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right), \\ |B_1 - B_2| &\le C_\alpha\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right).\end{aligned} \tag{233}$$

Since $|z_{1,2} + z_{2,2}| \le a$, substituting (230) and (233) into (229) gives

$$\left|\dot{a} - \frac{2\sin(\alpha/2)}{\pi R}a\right| \le C_\alpha a\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right). \tag{234}$$

Dividing (234) by $a > 0$ gives (213).

The last two lines of (211), written for coordinates $3, \dots, d$, give

$$\dot{o}_1 = \frac{1}{2\pi}(B_1 o_1 - (\pi - \gamma)o_2), \qquad \dot{o}_2 = \frac{1}{2\pi}(B_2 o_2 - (\pi-\gamma)o_1).$$

Therefore

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}Y^2 &= \frac{1}{2\pi}\Big(B_1\|o_1\|_2^2 + B_2\|o_2\|_2^2 - 2(\pi-\gamma)\langle o_1, o_2\rangle\Big) \\ &\le \frac{1}{2\pi}\Bigg[\left(\frac{2\sin(\alpha/2)}{R} - \pi + C\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right)\right)\left(\|o_1\|_2^2 + \|o_2\|_2^2\right) \\ &\qquad + (\pi-\gamma)\left(\|o_1\|_2^2 + \|o_2\|_2^2\right)\Bigg] \\ &\le \left[\frac{\sin(\alpha/2)}{\pi R} + C\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right)\right]Y^2.\end{aligned} \tag{235}$$

In deriving the second inequality in (235), we used $-2\langle o_1, o_2\rangle \le \|o_1\|_2^2 + \|o_2\|_2^2$ and $0 \le \gamma < \pi$. The terms $-\pi Y^2$ and $(\pi-\gamma)Y^2$ leave only a nonpositive remainder, so they may be dropped in the upper bound. Thus (235) gives

$$\frac{d}{dt}Y^2 \le 2\left[\frac{\sin(\alpha/2)}{\pi R} + C\left(1 + \frac{|\delta|}{R^2} + \frac{a^2+Y^2}{R^3}\right)\right]Y^2.$$

For $\xi > 0$,

$$\begin{aligned}\frac{d}{dt}(Y^2+\xi)^{1/2} &= \frac{\frac{d}{dt}Y^2}{2(Y^2+\xi)^{1/2}} \\ &\le \left[\frac{\sin(\alpha/2)}{\pi R} + C\left(1+\frac{|\delta|}{R^2}+\frac{a^2+Y^2}{R^3}\right)\right]\frac{Y^2}{(Y^2+\xi)^{1/2}} \\ &\le \left[\frac{\sin(\alpha/2)}{\pi R} + C\left(1+\frac{|\delta|}{R^2}+\frac{a^2+Y^2}{R^3}\right)\right](Y^2+\xi)^{1/2}.\end{aligned} \tag{236}$$

This proves (214); we use it for fixed $\xi > 0$ and then let $\xi \downarrow 0$ after integration. ■

Lemma 65 (Cone transport to a fixed positive scale) *There is a constant $\eta_* = \eta_*(\alpha) > 0$ with the following property. For initial data satisfying the bounds in theorem 63, assume in addition that*

$$a_0 + Y_0 + |\delta_0| \le \eta_* R_0.$$

Let t_* be the first time when $R(t_*) = r_s/2$. Then t_* is finite and there is a constant $C_{\alpha,m,M} \ge 1$ such that

$$C_{\alpha,m,M}^{-1} a_0 R_0^{-\mu_\alpha} \le a(t_*) \le C_{\alpha,m,M} a_0 R_0^{-\mu_\alpha}, \qquad Y(t_*) \le C_{\alpha,m,M} Y_0 R_0^{-\mu_\alpha/2},$$
$$|\delta(t_*)| \le C_{\alpha,m,M}\left(\sigma + a(t_*)^2 \Lambda_\alpha(R_0)\right),$$

where

$$\Lambda_\alpha(R_0) = \begin{cases} 1, & \mu_\alpha \ne \frac{1}{2}, \\ 1 + \log\frac{1}{R_0}, & \mu_\alpha = \frac{1}{2}. \end{cases}$$

Proof Choose $\eta_* > 0$ small enough depending only on α. Let τ be the first time when either $R = r_s/2$ or

$$\frac{a + Y + |\delta|}{R} = \eta_0.$$

On $[0, \tau]$, the stopped definition gives $0 < R \le r_s/2$ and $(a + Y + |\delta|)/R \le \eta_0$, so lemma 64 applies. After decreasing η_0 if needed,

$$\dot{R} \ge r_s - R - C\eta_0^2 \ge \frac{r_s}{4}, \qquad R \le r_s/2,$$

so R is strictly increasing and $dt \le C\, dR$ for $0 \le t \le \tau$.

For fixed α and the initial data a_0, Y_0, δ_0, R_0, define the comparison quantity, for $x \in [R_0, r_s/2]$, by

$$\mathcal{D}(x) = |\delta_0| + \int_{R_0}^{x} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^2}\, du.$$

Since $a_0 + Y_0 + |\delta_0| \leq \eta_* R_0$ and $0 < \mu_\alpha < 1$, we first compute the integral that controls the error terms in the transport estimates:

$$\begin{aligned}
&\int_{R_0}^{r_s/2} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^3}\, du \\
&\qquad = a_0^2 R_0^{-2\mu_\alpha} \int_{R_0}^{r_s/2} u^{2\mu_\alpha - 3}\, du + Y_0^2 R_0^{-\mu_\alpha} \int_{R_0}^{r_s/2} u^{\mu_\alpha - 3}\, du \\
&\qquad \leq C \frac{a_0^2}{R_0^2} + C \frac{Y_0^2}{R_0^2} \leq C \eta_*^2.
\end{aligned} \tag{237}$$

Next, for $x \in [R_0, r_s/2]$,

$$\begin{aligned}
&\frac{1}{x} \int_{R_0}^{x} \frac{a_0^2 (u/R_0)^{2\mu_\alpha}}{u^2}\, du \\
&\qquad = \frac{a_0^2 R_0^{-2\mu_\alpha}}{x} \int_{R_0}^{x} u^{2\mu_\alpha - 2}\, du \\
&\qquad \leq \begin{cases} C a_0^2 / (R_0 x), & \mu_\alpha < \frac{1}{2}, \\ C a_0^2 R_0^{-1} \log(x/R_0)/x, & \mu_\alpha = \frac{1}{2}, \\ C a_0^2 R_0^{-2\mu_\alpha} x^{2\mu_\alpha - 2}, & \mu_\alpha > \frac{1}{2}, \end{cases} \\
&\qquad \leq C \eta_*^2.
\end{aligned}$$

Here the middle case uses $(R_0/x) \log(x/R_0) \leq C$, and the last case uses $(R_0/x)^{2-2\mu_\alpha} \leq 1$. Similarly,

$$\frac{1}{x} \int_{R_0}^{x} \frac{Y_0^2 (u/R_0)^{\mu_\alpha}}{u^2}\, du = \frac{Y_0^2 R_0^{-\mu_\alpha}}{x} \int_{R_0}^{x} u^{\mu_\alpha - 2}\, du \leq C \frac{Y_0^2}{R_0 x} \leq C \eta_*^2.$$

Therefore

$$\begin{aligned}
\frac{\mathcal{D}(x)}{x} &\leq \frac{|\delta_0|}{x} + \frac{1}{x} \int_{R_0}^{x} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^2}\, du \leq C \eta_*, \\
\int_{R_0}^{r_s/2} \frac{\mathcal{D}(u)}{u^2}\, du &\leq \frac{|\delta_0|}{R_0} + \int_{R_0}^{r_s/2} \left(\int_{v}^{r_s/2} \frac{du}{u^2} \right) \frac{a_0^2 (v/R_0)^{2\mu_\alpha} + Y_0^2 (v/R_0)^{\mu_\alpha}}{v^2}\, dv \\
&\leq \eta_* + \int_{R_0}^{r_s/2} \frac{a_0^2 (v/R_0)^{2\mu_\alpha} + Y_0^2 (v/R_0)^{\mu_\alpha}}{v^3}\, dv \leq C \eta_*.
\end{aligned} \tag{238}$$

Before setting up the bootstrap, note that if $Y_0 = 0$, then the equations for the out-of-plane vectors o_1, o_2 form a homogeneous linear system with locally bounded coefficients on the stopped cone interval. By uniqueness, $Y(t) = 0$ on that interval. Thus the Y bootstrap below is automatic when $Y_0 = 0$; when $Y_0 > 0$, it has positive initial slack. Choose a large fixed K. Let τ_K be the first time before τ when one of

$$a(t) \leq K a_0 \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha}, \qquad Y(t) \leq K Y_0 \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha/2},$$

or

$$|\delta(t)| \leq K^3 \mathcal{D}(R(t))$$

fails, with the convention $\tau_K = \tau$ if no such failure occurs. These inequalities hold at $t = 0$. On $[0, \tau_K]$, (212) and $dt \leq C\, dR$ give

$$|\delta(t)| \leq |\delta_0| + CK^2 \int_{R_0}^{R(t)} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^2}\, du \leq CK^2 \mathcal{D}(R(t)). \tag{239}$$

After increasing K, this improves the δ bootstrap. We also compare the time integral with the R integral. From the estimate for $\dot{R}$ and the lower bound $\dot{R} \geq r_s/4$,

$$\begin{aligned} \left| \frac{dt}{dR} - \frac{1}{r_s - R} \right| &\leq C \frac{a^2 + Y^2}{R^2}, \\ \left| \int_0^t \frac{ds}{R(s)} - \int_{R_0}^{R(t)} \frac{du}{u(r_s - u)} \right| &\leq C \int_{R_0}^{R(t)} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^3}\, du \\ &\leq C. \end{aligned} \tag{240}$$

Moreover,

$$\left| \int_{R_0}^{R(t)} \frac{du}{u(r_s - u)} - \frac{1}{r_s} \log \frac{R(t)}{R_0} \right| \leq C. \tag{241}$$

Using (240) and (241) in (213) gives

$$\begin{aligned} \left| \log \frac{a(t)}{a_0} - \mu_\alpha \log \frac{R(t)}{R_0} \right| &\leq C + CK^3 \int_{R_0}^{r_s/2} \frac{\mathcal{D}(u)}{u^2}\, du \\ &\quad + CK^2 \int_{R_0}^{r_s/2} \frac{a_0^2 (u/R_0)^{2\mu_\alpha} + Y_0^2 (u/R_0)^{\mu_\alpha}}{u^3}\, du \leq C, \end{aligned} \tag{242}$$

where we used (237) and (238), and where η_* is chosen small after K is fixed. Thus

$$e^{-C} a_0 \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha} \leq a(t) \leq e^{C} a_0 \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha}. \tag{243}$$

Applying (214) gives the following explicit bound. If $Y_0 > 0$, then

$$\begin{aligned} \log \frac{Y(t)}{Y_0} &\leq \int_0^t \left[\frac{\sin(\alpha/2)}{\pi R(s)} + C \left(1 + \frac{|\delta(s)|}{R(s)^2} + \frac{a(s)^2 + Y(s)^2}{R(s)^3} \right) \right] ds \\ &\leq \frac{\mu_\alpha}{2} \log \frac{R(t)}{R_0} + C. \end{aligned} \tag{244}$$

If $Y_0 = 0$, then the linear equations for o_1, o_2 give $Y(t) = 0$ on the same interval. Hence in all cases

$$Y(t) \leq e^{C} Y_0 \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha/2}. \tag{245}$$

Increasing K once more improves the a and Y bootstrap bounds. Also,

$$\begin{aligned} \frac{a(t) + Y(t) + |\delta(t)|}{R(t)} &\leq C\eta_* \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha - 1} + C\eta_* \left(\frac{R(t)}{R_0} \right)^{\mu_\alpha/2 - 1} + C\eta_* \\ &\leq C\eta_* < \eta_0, \end{aligned}$$

so the cone cannot fail before $R = r_s/2$. Hence $\tau = t_*$, and t_* is finite. At $R(t_*) = r_s/2$, (243) and (245) give the two estimates for $a(t_*)$ and $Y(t_*)$.

It remains to estimate $\delta(t_*)$. From (212), (243), and (245),

$$|\delta(t_*)| \leq C\left(|\delta_0| + \int_{R_0}^{r_s/2} \frac{a_0^2(u/R_0)^{2\mu_\alpha} + Y_0^2(u/R_0)^{\mu_\alpha}}{u^2}\, du\right). \tag{246}$$

The integral in (246) that contains Y_0^2 satisfies

$$Y_0^2 R_0^{-\mu_\alpha} \int_{R_0}^{r_s/2} u^{\mu_\alpha - 2}\, du \leq C\frac{Y_0^2}{R_0} \leq C\sigma.$$

The integral in (246) that contains a_0^2 satisfies

$$\begin{aligned} a_0^2 R_0^{-2\mu_\alpha} \int_{R_0}^{r_s/2} u^{2\mu_\alpha - 2}\, du &\leq \begin{cases} Ca_0^2/R_0, & \mu_\alpha < \frac{1}{2}, \\ Ca_0^2 R_0^{-1}\left(1 + \log\frac{1}{R_0}\right), & \mu_\alpha = \frac{1}{2}, \\ Ca_0^2 R_0^{-2\mu_\alpha}, & \mu_\alpha > \frac{1}{2}, \end{cases} \\ &\leq C\left(\sigma + a(t_*)^2 \Lambda_\alpha(R_0)\right). \end{aligned}$$

Together with $|\delta_0| \leq M\sigma$, this proves the claimed bound for $\delta(t_*)$. ■

The initialization bounds in theorem 63 give

$$\frac{a_0 + Y_0 + |\delta_0|}{R_0} \leq \frac{C_{m,M}}{\sqrt{d}} \to 0.$$

Thus, after fixing η_* in lemma 65, the hypothesis $a_0 + Y_0 + |\delta_0| \leq \eta_* R_0$ holds for all $d \geq D_0$, after increasing D_0 if needed, and the preceding lemma applies to the present trajectory.

Choose $c_0 = c_0(\alpha) > 0$ inside the coordinate range of the saddle segment. After c_0 is fixed, choose the local saddle-chart radius. Then choose $0 < \varepsilon_{\text{in}} < \varepsilon$, with ε_{in} also smaller than $c_0/4$, the local saddle-chart radius, and $r_s/\sqrt{2}$. We take ε_{in} small enough depending only on α and ε so that the local saddle chart below applies in the ε_{in}-neighborhood of

$$\mathcal{C}_{c_0} = \{((r_s + c)u_1, (r_s - c)u_1) : |c| \leq c_0\}.$$

Shrink $\varepsilon_0 = \varepsilon_0(\alpha)$, if needed, so that for every $0 < \varepsilon \leq \varepsilon_0$ one can later choose $\eta_+ > \varepsilon$ while keeping the fixed constant multiple of η_+ inside the local saddle-chart smallness range and inside $|\delta| < c_0/2$. Let $\mathcal{U}_{\varepsilon_{\text{in}}}$ denote this neighborhood, and define

$$t_{\text{in}} = \inf\{t \geq 0 : \operatorname{dist}((w_1(t), w_2(t)), \mathcal{C}_{c_0}) < \varepsilon_{\text{in}}\}.$$

For all $t < t_*$, the cone phase has $R(t) < r_s/2$. For any $c \in [-c_0, c_0]$,

$$\begin{aligned} &\|(R(t) + \delta(t))u_1 - (r_s + c)u_1\|_2^2 + \|(R(t) - \delta(t))u_1 - (r_s - c)u_1\|_2^2 \\ &\quad = 2(R(t) - r_s)^2 + 2(\delta(t) - c)^2 \geq \frac{r_s^2}{2}. \end{aligned}$$

The remaining u_2 and out-of-plane coordinates only increase the distance to $\mathcal{C}_{c_0}$. Since $\varepsilon_{\text{in}} < r_s/\sqrt{2}$, this shows $t_{\text{in}} \geq t_*$.

Lemma 66 (From $R = r_s/2$ to $\mathcal{U}_{\varepsilon_{\rm in}}$) *Under the conclusions obtained at t_*, the trajectory reaches $\mathcal{U}_{\varepsilon_{\rm in}}$ after a bounded amount of time. Moreover, the estimates displayed in (255) and the following bounds for $a(t_{\rm in})$, $Y(t_{\rm in})$, and $\delta(t_{\rm in})$ hold with constants of the form $C_{\alpha,m,M,\varepsilon_{\rm in}}$, and these constants are independent of σ and d.*

Proof We now consider the part of the trajectory that starts at t_*, where $R(t_*) = r_s/2$, and ends when the trajectory first enters $\mathcal{U}_{\varepsilon_{\rm in}}$. The preceding estimates show that

$$a(t_*) + Y(t_*) + |\delta(t_*)| \to 0, \qquad R(t_*) = r_s/2.$$

Indeed, the assumptions $a_0 \le M\sigma$ and $m\sigma\sqrt{d} \le R_0 \le M\sigma\sqrt{d}$ give

$$\begin{aligned} a(t_*) &\le C\sigma(\sigma\sqrt{d})^{-\mu_\alpha} \le C(\sigma\sqrt{d})^{1-\mu_\alpha} \to 0, \\ Y(t_*) &\le C\sigma(\sigma\sqrt{d})^{-\mu_\alpha/2} \le C(\sigma\sqrt{d})^{1-\mu_\alpha/2} \to 0. \end{aligned}$$

In both inequalities we used $d \ge 2$, hence $\sigma \le \sigma\sqrt{d}$. The bound for $\delta(t_*)$ also tends to zero. If $\mu_\alpha \ne 1/2$, then

$$a(t_*)^2 \le C\sigma^2(\sigma\sqrt{d})^{-2\mu_\alpha} \le C(\sigma\sqrt{d})^{2-2\mu_\alpha} \to 0.$$

If $\mu_\alpha = 1/2$, then

$$\begin{aligned} a(t_*)^2\Lambda_\alpha(R_0) &\le C\sigma^2 R_0^{-1}\left(1 + \log\frac{1}{R_0}\right) \\ &\le C\frac{\sigma}{\sqrt{d}}\left(1 + \log\frac{1}{\sigma\sqrt{d}}\right) \le C\sigma\sqrt{d}\left(1 + \log\frac{1}{\sigma\sqrt{d}}\right) \to 0. \end{aligned}$$

For this part of the trajectory, we work on the stopped region where

$$\frac{r_s}{4} \le R \le 2r_s, \qquad a + Y + |\delta| \le \min\left\{\frac{r_s}{8}, 1\right\}.$$

Let $\tau_{\rm tr}$ be the first time after t_* when either $R \notin [r_s/4, 2r_s]$ or $a + Y + |\delta| > \min\{r_s/8, 1\}$, with $\tau_{\rm tr} = \infty$ if no such time occurs. On this region, for $i = 1, 2$,

$$\begin{aligned} &\frac{r_s}{8} \le R - |\delta| \le z_{i,1} \le R + |\delta| \le \frac{17r_s}{8}, \\ &\frac{r_s}{8} \le z_{i,1} \le \|w_i\|_2 \le z_{i,1} + |z_{i,2}| + \|o_i\|_2 \le \frac{9r_s}{4}. \end{aligned} \tag{247}$$

Thus every factor $z_{i,1}$, $\|w_i\|_2$, $1/z_{i,1}$, and $1/\|w_i\|_2$ is bounded by a constant depending only on α. The angle estimates then take the following form. For $i = 1, 2$,

$$\begin{aligned} \left|\theta_{i2} - \theta_{i1} - \frac{2z_{i,2}}{z_{i,1}}\right| &\le C_\alpha|z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2), \\ |\theta_{i1} + \theta_{i2} - \alpha| &\le C_\alpha(z_{i,2}^2 + \|o_i\|_2^2), \\ |\sin\theta_{i1} + \sin\theta_{i2} - 2\sin(\alpha/2)| &\le C_\alpha(z_{i,2}^2 + \|o_i\|_2^2), \\ |\|w_i\|_2 - z_{i,1}| &\le C(z_{i,2}^2 + \|o_i\|_2^2), \\ \left|\frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} z_{i,1} - 2\sin(\alpha/2)\right| &\le C_\alpha(z_{i,2}^2 + \|o_i\|_2^2). \end{aligned} \tag{248}$$

The first two lines are (219) and (223) with the powers of R replaced by constants depending only on α. The sine bound follows from the sine expansion used in (224). The norm bound follows from

$$\|w_i\|_2 - z_{i,1} = \frac{z_{i,2}^2 + \|o_i\|_2^2}{\|w_i\|_2 + z_{i,1}},$$

and the final line of (248) follows from the sine bound and the norm bound. All constants remain dimension-free because the bounds in (247) are dimension-free. Also,

$$\begin{aligned}
\left\|\frac{w_i}{\|w_i\|_2} - u_1\right\|_2 &\le \frac{\|w_i\|_2 - z_{i,1}}{\|w_i\|_2} + \frac{|z_{i,2}| + \|o_i\|_2}{\|w_i\|_2} \le C(a+Y), \qquad i = 1,2, \\
\gamma &\le C\left\|\frac{w_1}{\|w_1\|_2} - \frac{w_2}{\|w_2\|_2}\right\|_2 \le C(a+Y), \\
\left|z_{2,1} - \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| &= \left|\frac{z_{2,1}(\|w_1\|_2 - z_{1,1}) - z_{1,1}(\|w_2\|_2 - z_{2,1})}{\|w_1\|_2}\right| \\
&\le C(a^2+Y^2), \\
\left|z_{1,1} - \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| &\le C(a^2+Y^2).
\end{aligned} \tag{249}$$

From the first two lines of (211), for $\{i,j\} = \{1,2\}$, we have

$$\begin{aligned}
\dot z_{i,1} &= \frac{1}{2\pi}\left[(2\pi - \theta_{i1} - \theta_{i2})\cos(\alpha/2) + \frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} z_{i,1} - \pi(z_{i,1} + z_{j,1})\right] \\
&\quad + \frac{1}{2\pi}\left[\gamma z_{j,1} - \sin\gamma \frac{\|w_j\|_2 z_{i,1}}{\|w_i\|_2}\right], \qquad \{i,j\} = \{1,2\}.
\end{aligned} \tag{250}$$

The second square-bracketed expression in (250), $\gamma z_{j,1} - \sin\gamma\|w_j\|_2 z_{i,1}/\|w_i\|_2$, comes from the two terms $-\sin\gamma(\|w_j\|_2/\|w_i\|_2)z_{i,1}$ and $-(\pi-\gamma)z_{j,1}$ in the i-th one of the first two lines of (211), after expanding $-(\pi-\gamma)z_{j,1} = -\pi z_{j,1} + \gamma z_{j,1}$ and keeping $-\pi z_{i,1} - \pi z_{j,1}$ in the first bracket of (250). Writing it this way leaves the factor $z_{j,1} - \|w_j\|_2 z_{i,1}/\|w_i\|_2$ visible, and (249) bounds this factor by $C(a^2+Y^2)$. For $(i,j) = (1,2)$, we therefore get

$$\begin{aligned}
\left|\gamma z_{2,1} - \sin\gamma \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| &\le |\gamma|\left|z_{2,1} - \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2}\right| \\
&\quad + |\gamma - \sin\gamma| \frac{\|w_2\|_2 z_{1,1}}{\|w_1\|_2} \\
&\le C(a+Y)(a^2+Y^2) + C(a+Y)^3 \\
&\le C(a^2+Y^2),
\end{aligned} \tag{251}$$

because $a + Y \le 1$ on the stopped region and $(a+Y)^2 \le 2(a^2+Y^2)$. For $(i,j) = (2,1)$, the same calculation gives

$$\begin{aligned}
\left|\gamma z_{1,1} - \sin\gamma \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| &\le |\gamma|\left|z_{1,1} - \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2}\right| \\
&\quad + |\gamma - \sin\gamma| \frac{\|w_1\|_2 z_{2,1}}{\|w_2\|_2} \\
&\le C(a^2+Y^2),
\end{aligned}$$

using (249).

Substituting (248), (251), and its $(i,j) = (2,1)$ analogue into (250) gives

$$\left|\dot{z}_{i,1} - \frac{1}{2\pi}\left[(2\pi-\alpha)\cos(\alpha/2) + 2\sin(\alpha/2) - \pi(z_{1,1}+z_{2,1})\right]\right| \le C_\alpha(a^2+Y^2), \\ |\dot{z}_{i,1} - (r_s - R)| \le C_\alpha(a^2+Y^2). \tag{252}$$

In particular,

$$|\dot{R} - (r_s - R)| + |\dot{\delta}| \le C(a^2+Y^2).$$

To prove the bounds for $|\dot{a}|$ and $(d/dt)(Y^2+\xi)^{1/2}$, first note that the definition of B_i and (247) give

$$|B_i| \le \frac{|\sin\theta_{i1}| + |\sin\theta_{i2}|}{\|w_i\|_2} + \pi + |\sin\gamma|\frac{\|w_j\|_2}{\|w_i\|_2} \le C, \qquad \{i,j\} = \{1,2\}. \tag{253}$$

Also, (248) gives $|\theta_{12} - \theta_{11}| \le C|z_{1,2}|$ and $|\theta_{22} - \theta_{21}| \le C|z_{2,2}|$. Therefore the exact identity (229) gives

$$|\dot{a}| \le C\left(|z_{1,2}| + |z_{2,2}| + (|B_1|+|B_2|+1)a + |B_1 - B_2|\,|z_{1,2}+z_{2,2}|\right) \\ \le Ca,$$

where we used (253) and $|z_{1,2}+z_{2,2}| \le a$. The equations for o_1, o_2 and (253) give

$$\frac{1}{2}\frac{d}{dt}Y^2 = \frac{1}{2\pi}\left(B_1\|o_1\|_2^2 + B_2\|o_2\|_2^2 - 2(\pi-\gamma)\langle o_1, o_2\rangle\right) \\ \le C(\|o_1\|_2^2 + \|o_2\|_2^2) = CY^2.$$

Therefore, for every $\xi > 0$,

$$\frac{d}{dt}(Y^2+\xi)^{1/2} = \frac{\frac{d}{dt}Y^2}{2(Y^2+\xi)^{1/2}} \le C(Y^2+\xi)^{1/2}.$$

Therefore, whenever $R \ge r_s/2$ and $a + Y + |\delta|$ is small enough for (247) to hold, we have

$$|\dot{R} - (r_s - R)| + |\dot{\delta}| \le C(a^2 + Y^2 + |\delta|^2), \\ |\dot{a}| \le Ca, \\ \frac{d}{dt}(Y^2+\xi)^{1/2} \le C(Y^2+\xi)^{1/2}, \quad \text{for every } \xi > 0. \tag{254}$$

Choose a fixed time length, depending only on α and ε_{in}, such that the scalar solution $\dot{R} = r_s - R$ starting from $R = r_s/2$ is within $\varepsilon_{\text{in}}/4$ of r_s at that time. We first work for times between t_* and the earlier of τ_{tr} and t_* plus this fixed time length. For those times, the last two inequalities in (254) give, for every such stopped time,

$$a(t) \le e^{C(t-t_*)}a(t_*), \\ (Y(t)^2+\xi)^{1/2} \le e^{C(t-t_*)}(Y(t_*)^2+\xi)^{1/2}.$$

Letting $\xi \downarrow 0$ in the second estimate and using that the chosen transfer time length is fixed,

$$\begin{aligned} a(t)+Y(t) &\le C(a(t_*)+Y(t_*)),\\ |\delta(t)| &\le C\left(|\delta(t_*)|+\int_{t_*}^{t}(a(u)^2+Y(u)^2)\,du\right),\\ |R(t)-R_{\mathrm{bis}}(t)| &\le C\int_{t_*}^{t} e^{-(t-u)}(a(u)^2+Y(u)^2+|\delta(u)|^2)\,du, \end{aligned} \tag{255}$$

where R_{bis} denotes the scalar bisector solution, defined by $\dot{R}_{\mathrm{bis}} = r_s - R_{\mathrm{bis}}$ and $R_{\mathrm{bis}}(t_*) = r_s/2$. The bound for δ follows because (254) implies $|\dot{\delta}| \le C|\delta| + C(a^2+Y^2)$ for those stopped times, so

$$\begin{aligned} |\delta(t)| &\le e^{C(t-t_*)}|\delta(t_*)| + C\int_{t_*}^{t} e^{C(t-u)}(a(u)^2+Y(u)^2)\,du\\ &\le C\left(|\delta(t_*)|+\int_{t_*}^{t}(a(u)^2+Y(u)^2)\,du\right). \end{aligned}$$

We use the following elementary form of the variation-of-constants formula: if $Z'(s) = -Z(s)+F(s)$ on $[t_*, t]$, then

$$Z(t) = e^{-(t-t_*)}Z(t_*) + \int_{t_*}^{t} e^{-(t-u)}F(u)\,du. \tag{256}$$

Here the first inequality in (254) and the definition of R_{bis} give

$$\frac{d}{dt}(R-R_{\mathrm{bis}}) = -(R-R_{\mathrm{bis}}) + E_R \tag{257}$$

with $|E_R| \le C(a^2+Y^2+|\delta|^2)$. Applying (256) to (257), with $Z = R - R_{\mathrm{bis}}$ and $F = E_R$, and using $R(t_*) = R_{\mathrm{bis}}(t_*) = r_s/2$, gives

$$|R(t)-R_{\mathrm{bis}}(t)| \le C\int_{t_*}^{t} e^{-(t-u)}\left(a(u)^2+Y(u)^2+|\delta(u)|^2\right)du.$$

The scalar bisector solution is within $\varepsilon_{\mathrm{in}}/4$ of r_s at the chosen transfer time. Since the chosen transfer time is fixed, (255) and the bounds at t_* give

$$\begin{aligned} &a(t)+Y(t)+|\delta(t)|+|R(t)-R_{\mathrm{bis}}(t)|\\ &\le C\left(a(t_*)+Y(t_*)+|\delta(t_*)|+a(t_*)^2+Y(t_*)^2+|\delta(t_*)|^2\right). \end{aligned} \tag{258}$$

The right side of (258) is smaller than $\varepsilon_{\mathrm{in}}/4$ for every stopped time before the chosen transfer time and for all sufficiently small σ, by the estimates in lemma 65. It is also smaller than a fixed fraction of $\min\{r_s/8, 1\}$, and the $R - R_{\mathrm{bis}}$ estimate obtained from (256) keeps R within $r_s/4$ of the scalar bisector solution on the chosen fixed interval. Since the scalar bisector solution started at $r_s/2$ remains in $[r_s/2, r_s]$, neither boundary in the definition of τ_{tr} is reached before the chosen fixed transfer time. Thus the stopped estimates above apply on the whole transfer interval. Therefore the distance to $\mathcal{C}_{c_0}$ is less than $\varepsilon_{\mathrm{in}}$ at the chosen transfer time, so the trajectory enters $\mathcal{U}_{\varepsilon_{\mathrm{in}}}$. Also,

$$0 \le t_* \le \frac{4}{r_s}\left(\frac{r_s}{2} - R_0\right) \le 2$$

for all sufficiently small σ, by $\dot{R} \geq r_s/4$ before t_*. Hence $t_{\text{in}} < \infty$ and $0 \leq t_{\text{in}} - t_* \leq C_{\alpha,m,M,\varepsilon_{\text{in}}}$. Moreover,

$$C^{-1}a(t_*) \leq a(t_{\text{in}}) \leq Ca(t_*), \qquad Y(t_{\text{in}}) \leq CY(t_*). \tag{259}$$

Combining this with (243) at $R(t_*) = r_s/2$, and changing C if needed, gives

$$C^{-1}a_0 R_0^{-\mu_\alpha} \leq a(t_{\text{in}}) \leq Ca_0 R_0^{-\mu_\alpha}. \tag{260}$$

For every time between t_* and the end of the chosen fixed transfer interval, (255) and the estimates in lemma 65 give

$$a(t) + Y(t) + |\delta(t)| \leq C\left(a(t_*) + Y(t_*) + |\delta(t_*)| + a(t_*)^2 + Y(t_*)^2\right) \leq 1$$

for all sufficiently small σ. Thus the term $C|\delta|^2$ in (254) is bounded by $C|\delta|$. Applying the inequality to $(\delta(t)^2 + \xi)^{1/2}$ and then letting $\xi \downarrow 0$ gives

$$\begin{aligned} |\delta(t_{\text{in}})| &\leq C\left(|\delta(t_*)| + \int_{t_*}^{t_{\text{in}}} (a(u)^2 + Y(u)^2)\, du\right) \\ &\leq C\left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0)\right). \end{aligned} \tag{261}$$

Also,

$$\begin{aligned} \frac{Y(t_{\text{in}})^2}{a(t_{\text{in}})} &\leq C\frac{Y_0^2 R_0^{-\mu_\alpha}}{a_0 R_0^{-\mu_\alpha}} \leq C\sigma, \\ a(t_{\text{in}}) \leq Ca_0 R_0^{-\mu_\alpha} \to 0, &\qquad \frac{a(t_{\text{in}})}{\sigma} \geq C^{-1} R_0^{-\mu_\alpha} \to \infty. \end{aligned} \tag{262}$$

Finally,

$$\left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0)\right) \log \frac{1}{a(t_{\text{in}})} \to 0. \tag{263}$$

Indeed, $a_0 \geq m\sigma$ and $R_0 \to 0$ give $a(t_{\text{in}}) \geq c\sigma$ for all sufficiently small σ, and hence

$$\log(1/a(t_{\text{in}})) \leq C \log(1/\sigma).$$

If $\mu_\alpha \neq 1/2$, then

$$a(t_{\text{in}})^2 \leq C\sigma^{2-2\mu_\alpha},$$

and the preceding display follows. If $\mu_\alpha = 1/2$, then

$$a(t_{\text{in}})^2 \Lambda_\alpha(R_0) \leq C\sigma^2 R_0^{-1}\left(1 + \log \frac{1}{R_0}\right) \leq C\sigma\left(1 + \log \frac{1}{\sigma}\right),$$

and the product with $\log(1/a(t_{\text{in}}))$ still tends to zero. ■

It remains to estimate the time after the trajectory enters $\mathcal{U}_{\varepsilon_{\text{in}}}$. The definition of $N_\perp$ and a gives

$$N_\perp(t)^2 = \frac{(z_{1,2}(t) + z_{2,2}(t))^2}{2} + \frac{a(t)^2}{2} + Y(t)^2.$$

The local-passage argument below first checks that the theorem's stopping time is not reached before t_{in}. After that check, T_ε is the first time after t_{in} at which $N_\perp(t) \geq \varepsilon$. The distance used to define the entry time is related to the same coordinates as follows: if $|\delta(t)| < c_0/2$, the closest point on $\mathcal{C}_{c_0}$ has parameter $\delta(t)$, and

$$\operatorname{dist}((w_1(t), w_2(t)), \mathcal{C}_{c_0})^2 = 2(R(t) - r_s)^2 + N_\perp(t)^2. \tag{264}$$

Lemma 67 (Local saddle estimates) *There is a constant $C_{\alpha,\varepsilon}$ such that the estimates used in the local chart are*

$$|\dot{a} - \mu_\alpha a| \le C_{\alpha,\varepsilon}\left((|R - r_s| + |\delta| + a + Y)a + Y^2\right), \tag{265}$$

$$\frac{d}{dt}Y^2 \le (\mu_\alpha + C_{\alpha,\varepsilon}(|R - r_s| + |\delta| + a + Y))\, Y^2, \tag{266}$$

$$|\dot{\delta}| \le C_{\alpha,\varepsilon}\left(|\delta|(|R - r_s| + a + Y) + a^2 + Y^2\right), \tag{267}$$

and

$$\dot{R} - (r_s - R) = E_R, \qquad |E_R| \le C_{\alpha,\varepsilon}\left(a^2 + Y^2 + |\delta|(a + Y)\right). \tag{268}$$

Proof We now derive (265)–(268). The estimates are used only when $|R - r_s| + |\delta| + a + Y \le r_s/8$. Hence

$$\begin{aligned} &\frac{3}{4}r_s \le r_s - |R - r_s| - |\delta| \le z_{i,1} \le r_s + |R - r_s| + |\delta| \le \frac{5}{4}r_s, \\ &\frac{3}{4}r_s \le z_{i,1} \le \|w_i\|_2 \le z_{i,1} + |z_{i,2}| + \|o_i\|_2 \le \frac{11}{8}r_s. \end{aligned} \tag{269}$$

In particular, every occurrence of $z_{i,1}$, $\|w_i\|_2$, $1/z_{i,1}$, and $1/\|w_i\|_2$ is bounded by a constant depending only on α. Hence the angle calculations (219), (223), and (224) become

$$\begin{aligned} \left|\theta_{i2} - \theta_{i1} - \frac{2z_{i,2}}{z_{i,1}}\right| &\le C_\alpha |z_{i,2}|(z_{i,2}^2 + \|o_i\|_2^2), \\ |\theta_{i1} + \theta_{i2} - \alpha| &\le C_\alpha(z_{i,2}^2 + \|o_i\|_2^2), \\ \left|\frac{\sin\theta_{i1} + \sin\theta_{i2}}{\|w_i\|_2} z_{i,1} - 2\sin(\alpha/2)\right| &\le C_\alpha(z_{i,2}^2 + \|o_i\|_2^2). \end{aligned} \tag{270}$$

These estimates are dimension-free because o_i enters only through $\|o_i\|_2^2$.

We first prove (267) and (268). Using (250), the last two lines of (270), and the bound (251) for $\gamma z_{2,1} - \sin\gamma \|w_2\|_2 z_{1,1}/\|w_1\|_2$, we get

$$\begin{aligned} \left|\dot{z}_{i,1} - \frac{1}{2\pi}\left[(2\pi - \alpha)\cos(\alpha/2) + 2\sin(\alpha/2) - \pi(z_{1,1} + z_{2,1})\right]\right| &\le C_\alpha(a^2 + Y^2), \\ |\dot{z}_{i,1} - (r_s - R)| &\le C_\alpha(a^2 + Y^2), \qquad i = 1, 2. \end{aligned} \tag{271}$$

Therefore

$$\begin{aligned} \left|\dot{R} - (r_s - R)\right| &= \left|\frac{\dot{z}_{1,1} + \dot{z}_{2,1}}{2} - (r_s - R)\right| \le C_\alpha(a^2 + Y^2), \\ |\dot{\delta}| &= \left|\frac{\dot{z}_{1,1} - \dot{z}_{2,1}}{2}\right| \le C_\alpha(a^2 + Y^2). \end{aligned}$$

These two inequalities imply (267) and (268).

For $a = z_{1,2} - z_{2,2}$, we use the exact identity (229). The first line of (270) gives

$$\left| \big((\theta_{12} - \theta_{11}) - (\theta_{22} - \theta_{21}) \big) \sin(\alpha/2) - 2 \sin(\alpha/2) \left(\frac{z_{1,2}}{z_{1,1}} - \frac{z_{2,2}}{z_{2,1}} \right) \right| \le C_\alpha \left(a^3 + aY^2 \right).$$

The fraction in the leading term satisfies

$$\frac{z_{1,2}}{z_{1,1}} - \frac{z_{2,2}}{z_{2,1}} = \frac{R(z_{1,2} - z_{2,2}) - \delta(z_{1,2} + z_{2,2})}{R^2 - \delta^2},$$
$$\left| \frac{z_{1,2}}{z_{1,1}} - \frac{z_{2,2}}{z_{2,1}} - \frac{a}{r_s} \right| \le C(|R - r_s| + |\delta|)a.$$

Since $a + Y \le r_s/8$ in the local chart, after decreasing the chart radius if needed we have $a \le 1$ and $Y \le 1$. Therefore

$$a^3 + aY^2 \le a^2 + Y^2.$$

Thus

$$\begin{aligned} &\left| \big((\theta_{12} - \theta_{11}) - (\theta_{22} - \theta_{21}) \big) \sin(\alpha/2) - \frac{2\sin(\alpha/2)}{r_s} a \right| \\ &\quad \le C_\alpha \left((|R - r_s| + |\delta| + a + Y)a + Y^2 \right). \end{aligned} \tag{272}$$

Next, from the definition of B_i and (270),

$$\left| B_i - \left(\frac{2\sin(\alpha/2)}{z_{i,1}} - \pi - \sin\gamma \frac{\|w_j\|_2}{\|w_i\|_2} \right) \right| \le C_\alpha (z_{i,2}^2 + \|o_i\|_2^2), \qquad \{i, j\} = \{1, 2\}.$$

Using $|1/z_{1,1} + 1/z_{2,1} - 2/r_s| \le C(|R - r_s| + |\delta|)$, $|1/z_{1,1} - 1/z_{2,1}| \le C|\delta|$, $|\gamma| \le C(a + Y)$, and $\|w_1\|_2/\|w_2\|_2 + \|w_2\|_2/\|w_1\|_2 \le C$, we obtain

$$\begin{aligned} \left| \frac{B_1 + B_2}{2} + \pi - \gamma - \frac{2\sin(\alpha/2)}{r_s} \right| &\le C_\alpha (|R - r_s| + |\delta| + a + Y), \\ |B_1 - B_2| &\le C_\alpha (|\delta| + a + Y). \end{aligned} \tag{273}$$

Substituting (272) and (273) into (229), and using $|z_{1,2} + z_{2,2}| \le a$, gives

$$\begin{aligned} \left| \dot{a} - \frac{1}{2\pi} \left[\frac{2\sin(\alpha/2)}{r_s} a + \frac{2\sin(\alpha/2)}{r_s} a \right] \right| &\le C_\alpha ((|R - r_s| + |\delta| + a + Y)a + Y^2), \\ |\dot{a} - \mu_\alpha a| &\le C_\alpha ((|R - r_s| + |\delta| + a + Y)a + Y^2), \end{aligned} \tag{274}$$

which proves (265).

For $Y^2 = \|o_1\|_2^2 + \|o_2\|_2^2$, summing the last two lines of (211) over coordinates $3, \dots, d$ gives

$$\frac{1}{2} \frac{d}{dt} Y^2 = \frac{1}{2\pi} \left(B_1 \|o_1\|_2^2 + B_2 \|o_2\|_2^2 - 2(\pi - \gamma) \langle o_1, o_2 \rangle \right).$$

The estimates used to prove (273) also give

$$\left| B_i - \left(\frac{2\sin(\alpha/2)}{r_s} - \pi \right) \right| \le C_\alpha (|R - r_s| + |\delta| + a + Y), \qquad i = 1, 2.$$

Since $-2\langle o_1, o_2\rangle \le \|o_1\|_2^2 + \|o_2\|_2^2$ and $|\gamma| \le C(a+Y)$,

$$\frac{1}{2}\frac{d}{dt}Y^2 \le \frac{1}{2\pi}\left[\left(\frac{2\sin(\alpha/2)}{r_s} - \pi + C_\alpha(|R-r_s|+|\delta|+a+Y)\right)Y^2 + (\pi + C_\alpha(a+Y))\,Y^2\right]$$
$$= \frac{\mu_\alpha}{2}Y^2 + C_\alpha(|R-r_s|+|\delta|+a+Y)Y^2.$$

Multiplying by 2 proves (266). ■

Lemma 68 (Local saddle passage) *There is a constant $C_{\alpha,m,M,\varepsilon_{\text{in}},\varepsilon}$ such that, for all sufficiently small σ,*

$$\left|T_\varepsilon - t_{\text{in}} - \frac{1}{\mu_\alpha}\log\frac{1}{a(t_{\text{in}})}\right| \le C_{\alpha,m,M,\varepsilon_{\text{in}},\varepsilon}.$$

Proof First note that the stopping time in the theorem is not reached before t_{in}. Since $z_{1,2}(t) > 0 > z_{2,2}(t)$ on the times considered, $N_\perp(t) \le a(t) + Y(t)$. On $[0, t_*]$, (243) and (245) give $\sup_{0\le t\le t_*} N_\perp(t) \to 0$. By lemma 66, $t_{\text{in}} \ge t_*$. On $[t_*, t_{\text{in}}]$, (255) and the estimates at t_* give $\sup_{t_*\le t\le t_{\text{in}}} N_\perp(t) \to 0$. Hence, for all sufficiently small σ, $N_\perp(t) < \varepsilon$ for every $0 \le t \le t_{\text{in}}$, and the first time in the theorem may be searched for over $t \ge t_{\text{in}}$.

Choose fixed numbers $0 < \eta_- < \eta_+$ so that $4\eta_- < \varepsilon < \eta_+$. We choose η_+ only after c_0, the local saddle-chart radius, and ε_{in} have been fixed, and we make it small enough that a fixed constant multiple of η_+ keeps $|R-r_s|+|\delta|+a+Y$ inside the smallness range required in lemma 67 and keeps $|\delta| < c_0/2$. For $\eta \in [\eta_-, \eta_+]$, set

$$H_\eta = \inf\left\{t \ge t_{\text{in}} : \frac{a(t)}{\sqrt{2}} = \eta\right\}.$$

We prove that, uniformly for $\eta \in [\eta_-, \eta_+]$,

$$\left|H_\eta - t_{\text{in}} - \frac{1}{\mu_\alpha}\log\frac{1}{a(t_{\text{in}})}\right| \le C_{\alpha,m,M,\varepsilon_{\text{in}},\varepsilon}.$$

For fixed η, let τ_η be the first time when $a/\sqrt{2} = \eta$, or $|R-r_s|+|\delta|+a+Y$ leaves the smallness range in lemma 67, or $|\delta| = c_0/2$, or one of the bootstrap bounds (275)–(277) fails. First choose a large constant K_R. Given K_R, choose K_δ larger than the Gronwall constant $C_{\alpha,\varepsilon,K_R}$. Given K_R and K_δ, choose K_Y large enough. On $[t_{\text{in}}, \tau_\eta]$ bootstrap

$$Y(t)^2 \le K_Y\sigma a(t), \tag{275}$$

$$|\delta(t)| \le K_\delta\left(\sigma + a(t_{\text{in}})^2\Lambda_\alpha(R_0) + a(t)^2\right), \tag{276}$$

and

$$|R(t)-r_s| \le \eta_+, \qquad \int_{t_{\text{in}}}^t |R(u)-r_s|\,du \le K_R. \tag{277}$$

These inequalities hold initially after choosing ε_{in} small and then taking σ small. Under (275)–(277), the local estimate (265) gives

$$\begin{aligned}\frac{\dot a}{a} &\ge \mu_\alpha - C\left(|R-r_s|+|\delta|+a+Y\right) - C\frac{Y^2}{a}\\ &\ge \mu_\alpha - C\left(\eta_+ + K_\delta\left(\sigma + a(t_{\text{in}})^2\Lambda_\alpha(R_0) + a^2\right) + a + Y + K_Y\sigma\right)\\ &\ge \frac{\mu_\alpha}{2}\end{aligned} \tag{278}$$

after decreasing η_+ and taking σ small. Hence

$$\dot{a} \geq \frac{\mu_\alpha}{2} a \tag{279}$$

and a is increasing on the stopped interval. Therefore

$$\begin{aligned} t - t_{\text{in}} &\leq \frac{2}{\mu_\alpha} \log \frac{a(t)}{a(t_{\text{in}})} \leq C \log \frac{1}{a(t_{\text{in}})}, \\ \int_{t_{\text{in}}}^{t} a(u)\, du &\leq \frac{2}{\mu_\alpha} \int_{a(t_{\text{in}})}^{a(t)} dv \leq C a(t) \leq C, \\ \int_{t_{\text{in}}}^{t} a(u)^2\, du &\leq \frac{2}{\mu_\alpha} \int_{a(t_{\text{in}})}^{a(t)} v\, dv \leq C a(t)^2, \\ \int_{t_{\text{in}}}^{t} Y(u)^2\, du &\leq K_Y \sigma \int_{t_{\text{in}}}^{t} a(u)\, du \leq C K_Y \sigma a(t), \\ \int_{t_{\text{in}}}^{t} Y(u)\, du &\leq C\sqrt{K_Y \sigma} \int_{a(t_{\text{in}})}^{a(t)} v^{-1/2}\, dv \leq C\sqrt{\sigma}. \end{aligned} \tag{280}$$

Also

$$Y(t) \leq \sqrt{K_Y \sigma a(t)} = a(t)\sqrt{\frac{K_Y \sigma}{a(t)}} \leq a(t)\sqrt{\frac{K_Y \sigma}{a(t_{\text{in}})}} \leq a(t), \tag{281}$$

for all sufficiently small σ, where the final inequality uses the bound $a(t_{\text{in}})/\sigma \to \infty$ in (262). Using (267), (277), and (280) gives

$$\begin{aligned} |\delta(t)| &\leq \exp\left(C \int_{t_{\text{in}}}^{t} (|R(u) - r_s| + a(u) + Y(u))\, du \right) \cdot \left[|\delta(t_{\text{in}})| + C \int_{t_{\text{in}}}^{t} (a(u)^2 + Y(u)^2)\, du \right] \\ &\leq C_{\alpha,\varepsilon,K_R} \left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0) + \int_{t_{\text{in}}}^{t} (a(u)^2 + Y(u)^2)\, du \right) \\ &\leq C_{\alpha,\varepsilon,K_R} \left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0) + a(t)^2 \right). \end{aligned} \tag{282}$$

Choosing K_δ larger than the constant $C_{\alpha,\varepsilon,K_R}$ in (282) makes (282) stronger than the bootstrap bound (276). Next, (268), (282), and (281) give

$$\begin{aligned} |R(t) - r_s| &\leq e^{-(t-t_{\text{in}})/2} |R(t_{\text{in}}) - r_s| \\ &\quad + C \left(\int_{t_{\text{in}}}^{t} e^{-(t-u)/2} \left[a(u)^2 + Y(u)^2 + |\delta(u)|(a(u) + Y(u)) \right]\, du \right) \\ &\leq e^{-(t-t_{\text{in}})/2} |R(t_{\text{in}}) - r_s| + C \left(a(t)^2 + \sigma a(t) + \left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0) \right) a(t) \right). \end{aligned} \tag{283}$$

The final inequality in (283) uses

$$\begin{aligned} Y(u)^2 &\leq K_Y \sigma a(u), \\ |\delta(u)|(a(u) + Y(u)) &\leq C \left(\sigma + a(t_{\text{in}})^2 \Lambda_\alpha(R_0) + a(u)^2 \right) a(u), \end{aligned}$$

which follows from (282) and (281). Since (279) gives $\int_{t_{\text{in}}}^{t} e^{-(t-u)/2} a(u)^2\, du \leq C a(t)^2$ and $\int_{t_{\text{in}}}^{t} e^{-(t-u)/2} a(u)\, du \leq C a(t)$, the final bound in (283) follows. After choosing ε_{in} and η_+ small, and then taking σ small, this improves the pointwise and integral bounds for $R - r_s$.

It remains to prove that the bootstrap bound (275), namely $Y(t)^2 \le K_Y \sigma a(t)$, also improves. For $\xi > 0$, (266) and (265) give

$$\begin{aligned}
\frac{d}{dt}\log(Y(t)^2+\xi) &= \frac{\frac{d}{dt}Y(t)^2}{Y(t)^2+\xi} \\
&\le \left(\mu_\alpha + C(|R(t)-r_s| + |\delta(t)| + a(t) + Y(t))\right)\frac{Y(t)^2}{Y(t)^2+\xi} \\
&\le \mu_\alpha + C(|R(t)-r_s| + |\delta(t)| + a(t) + Y(t)), \qquad (284)\\
\frac{d}{dt}\log a(t) &\ge \mu_\alpha - C(|R(t)-r_s| + |\delta(t)| + a(t) + Y(t)) - C\frac{Y(t)^2}{a(t)},\\
\frac{d}{dt}\log\frac{Y(t)^2+\xi}{a(t)} &\le C\left(|R(t)-r_s| + |\delta(t)| + a(t) + Y(t)\right) + C\frac{Y(t)^2}{a(t)}.
\end{aligned}$$

The integral of $C(|R-r_s| + |\delta| + a + Y)$ in (284) is bounded because

$$\begin{aligned}
\int_{t_{\text{in}}}^{t} |R(u)-r_s|\,du &\le K_R,\\
\int_{t_{\text{in}}}^{t} a(u)\,du &\le C,\\
\int_{t_{\text{in}}}^{t} Y(u)\,du &\le C\sqrt{\sigma},\\
\int_{t_{\text{in}}}^{t} |\delta(u)|\,du &\le K_\delta\left(\sigma + a(t_{\text{in}})^2\Lambda_\alpha(R_0)\right)(t-t_{\text{in}}) + K_\delta\int_{t_{\text{in}}}^{t} a(u)^2\,du\\
&\le C + CK_\delta\left(\sigma + a(t_{\text{in}})^2\Lambda_\alpha(R_0)\right)\log\frac{1}{a(t_{\text{in}})} \le C,
\end{aligned}$$

where (263) gives the final bound. The integral of $C\,Y(t)^2/a(t)$ in (284) satisfies

$$\int_{t_{\text{in}}}^{t}\frac{Y(u)^2}{a(u)}\,du \le K_Y\sigma(t-t_{\text{in}}) \le CK_Y\sigma\log\frac{1}{a(t_{\text{in}})} \to 0. \qquad (285)$$

Integrating (284), using (285), and letting $\xi \downarrow 0$ gives

$$\frac{Y(t)^2}{a(t)} \le C\frac{Y(t_{\text{in}})^2}{a(t_{\text{in}})} \le C\sigma, \qquad (286)$$

and the choice of K_Y closes the final bootstrap. The estimates above also exclude the non-bootstrap stopping events. The bound (282) and the choices of ε_{in} and η_+ keep $|\delta(t)| < c_0/2$. The improved pointwise bound for $R - r_s$, the inequality $Y(t) \le a(t)$ from (281), and $a(t) \le \sqrt{2}\eta_+$ before H_η keep $|R-r_s| + |\delta| + a + Y$ inside the local smallness range. Finally, (278) gives $\dot{a} \ge (\mu_\alpha/2)a$, so the level $a/\sqrt{2} = \eta$ is reached in finite time if no earlier stopping event occurs. Therefore the first stopping event is H_η, and $\tau_\eta = H_\eta$.

On $[t_{\text{in}}, H_{\eta_-})$, the definition of $N_\perp$ and the closed estimates give

$$\begin{aligned}
N_\perp(t)^2 &= \frac{(z_{1,2}(t)+z_{2,2}(t))^2}{2} + \frac{a(t)^2}{2} + Y(t)^2\\
&\le 2\eta_-^2 + C\sigma\eta_+ < \varepsilon^2,
\end{aligned}$$

after choosing η_- small and then taking σ small. Indeed, $a(t) < \sqrt{2}\eta_-$ before H_{η_-}, so $(z_{1,2}(t) + z_{2,2}(t))^2/2 + a(t)^2/2 \le 2\eta_-^2$. Also (286) gives $Y(t)^2 \le C\sigma a(t) \le C\sigma\eta_+$. Thus

$$N_\perp(t) < \varepsilon$$

for all $t \in [t_{\text{in}}, H_{\eta_-})$, and hence $H_{\eta_-} \le T_\varepsilon$. At H_{η_+},

$$N_\perp(H_{\eta_+}) \ge \frac{a(H_{\eta_+})}{\sqrt{2}} = \eta_+ > \varepsilon,$$

so $T_\varepsilon \le H_{\eta_+}$. Therefore

$$H_{\eta_-} \le T_\varepsilon \le H_{\eta_+}.$$

With the bootstrap closed, (265) and (286) give

$$\left|\frac{d}{dt}\log a(t) - \mu_\alpha\right| \le C\left(|R(t) - r_s| + |\delta(t)| + a(t) + Y(t)\right) + C\frac{Y(t)^2}{a(t)}. \tag{287}$$

The right side of (287) has uniformly bounded integral on $[t_{\text{in}}, H_\eta]$ by (277), (280), (282), and (285). Since $a(H_\eta) = \sqrt{2}\eta$, and η stays in the fixed interval $[\eta_-, \eta_+]$,

$$\begin{aligned} \left|H_\eta - t_{\text{in}} - \frac{1}{\mu_\alpha}\log\frac{\sqrt{2}\eta}{a(t_{\text{in}})}\right| &\le C, \\ \left|H_\eta - t_{\text{in}} - \frac{1}{\mu_\alpha}\log\frac{1}{a(t_{\text{in}})}\right| &\le C. \end{aligned} \tag{288}$$

The sandwich $H_{\eta_-} \le T_\varepsilon \le H_{\eta_+}$ and (288) give

$$\left|T_\varepsilon - t_{\text{in}} - \frac{1}{\mu_\alpha}\log\frac{1}{a(t_{\text{in}})}\right| \le C. \tag{289}$$

■

Before t_*, the cone estimate gives $\dot{R} \ge r_s/4$, hence $0 \le t_* \le 2$ for all sufficiently small σ. Also, lemma 66 gives $0 \le t_{\text{in}} - t_* \le C_{\alpha,m,M,\varepsilon_{\text{in}}}$. Therefore $0 \le t_{\text{in}} \le C_{\alpha,m,M,\varepsilon_{\text{in}}}$. Also, (260) gives, after increasing $C_{\alpha,m,M,\varepsilon_{\text{in}}}$ if needed,

$$C^{-1}_{\alpha,m,M,\varepsilon_{\text{in}}} a_0 R_0^{-\mu_\alpha} \le a(t_{\text{in}}) \le C_{\alpha,m,M,\varepsilon_{\text{in}}} a_0 R_0^{-\mu_\alpha},$$

and combining this with lemma 68 gives, since ε_{in} was fixed in terms of α and ε,

$$\left|T_\varepsilon - \frac{1}{\mu_\alpha}\log\frac{1}{a_0 R_0^{-\mu_\alpha}}\right| \le C_{\alpha,m,M,\varepsilon}. \tag{290}$$

Finally, from $m\sigma \le a_0 \le M\sigma$ and $m\sigma\sqrt{d} \le R_0 \le M\sigma\sqrt{d}$,

$$\left|\log\frac{1}{a_0 R_0^{-\mu_\alpha}} - \left(\log\frac{1}{\sigma} + \mu_\alpha \log(\sigma\sqrt{d})\right)\right| \le C_{\alpha,m,M},$$

$$\left|\log\frac{1}{a_0 R_0^{-\mu_\alpha}} - \left((1-\mu_\alpha)\log\frac{1}{\sigma} + \frac{\mu_\alpha}{2}\log d\right)\right| \le C_{\alpha,m,M}.$$

Dividing by μ_α proves

$$\left| T_\varepsilon - \left[\left(\frac{1}{\mu_\alpha} - 1 \right) \log \frac{1}{\sigma} + \frac{1}{2} \log d \right] \right| \le C_{\alpha,m,M,\varepsilon}.$$

This is the asserted escape-time estimate. ■