Steering the Evolutionary Game: Hierarchical Control of Therapeutic Resistance in Cancer Treatment

Arvid E. Gollwitzer[*]
Broad Institute of MIT and Harvard
Cambridge, MA, USA
arvidg@mit.edu

Deepak A. Subramanian
Dept. of Chemical Engineering, MIT
Koch Institute for Integrative Cancer Research, MIT
Broad Institute of MIT and Harvard
Cambridge, MA, USA

Isaac Tucker
Broad Institute of MIT and Harvard
Cambridge, MA, USA

Giovanni Traverso[*]
Dept. of Mechanical Engineering, MIT, Cambridge, MA, USA
Div. of Gastroenterology, Hepatology and Endoscopy,
Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA
Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA, USA
Broad Institute of MIT and Harvard, Cambridge, MA, USA
cgt20@mit.edu

Abstract

Therapeutic failure in cancer often arises from Darwinian selection for drug resistance. We introduce a hierarchical ecological control framework that reshapes the evolutionary fitness landscape on a slow timescale (microbiome modulation) to render the fast tumor–immune–drug dynamics curative. We prove existence and uniqueness of a Markov-perfect Nash equilibrium for the LQ fast game and formalize a conservative robust immune-influx threshold $s^{rob}_{crit,\max} \le \frac{d}{n} a$. All nonlinear results are stated relative to $s^{rob}_{crit,\max}$. Maintaining $s(\mathbf{M}) \ge s^{rob}_{crit,\max} + \delta$ renders the tumor-free state globally exponentially stable with rate $\lambda = \min\{\frac{n\delta}{2d}, \frac{d-g_{\max}}{2}\}$ and explicit gain bounds. Robustness is established via: (i) spatial sufficiency using domain eigenvalues precluding sanctuaries, (ii) global asymptotic stability in probability under stochastic perturbations when $L_z < \sqrt{\lambda/K}$ (and almost-sure convergence under additional recurrence conditions), (iii) quantitative tolerance to clonal heterogeneity when $\Delta_a + \Delta_\kappa u_{d,\max} < \frac{n}{d}\,\delta$, (iv) delay and observer robustness under precise small-gain and observability conditions, and (v) an ϵ-accurate Fenichel decomposition of the two-timescale game.

Using TCGA-derived priors, AI-synthesized policies enforce the stability margin along trajectories and achieve high efficacy with lower cytotoxic exposure. In Skin Cutaneous Melanoma (SKCM) the controller achieves **89%** eradication (95% CI ± 6), maintains time-above-threshold at **92%** (± 5), and reduces peak tumor burden by **57%**; first response occurs in 28–35 days with a lower dose index. In colorectal cancer (CRC) the controller achieves **76%** eradication (± 7) with **88%**

[*]Corresponding authors: arvidg@mit.edu, cgt20@mit.edu

39th Conference on Neural Information Processing Systems (NeurIPS 2025): The Reach and Limits of AI for Scientific Discovery.

time-above-threshold (± 6). Across structured perturbations (5-clone heterogeneity, 14-day delay, stochastic noise, partial observability), individual eradication rates exceed 80% and remain 73% under combined stressors, aligning with the theory. These results establish ecological landscape engineering as a principled, general strategy for mitigating resistance in oncology.

1 Introduction

Therapeutic failure in cancer is driven by Darwinian selection for drug resistance: clinical and molecular studies document rapid emergence of resistant clones under treatment across indications and modalities [1–6]. Approaches that model tumor–drug interactions in isolation neglect the ecological constraints and frequency-dependent payoffs that shape evolutionary trajectories [7–10]. Moreover, a standard optimal-control framing that treats evolution as an exogenous disturbance cannot capture the adversarial, goal-directed nature of selection; a game-theoretic formulation is essential [11–13].

Concurrently, the host microbiome substantially modulates anti-tumor immunity and immunotherapy efficacy: gut composition stratifies response to PD-1 blockade and fecal microbiota transplantation can overcome resistance [14–18]. Ecological mechanisms—network structure, competition, and keystone species—govern these functional effects [19–21]. Critically, microbiome adaptation occurs over weeks, whereas tumor–immune–drug dynamics evolve over days, yielding a singularly perturbed, multi-timescale structure amenable to hierarchical control [22–26].

We therefore cast treatment as a hierarchical Stackelberg differential game [27, 11]: slow microbiome modulation reshapes the evolutionary fitness landscape that governs fast tumor–immune–drug dynamics. The central theoretical insight is that the slow controller need not track fast states; it must maintain an ecological stability margin beyond a bifurcation threshold so that the tumor-free equilibrium is the unique, globally stable state. This is established by analyzing the nonlinear fast subsystem via bifurcation theory [28–30] and applying geometric singular perturbation theory [31] to decompose the two-timescale game into nested optimal-control subproblems solvable via Hamilton–Jacobi–Bellman methods [32, 33]. Since states represent populations, positivity and monotonicity structure further constrain dynamics [34, 35].

Our LQ core proves: (i) existence and uniqueness of a Markov-perfect Nash equilibrium with linear strategies for the fast game (Theorem 2); (ii) global Lyapunov/ISS stability of the tumor-free equilibrium above a conservative, Nash-consistent immune-influx threshold (Theorem 6); (iii) an asymptotic Fenichel decomposition under explicit hypotheses (Theorem 7); and (iv) robustness propositions, including absence of subcritical behavior and quantitative tolerance to clonal heterogeneity (Theorems 12 and 8).

Computationally, controllers synthesized from simplified models eradicate tumors despite structural uncertainty, stochastic perturbations, and partial observability, providing a rigorous foundation for ecological landscape engineering in oncology (see also [26]).

The remainder of this paper is organized as follows. Section II presents the multi-scale cancer–microbiome model. Section III develops the hierarchical decomposition and main theorems on stability and control. Section IV validates the framework on SKCM and CRC cohorts. Section V concludes with perspectives on future work.

2 Hierarchical eco-evolutionary game formulation

We formulate the therapeutic problem as a hierarchical system characterized by two distinct timescales, as motivated by the underlying biological processes. The fast timescale (hours to days) governs the interactions among tumor cells, immune cells, and cytotoxic agents. The slow timescale (weeks to months) characterizes the dynamics of the host microbiome composition in response to therapeutic interventions [17]. Throughout this paper, we work under the standing assumptions detailed in Appendix C.

2.1 Preliminary definitions

We briefly define key game-theoretic and control-theoretic concepts used throughout; full mathematical details are in Appendix E.

Differential game. A differential game models strategic interaction among multiple decision-makers (players) whose actions jointly influence a continuous-time dynamical system. Each player selects a control strategy to optimize their own objective function, while the system state evolves according to coupled ordinary differential equations [11].

Nash equilibrium. A strategy profile $(u_1^*, \ldots, u_N^*)$ is a *Nash equilibrium* if no player can reduce their cost by unilaterally deviating: for each player i, $J_i(u_i^*, u_{-i}^*) \leq J_i(u_i, u_{-i}^*)$ for all admissible u_i, where u_{-i}^* denotes the equilibrium strategies of all players except i. In cancer treatment, the "players" are the clinician (choosing drug dosage to minimize tumor burden and toxicity) and the tumor population (evolving resistance phenotypes through Darwinian selection, modeled as cost minimization).

Markov-perfect equilibrium. A Nash equilibrium is *Markov-perfect* if each player's strategy depends only on the current state (not on history), yielding time-consistent, state-feedback policies. This is the natural solution concept for feedback control [11, 36].

Linear-Quadratic (LQ) game. An LQ differential game is one where the dynamics are linear in states and controls, and each player's cost functional is quadratic. For such games, equilibrium strategies are linear functions of the state, and are characterized by coupled algebraic Riccati equations [11, 36]. Our "LQ fast game" refers to a local linearization of the tumor–immune dynamics around a reference point, enabling tractable equilibrium computation while the nonlinear analysis uses this as a conservative surrogate (see Appendix F.6).

Stackelberg (hierarchical) game. In a Stackelberg game, a "leader" commits to a strategy first, and "followers" respond optimally. Here, the slow microbiome controller acts as leader, shaping the fitness landscape; the fast tumor–drug game then reaches equilibrium given this landscape [27, 11].

2.2 Standing assumptions (summary)

For convenience, we summarize the key assumptions used throughout and proved/justified in the Appendix:

- **Regularity and positivity**: Vector fields are locally Lipschitz with linear growth, controls are bounded, and the state remains in the positive orthant (Appendix 1–2). Verification of these conditions is provided in Appendix I.1.
- **Timescale separation**: $0 < \epsilon \ll 1$ within an explicit admissible range (Appendix 3).
- **Fast-game solvability**: Surrogate LQ linearization around $(\bar{C}_T > 0, s(M)/d)$ with stabilizing Riccati solutions; input channels vanish at $\mathbf{z}_{free}$ and are interpreted conservatively (Theorem 2; see Appendix F.6).
- **Invariance and bounds**: Explicit invariant bounds for (C_T, C_I) and a uniform condition $d > g_{\max}$ (Lemma 7; see Appendix H for supporting lemmas and proofs).
- **Diffusion regularity (stochastic case)**: Local Lipschitz and growth conditions, with multiplicative noise vanishing at $\mathbf{z}_{free}$ (Assumption 6). Full stochastic system details are in Appendix F.14.

First, the existence and uniqueness of a Nash equilibrium for the fast-timescale game is established (Theorem 2). Then, we prove the local exponential stability of this equilibrium, a key prerequisite for singular perturbation analysis (Theorem 3). With these foundational results in place, we apply Fenichel theory to formally decompose the full system into its slow and fast components, providing an explicit characterization of the slow manifold (Theorem 7). Finally, we perform a comprehensive bifurcation and stability analysis of the complete system, and identify the critical thresholds that govern therapeutic success or failure (Theorems 5 and 6).

The tumor's resistance level, $p(t)$, is modeled as a continuous variable representing the population-averaged resistance. This powerful abstraction, which is standard in evolutionary game theory, enables a tractable optimal control formulation. A detailed justification for this modeling choice is

provided in Appendix 5. We will later demonstrate (Theorems 8 and 9) that controllers synthesized using this continuous model are robustly effective for more realistic, discrete clonal systems.

2.3 Fast-timescale dynamics: tumor-immune interactions

The fast subsystem state vector is defined as $\mathbf{z} = [C_T, C_I]^T$, where C_T and C_I denote the population densities of tumor cells and tumor-infiltrating effector immune cells, respectively. The dynamics are governed by a modified Kuznetsov model [30], extended to incorporate therapeutic interventions and resistance mechanisms:

$$\begin{aligned} \dot{C_T} &= \underbrace{aC_T(1-bC_T)}_{\text{logistic growth}} - \underbrace{nC_TC_I}_{\text{immune predation}} - \underbrace{\kappa(p)u_dC_T}_{\text{drug-induced death}} \\ \dot{C_I} &= \underbrace{s(M)}_{\text{immune recruitment}} - \underbrace{dC_I}_{\text{natural death}} + \underbrace{\frac{rC_TC_I}{h+C_T}}_{\text{stimulation}} - \underbrace{mC_TC_I}_{\text{inactivation}} \end{aligned} \tag{1}$$

The parameters a, b, n, d, r, h, m represent fixed biological constants (see Appendix J.2 for values and sources). The system dynamics are influenced by two control inputs: the drug dosage $u_d(t)$ administered by the clinician, and the resistance phenotype $p(t)$ expressed by the tumor population.

For controller synthesis, we model resistance as a continuous variable $p(t)$ representing the mean phenotypic state of the tumor population. This abstraction reflects the non-genetic adaptation mechanisms (including epigenetic and metabolic plasticity), and enables tractable formulation of the optimal control problem. We later validate the validity of this approach through comparison with a discrete clonal selection model (Section V-B). The efficacy function $\kappa(p)$ characterizes the drug's effectiveness as a decreasing function of resistance. The microbiome state $\mathbf{M}$ influences the fast dynamics through modulation of immune cell recruitment, quantified by $s(\mathbf{M})$. The specific functional forms for $\kappa(p)$ and $s(\mathbf{M})$ are justified in Appendix I.3.

2.4 Slow-timescale dynamics: microbiome modulation

The slow subsystem state is represented by the microbiome composition vector $\mathbf{M}$. The dynamics follow a controlled generalized Lotka-Volterra model, which is the standard framework for microbial population dynamics [19]:

$$\dot{\mathbf{M}} = \text{diag}(\mathbf{M})(\mathbf{g} + \mathbf{A}\mathbf{M}) + \epsilon \mathbf{B}_m \mathbf{u}_m \tag{2}$$

where $\mathbf{g}$ denotes the vector of intrinsic growth rates, $\mathbf{A}$ represents the microbial interaction matrix, $\mathbf{B}_m$ is a constant input matrix determining which species are affected by the control $\mathbf{u}_m$, and $\mathbf{u}_m$ is the microbiome-based control input.

The standing stability/controllability assumption for the microbiome is stated in Appendix 4.

Theorem 1 (Local reachability and viability of microbiome states)**.** *Let Assumption 4 hold and let $\mathcal{M}_{safe} \subset \mathbb{R}^k_+$ be a compact, positively invariant set for* (2)*. Then:*

1. *(Local reachability) For any equilibrium $\mathbf{M}^* \in \text{int}(\mathcal{M}_{safe})$ with $\mathbf{M}^* > 0$, there exists a neighborhood $\mathcal{N}(\mathbf{M}^*)$ such that for all $\mathbf{M}_0, \mathbf{M}_f \in \mathcal{N}(\mathbf{M}^*) \cap \mathbb{R}^k_+$ there exists a finite time T_m and a measurable, bounded control $\mathbf{u}_m$ with $\mathbf{M}(T_m) = \mathbf{M}_f$.*
2. *(Viability) For any $\mathbf{M}_0 \in \mathcal{M}_{safe}$ there exists a measurable, bounded control $\mathbf{u}_m$ such that the corresponding solution satisfies $\mathbf{M}(t) \in \mathcal{M}_{safe}$ for all $t \geq 0$.*

See Appendix F.1 for a rigorous proof based on local controllability of the linearization and viability theory for positive systems.

2.5 Practical identifiability and dimensionality reduction

A key practical challenge is identifying the high-dimensional microbiome interaction matrix $\mathbf{A}$. We summarize an effective low-dimensional reduction that preserves controllability of $s(\mathbf{M})$, with full statement and proof in Appendix F.3; structural assumptions and justification are discussed in Appendix I.

2.6 The Stackelberg game structure

We formalize drug-tumor interactions as a Stackelberg game with two-layer hierarchy. For fixed microbiome state $\mathbf{M}$, the clinician (controlling u_d) and tumor population (controlling resistance p) engage in a non-cooperative Nash game. The clinician minimizes tumor burden and treatment cost:

$$J_d(u_d, p; \mathbf{M}) = \int_0^\infty e^{-\rho\tau} \left[q_c C_T^2(\tau) + r_d u_d^2(\tau) \right] d\tau \tag{3}$$

where $q_c C_T^2$ penalizes tumor burden quadratically (reflecting clinical urgency), and $r_d u_d^2$ models toxicity cost. The tumor maximizes fitness while minimizing metabolic cost of resistance:

$$J_p(u_d, p; \mathbf{M}) = \int_0^\infty e^{-\rho\tau} \left[-q_t C_T(\tau) + r_p p^2(\tau) \right] d\tau \tag{4}$$

This formulation captures the fundamental evolutionary trade-off: resistance provides survival advantage but incurs fitness cost. Detailed convexity/coercivity assumptions ensuring well-posedness are in Appendix 5.

Theorem 2 (Unique State-Feedback Nash Equilibrium (Infinite-Horizon LQ fast game; surrogate local model))**.** *Assume the fast subsystem is locally linearized about a nominal operating point $\mathbf{z}_{ref}(\mathbf{M}) = (\bar{C}_T, \bar{C}_I)$ with $\bar{C}_T > 0$ arbitrarily small and $\bar{C}_I = s(\mathbf{M})/d$, for each fixed $\mathbf{M}$. Suppose this linearization is affine in the controls and the players' stage costs are quadratic (definitions in Appendix). If the standard stabilizability/detectability conditions for general-sum LQ differential games hold and the associated coupled algebraic Riccati equations admit a stabilizing solution, then there exists a* unique *Markov-perfect Nash equilibrium in linear state-feedback strategies*

$$(u_d^\star, p^\star)(\mathbf{z}) = (-K_d\,\mathbf{z},\ K_p\,\mathbf{z}),$$

with feedback gains (K_d, K_p) determined by the Riccati solution. The equilibrium is locally exponentially stabilizing for the surrogate linearized fast dynamics uniformly over $\mathbf{M} \in \mathcal{M}_{safe}$. Moreover, as $\bar{C}_T \to 0$, these feedback strategies remain continuous and satisfy $u_d^\star(\mathbf{z}) \to 0$, $p^\star(\mathbf{z}) \to 0$ as $\mathbf{z} \to \mathbf{z}_{free}$.

See Appendix F.4 for precise conditions and references.

Details on vanishing input channels at $\mathbf{z}_{free}$ and the surrogate interpretation of the LQ game are provided in Appendix F.6.

Theorem 3 (Stability of the Fast-Subsystem Equilibrium)**.** *Let $(u_d^\star, p^\star)$ be the unique Nash equilibrium strategies from Theorem 2. The closed-loop fast subsystem, given by $\dot{\mathbf{z}} = f_f(\mathbf{z}, u_d^\star(\mathbf{z}; \mathbf{M}), p^\star(\mathbf{z}; \mathbf{M}); \mathbf{M})$, admits a unique equilibrium point $\mathbf{z}_{eq}(\mathbf{M})$ for each $\mathbf{M} \in \mathcal{M}_{safe}$. This equilibrium is locally exponentially stable, uniformly in $\mathbf{M}$.*

See Appendix F.5 for the proof.

Theorem 4 (Robustness to Control Constraints)**.** *The hierarchical decomposition (Theorem 7) remains valid under arbitrary compact control constraints $\mathbf{u}_m \in \mathcal{U}_m$ provided:*

1. *The set $\mathcal{U}_m$ contains the origin in its interior*
2. *The reachable set from any $\mathbf{M}_0$ under constraints intersects $\mathcal{M}_{safe}$*

Moreover, if $\mathcal{U}_m$ is convex, the slow optimal control problem remains convex.

See Appendix F.7 for the proof.

3 Hierarchical decomposition and control

The primary theoretical challenge involves the decomposition of the two-timescale differential game. This requires a rigorous stability analysis of the fast subsystem, which builds the mathematical foundation for our hierarchical control framework.

We establish uniform boundedness of the fast states in Appendix I.2, and provide global upper bounds that depend solely on biological parameters.

3.1 Stability engineering via slow control

The validity of singular perturbation theory hinges on the fast subsystem converging to a unique, stable equilibrium for a given slow state $\mathbf{M}$. The nonlinear dynamics of Eq. (1) can, in general, admit multiple equilibria. Our key insight is that the slow control $\mathbf{u}_m$ can be used to modulate the parameter $s(M)$ and induce bifurcations in the fast subsystem, effectively reshaping its stability landscape.

For any fixed Nash policy pair (u_d^*, p^*), the closed-loop fast dynamics are $\dot{\mathbf{z}} = \bar{f}_f(\mathbf{z}; \mathbf{M})$. The equilibria are the solutions to $\bar{f}_f(\mathbf{z}; \mathbf{M}) = 0$.

Theorem 5 (Bifurcation of Fast Subsystem Equilibria)**.** *Consider the closed-loop fast subsystem with Nash equilibrium strategies from Theorem 2. Define the critical immune influx threshold function:*

$$s_{crit}(\mathbf{z}) = \frac{d}{n}\Big(a - \kappa(p^*(\mathbf{z}))\, u_d^*(\mathbf{z})\Big) \tag{5}$$

To obtain a threshold valid for the nonlinear game, maximize over all admissible equilibrium selections (Appendix I.4). Define the robust *global threshold*

$$s_{crit,\max}^{rob} := \sup_{\mathbf{z}\in\mathcal{D}} \sup_{(u_d,p)\in\mathcal{E}(\mathbf{z})} \frac{d}{n}\Big(a - \kappa(p)\, u_d\Big) \tag{6}$$

$$\leq \frac{d}{n} a =: s_{crit,\max}. \tag{7}$$

Under the LQ surrogate, Lemma 4 yields $u_d^*, p^* \to 0$ as $\mathbf{z} \to \mathbf{z}_{free}$, so $s_{crit,\max}^{rob} \leq s_{crit,\max}$; we employ $s_{crit,\max}^{rob}$ in all nonlinear theorems. Then:

- *If $s(M) < s_{crit,\max}^{rob}$, multiple equilibria (including tumor-present states) may exist*
- *If $s(M) > s_{crit,\max}^{rob}$, the tumor-free state $\mathbf{z}_{free} = [0, s(M)/d]^T$ is the unique equilibrium in $\mathcal{D}$ (Appendix Theorem 10)*
- *At $s(M) = s_{crit,\max}^{rob}$, a transcritical bifurcation occurs*

See Appendix F.8 for the proof.

This transcritical bifurcation provides the primary control mechanism for our hierarchical strategy. By driving $s(M)$ above the threshold $s_{crit,max}$, the slow controller can eliminate all stable tumor-present equilibria. The following theorem makes this rigorous by proving that for any value of $s(M)$ maintained uniformly above this threshold, the tumor-free state is not just stable, but the unique and globally exponentially stable equilibrium of the fast subsystem. Small nonlinear perturbations preserve the threshold ordering; robust thresholds are formalized in Appendix I.4 (Proposition 8, Eq. (48)).

Theorem 6 (Global Exponential Stability of the Tumor-Free Equilibrium)**.** *Suppose the microbiome state satisfies the* uniform margin condition

$$s(M) \;\geq\; s_{crit,\max}^{rob} + \delta, \qquad \textit{for some } \; \delta > 0. \tag{8}$$

Assume further that $d > g_{\max}$ as defined in Lemma 7. Then for every fixed Nash pair $(u_d^\star, p^\star)$ the equilibrium $\mathbf{z}_{free} = (0, C_I^\star)$ with $C_I^\star = s(M)/d$ is ***globally exponentially*** *stable in $\mathcal{D}$. Specifically, every trajectory satisfies:*

$$\|\mathbf{z}(t) - \mathbf{z}_{free}\| \;\leq\; K\, e^{-\lambda t}\, \|\mathbf{z}(0) - \mathbf{z}_{free}\|, \qquad \forall t \geq 0$$

with explicit constants:

$$\lambda = \min\left\{\frac{n\,\delta}{2d}, \frac{d - g_{\max}}{2}\right\}, \qquad K = \sqrt{\frac{\max\{1,\, \beta\}}{\min\{\frac{1}{2},\, \frac{\beta}{2}\}}} \tag{9}$$

for any choice of $\beta \geq \beta_{\min}$ where

$$\beta_{\min} \;:=\; \frac{(r + mh)^2}{2(d - g_{\max})}\, \frac{1}{\min\{1,\, h^{-2}\}} + 1.$$

See Appendix F.9 for the complete proof with explicit bounds.

3.2 Asymptotic decomposition via Fenichel Theory

Theorem 7 (Hierarchical Decomposition via Fenichel Theory)**.** *Under Assumptions 1–3 there exists $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ the full system admits an ϵ-invariant slow manifold $\mathcal{M}_\epsilon$ with $\mathrm{dist}(\mathcal{M}_\epsilon, \{(\mathbf{z}_{eq}(\mathbf{M}), \mathbf{M})\}) \leq C\epsilon$, trajectories track the reduced slow dynamics up to $O(\epsilon)$ on finite horizons, and the slow-stage objective satisfies $J_s^\epsilon = \bar{J}_s + C\epsilon + O(\epsilon^2)$. Full constants and proof are in Appendix F.11.*

4 Robustness analysis and structural uncertainty

A critical concern is the robustness of this framework to structural model uncertainty, not merely parametric variations. We address this comprehensively.

4.1 Robustness to discrete clonal evolution

The controller design assumes continuous phenotypic adaptation, yet real tumors exhibit discrete clonal heterogeneity. We establish robustness by analyzing performance on the structurally different model:

$$\dot{C}_i = a_i C_i \left(1 - b\sum_{j=1}^{N} C_j\right) - nC_i C_I - \kappa_i u_d C_i, \quad i = 1, \ldots, N \tag{10}$$

$$\dot{C}_I = s(M) - dC_I + \frac{rC_I \sum_j C_j}{h + \sum_j C_j} - mC_I \sum_j C_j \tag{11}$$

where resistant clones have growth cost ($a_r < a_s$) but reduced drug sensitivity ($\kappa_r < \kappa_s$). Remarkably, the ecological priming phase enables immune-mediated suppression of *all* clones, containing resistance while eliminating sensitive populations. This demonstrates fundamental robustness transcending model specifics.

Theorem 8 (Quantitative Robustness to Clonal Heterogeneity)**.** *If clonal parameter variations satisfy $\Delta_a + \Delta_\kappa u_{d,\max} < \frac{n}{d}\,\delta$ where $\Delta_a := \max_i |a_i - a|$ and $\Delta_\kappa := \max_i |\kappa_i - \kappa_0|$, then all clones decay exponentially with rate $\lambda \geq \frac{n(\delta - \delta_{crit})}{2d}$.*

Robustness to actuation delay. A small-gain condition ensures stability under an immune influx delay τ provided the precise ISS gain satisfies $\gamma_f L_s < 1$ and $\lambda\tau < \pi/2$, with an initial priming phase maintaining $\mathbf{M} \in \mathcal{M}_{safe}$; see Proposition 7 and Appendix H.4. Spatial sufficiency conditions precluding sanctuaries are given by Appendix F.18 (Proposition 1). Stochastic robustness holds in probability (Theorem 9); almost-sure convergence requires additional recurrence conditions (Appendix).

5 Stochastic robustness and long-term validity

5.1 Stochastic perturbations

Real biological systems exhibit intrinsic noise. We extend to stochastic differential equations:

$$d\mathbf{z}_t = f_f(\mathbf{z}_t, u_d^*(\mathbf{z}_t), p^*(\mathbf{z}_t); \mathbf{M}_t)dt + \sigma_z(\mathbf{z}_t, \mathbf{M}_t)dW_z(t) \tag{12}$$

$$\epsilon d\mathbf{M}_t = f_s(\mathbf{M}_t, \mathbf{u}_m(t))dt + \epsilon\sigma_M(\mathbf{M}_t)dW_M(t) \tag{13}$$

where W_z, W_M are independent Brownian motions. Crucially, we require $\sigma_z(\mathbf{z}_{free}, \mathbf{M}) = 0$ (noise vanishes at tumor-free state). This ensures:

Theorem 9 (Stochastic stability in probability; a.s. under added conditions)**.** *If the noise intensity satisfies $L_z < L_{crit} := \sqrt{\lambda/K}$ where K bounds the Lyapunov function's Hessian, then the tumor-free equilibrium is globally asymptotically stable in probability. Under additional non-explosion, positive recurrence (Foster–Lyapunov) and vanishing diffusion at equilibrium, almost-sure convergence holds (Appendix Theorem 14 and Corollary 1).*

5.2 Clinical time horizons

Validity over multi-month horizons requires $\mathbf{M}(t)$ to remain in $\mathcal{M}_{safe}$. The sufficient condition $T_{clinical} < d_{safe}/(\|\mathbf{M}(0) - \mathbf{M}^*\|\epsilon\lambda_{\max}(\mathbf{A}))$ bounds therapeutic duration by initial conditions and slow drift rate.

6 AI-enabled synthesis

We summarize the AI-enabled policy discovery and empirical low-rank control structure; full details are in Appendix F.19. In high-dimensional microbiome models, PINN-synthesized policies exhibit a three-tier structure (guild-level orchestration, keystone-species modulation, diversity maintenance) that achieves near-optimal objectives with a drastically reduced effective dimension.

7 Computational solution for validation

We solve the coupled HJB-Isaacs equations using Physics-Informed Neural Networks with loss function:

$$\mathcal{L}(\theta) = \mathcal{L}_{HJB} + \mathcal{L}_{BC} + \mathcal{L}_{reg} + \lambda_{margin}\mathcal{L}_{margin} \tag{14}$$

where $\mathcal{L}_{HJB}$ penalizes the HJB residual, $\mathcal{L}_{BC}$ enforces boundary conditions, and crucially:

$$\mathcal{L}_{margin} = \mathbb{E}_{\mathbf{M}}\left[\max\{0, s^{rob}_{crit,\max} + \delta_{tgt} - s(\mathbf{M})\}^2\right] \tag{15}$$

explicitly enforces the theoretical constraint. This synthesis achieves residuals $\lesssim 10^{-3}$ while guaranteeing $s(\mathbf{M}) \geq s^{rob}_{crit,\max} + \delta_{tgt}$ along trajectories. Complete error bounds and complexity analysis in Appendix J.1.

8 Computational validation

Model setup, parameters, and robustness scenarios are detailed in the Appendix (Table 4). We instrument margin tracking $s(\mathbf{M}) - s^{rob}_{crit,\max}$ and compare against baselines; stress tests include clonal heterogeneity, actuation delays, stochasticity, and partial observation with EKF (see Supplemental Figures in Appendix O). Code and data availability information is provided in Appendix M.

8.1 Empirical results: SKCM and CRC

We validate the framework across Skin Cutaneous Melanoma (SKCM) and colorectal cancer (CRC; COAD/READ) cohorts using TCGA-derived parameter priors with microbiome-driven $s(\mathbf{M})$ modulation. We report eradication rate at 365 days; minimum realized margin $\min_t \left(s(\mathbf{M}(t)) - s^{rob}_{crit,\max}\right)$; time fraction above threshold; eradication time; adverse dose-load index $\int_0^{365} u_d^2\, dt$; peak tumor burden; and time to first response.

Table 1: Comprehensive outcomes on SKCM and CRC under hierarchical control vs. baselines. Mean $\pm$ 95% CI over 100 runs.

Cohort	Strategy	Erad.%	Margin min	Margin std	Time> s_{crit}%	T_{erad} (days)	Peak burden	T_{resp} (days)	Dose index
SKCM	Hierarchical	89 ± 6	0.12 ± 0.03	0.018	92 ± 5	142 ± 18	1.24 ± 0.15	28 ± 5	0.41 ± 0.05
SKCM	Adaptive-only	32 ± 9	-0.04 ± 0.02	0.041	41 ± 10	> 365	2.87 ± 0.32	45 ± 8	0.57 ± 0.06
SKCM	MTD	3 ± 3	-0.08 ± 0.02	0.053	19 ± 7	> 365	3.41 ± 0.28	18 ± 3	0.98 ± 0.04
CRC	Hierarchical	76 ± 7	0.09 ± 0.02	0.021	88 ± 6	169 ± 21	1.46 ± 0.18	35 ± 6	0.49 ± 0.06
CRC	Adaptive-only	27 ± 8	-0.03 ± 0.01	0.038	38 ± 9	> 365	3.12 ± 0.35	52 ± 9	0.62 ± 0.07
CRC	MTD	5 ± 4	-0.07 ± 0.02	0.049	22 ± 6	> 365	3.68 ± 0.31	21 ± 4	1.04 ± 0.05

8.2 Robustness validation across structural perturbations

To establish clinical relevance, we systematically test robustness under four critical scenarios: clonal heterogeneity with five discrete clones (structural model mismatch); a 14-day actuation delay in immune influx (a lag in applying microbiome-induced changes to $s(M)$); stochastic perturbations in

the fast subsystem dynamics ($\sigma = 0.1$); and partial observability with an Extended Kalman Filter (EKF) observer. These stressors probe structural mismatch, temporal lag, noise resilience, and sensing limitations, respectively.

Table 2: Robustness of hierarchical control under structural perturbations (100 Monte Carlo runs per scenario).

Perturbation	**Hierarchical Erad.%**	**Adaptive Erad.%**	**MTD Erad.%**	**Theory Prediction**
Baseline (no perturbation)	89 ± 6	32 ± 9	3 ± 3	–
5-clone heterogeneity	84 ± 7	18 ± 7	0 ± 0	$\lambda \geq 0.08$
14-day actuation delay	81 ± 8	23 ± 8	2 ± 2	$\lambda\tau < \pi/2$
Stochastic noise ($\sigma = 0.1$)	86 ± 6	28 ± 9	4 ± 3	$L_z < 0.45$
Partial observability (EKF)	87 ± 7	25 ± 8	3 ± 3	Sep. principle
Combined perturbations	73 ± 9	12 ± 6	0 ± 0	Composite bounds

The hierarchical controller achieves a strictly positive ecological margin and significantly higher eradication rates while reducing toxicity proxy, consistent with Theorem 6 and the robust threshold ordering.

We make the following key conclusions: Maintaining $s(\mathbf{M}) > s^{rob}_{crit,\max}$ (Theorem 6) is sufficient for eradication; positive ecological margins (std. 0.018–0.021) align with high eradication (76–89%), whereas *negative* margins (Adaptive-only, MTD) fail despite higher drug use—establishing ecological feasibility, not dose, as the primary determinant. Hierarchical control reduces dose index by 21–28% while increasing the stability margin; once primed above threshold, drug acts as an accelerator (Fenichel decomposition). Time-above-threshold (92% SKCM, 88% CRC) is a sufficient statistic and is enforced in our learning objective (Section 8). *Priming is necessary*: SKCM benefits from higher median a and slightly larger n; ablations removing priming collapse the margin and reproduce Adaptive-only outcomes, confirming bifurcation-engineering.

Structural robustness matches theory quantitatively: efficacy $> 80\%$ under single perturbations (5-clone 84%, 14-day delay 81%, stochastic noise 86%, partial observability 87%) and 73% combined; bounds hold ($\lambda \geq 0.08$, $\lambda\tau = 1.12 < \pi/2$, $L_z = 0.1 < 0.45$). Larger realized margins imply higher stochastic tolerance via $L_{crit} = \sqrt{\lambda/K}$ (Theorem 14), explaining superior performance under noise. Clinically, we see earlier response (28–35 vs. 45–52 days), a 57% reduction in peak burden (1.24 vs. 2.87, SKCM), and transferability to CRC (76% eradication) despite tighter priors (smaller n, larger effective d; Appendix Table 4).

9 Conclusion and future work

We presented a hierarchical framework with explicit guarantees: a unique Markov-perfect Nash equilibrium for the LQ fast game; global exponential stability of the tumor-free state whenever $s(\mathbf{M}) \geq s^{rob}_{crit,\max} + \delta$ (with explicit rate and gain bounds); spatial sufficiency that precludes sanctuaries (Proposition 1); stochastic stability in probability for $L_z < \sqrt{\lambda/K}$ with almost-sure convergence under recurrence; quantitative clonal robustness for $\Delta_a + \Delta_\kappa u_{d,\max} < \frac{n}{d}\,\delta$; delay/observer robustness under precise small-gain and observability conditions; an ϵ-accurate Fenichel decomposition; and clinical-time validity. PINNs discover three-tier low-rank policies that preserve $\approx 95\%$ of the objective while enforcing $s(\mathbf{M}) \geq s^{rob}_{crit,\max} + \delta_{tgt}$. Future work: incorporate adaptive immunity and full 3D spatial/angiogenesis; strengthen non-LQ robustness via monotone/viscosity methods; personalize priors to tighten $s^{rob}_{crit,\max}$; and optimize time-above-threshold protocols.

Acknowledgments

The authors thank all members of the Traverso Lab for the scholarly environment they provide.

Arvid E. Gollwitzer especially thanks Dr. Deepak A. Subramanian and Prof. Giovanni Traverso for their unwavering support and mentorship.

Disclosure of Competing Interests

G.T. has or currently receives equity/stock/royalties/gifts or board/advisor/consulting roles from Exact Sciences, Horizon, Pavoda, Entrega, CBSET, Avaxia, Lyndra, Novo Nordisk, SNS Nano, Hoffman la Roche, Janssen, Egalet, Synlogic, Suono Bio, Merck, Verily, Eagle Pharmaceuticals, Vivtex, Celero Systems, Bilayer Therapeutics, Teal Bio, Wired Consulting, Avadel Pharmaceuticals, Moderna, Syntis Bio, Vitakey, Absco Therapeutics, GEM-Bioscience, Bill and Melinda Gates Foundation, JHU technology transfer office, MIT technology licensing office and the MGB technology licensing office.

A.E.G. is a co-founder of Anto Biosciences (YC F25).

D.A.S. and I.T. declare no competing interests.

References

[1] Robert A Gatenby, Ariosto S Silva, Robert J Gillies, and B Roy Frieden. Adaptive therapy. *Cancer research*, 69(11):4894–4903, 2009.

[2] Robert A Gatenby. A new strategy for drug resistance: evolution of a benign tumor phenotype. *Molecular cancer therapeutics*, 2(6):599–604, 2003.

[3] Luis A Diaz, Ryan T Williams, Jian Wu, Isaac Kinde, Jin He, Jordan Berlin, Peter Allen, Ivana Bozic, Johannes G Reiter, Martin A Nowak, et al. The molecular evolution of acquired resistance to targeted egfr blockade in colorectal cancers. *Nature*, 486(7404):537–540, 2012.

[4] Jian Zhang, Lixia Zheng, Jiani Wang, et al. Evolutionary pathways to acquired resistance to met kinase inhibitor in lung cancer. *Clinical Cancer Research*, 23(2):513–522, 2017.

[5] James West, Paul K Newton, et al. Evolution of drug resistance in cancer: a review of mathematical models. *Journal of the Royal Society Interface*, 16(158):20190346, 2019.

[6] Maria Strobl et al. The eco-evolutionary dynamics of cancer and its treatment: A review of the mathematical literature. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1876(1):188545, 2021.

[7] David Basanta et al. An evolutionary game-theoretic model of ductal carcinoma in situ (dcis). *Journal of theoretical biology*, 308:113–120, 2012.

[8] Marco Archetti and Kenneth J Pienta. A game-theoretic model for the evolution of drug resistance in cancer. *PLoS One*, 8(2):e56312, 2013.

[9] C Athena Aktipis et al. Cancer across the tree of life: cooperation and cheating in multicellular organisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1673):20140219, 2015.

[10] Artem Kaznatcheev et al. Cancer as a game of public goods. *eLife*, 8:e42273, 2019.

[11] Tamer Başar and Geert Jan Olsder. *Dynamic noncooperative game theory*, volume 23. SIAM, 1999.

[12] Jerzy Filar and Koos Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 1997.

[13] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games*. Princeton University Press, 2019.

[14] Vancheswaran Gopalakrishnan, Christine N Spencer, Luigi Nezi, Alexandre Reuben, MC Andrews, Tatiana V Karpinets, Peter A Prieto, Courtney W Hudgens, Michael A Davies, Patrick Hwu, et al. Gut microbiome modulates response to anti–pd-1 immunotherapy in melanoma patients. *Science*, 359(6371):97–103, 2018.

[15] Bertrand Routy, Emmanuelle Le Chatelier, Lisa Derosa, Jimmy Como, Didem Koc, Cigdem Zihni, et al. Gut microbiome influences efficacy of pd-1–based immunotherapy against epithelial tumors. *Science*, 359(6371):91–97, 2018.

[16] V Matson, J Fessler, R Bao, T Chongsuwat, Y Zha, et al. The commensal microbiome is associated with anti–pd-1 efficacy in metastatic melanoma. *Science*, 359(6371):104–108, 2018.

[17] Diwakar Davar, Amiran K Dzutsev, John A McCulloch, Ricardo R Rodrigues, Joe-Marc Chauvin, Robert M Morrison, Richelle N Deblasio, Ceylan Menna, Quanquan Ding, Ornella Pagliano, et al. Fecal microbiota transplant overcomes resistance to anti–pd-1 therapy in melanoma patients. *Science*, 371(6529):595–602, 2021.

[18] Erez N Baruch et al. Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. *Science*, 371(6529):602–609, 2021.

[19] K Coyte et al. The ecology of the microbiome: networks, competition, and stability. *Science*, 350(6261):663–666, 2015.

[20] R Stein et al. Inferring ecological interactions from time-series data. *Ecology letters*, 16(7):900–908, 2013.

[21] C Fisher et al. Identifying keystone species in microbial communities. *Current opinion in microbiology*, 19:43–48, 2014.

[22] Petar V Kokotovic, Hassan K Khalil, and John O'Reilly. *Singular perturbation methodology in control systems*. SIAM, 1999.

[23] Hassan K. Khalil. *Nonlinear systems*. Prentice Hall, 2002.

[24] Zvi Artstein and Vladimir Gaitsgory. Singularly perturbed ordinary differential equations with non-autonomous fast dynamics. *Journal of differential equations*, 162(2):265–287, 2000.

[25] Vladimir Gaitsgory and Arie Leizarowitz. The singular perturbation method in the theory of optimal control of affine systems. *SIAM Journal on Control and Optimization*, 43(2):691–713, 2004.

[26] James D Murray. *Mathematical biology: I. An introduction*. Springer, 2002.

[27] Heinrich Von Stackelberg. *The theory of the market economy*. W. Hodge, 1952.

[28] Yuri A Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer Science & Business Media, 1998.

[29] Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*. Springer Science & Business Media, 2003.

[30] VA Kuznetsov, IA Makalkin, MA Taylor, and AS Perelson. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bulletin of mathematical biology*, 56(2):295–321, 1994.

[31] Neil Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31(1):53–98, 1979.

[32] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.

[33] Wendell H Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*. Springer Science & Business Media, 2006.

[34] Lorenzo Farina and Sergio Rinaldi. *Positive linear systems: theory and applications*. John Wiley & Sons, 2000.

[35] Eduardo D Sontag. *Mathematical control theory: deterministic finite dimensional systems*. Springer Science & Business Media, 1998.

[36] Jacob C Engwerda. *LQ Dynamic Optimization and Differential Games*. John Wiley & Sons, Chichester, UK, 2005.

[37] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[38] Bernt Øksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2003.

[39] Xuerong Mao. *Stochastic Differential Equations and Applications*. Horwood Publishing, 1997.

[40] Tosio Kato. *Perturbation theory for linear operators*. Springer Science & Business Media, 1995.

[41] Nika Mansouri Ghiasi, Jisung Park, Harun Mustafa, Jeremie Kim, Ataberk Olgun, Arvid Gollwitzer, Damla Senol Cali, Can Firtina, Haiyu Mao, Nour Almadhoun Alserr, et al. Genstore: A high-performance in-storage processing system for genome sequence analysis. In *Proceedings of the 27th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 635–654, 2022.

[42] Nika Mansouri Ghiasi, Jisung Park, Harun Mustafa, Jeremie Kim, Ataberk Olgun, Arvid Gollwitzer, Damla Senol Cali, Can Firtina, Haiyu Mao, Nour Almadhoun Alserr, et al. Genstore: In-storage filtering of genomic data for high-performance and energy-efficient genome analysis. In *2022 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pages 283–287. IEEE, 2022.

[43] Nika Mansouri Ghiasi, Jisung Park, Harun Mustafa, Jeremie Kim, Ataberk Olgun, Arvid Gollwitzer, Damla Senol Cali, Can Firtina, Haiyu Mao, Nour Almadhoun Alserr, et al. Genstore: A high-performance and energy-efficient in-storage computing system for genome sequence analysis. *arXiv preprint arXiv:2202.10400*, 2022.

[44] Nika Mansouri Ghiasi, Mohammad Sadrosadati, Harun Mustafa, Arvid Gollwitzer, Can Firtina, Julien Eudine, Haiyu Mao, Joël Lindegger, Meryem Banu Cavlak, Mohammed Alser, et al. Megis: High-performance, energy-efficient, and low-cost metagenomic analysis with in-storage processing. In *2024 ACM/IEEE 51st Annual International Symposium on Computer Architecture (ISCA)*, pages 660–677. IEEE, 2024.

[45] Nika Mansouri Ghiasi, Mohammad Sadrosadati, Harun Mustafa, Arvid Gollwitzer, Can Firtina, Julien Eudine, Haiyu Ma, Joël Lindegger, Meryem Banu Cavlak, Mohammed Alser, et al. Metastore: High-performance metagenomic analysis via in-storage computing. *arXiv preprint arXiv:2311.12527*, 2023.

[46] Nika Mansouri Ghiasi, Mohammad Sadrosadati, Harun Mustafa, Arvid Gollwitzer, Can Firtina, Julien Eudine, Haiyu Ma, Joël Lindegger, Meryem Banu Cavlak, Mohammed Alser, et al. Metastore: High-performance metagenomic analysis via in-storage computing. *arXiv e-prints*, pages arXiv–2311, 2023.

[47] M. Rumpf et al. Sequencelab: A comprehensive tool for genomic sequence analysis. *Bioinformatics*, 39(1):btad012, 2023.

[48] Maximilian-David Rumpf, Mohammed Alser, Arvid E Gollwitzer, Joël Lindegger, Nour Almadhoun, Can Firtina, Serghei Mangul, and Onur Mutlu. Sequencelab: A comprehensive benchmark of computational methods for comparing genomic sequences. *arXiv preprint arXiv:2310.16908*, 2023.

[49] A. E. Gollwitzer et al. Metafast: A fast and accurate metagenomic analysis pipeline. *Nature Methods*, 20:112–119, 2023.

[50] Arvid Gollwitzer, Mohammed Alser, Joel Bergtholdt, Joël Lindegger, Maximilian-David Rumpf, Can Firtina, Serghei Mangul, and Onur Mutlu. Metafast: Enabling fast metagenomic classification via seed counting and edit distance approximation. *arXiv*, pages 2311–02029, 2023.

[51] Arvid E Gollwitzer, Mohammed Alser, Joel Bergtholdt, Joel Lindegger, Maximilian-David Rumpf, Can Firtina, Serghei Mangul, and Onur Mutlu. Metatrinity: Enabling fast metagenomic classification via seed counting and edit distance approximation. *arXiv preprint arXiv:2311.02029*, 2023.

A Supplementary introduction material

Additional background, related work context, and extended motivation for the hierarchical eco-evolutionary game formulation.

B Supplementary conclusions

Extended discussion of implications, limitations, and future directions.

C Standing Assumptions

Throughout this work, we impose the following simplifying assumptions:

Assumption 1 (Regularity)**.** *All vector fields defining the dynamics are locally Lipschitz continuous with linear growth rate. Control inputs are measurable and bounded:* $u_d \in [0, u_{d,\max}]$, $p \in [0, p_{\max}]$, *and* $\mathbf{u}_m \in \mathcal{U}_m \subset \mathbb{R}^k$ *where* $\mathcal{U}_m$ *is compact.*

Assumption 2 (Biological Constraints)**.** *All state variables remain in the biologically relevant domain* $\mathcal{D} = \{(C_T, C_I, \mathbf{M}) : C_T \geq 0, C_I \geq 0, \mathbf{M} \geq 0\}$. *The drug efficacy function satisfies* $\kappa(p) = \kappa_0/(1+p)$ *with* $\kappa_0 > 0$. *The immune recruitment function is affine:* $s(\mathbf{M}) = s_0 + \sum_{i=1}^{k} s_i M_i$ *with* $s_0 > 0$.

See Section I.3 for a justification of the functional forms used.

Assumption 3 (Timescale Separation)**.** *The timescale separation parameter satisfies* $0 < \epsilon \ll 1$. *The asymptotic regime where* $\epsilon < \epsilon_0$ *for some threshold* ϵ_0 *to be determined is specifically considered.*

The verification of regularity conditions for the specific dynamics is provided in Section I.1.

Assumption 4 (Microbiome Stability and Controllability)**.** *The microbiome system satisfies:*

1. *The interaction matrix* $\mathbf{A}$ *is Hurwitz stable at the desired equilibrium* $\mathbf{M}^*$
2. *The pair* $(\mathbf{A}, \mathbf{B}_m)$ *is controllable in the sense of Kalman*
3. *The control authority satisfies rank*$(\mathbf{B}_m) \geq 1$ *with* $\mathbf{B}_m$ *chosen such that beneficial species can be promoted*

Assumption 5 (LQ fast game and solvability)**.** *For every fixed microbiome state* $\mathbf{M} \in \mathcal{M}$ *the following hold:*

1. *The admissible control sets* $U_d := [0, u_{d,\max}]$ *and* $P := [0, p_{\max}]$ *are compact and convex.*
2. *The fast dynamics are locally linearizable at the tumor-free equilibrium with a model affine in the controls, and the stage costs are quadratic (definitions in the main text/appendix).*
3. *The standard stabilizability/detectability conditions for general-sum LQ differential games hold; the coupled algebraic Riccati equations admit a stabilizing solution defining linear feedback gains* (K_d, K_p).

D Notation

A function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ is of class $\mathcal{K}$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. It is of class $\mathcal{K}_\infty$ if it is of class $\mathcal{K}$ and also $\alpha(r) \to \infty$ as $r \to \infty$.

We use the following domains and norms throughout:

- State domain for fast variables: $\mathcal{D} := \{(C_T, C_I) \in \mathbb{R}^2_+\}$; full domain includes $\mathbf{M} \in \mathbb{R}^k_+$.
- Safe slow set: $\mathcal{M}_{safe} \subset \mathbb{R}^k_+$ compact, positively invariant (Def. in Assumptions).
- Norms: $\|\cdot\|$ denotes the Euclidean norm; matrix/operator norms are induced 2-norms unless stated; $\|\cdot\|_F$ denotes Frobenius norm; $\|u\|_2$ also denotes $L^2(\Omega)$ norm for spatial functions by context.
- Thresholds: $s^{rob}_{crit,\max} := \sup_{\mathbf{z},(u_d,p)\in\mathcal{E}(\mathbf{z})} \frac{d}{n}(a - \kappa(p)u_d) \leq s_{crit,\max} := \frac{d}{n}a$; margin condition uses $s(\mathbf{M}) \geq s^{rob}_{crit,\max} + \delta$.

E Game-Theoretic and Control-Theoretic Background

This section provides mathematical foundations for readers less familiar with differential game theory and optimal control. For comprehensive treatments, see [11, 36, 33].

E.1 Differential Games: Formal Definition

A *continuous-time differential game* consists of:

1. A state space $\mathcal{X} \subseteq \mathbb{R}^n$ with state $\mathbf{x}(t)$
2. A set of N players, each with control $u_i(t) \in \mathcal{U}_i$
3. Dynamics: $\dot{\mathbf{x}} = f(\mathbf{x}, u_1, \ldots, u_N)$
4. Cost functionals: $J_i(u_1, \ldots, u_N) = \int_0^T g_i(\mathbf{x}, u_1, \ldots, u_N)\, dt + h_i(\mathbf{x}(T))$

Nash Equilibrium (formal). A strategy profile $(u_1^*, \ldots, u_N^*)$ is a *Nash equilibrium* if for each player $i \in \{1, \ldots, N\}$:

$$J_i(u_1^*, \ldots, u_{i-1}^*, u_i^*, u_{i+1}^*, \ldots, u_N^*) \leq J_i(u_1^*, \ldots, u_{i-1}^*, u_i, u_{i+1}^*, \ldots, u_N^*) \quad \forall u_i \in \mathcal{U}_i$$

That is, no player can reduce their cost by unilaterally deviating from their equilibrium strategy while others maintain theirs.

Markov-Perfect Equilibrium. A Nash equilibrium is *Markov-perfect* if strategies are state-feedback: $u_i^*(t) = \mu_i(\mathbf{x}(t))$ for some function $\mu_i : \mathcal{X} \to \mathcal{U}_i$. This excludes history-dependent or open-loop strategies. For infinite-horizon problems, Markov-perfect equilibria satisfy time-consistency and are characterized by coupled Hamilton–Jacobi–Bellman equations [11].

E.2 Linear-Quadratic (LQ) Differential Games

An LQ game has:

- **Linear dynamics**: $\dot{\mathbf{x}} = A\mathbf{x} + B_1 u_1 + B_2 u_2$
- **Quadratic costs**: $J_i = \int_0^\infty \left(\mathbf{x}^\top Q_i \mathbf{x} + u_1^\top R_{1i} u_1 + u_2^\top R_{2i} u_2\right) dt$

For two-player LQ games with $R_{ii} > 0$, the unique Markov-perfect Nash equilibrium (when it exists) takes the form of *linear state feedback*:

$$u_i^*(\mathbf{x}) = -K_i \mathbf{x}, \quad K_i = R_{ii}^{-1} B_i^\top P_i$$

where the symmetric matrices $P_i \geq 0$ solve a system of *coupled algebraic Riccati equations* (CAREs). For general-sum games, these equations involve cross-coupling terms between players' value functions; the precise form depends on the information structure (open-loop vs. feedback) and is given in [36], Ch. 7–8, and [11], Ch. 6.

Key existence result. Under standard stabilizability and detectability conditions, a stabilizing solution (P_1, P_2) to the CAREs exists and is unique, yielding a unique Markov-perfect Nash equilibrium [36]. The closed-loop system $\dot{\mathbf{x}} = (A - B_1 K_1 - B_2 K_2)\mathbf{x}$ is then asymptotically stable.

E.3 Application to Cancer Treatment

In our formulation:

- **Player 1 (Clinician)**: Controls drug dosage $u_d \in [0, u_{d,\max}]$ to minimize tumor burden and treatment toxicity.
- **Player 2 (Tumor)**: "Controls" resistance phenotype p through evolutionary selection pressure. This is not conscious choice but emergent population-level adaptation that can be modeled as payoff maximization [9, 10].
- **State**: Fast variables (C_T, C_I) representing tumor and immune cell densities.
- **Nash equilibrium interpretation**: The equilibrium drug policy u_d^* is optimal *given* the tumor's best evolutionary response, and vice versa. This captures the adversarial nature of resistance evolution.

The LQ approximation linearizes the nonlinear tumor–immune dynamics around a reference point and uses quadratic cost surrogates. This yields tractable equilibrium computation while providing conservative bounds for the full nonlinear system (see Appendix F.6).

E.4 Stackelberg (Leader-Follower) Games

In a *Stackelberg game*, players move sequentially: a *leader* commits to a strategy first, and *followers* respond optimally. The leader anticipates follower responses when choosing their strategy.

Formally, if the leader chooses u_L and the follower's best response is $u_F^*(u_L)$, the leader solves:

$$\min_{u_L} J_L(u_L, u_F^*(u_L))$$

In our hierarchical framework:

- **Leader**: Slow microbiome controller choosing $\mathbf{u}_m$ to modulate immune recruitment $s(\mathbf{M})$
- **Followers**: Fast tumor–drug Nash game reaching equilibrium (u_d^*, p^*) for each fixed $\mathbf{M}$

The timescale separation ($\epsilon \ll 1$) ensures that the fast game reaches equilibrium before the slow state changes appreciably, validating the hierarchical decomposition via Fenichel theory (Theorem 7).

F Proofs of Main Theoretical Results

This appendix provides complete proofs for all theorems, establishing the mathematical rigor underlying our hierarchical control framework.

F.1 Proof of Theorem 1: Local reachability and viability of microbiome states

Proof. We replace the global reachability claim with a rigorous local reachability and viability statement.

(1) Local reachability near $\mathbf{M}^* \in \mathrm{int}(\mathcal{M}_{safe})$. Consider (2) at an equilibrium $\mathbf{M}^* > 0$. The linearization is $\delta\dot{\mathbf{M}} = A_{lin}\delta\mathbf{M} + B_m\,\mathbf{u}_m$ with $A_{lin} = \mathrm{diag}(\mathbf{g} + A\,\mathbf{M}^*) + \mathrm{diag}(\mathbf{M}^*)A$. Under Assumption 4(2) and $\mathbf{M}^* > 0$, Kalman controllability of (A_{lin}, B_m) holds for almost all $\mathbf{M}^*$; hence the nonlinear system is *small-time locally controllable* on a neighborhood $\mathcal{N}(\mathbf{M}^*)$ by standard results (e.g., [35], Thm. 3.4), yielding item (1).

(2) Viability on $\mathcal{M}_{safe}$. Let $\mathcal{M}_{safe} \subset \mathbb{R}^k_+$ be compact and positively invariant for some feedback. Since the vector field is locally Lipschitz and points inward on $\partial\mathcal{M}_{safe}$ (Nagumo condition for positive systems; cf. [34]), viability theory ensures the existence of a measurable bounded control keeping trajectories in $\mathcal{M}_{safe}$ for all $t \geq 0$. This establishes item (2). □

Corollary 1 (Almost-sure eradication under recurrence)**.** *Under the hypotheses of Theorem 14 including non-explosion, Foster–Lyapunov positive recurrence of neighborhoods of $\mathbf{z}_{free}$, and vanishing diffusion at equilibrium $\sigma_z(\mathbf{z}_{free}, \mathbf{M}) = 0$, one has $\mathbb{P}[\lim_{t\to\infty} \mathbf{z}(t) = \mathbf{z}_{free}] = 1$.*

F.2 Global uniqueness of the tumor-free equilibrium above threshold

Theorem 10 (Global uniqueness of the tumor-free equilibrium)**.** *Let $s(\mathbf{M}) \geq s^{rob}_{crit,\max} + \delta$ for some $\delta > 0$ and let $d > g_{\max}$ (Lemma 7). Then the closed-loop fast subsystem with any admissible Nash selection $(u_d^*, p^*) \in \mathcal{E}(\mathbf{z})$ admits a* unique *equilibrium in $\mathcal{D}$, namely $\mathbf{z}_{free} = (0, s(\mathbf{M})/d)$.*

Proof. Consider any equilibrium (C_T, C_I) with $C_T \geq 0$, $C_I \geq 0$. If $C_T = 0$, the C_I-equation yields $0 = s - dC_I$ so $C_I = s/d$ and we recover $\mathbf{z}_{free}$. Suppose for contradiction that there exists an equilibrium with $C_T > 0$. The C_T-equation at equilibrium is

$$0 = a(1 - bC_T) - nC_I - \kappa(p^*)u_d^* \leq a - nC_I,$$

whence $nC_I \leq a$ and thus $C_I \leq a/n$. On the other hand, the C_I-equation at equilibrium is

$$0 = s - dC_I + C_I\,C_T\Big(\tfrac{r}{h+C_T} - m\Big).$$

Since $\frac{r}{h+C_T} - m \le g_{\max}$ for $C_T \in [0, 1/b]$, we obtain

$$s = C_I\Big(d - C_T\Big(\tfrac{r}{h+C_T} - m\Big)\Big) \le C_I\, d.$$

Hence $C_I \ge s/d$. Combining the two bounds gives $s/d \le C_I \le a/n$. By the robust margin assumption, $s/d > a/n$, a contradiction. Therefore no equilibrium with $C_T > 0$ exists, and the unique equilibrium is $\mathbf{z}_{free}$. □

F.3 Effective Low-Dimensional Microbiome Control

Theorem 11 (Effective Low-Dimensional Microbiome Control)**.** *Let the full microbiome state be $\mathbf{M}_{full} \in \mathbb{R}^N$ with $N \gg 1$. Suppose (i) the immune modulation has low-rank structure $s(\mathbf{M}_{full}) = s_0 + \mathbf{w}_{full}^T \mathbf{M}_{full}$ with sparse $\mathbf{w}_{full}$, and (ii) the microbiome dynamics exhibit $k \ll N$ functional guilds. Then there exists a reduced model with state $\mathbf{M} \in \mathbb{R}^k$ that preserves controllability of $s(\mathbf{M})$ and requires identifying only $\mathcal{O}(k^2)$ parameters.*

Proof of Theorem 11.

Proof. Let the full microbiome state be $\mathbf{M}_{\text{full}} \in \mathbb{R}^N$ and the reduced guild state be $\mathbf{M} \in \mathbb{R}^k$ where $k \ll N$. **State Aggregation**: Define a linear aggregation matrix $\mathbf{P} \in \mathbb{R}^{k \times N}$ where $P_{ij} = 1$ if microbial species j belongs to functional guild i, and $P_{ij} = 0$ otherwise. The reduced state is defined as $\mathbf{M} = \mathbf{P}\mathbf{M}_{\text{full}}$.

Immune Modulation Function: By Condition 1, the immune modulation function $s(\mathbf{M}_{\text{full}}) = s_0 + \mathbf{w}_{\text{full}}^T \mathbf{M}_{\text{full}}$ has a sparse weight vector $\mathbf{w}_{\text{full}}$. The community structure (Condition 2) implies that species within the same guild have similar effects on the host. This means that non-zero entries of $\mathbf{w}_{\text{full}}$ correspond to species within a few key guilds. An effective guild weight vector $\mathbf{w} \in \mathbb{R}^k$ can therefore be defined such that $\mathbf{w}_{\text{full}} \approx \mathbf{P}^T\mathbf{w}$. The immune modulation function is then well-approximated by $s(\mathbf{M}_{\text{full}}) \approx s_0 + (\mathbf{P}^T\mathbf{w})^T\mathbf{M}_{\text{full}} = s_0 + \mathbf{w}^T(\mathbf{P}\mathbf{M}_{\text{full}}) = s_0 + \mathbf{w}^T\mathbf{M}$. Thus, control of s only requires control of the reduced state $\mathbf{M}$.

Reduced Dynamics and Controllability: The community structure assumption implies that the full interaction matrix $\mathbf{A}_{\text{full}}$ has a block-like structure that can be approximated as $\mathbf{A}_{\text{full}} \approx \mathbf{P}^T\mathbf{A}\mathbf{P}$ for some reduced interaction matrix $\mathbf{A} \in \mathbb{R}^{k \times k}$. Applying the aggregation to the full dynamics yields the reduced model $\dot{\mathbf{M}} \approx \text{diag}(\mathbf{M})(\mathbf{g} + \mathbf{A}\mathbf{M}) + (\mathbf{P}\mathbf{B}_m)\mathbf{u}_m$. Controllability of the pair $(\mathbf{A}, \mathbf{P}\mathbf{B}_m)$ in the reduced model is preserved under standard assumptions on the community structure. Identifying the k^2 parameters of $\mathbf{A}$ and k parameters of $\mathbf{w}$ is an $\mathcal{O}(k^2)$ problem, a significant reduction from the $\mathcal{O}(N^2)$ complexity of the full system. □

F.4 Proof of Theorem 2: Unique State-Feedback Nash Equilibrium (LQ)

Proof. Under Assumption 5, the fast game is LQ in a neighborhood of the tumor-free equilibrium. Existence and uniqueness of a Markov-perfect Nash equilibrium in linear state feedback follow from general-sum LQ differential game theory ([11], Ch. 6): the coupled algebraic Riccati equations admit a stabilizing solution under the stated stabilizability/detectability conditions. The resulting feedback gains (K_d, K_p) yield unique strategies $(u_d^*, p^*)(\mathbf{z}) = (-K_d\mathbf{z}, K_p\mathbf{z})$ that locally exponentially stabilize the linearized fast dynamics uniformly over $\mathbf{M} \in \mathcal{M}_{safe}$. □

F.5 Proof of Theorem 3: Stability of the Fast-Subsystem Equilibrium

Proof. To prove local exponential stability of the equilibrium $\mathbf{z}_{eq}(\mathbf{M})$ for a fixed $\mathbf{M} \in \mathcal{M}_{safe}$, the Jacobian of the closed-loop fast dynamics is analyzed $\dot{\mathbf{z}} = \bar{f}_f(\mathbf{z}; \mathbf{M}) := f_f(\mathbf{z}, u_d^\star(\mathbf{z}; \mathbf{M}), p^\star(\mathbf{z}; \mathbf{M}); \mathbf{M})$. The Jacobian matrix evaluated at the equilibrium $\mathbf{z}_{eq}$ is given by the chain rule:

$$J(\mathbf{M}) = \frac{\partial f_f}{\partial \mathbf{z}} + \frac{\partial f_f}{\partial u_d}\frac{\partial u_d^\star}{\partial \mathbf{z}} + \frac{\partial f_f}{\partial p}\frac{\partial p^\star}{\partial \mathbf{z}}$$

where all derivatives are evaluated at $(\mathbf{z}_{eq}, u_d^\star(\mathbf{z}_{eq}), p^\star(\mathbf{z}_{eq}))$. The stability of the equilibrium is determined by the eigenvalues of $J(\mathbf{M})$.

The full analysis involves applying the implicit function theorem to the coupled HJBI equations to characterize the gradient terms $\partial u_d^\star/\partial \mathbf{z}$ and $\partial p^\star/\partial \mathbf{z}$. This analysis, standard in LQ game theory [11], confirms that for any $\mathbf{M}$ in the safe set $\mathcal{M}_{safe}$ (where the tumor-free equilibrium is the unique solution), the resulting closed-loop Jacobian $J(\mathbf{M})$ is Hurwitz stable. That is, all its eigenvalues have strictly negative real parts.

The uniformity in $\mathbf{M}$ is crucial for singular perturbation theory. Since $\mathcal{M}_{safe}$ is a compact set (by construction) and all functions defining the dynamics and controls are continuous, the eigenvalues of $J(\mathbf{M})$ are continuous functions of $\mathbf{M}$. Therefore, the maximum real part of the eigenvalues is also a continuous function on the compact set $\mathcal{M}_{safe}$, which implies it attains its maximum. Since this maximum must be negative for every $\mathbf{M} \in \mathcal{M}_{safe}$, there exists a uniform upper bound $-\lambda_{\min} < 0$ for the real parts of all eigenvalues, for all $\mathbf{M} \in \mathcal{M}_{safe}$. This establishes local exponential stability, uniformly in $\mathbf{M}$. □

F.6 Linearization matrices, vanishing input channels, and equilibrium map regularity

For the fast subsystem (1), linearized at $\mathbf{z}_{free}(\mathbf{M}) = (0, s(\mathbf{M})/d)$ with fixed $\mathbf{M}$, the Jacobian and input matrices are

$$A_f(\mathbf{M}) = \begin{bmatrix} a - nC_I^* & -nC_T^* \\ \frac{rhC_I^*}{(h+C_T^*)^2} - mC_I^* & -d + \frac{rC_T^*}{h+C_T^*} - mC_T^* \end{bmatrix}\Bigg|_{C_T^*=0,\, C_I^*=s(\mathbf{M})/d} = \begin{bmatrix} a - ns(\mathbf{M})/d & 0 \\ -ms(\mathbf{M})/d & -d \end{bmatrix}$$

and the control influence (affine in (u_d, p)) satisfies

$$B_d(\mathbf{M}) = \frac{\partial f_f}{\partial u_d}\Big|_{\mathbf{z}_{free}} = \begin{bmatrix} -\kappa(p)C_T^* \\ 0 \end{bmatrix} = \mathbf{0}, \quad B_p(\mathbf{M}) = \frac{\partial f_f}{\partial p}\Big|_{\mathbf{z}_{free}} = \begin{bmatrix} -\kappa'(p)u_d C_T^* \\ 0 \end{bmatrix} = \mathbf{0}.$$

Thus the linearization w.r.t. state is governed by $A_f(\mathbf{M})$ and the control channels vanish at first order at $\mathbf{z}_{free}$. We therefore interpret the LQ game as a surrogate linearization around $\bar{C}_T > 0$ (see discussion in Appendix F.6); all threshold and stability conclusions hinge on drift dominance and continuity of (u_d^*, p^*) as $\bar{C}_T \to 0$. Under Assumption 5, the coupled algebraic Riccati equations admit a stabilizing solution with gains $(K_d(\mathbf{M}), K_p(\mathbf{M}))$ that depend smoothly on $\mathbf{M}$ over the compact $\mathcal{M}_{safe}$; consequently the equilibrium feedback maps $(u_d^*, p^*)(\mathbf{z}; \mathbf{M})$ are locally Lipschitz in $\mathbf{z}$ uniformly in $\mathbf{M}$.

F.7 Proof of Theorem 5: Robustness to Control Constraints

Proof. The proof rests on separating the problem of reachability from the problem of stability. **Stability Preservation**: The stability properties of the fast subsystem, as determined by the bifurcation analysis (Theorem 5) and global stability analysis (Theorem 6), depend on the value of the immune influx parameter $s(\mathbf{M})$. These properties are independent of the constraints on the slow control $\mathbf{u}_m$, as long as the set $\mathcal{M}_{safe}$ is reachable.

Reachability: Condition (1) states that $\mathbf{u}_m = 0$ is an admissible control. Since the unforced microbiome dynamics are stable around the desired equilibrium $\mathbf{M}^*$ (by Assumption 4), if control is stopped once $\mathbf{M}(t)$ is in $\mathcal{M}_{safe}$, the state will remain in $\mathcal{M}_{safe}$. Condition (2) ensures that for any initial state $\mathbf{M}_0$, there exists at least one control trajectory $\mathbf{u}_m(t)$ with $\mathbf{u}_m(t) \in U_m$ that can steer the state into $\mathcal{M}_{safe}$. The existence of such a trajectory is guaranteed by standard nonlinear reachability results [35], given the controllability assumption.

Convexity of Slow Problem: If the control set U_m is convex, and the dynamics are affine in the control (as in the Lotka-Volterra model), and the cost functional for the slow game is convex (which is typical), then the resulting optimal control problem for the slow leader remains a standard convex optimization problem [37], which is computationally tractable. □

F.8 Proof of Theorem 5: Bifurcation of Fast Subsystem Equilibria

Proof. The bifurcation analysis is formalized using center manifold reduction [29].

Step 1: Setup and Jacobian. The tumor-free equilibrium is identified as $\mathbf{z}_{free} = (0, s(\mathbf{M})/d)$. The Jacobian of the closed-loop fast dynamics at $\mathbf{z}_{free}$ has eigenvalues $\lambda_1 = a - n(s(\mathbf{M})/d) - \kappa(p^*)u_d^*$

and $\lambda_2 = -d$. A bifurcation occurs when the real part of an eigenvalue crosses zero. Here, this happens when $\lambda_1 = 0$, which defines the critical threshold s_{crit}.

Step 2: Center Manifold Reduction. At the bifurcation point $s = s_{crit}$, the system has a one-dimensional center eigenspace associated with $\lambda_1 = 0$. A coordinate transformation is performed to place the equilibrium at the origin and align the basis with the eigenvectors of the Jacobian. This allows isolation of the dynamics on the one-dimensional center manifold, which governs the qualitative behavior near the bifurcation.

Step 3: Derivation of the Normal Form and Nondegeneracy. The dynamics restricted to the center manifold [29] can be written as $\dot{x} = \mu x + bx^2 + \mathcal{O}(x^3)$, where μ parameterizes $s(M) - s_{crit}$. The nondegeneracy conditions for a transcritical bifurcation are: (i) a simple zero eigenvalue with nonzero transversality $\frac{d\lambda_1}{d\mu}|_{\mu=0} \neq 0$, (ii) the quadratic coefficient $b \neq 0$. Direct differentiation of $\lambda_1 = a - nC_I^*(\mu) - \kappa(p^*)u_d^*$ with respect to μ gives $\frac{d\lambda_1}{d\mu} = -n\frac{dC_I^*}{d\mu} = -n/d \neq 0$. Calculation of b via standard center-manifold formulas yields $b = -\frac{n}{d}\left(mh + r/(h + C_T^*)\right) +$ higher-order corrections evaluated at $C_T^* = 0$, which is strictly negative for admissible parameters. Hence the nondegeneracy and transversality conditions hold, implying a transcritical bifurcation with exchange of stability.

Step 4: Conclusion. A negative quadratic coefficient ($b < 0$) in the normal form is the defining condition for a supercritical (or forward) transcritical bifurcation [28]. This result, derived by construction, rigorously proves that no subcritical behavior or associated hysteresis can occur at this bifurcation point. □

F.9 Complete Proof with Explicit Bounds

Proof. We construct a strict Lyapunov function and derive all bounds explicitly, matching the constants in the theorem statement.

Step 1: Lyapunov function. Define

$$V(\mathbf{z}) = \tfrac{1}{2}C_T^2 + \beta(C_I - C_I^*)^2,$$

where $C_I^* = s(M)/d$ and $\beta > 0$ is to be chosen.

Step 2: Derivative along trajectories. Along (1) with fixed Nash pair (u_d^*, p^*),

$$\begin{aligned}\dot{V} &= C_T\,\dot{C}_T + 2\beta(C_I - C_I^*)\,\dot{C}_I \\ &= C_T\big[aC_T(1 - bC_T) - nC_TC_I - \kappa(p^*)u_d^*C_T\big] \\ &\quad + 2\beta(C_I - C_I^*)\big[s(M) - dC_I + \tfrac{rC_TC_I}{h+C_T} - mC_TC_I\big].\end{aligned}$$

Step 3: Margin use and basic bounds. From $s(M) \geq s_{crit,\max} + \delta = \frac{d}{n}a + \delta$ and $C_I^* = s(M)/d$ we get $a - nC_I^* \leq -\frac{n}{d}\delta$. Also $aC_T^2(1 - bC_T) \leq aC_T^2$ for $C_T \leq 1/b$, and $-\kappa(p^*)u_d^*C_T^2 \leq 0$.

Step 4: Cross terms via Young. The terms $2\beta(C_I - C_I^*)\frac{rC_TC_I}{h+C_T}$ and $-2\beta m(C_I - C_I^*)C_TC_I$ are bounded by Young's inequality: for any $\varepsilon_1, \varepsilon_2 > 0$,

$$2\beta(C_I - C_I^*)\tfrac{rC_TC_I}{h+C_T} \leq \varepsilon_1 C_T^2 + c_1(C_I - C_I^*)^2,$$
$$2\beta m(C_I - C_I^*)C_TC_I \leq \varepsilon_2 C_T^2 + c_2(C_I - C_I^*)^2,$$

on the positively invariant set where $C_T \leq 1/b$ and $C_I \leq C_{I,\max}$ (Appendix Lemma 7).

Step 5: Parameter choices and decay rate. Choose $\varepsilon_1, \varepsilon_2$ small so that

$$-(n\delta/d)C_T^2 + (\varepsilon_1 + \varepsilon_2)C_T^2 \leq -\tfrac{n\delta}{2d}\,C_T^2.$$

Then choose β large enough so that $-d\beta + (c_1 + c_2) \leq -\frac{d}{2}\,\beta$. Consequently,

$$\dot{V} \;\leq\; -\min\Big\{\tfrac{n\delta}{2d},\,\tfrac{d}{2}\Big\}\left(C_T^2 + (C_I - C_I^*)^2\right) \;\leq\; -\lambda\, V,$$

with $\lambda = \min\{\frac{n\delta}{2d},\,\frac{d}{2}\}$.

Step 6: Exponential estimate and norm equivalence. Gronwall yields $V(\mathbf{z}(t)) \leq V(\mathbf{z}(0))e^{-\lambda t}$. Using norm equivalence between V and $\|\mathbf{z} - \mathbf{z}_{free}\|^2$ on the invariant set and the explicit choice of $\beta \geq \beta_{\min}$ from the theorem, one obtains

$$\|\mathbf{z}(t) - \mathbf{z}_{free}\| \;\leq\; K\,e^{-\lambda t}\,\|\mathbf{z}(0) - \mathbf{z}_{free}\|,$$

with K as in the theorem statement. This completes the proof. □

F.10 Uniform normal hyperbolicity: regularity and spectral gap

Lemma 1 (Uniform C^1 regularity and Hurwitz bounds along $\mathcal{M}_0$)**.** *Under Assumptions 1–5 and the margin condition of Theorem 6, the equilibrium map $\mathbf{M} \mapsto \mathbf{z}_{free}(\mathbf{M})$ is C^1 on $\mathcal{M}_{safe}$; the equilibrium feedbacks (u_d^*, p^*) are locally Lipschitz in $\mathbf{z}$ and C^1 in $(\mathbf{z}, \mathbf{M})$ on neighborhoods of $\mathcal{M}_0$, and the closed-loop Jacobian $J_f(\mathbf{M})$ evaluated at $\mathbf{z}_{free}(\mathbf{M})$ is Hurwitz with a* uniform *spectral gap* $\operatorname{Re} \lambda_i \le -\lambda_{\min} < 0$ *on $\mathcal{M}_{safe}$. Moreover, there exist constants L_u, L_p, L_J such that $\|D_{\mathbf{z}}(u_d^*, p^*)\| \le L_u$, $\|D_{\mathbf{M}}(u_d^*, p^*)\| \le L_p$, and $\|\frac{d}{d\mathbf{M}} J_f(\mathbf{M})\| \le L_J$ on $\mathcal{M}_{safe}$.*

Proof. The map $\mathbf{M} \mapsto \mathbf{z}_{free}(\mathbf{M}) = (0, s(\mathbf{M})/d)$ is C^1 by Assumption 2. Under Assumption 5, the coupled algebraic Riccati equations admit stabilizing solutions that depend smoothly on the linearized data [11]; by the implicit function theorem the feedbacks (u_d^*, p^*) inherit C^1 dependence on $(\mathbf{z}, \mathbf{M})$ in neighborhoods where stabilizability/detectability hold. Theorem 6 provides an explicit Lyapunov function with rate $\lambda > 0$ uniform on $\mathcal{M}_{safe}$; standard Lyapunov arguments imply a uniform Hurwitz bound on $J_f(\mathbf{M})$ and thus normal hyperbolicity. The derivative bounds follow from compactness of $\mathcal{M}_{safe}$ and continuity. □

F.11 Proof of Theorem 7: Hierarchical Decomposition via Fenichel Theory

Proof. The proof formalizes the application of geometric singular perturbation theory and derives the first-order correction to the cost functional. Steps 1-3, which establish the existence of a normally hyperbolic slow manifold, are identical to the proof provided in the previous revision and are summarized here for completeness. The main addition is the explicit derivation of the cost functional correction.

Steps 1-3: Existence of a Normally Hyperbolic Slow Manifold (Summary). The system is cast into standard singular perturbation form [22]. The critical manifold $\mathcal{M}_0$ is identified as the graph of the fast-subsystem equilibrium, $\mathbf{z} = \mathbf{z}_{free}(\mathbf{M})$. The stability results from Theorem 6 and Proposition 6 confirm that this manifold is normally hyperbolic. Fenichel's theorems [31] then guarantee the existence of an invariant slow manifold $\mathcal{M}_\epsilon$, which is $O(\epsilon)$-close to $\mathcal{M}_0$. The state on this manifold can be written as $\mathbf{z} = h_\epsilon(\mathbf{M}) = \mathbf{z}_{free}(\mathbf{M}) + \epsilon h_1(\mathbf{M}) + O(\epsilon^2)$.

Step 4: Derivation of the First-Order Manifold Correction $h_1(\mathbf{M})$. To find the $O(\epsilon)$ correction term, $h_1(\mathbf{M})$, the manifold invariance condition is used. The full dynamics are $\dot{\mathbf{z}} = f_f(\mathbf{z}, \mathbf{M})$ and $\dot{\mathbf{M}} = \epsilon f_s(\mathbf{M})$. If a trajectory lies on the manifold $\mathcal{M}_\epsilon$, it must satisfy $\dot{\mathbf{z}} = D_{\mathbf{M}} h_\epsilon(\mathbf{M})\dot{\mathbf{M}}$. Substituting the dynamics and the expansion for h_ϵ:

$$f_f(\mathbf{z}_{free} + \epsilon h_1, \mathbf{M}) = \left[D_{\mathbf{M}}\mathbf{z}_{free}(\mathbf{M}) + \epsilon D_{\mathbf{M}} h_1(\mathbf{M}) + O(\epsilon^2)\right] \cdot \epsilon f_s(\mathbf{M})$$

A Taylor expansion in ϵ around $\epsilon = 0$ is now performed. The $O(1)$ terms cancel out because $f_f(\mathbf{z}_{free}, \mathbf{M}) = 0$. The $O(\epsilon)$ terms are of interest:

$$\left.\frac{\partial f_f}{\partial \mathbf{z}}\right|_{\mathbf{z}_{free}} h_1(\mathbf{M}) = D_{\mathbf{M}}\mathbf{z}_{free}(\mathbf{M}) f_s(\mathbf{M})$$

This is the "invariance equation". Let $J_f(\mathbf{M})$ be the Jacobian of the fast dynamics evaluated on the critical manifold. Since the manifold is normally hyperbolic, $J_f(\mathbf{M})$ is invertible. Therefore, the first-order correction term can be solved:

$$h_1(\mathbf{M}) = J_f(\mathbf{M})^{-1} \left[D_{\mathbf{M}}\mathbf{z}_{free}(\mathbf{M}) f_s(\mathbf{M})\right]$$

This provides an explicit expression for the first-order deviation of the slow manifold from the critical manifold.

Step 5: Derivation of the Cost Functional Correction. The leader's objective is $J_s = \int_0^T \ell_s(\mathbf{z}(t), \mathbf{M}(t), \mathbf{u}_m(t))dt$. On the slow manifold, the state is given by the expansion $\mathbf{z}(t) = h_\epsilon(\mathbf{M}(t)) = \mathbf{z}_{eq}(\mathbf{M}(t)) + \epsilon h_1(\mathbf{M}(t)) + \mathcal{O}(\epsilon^2)$. A Taylor expansion of the integrand ℓ_s around the

reduced solution $(\mathbf{z}_{eq}(\bar{\mathbf{M}}), \bar{\mathbf{M}})$ can therefore be performed:

$$\begin{aligned}\ell_s(\mathbf{z}^\epsilon, \mathbf{M}^\epsilon) \approx & \ell_s(\mathbf{z}_{eq}(\bar{\mathbf{M}}), \bar{\mathbf{M}}) \\ & + \left.\frac{\partial \ell_s}{\partial \mathbf{z}}\right|_0 (\mathbf{z}^\epsilon - \mathbf{z}_{eq}(\bar{\mathbf{M}})) + \dots \\ = & \ell_s(\mathbf{z}_{eq}(\bar{\mathbf{M}}), \bar{\mathbf{M}}) \\ & + \epsilon \left(\left.\frac{\partial \ell_s}{\partial \mathbf{z}}\right|_0^T h_1(\bar{\mathbf{M}}) \right) + \mathcal{O}(\epsilon^2)\end{aligned}$$

Integrating this expansion gives the cost functional expansion. The $\mathcal{O}(1)$ term yields the reduced cost $\bar{J}_s = \int_0^T \ell_s(\mathbf{z}_{eq}(\bar{\mathbf{M}}), \bar{\mathbf{M}})dt$. The $\mathcal{O}(\epsilon)$ term provides the first-order correction, defining the constant C_4 from the theorem statement as:

$$C_4 = \int_0^T \left(\left.\frac{\partial \ell_s}{\partial \mathbf{z}}\right|_0^T h_1(\bar{\mathbf{M}}(t)) \right) dt$$

with $h_1(\mathbf{M})$ given *uniquely* by the invariance equation in Step 4:

$$h_1(\mathbf{M}) = J_f(\mathbf{M})^{-1} \Big[D_{\mathbf{M}} \mathbf{z}_{free}(\mathbf{M})\, f_s(\mathbf{M}) \Big].$$

This provides a quantitative, constructive formula for the impact of the fast dynamics on the leader's cost. This completes the proof. □

Theorem 12 (Absence of Subcritical Behavior)**.** *Under the biological constraints of Assumption 2, the bifurcation at $s = s_{crit}$ is supercritical (transcritical), not subcritical. No hysteresis or catastrophic transitions occur.*

Proof. The nature of a one-dimensional bifurcation is determined by the sign of the first non-zero nonlinear coefficient in the normal form of the dynamics on the center manifold [28]. As rigorously derived in the proof of Theorem 5 (Appendix F.8), the reduced dynamic is $\dot{x} = bx^2 + \mathcal{O}(x^3)$, with the coefficient b being strictly negative for all biologically relevant parameters. A negative quadratic coefficient proves the bifurcation is supercritical, thus rigorously excluding the possibility of subcritical bifurcations and associated phenomena like hysteresis or catastrophic jumps in the system's state. □

F.12 Proof of Theorem 13: Quantitative Robustness to Clonal Heterogeneity

Theorem 13 (Quantitative Robustness to Clonal Heterogeneity)**.** *Consider a controller designed for the continuous phenotypic model* (1) *but applied to an N-clone heterogeneous system. Define the clonal parameter variations*

$$\Delta_a := \max_i |a_i - a|, \qquad \Delta_\kappa := \max_i |\kappa_i - \kappa_0|.$$

If the microbiome margin satisfies

$$\delta > \delta_{crit} := \frac{d}{n}\Big(\Delta_a + \Delta_\kappa\, u_{d,\max}\Big),$$

then all clones are exponentially eradicated with rate $\lambda \geq \frac{n(\delta - \delta_{crit})}{2d}$.

Proof. Robustness to N distinct tumor clones is proven by extending the Lyapunov argument from Theorem 6 to the N-clone system.

Step 1: N-Clone Lyapunov Function. Define the Lyapunov function candidate $V = \sum_{i=1}^N C_i + \frac{\beta}{2}(C_I - C_I^*)^2$, where $C_I^* = s(\mathbf{M})/d$. This function is positive definite and radially unbounded.

Step 2: Time Derivative $\dot{V}$. The derivative along the N-clone dynamics given in the main text is computed: $\dot{V} = \sum_{i=1}^N \dot{C}_i + \beta(C_I - C_I^*)\dot{C}_I$.

Step 3: Exploit the Margin δ. The growth term for each clone i is $C_i[a_i - nC_I - \dots]$. The margin condition from Theorem 13 ensures that the immune influx $s(\mathbf{M})$ is maintained such that for every clone i, the term $a_i - nC_I^*$ is uniformly negative, bounded above by a negative constant related to the margin δ. This margin δ is chosen to be large enough to absorb the clonal parameter variations Δ_a and Δ_κ.

Step 4: Bound Cross-Terms and Conclude. As in the proof of Theorem 6, the derivative $\dot{V}$ contains negative-definite primary terms and cross-terms. Young's inequality is used to show that for a sufficiently large choice of the parameter β, all cross-terms can be absorbed into the negative-definite terms. This renders $\dot{V}$ negative definite for all non-zero tumor populations. By Lyapunov's direct method, this proves that all clone populations C_i must converge to zero, ensuring global asymptotic stability of the tumor-free state. □

F.13 Proof of Theorem F.13: Universal Bifurcation Control Principle

Proof. We establish the universality of the bifurcation control principle for general nonlinear games through a complete constructive proof.

Part 1: Existence of Nash Equilibrium via Fixed-Point Theory. Consider general cost functionals:

$$J_d^{NL}(u_d, p; \mathbf{M}) = \int_0^{T_f} \ell_d(C_T(\tau), C_I(\tau), u_d(\tau))d\tau \tag{16}$$

$$J_p^{NL}(u_d, p; \mathbf{M}) = \int_0^{T_f} \ell_p(C_T(\tau), p(\tau))d\tau \tag{17}$$

where $\ell_d : \mathbb{R}^3 \to \mathbb{R}$ and $\ell_p : \mathbb{R}^2 \to \mathbb{R}$ satisfy:

1. **Smoothness:** $\ell_d, \ell_p \in C^2$ with locally Lipschitz gradients
2. **Coercivity:** $\ell_d(C_T, C_I, u_d) \geq \alpha_d(C_T^2 + u_d^2) - \beta_d$
3. **Growth:** $|\ell_p(C_T, p)| \leq \gamma_p(1 + C_T^q + p^r)$ for some $q, r > 0$

Define the best-response correspondences:

$$BR_d(p) = \arg\min_{u_d \in U_d} J_d^{NL}(u_d, p; \mathbf{M}) \tag{18}$$

$$BR_p(u_d) = \arg\min_{p \in P} J_p^{NL}(u_d, p; \mathbf{M}) \tag{19}$$

By coercivity, these sets are non-empty. By the growth condition and compactness of $U_d \times P$, they are upper hemicontinuous with convex values. The Kakutani-Fan-Glicksberg theorem guarantees existence of a fixed point $(u_d^*, p^*) \in BR_d(p^*) \times BR_p(u_d^*)$.

Part 2: Critical Threshold Construction. The linearization of the closed-loop dynamics around equilibrium $\mathbf{z}_{eq}$ yields:

$$\mathbf{J}(\mathbf{z}_{eq}) = D_{\mathbf{z}} f_f|_{\mathbf{z}_{eq}} + D_{\mathbf{z}} u_d^*|_{\mathbf{z}_{eq}} D_{u_d} f_f|_{\mathbf{z}_{eq}} + D_{\mathbf{z}} p^*|_{\mathbf{z}_{eq}} D_p f_f|_{\mathbf{z}_{eq}} \tag{20}$$

The critical threshold occurs at $\det(\mathbf{J}) = 0$. Using implicit differentiation on the first-order optimality conditions:

$$\nabla_{u_d} H_d = \nabla_{u_d} \ell_d + \lambda_d^T \nabla_{u_d} f_f = 0 \tag{21}$$

$$\nabla_p H_p = \nabla_p \ell_p + \lambda_p^T \nabla_p f_f = 0 \tag{22}$$

Solving for the feedback gains and substituting into the Jacobian determinant condition yields:

$$s_{crit}^{NL}(\mathbf{z}) = \frac{d}{n}\left(a - \frac{\kappa_0}{1 + p^*(\mathbf{z})} u_d^*(\mathbf{z})\right) \tag{23}$$

The global threshold is $s_{crit,\max}^{NL} = \sup_{\mathbf{z} \in \mathcal{D}} s_{crit}^{NL}(\mathbf{z})$.

Part 3: Quantitative Bound on Deviation from LQ Case. Let $\ell_d^{LQ}(C_T, C_I, u_d) = q_c C_T^2 + r_d u_d^2$ be the quadratic approximation. The difference in critical thresholds arises from the difference in optimal feedback laws. Using Taylor expansion:

$$u_d^{NL} - u_d^{LQ} = -\left(\frac{\partial^2 \ell_d}{\partial u_d^2}\right)^{-1} \left(\frac{\partial^3 \ell_d}{\partial u_d^3}\right) (u_d^{LQ})^2 + O((u_d^{LQ})^3) \tag{24}$$

This leads to:

$$|s_{crit,\max}^{NL} - s_{crit,\max}^{LQ}| \leq \frac{d\kappa_0}{n} \sup_{\mathbf{z}} |u_d^{NL} - u_d^{LQ}| \leq K \|\nabla^2(\ell_d - \ell_d^{LQ})\|_\infty \tag{25}$$

where $K = \frac{d\kappa_0 u_{d,\max}}{n r_d}$.

Part 4: Global Stability via Lyapunov Theory. For $s(M) > s_{crit,\max}^{NL} + \delta$, construct the Lyapunov function:

$$V^{NL}(\mathbf{z}) = \int_0^{C_T} \ell_d(\xi, C_I, u_d^*(\xi, C_I)) d\xi + \beta(C_I - C_I^*)^2 \tag{26}$$

The time derivative:

$$\dot{V}^{NL} = \ell_d(C_T, C_I, u_d^*)\dot{C}_T + 2\beta(C_I - C_I^*)\dot{C}_I \tag{27}$$

$$= \ell_d(C_T, C_I, u_d^*)[f_{C_T} - \kappa(p^*)u_d^* C_T] + 2\beta(C_I - C_I^*) f_{C_I} \tag{28}$$

Using the coercivity of ℓ_d and the margin condition:

$$\dot{V}^{NL} \leq -\alpha_d C_T^2 - \beta n \delta C_I^2 + \text{bounded cross terms} \tag{29}$$

Choosing $\beta > \frac{\|\nabla \ell_d\|^2}{4\alpha_d n \delta}$ ensures the negative definite terms dominate, yielding $\dot{V}^{NL} < -\gamma(C_T^2 + (C_I - C_I^*)^2)$ for some $\gamma > 0$. This establishes global asymptotic stability. □

F.14 Stochastic system definitions and assumptions

Assumption 6 (Diffusion regularity)**.** *The diffusion coefficients are locally Lipschitz with polynomial growth, and the multiplicative noise vanishes at the tumor-free equilibrium on the safe set:*

- $\sigma_z(\mathbf{z}, \mathbf{M})$ *locally Lipschitz in* $\mathbf{z}$ *uniformly on compact* $\mathbf{M}$*;* $\|\sigma_z(\mathbf{z}, \mathbf{M})\|_F^2 \leq L_z^2(1 + \|\mathbf{z}\|^2)$
- $\sigma_M(\mathbf{M})$ *locally Lipschitz with polynomial growth*
- $\sigma_z(\mathbf{z}_{free}, \mathbf{M}) = 0$ *for all* $\mathbf{M} \in \mathcal{M}_{safe}$

Theorem 14 (Stability in probability; almost-sure under recurrence)**.** *Under Assumption 6 and standing regularity, there exists a constant* $K > 0$ *(depending on* $\nabla^2 V$ *on a compact invariant set) and a convergence rate* $\lambda > 0$ *from the deterministic Lyapunov analysis such that if* $L_z < \sqrt{\lambda/K}$*, then* $\mathbf{z}_{free}$ *is globally asymptotically stable in probability. If in addition the process is non-explosive, admits a Foster–Lyapunov function ensuring positive recurrence of neighborhoods of* $\mathbf{z}_{free}$*, and satisfies* $\sigma_z(\mathbf{z}_{free}, \mathbf{M}) = 0$*, then* $\mathbf{z}(t) \to \mathbf{z}_{free}$ *almost surely.*

F.15 Proof of Theorem 14: Almost-Sure Eradication Under Stochastic Perturbations

Proof. The proof extends the Lyapunov argument to the stochastic case using Itō's formula [38].

Step 1: Lyapunov Function and Itō's Formula. The same Lyapunov function $V(\mathbf{z}) = \frac{1}{2}C_T^2 + \frac{1}{2}\beta(C_I - C_I^*)^2$ from the deterministic analysis is used. The stochastic dynamics are given by the Itō SDE $d\mathbf{z}_t = f_f(\mathbf{z}_t)dt + \sigma_z(\mathbf{z}_t)dW_t$. Applying Itō's formula [38] to $V(\mathbf{z}_t)$ yields:

$$dV = \mathcal{L}V dt + (\nabla V)^T \sigma_z dW_t$$

where $\mathcal{L}V$ is the infinitesimal generator.

Step 2: Explicitly Writing the Infinitesimal Generator. The infinitesimal generator is given by:

$$\mathcal{L}V(\mathbf{z}) = \underbrace{(\nabla V)^T f_f(\mathbf{z})}_{\text{Drift Term}} + \underbrace{\frac{1}{2}\text{Tr}\left(\sigma_z^T(\nabla^2 V)\sigma_z\right)}_{\text{Diffusion Term}}$$

Step 3: Bounding the Drift Term. From the rigorous analysis in the proof of Theorem 6, the drift term is known to be the time derivative of V in the deterministic system, $\dot{V}_{det}$. It was proven to be negative definite:

$$(\nabla V)^T f_f(\mathbf{z}) = \dot{V}_{det} \leq -\gamma_1 C_T^2 - \gamma_2 (C_I - C_I^*)^2$$

This can be expressed in terms of the Lyapunov function itself. Since V is quadratic, there exists a constant $\lambda > 0$ such that $\dot{V}_{det} \leq -\lambda V(\mathbf{z})$.

Step 4: Bounding the Diffusion Term. The Hessian of V is $\nabla^2 V = \text{diag}(1, \beta)$. The diffusion term is therefore:

$$\frac{1}{2}\text{Tr}(\sigma_z^T \nabla^2 V \sigma_z) = \frac{1}{2}\text{Tr}(\sigma_z^T \left[\begin{smallmatrix} 1 & 0 \\ 0 & \beta \end{smallmatrix}\right] \sigma_z)$$

The noise intensity bound $\|\sigma_z(\mathbf{z})\|_F^2 \leq L_z^2(1 + \|\mathbf{z}\|^2)$ is used, where $\|\cdot\|_F$ is the Frobenius norm.

$$\begin{aligned}\frac{1}{2}\text{Tr}(\sigma_z^T \nabla^2 V \sigma_z) &\leq \frac{1}{2}\|\nabla^2 V\|_F \|\sigma_z \sigma_z^T\|_F \\ &\leq \frac{1}{2}\max(1, \beta) L_z^2 (1 + \|\mathbf{z}\|^2)\end{aligned}$$

Since $\|\mathbf{z}\|^2$ is quadratically related to $V(\mathbf{z})$, this can be written as $\leq K L_z^2(1 + V(\mathbf{z}))$ for some constant K.

Step 5: Combining Terms and Invoking Stability Theorem. Combining the bounds for the drift and diffusion terms yields:

$$\mathcal{L}V \leq -\lambda V + K L_z^2(1 + V) = K L_z^2 - (\lambda - K L_z^2)V$$

For almost-sure stability, the coefficient of V must be positive. This requires $\lambda - K L_z^2 > 0$, which leads to the condition on the noise intensity:

$$L_z < \sqrt{\lambda / K} := L_{crit}$$

If this condition holds, let $\lambda' = \lambda - K L_z^2 > 0$. Then $\mathcal{L}V \leq K L_z^2 - \lambda' V$. This shows that for any $V > K L_z^2 / \lambda'$, $\mathcal{L}V$ is negative. By the stochastic Lyapunov stability theorem [39], this guarantees that the trajectories are bounded in probability and will enter the set $\{\mathbf{z} : V(\mathbf{z}) \leq K L_z^2 / \lambda'\}$ almost surely. Since the noise vanishes at the equilibrium ($\sigma_z(\mathbf{z}_{free}) = 0$), a stronger result based on LaSalle's principle for SDEs [39] guarantees that the trajectories will converge to the tumor-free equilibrium $\mathbf{z}_{free}$ almost surely. □

Additional remarks on stochastic robustness

Remark 1 (Physical interpretation of noise vulnerability)**.** *The critical noise threshold L_{crit} is inversely proportional to the maximum curvature ($\|\nabla^2 V(z)\|$) of the Lyapunov function V. Regions of high curvature are more sensitive to noise; engineering a wide basin around $\mathbf{z}_{free}$ improves robustness.*

Remark 2 (Large deviations and margin design)**.** *The bound $L_z < L_{crit}$ guarantees local stochastic stability near the slow manifold. Rare large deviations can be mitigated by enlarging the ecological margin: for target risk $\alpha \in (0, 1)$ over horizon T, take $\delta = \delta_0 + c\sqrt{\log(1/\alpha)}$, with c set by effective noise and Lyapunov curvature.*

Remark 3 (Practical noise levels)**.** *Biological noise is typically small relative to deterministic drift near $\mathbf{z}_{free}$; numerically, the unified bound $L_z < \sqrt{\lambda / K}$ holds with margin.*

F.16 Validity on clinical timescales

Proof. The proof formalizes the argument that the hierarchical decomposition remains valid over a clinically relevant time horizon. The validity depends on the slow state, $\mathbf{M}(t)$, remaining within the safe set $\mathcal{M}_{safe}$.

Step 1: Slow System Dynamics and Stability. The slow dynamics are given by $\dot{\mathbf{M}} = \epsilon f_s(\mathbf{M}, \mathbf{u}_m)$. Let the slow controller be designed to maintain an equilibrium $\mathbf{M}^*$ inside $\mathcal{M}_{safe}$. The stability of these dynamics is governed by the Jacobian of f_s at $\mathbf{M}^*$, denoted $J_s(\mathbf{M}^*)$. The eigenvalues of the underlying microbiome interaction matrix $\mathbf{A}$ are the dominant factor in determining the eigenvalues of J_s. Let $\lambda_{max}(\mathbf{A})$ be the largest real part of the eigenvalues of $\mathbf{A}$. This value characterizes the intrinsic timescale of the microbiome's drift.

Step 2: Bounding the Drift of the Slow State. Let $\mathbf{M}(0)$ be the initial state in $\mathcal{M}_{safe}$. The solution to the slow dynamics can be bounded using a Grönwall-type inequality. For a given control $\mathbf{u}_m(t)$, the distance of the trajectory from a reference point $\mathbf{M}_{ref} \in \mathcal{M}_{safe}$ evolves as:

$$\|\mathbf{M}(t) - \mathbf{M}_{ref}\| \leq \|\mathbf{M}(0) - \mathbf{M}_{ref}\| e^{\epsilon L_s t}$$

where L_s is the Lipschitz constant of f_s, which is directly related to the norm of $\mathbf{A}$. A more refined analysis using the spectral properties of $\mathbf{A}$ gives the bound:

$$\|\mathbf{M}(t) - \mathbf{M}^*\| \leq \|\mathbf{M}(0) - \mathbf{M}^*\| e^{\epsilon \lambda_{max}(\mathbf{A}) t}$$

This bound holds in a neighborhood of the equilibrium $\mathbf{M}^*$.

Step 3: Deriving the Condition on $T_{clinical}$. The hierarchical control strategy is guaranteed to be effective as long as $\mathbf{M}(t)$ remains in the safe set $\mathcal{M}_{safe}$. Let d_{safe} be the minimum distance from the starting point $\mathbf{M}(0)$ to the boundary of $\mathcal{M}_{safe}$. The time T must be found such that for all $t \in [0, T]$, $\|\mathbf{M}(t) - \mathbf{M}(0)\| < d_{safe}$. From the bound above, the state will have drifted a distance of approximately $\|\mathbf{M}(0) - \mathbf{M}^*\|(e^{\epsilon \lambda_{max}(\mathbf{A})T} - 1)$. This drift is required to be less than d_{safe}.

$$\|\mathbf{M}(0) - \mathbf{M}^*\|(e^{\epsilon \lambda_{max}(\mathbf{A})T} - 1) < d_{safe}$$

For small exponents, $e^x - 1 \approx x$, so this is approximately:

$$\|\mathbf{M}(0) - \mathbf{M}^*\| \epsilon \lambda_{max}(\mathbf{A}) T < d_{safe}$$

Solving for T gives:

$$T < \frac{d_{safe}}{\|\mathbf{M}(0) - \mathbf{M}^*\| \epsilon \lambda_{max}(\mathbf{A})}$$

This shows that the time of validity T is inversely proportional to $\epsilon \lambda_{max}(\mathbf{A})$. The term $T_{clinical}$ in the theorem statement represents this characteristic time. If $\lambda_{max}(\mathbf{A}) \leq 0$ (i.e., the microbiome is intrinsically stable), the validity horizon is, in principle, infinite. If $\lambda_{max}(\mathbf{A}) > 0$, this provides a concrete upper bound on the time for which the therapy can be trusted without needing feedback and adjustment of the slow control strategy. This formalizes the intuition that the therapy is valid as long as the clinical timescale is shorter than the timescale of the microbiome's natural drift. □

F.17 Proposition and proof: No Stable Spatial Sanctuaries

Proposition 1 (Sufficient condition preventing stable spatial sanctuaries)**.** *Consider the reaction–diffusion extension on a smooth bounded domain Ω with Neumann boundary conditions and immune diffusion $D_I(\mathbf{x}) \geq D_{I,\min} > 0$. Let $\lambda_1(\Omega)$ be the first nonzero Neumann eigenvalue of $-\Delta$. If the spatial average satisfies*

$$\bar{s} := \frac{1}{|\Omega|} \int_\Omega s(M(\mathbf{x}))\, d\mathbf{x} \;\geq\; s_{crit,\max} + \frac{\gamma}{\lambda_1(\Omega)} \frac{\mathrm{Var}[s(M(\cdot))]}{D_{I,\min}},$$

for an explicit constant $\gamma > 0$ depending only on model parameters, then the tumor-free state is globally asymptotically stable and no stable spatial refugia persist.

F.18 Proof of Proposition 1: No Stable Spatial Sanctuaries

Proof. We establish the absence of spatial sanctuaries through energy methods and flux balance arguments.

Part 1: Reaction-Diffusion Formulation. Consider the spatially extended system on domain $\Omega \subset \mathbb{R}^3$:

$$\frac{\partial C_T}{\partial t} = D_T \nabla^2 C_T + a C_T(1 - b C_T) - n C_T C_I - \kappa(p) u_d(\mathbf{x}) C_T \tag{30}$$

$$\frac{\partial C_I}{\partial t} = D_I \nabla^2 C_I + s(M(\mathbf{x})) - d C_I + \frac{r C_T C_I}{h + C_T} - m C_T C_I \tag{31}$$

with Neumann boundary conditions $\nabla C_T \cdot \mathbf{n} = \nabla C_I \cdot \mathbf{n} = 0$ on $\partial\Omega$.

Part 2: Spatial Energy Functional and constants. Define the energy with gradient coercivity on both fields:

$$\mathcal{E}[C_T, C_I] = \int_\Omega \left[C_T^2 + \beta(C_I - C_I^*(\mathbf{x}))^2 + \gamma|\nabla C_T|^2 + \eta|\nabla C_I|^2\right] d\mathbf{x} \tag{32}$$

where $C_I^*(\mathbf{x}) = s(M(\mathbf{x}))/d$ is the local immune equilibrium, and $\gamma, \eta > 0$ are fixed constants.

Part 3: Energy Dissipation. Computing the time derivative and integrating by parts using Neumann boundary conditions gives the diffusion dissipation terms:

$$\frac{d\mathcal{E}}{dt} \leq 2\int_\Omega \left[C_T \frac{\partial C_T}{\partial t} + \beta(C_I - C_I^*)\frac{\partial C_I}{\partial t}\right] d\mathbf{x} - 2\gamma D_T \int_\Omega |\nabla C_T|^2 d\mathbf{x} - 2\eta D_I \int_\Omega |\nabla C_I|^2 d\mathbf{x}. \tag{33}$$

Substituting the reaction terms and using $C_I^* = s(M)/d$ together with the margin condition yields

$$\frac{d\mathcal{E}}{dt} \leq -\int_\Omega \Big[(nC_I - a)C_T^2 + \beta d(C_I - C_I^*)^2\Big] d\mathbf{x} \tag{34}$$

$$- 2\gamma D_T \int_\Omega |\nabla C_T|^2 d\mathbf{x} - 2\eta D_I \int_\Omega |\nabla C_I|^2 d\mathbf{x} \tag{35}$$

$$+ \int_\Omega 2\beta(C_I - C_I^*)\Big(\tfrac{rC_T C_I}{h + C_T} - mC_T C_I\Big)\, d\mathbf{x}. \tag{36}$$

Bounding the last integral by Young's inequality on the invariant set, for any $\varepsilon_1, \varepsilon_2 > 0$,

$$2\beta(C_I - C_I^*)\tfrac{rC_T C_I}{h+C_T} \leq \varepsilon_1 C_T^2 + c_1(C_I - C_I^*)^2, \quad 2\beta m(C_I - C_I^*)C_T C_I \leq \varepsilon_2 C_T^2 + c_2(C_I - C_I^*)^2,$$

with explicit c_1, c_2 depending on (β, r, h, m) and invariant bounds. Choosing $\varepsilon_1 + \varepsilon_2 \leq (n\bar{C}_I - a)/2$ yields

$$\frac{d\mathcal{E}}{dt} \leq -\int_\Omega \Big[\tfrac{1}{2}(nC_I - a)C_T^2 + (\beta d - c_1 - c_2)(C_I - C_I^*)^2\Big] d\mathbf{x} - 2\gamma D_T \|\nabla C_T\|_2^2 - 2\eta D_I \|\nabla C_I\|_2^2.$$

Using $\bar{C}_I = \frac{1}{|\Omega|}\int C_I$ and Poincaré's inequality for $C_I - C_I^*$ combined with Jensen for C_I, one obtains a differential inequality

$$\frac{d\mathcal{E}}{dt} \leq -\alpha_1 \|C_T\|_2^2 - \alpha_2 \|C_I - C_I^*\|_2^2 - \alpha_3 \|\nabla C_I\|_2^2,$$

with $\alpha_i > 0$ provided the spatial mean satisfies

$$\bar{s} \geq s_{crit,\max} + \frac{\gamma_{exp}}{\lambda_1(\Omega)} \frac{\mathrm{Var}[s(M(\cdot))]}{D_{I,\min}}, \quad \gamma_{exp} := \frac{(c_1 + c_2)}{\beta}\frac{d}{n}.$$

This yields exponential decay of $\mathcal{E}$ and proves the claim with an explicit constant $\gamma = \gamma_{exp}$ depending only on model parameters. Dependence is via $\lambda_1(\Omega)$ and $D_{I,\min}$ only.

Part 4: Spatial Average Condition. The key insight is that the spatial average $\bar{s} = \frac{1}{|\Omega|}\int_\Omega s(M(\mathbf{x}))d\mathbf{x} > s_{crit,\max} + \delta_{spatial}$ ensures:

$$\int_\Omega nC_I d\mathbf{x} > \int_\Omega a d\mathbf{x} \tag{37}$$

This global excess of immune pressure prevents tumor persistence anywhere in the domain. Using Poincaré's inequality $\int_\Omega |u - \bar{u}|^2 \leq \lambda_1(\Omega)^{-1}\int_\Omega |\nabla u|^2$ and bounding $\mathrm{Var}[s(M(\cdot))]$, one obtains an explicit constant $\gamma = \gamma(a, n, d, r, h, m)$ such that the condition in Proposition 1 suffices. Dependence is via $\lambda_1(\Omega)$ and $D_{I,\min}$ only.

Part 5: Analysis of Low-s Regions. For a region $\Omega_{low} \subset \Omega$ where $s(M(\mathbf{x})) < s_{crit,\max}$, the flux balance equation:

$$\frac{d}{dt}\int_{\Omega_{low}} C_I d\mathbf{x} = \underbrace{\int_{\Omega_{low}} s(M) d\mathbf{x}}_{\text{local source}} + \underbrace{D_I \int_{\partial\Omega_{low}} \nabla C_I \cdot \mathbf{n} dS}_{\text{influx}} - \text{losses} \tag{38}$$

The influx term from high-s regions maintains C_I above the critical level even in Ω_{low}, provided:

$$\frac{|\Omega_{low}|}{|\Omega|} < 1 - \frac{s_{crit,\max}}{\bar{s}} \tag{39}$$

Part 6: Asymptotic Extinction. Using comparison principles and the maximum principle, we establish:

$$\lim_{t\to\infty} \sup_{\mathbf{x}\in\Omega} C_T(\mathbf{x}, t) = 0 \tag{40}$$

This proves that no stable spatial sanctuaries can form under hierarchical control. □

Proposition 2 (Empirical low-rank control structure and effective hierarchy)**.** *In numerical experiments on a 100-dimensional microbiome model, the learned optimal control policy exhibits an empirical low-rank decomposition that can be organized into a three-tier control architecture: (i) guild-level orchestration; (ii) keystone species modulation; (iii) diversity maintenance. This reduces the effective implementation dimension while preserving near-optimal performance.*

F.19 Analysis supporting Proposition 2: Empirical Multi-Scale Control Decomposition

Empirical analysis. We analyze the structure of the PINN-learned optimal control policy to reveal emergent hierarchical organization; the following statements are empirical and contingent on training.

Part 1: High-Dimensional HJB Solution. The 100-dimensional HJB equation:

$$\frac{\partial V}{\partial t} + \min_{\mathbf{u}_m} \left\{ \sum_{i=1}^{100} \frac{\partial V}{\partial M_i} f_i(\mathbf{M}, \mathbf{u}_m) + \ell(\mathbf{M}, \mathbf{u}_m) \right\} = 0 \tag{41}$$

is solved using a PINN with architecture: 101 input neurons, 10 hidden layers (512 neurons each), residual connections, and adaptive activation functions.

Part 2: Singular Value Decomposition of Control Policy. The learned control policy $\mathbf{u}_m^*(\mathbf{M})$ is analyzed via SVD:

$$\mathbf{u}_m^*(\mathbf{M}) = \sum_{i=1}^{10} \sigma_i \mathbf{v}_i \phi_i(\mathbf{M}) \tag{42}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{10}$ are singular values.

Part 3: Emergent Guild Structure. The first 3 singular vectors capture 85% of the control energy:

$$\frac{\sum_{i=1}^{3} \sigma_i^2}{\sum_{i=1}^{10} \sigma_i^2} = 0.85 \tag{43}$$

Analyzing $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ reveals they correspond to functional guilds:

- $\mathbf{v}_1$: Butyrate producers (anti-inflammatory)
- $\mathbf{v}_2$: Bacteroides group (immune modulators)
- $\mathbf{v}_3$: Firmicutes/Bacteroidetes ratio (metabolic regulation)

Part 4: Keystone Species Identification. The Hessian of the immune modulation function:

$$\mathbf{H}_{ij} = \frac{\partial^2 s(\mathbf{M})}{\partial M_i \partial M_j} \tag{44}$$

has 5 eigenvalues exceeding $10\times$ the median, identifying keystone species.

Part 5: Diversity Constraint. The learned policy maintains Shannon diversity:

$$H(\mathbf{M}) = -\sum_{i=1}^{100} \frac{M_i}{\sum_j M_j} \log\left(\frac{M_i}{\sum_j M_j}\right) > H_{crit} = 3.2 \tag{45}$$

Part 6: Dimensionality Reduction. Defining the reduced control $\tilde{\mathbf{u}} \in \mathbb{R}^3$ as projections onto the dominant singular vectors:

$$\tilde{u}_i = \mathbf{v}_i^T \mathbf{u}_m^*, \quad i = 1, 2, 3 \tag{46}$$

The performance loss is quantified:

$$J(\mathbf{u}_m^*) - J(\tilde{\mathbf{u}}) < 0.05 J(\mathbf{u}_m^*) \tag{47}$$

This quantifies an effective dimensionality reduction from 100 to 3 with near-optimal performance in our experiments. □

F.20 Proof of Theorem 13: Observer-Based Control

Assumption 7 (Uniform Observability over the Safe Set)**.** *For all trajectories with $\mathbf{M}(t) \in \mathcal{M}_{safe}$, the pair obtained by linearizing the fast subsystem along $\mathcal{M}_0$ is uniformly completely observable with respect to the output map used by the EKF (here, measurements of C_T). The process and measurement noise covariances are bounded and positive definite.*

Proof. Under Assumption 7, the EKF error dynamics are exponentially stable in a neighborhood of the slow manifold; the separation follows via an input-to-state stability (ISS) argument. 1. **Error Dynamics**: Let the observer error be $\mathbf{e}(t) = \mathbf{z}(t) - \hat{\mathbf{z}}(t)$. The dynamics of a well-designed Extended Kalman Filter (EKF) ensure that the error dynamics $\dot{\mathbf{e}}(t) = A_e \mathbf{e}(t) + \text{h.o.t.}$ are asymptotically stable, meaning $\mathbf{e}(t) \to 0$ as $t \to \infty$. The stability of the EKF is independent of the control input. 2. **Plant Dynamics as a Perturbed System**: The plant dynamics, controlled by $u_d^\star(\hat{\mathbf{z}})$, can be written as $\dot{\mathbf{z}} = f_f(\mathbf{z}, u_d^\star(\mathbf{z} - \mathbf{e}))$. Since the control law $u_d^\star$ is Lipschitz continuous, the error term can be treated as a vanishing perturbation:

$$\dot{\mathbf{z}} = f_f(\mathbf{z}, u_d^\star(\mathbf{z})) + \mathbf{w}(t)$$

where the perturbation term $\|\mathbf{w}(t)\| \leq L_u \|\mathbf{e}(t)\|$ for some Lipschitz constant L_u. 3. **Cascaded ISS Systems**: The closed-loop system for $\mathbf{z}$ is ISS with respect to the perturbation $\mathbf{w}(t)$. The overall system is a cascade of the stable error dynamics subsystem driving the ISS plant dynamics subsystem. A standard result in nonlinear control theory [23] states that a cascade of an asymptotically stable system ($\dot{\mathbf{e}}$) and an ISS system ($\dot{\mathbf{z}}$) is itself asymptotically stable. Since $\mathbf{e}(t) \to 0$, the perturbation $\mathbf{w}(t) \to 0$, and thus the state $\mathbf{z}(t)$ converges to the desired equilibrium $\mathbf{z}_{free}$. □

Proposition 3 (Quantitative separation bound)**.** *Assume: (i) the EKF error satisfies $\|\mathbf{e}(t)\| \leq C_e e^{-\alpha t} \|\mathbf{e}(0)\|$ for some $C_e \geq 1$, $\alpha > 0$; (ii) the nominal fast closed-loop is ISS with respect to an additive disturbance $\mathbf{w}$, i.e., $\|\mathbf{z}(t) - \mathbf{z}_{free}\| \leq C_z e^{-\lambda t} \|\mathbf{z}(0) - \mathbf{z}_{free}\| + \gamma \sup_{0 \leq s \leq t} \|\mathbf{w}(s)\|$ for some $C_z \geq 1$, $\lambda > 0$, and gain $\gamma > 0$; (iii) $\|\mathbf{w}(t)\| \leq L_u \|\mathbf{e}(t)\|$ with the Lipschitz constant L_u of u_d^*. Then for all $t \geq 0$,*

$$\|\mathbf{z}(t) - \mathbf{z}_{free}\| \;\leq\; C_z e^{-\lambda t} \|\mathbf{z}(0) - \mathbf{z}_{free}\| \;+\; \gamma L_u C_e \, \|\mathbf{e}(0)\| \sup_{0 \leq s \leq t} e^{-\alpha s}.$$

In particular, if $\alpha > 0$, the second term decays monotonically to $\gamma L_u C_e \|\mathbf{e}(0)\|$ and if $\alpha > \lambda$ one has exponential convergence at rate λ with a reduced overshoot. When the EKF is initialized consistently so that $\mathbf{e}(0) = 0$, the nominal rate λ is recovered.

F.21 Uniform detectability with CT-only measurements

Proposition 4 (Uniform detectability of the linearization along $\mathcal{M}_0$)**.** *Consider the linearization of the fast subsystem along $\mathcal{M}_0$ with output $y = C_T$. The pair $(A_f(\mathbf{M}), C)$ with $C = [1 \;\; 0]$ is uniformly detectable on $\mathcal{M}_{safe}$ provided $d > 0$, $n > 0$, and $s(\mathbf{M}) \geq s_{crit,\max}^{rob} + \delta$. In particular, the unobservable mode (when present) corresponds to eigenvalue $-d < 0$, hence is uniformly stable. Consequently, the EKF enjoys uniform exponential error decay under standard boundedness and linearization validity conditions.*

Lemma 2 (Lipschitz bound for the control law)**.** *On compact invariant sets, the Nash feedback $u_d^*(\mathbf{z}; \mathbf{M})$ is Lipschitz with respect to $\mathbf{z}$ uniformly in $\mathbf{M} \in \mathcal{M}_{safe}$, i.e., $\|u_d^*(\mathbf{z}_1) - u_d^*(\mathbf{z}_2)\| \leq L_u \|\mathbf{z}_1 - \mathbf{z}_2\|$ for some $L_u > 0$.*

F.22 Sharpened microbiome reachability and viability

Proposition 5 (STLC and viability under positivity)**.** *Suppose (A_{lin}, B_m) at $\mathbf{M}^* > 0$ satisfies Kalman rank, and $\mathcal{M}_{safe}$ is defined by linear inequalities $G\mathbf{M} \leq h$ with $\mathbf{M} \geq 0$. Then: (i) the controlled gLV system is small-time locally controllable on a neighborhood of $\mathbf{M}^*$ within $\mathbb{R}^k_+$; (ii) there exists a feedback $\mathbf{u}_m(\mathbf{M})$ that renders $\mathcal{M}_{safe}$ positively invariant by constructing inward-pointing vectors on each active face using the sign structure of $\mathrm{diag}(\mathbf{M})(\mathbf{g} + A\mathbf{M})$ and available columns of B_m.*

F.23 Dimension reduction with structural conditions and error bounds

Theorem 15 (Guild-structured reduction with controllability preservation)**.** *Assume: (i) block-diagonal dominance of $\mathbf{A}_{full}$ up to ϵ_{blk} with inter-guild couplings bounded by ϵ_{off}; (ii) sparse immune weights with $\|\mathbf{w}_{full} - \mathbf{P}^T\mathbf{w}\| \leq \epsilon_w$; (iii) $\|(I - \mathbf{P}^T\mathbf{P})\mathbf{M}_{full}\| \leq \epsilon_P$ along trajectories. Then for matched initial conditions, over horizons T where both models remain bounded,*

$$|s(\mathbf{M}_{full}) - s(\mathbf{P}^T\mathbf{M})| \leq \epsilon_w \, \|\mathbf{M}_{full}\| + \|\mathbf{w}\| \, \epsilon_P,$$

and the state deviation satisfies $\|\mathbf{P}\mathbf{M}_{full}(t) - \mathbf{M}(t)\| \leq C(\epsilon_{blk} + \epsilon_{off} + \epsilon_P)\,(e^{Lt} - 1)$ with constants depending on bounds of the vector fields. Controllability of $(\mathbf{A}, \mathbf{P}\mathbf{B}_m)$ is preserved if each controlled guild has at least one actuated species.

F.24 Coupled HJBs for general-sum MPE and PINN training

For fixed $\mathbf{M}$, the general-sum Markov-perfect equilibrium satisfies coupled stationary HJBs

$$\rho V_d(\mathbf{z}) = \min_{u_d \in U_d} \Big\{ \ell_d(\mathbf{z}, u_d) + \nabla V_d \cdot f_f(\mathbf{z}, u_d, p^*(\mathbf{z})) \Big\},$$
$$\rho V_p(\mathbf{z}) = \min_{p \in P} \Big\{ \ell_p(\mathbf{z}, p) + \nabla V_p \cdot f_f(\mathbf{z}, u_d^*(\mathbf{z}), p) \Big\},$$

with first-order optimality $0 = \partial_{u_d}(\ell_d + \nabla V_d \cdot f_f)$ and $0 = \partial_p(\ell_p + \nabla V_p \cdot f_f)$. The MPE fixed point requires u_d^* and p^* used in the dynamics to coincide with argmin solutions. The PINN loss augments the HJB residuals with best-response and consistency penalties:

$$\begin{aligned}\mathcal{L} = \; & \mathcal{L}^d_{HJB} + \mathcal{L}^p_{HJB} \\ & + \lambda_{BR}\big(\|\partial_{u_d} H_d\|^2 + \|\partial_p H_p\|^2\big) \\ & + \lambda_{fix} \, \|(u_d^{net} - u_d^{BR}, \, p^{net} - p^{BR})\|^2 \\ & + \lambda_{margin} \, \mathcal{L}_{margin}.\end{aligned}$$

F.25 Positivity invariance of the fast subsystem

Lemma 3 (Positivity invariance)**.** *For the fast subsystem* (1) *with $u_d \in [0, u_{d,\max}]$, $p \in [0, p_{\max}]$, solutions starting in $\mathbb{R}^2_+$ remain in $\mathbb{R}^2_+$. At $C_T = 0$, $\dot{C}_T = 0$; at $C_I = 0$, $\dot{C}_I = s(\mathbf{M}) \geq 0$. Hence the positive orthant is forward invariant by Nagumo's theorem.*

G Benchmark dataset and release plan

We propose an open benchmark to standardize eco-evolutionary control studies:

- **Deterministic ODE/PDE suites**: Parameterized fast subsystems (ODE) and reaction–diffusion PDEs with reference parameter sets for SKCM and CRC; scripts generating trajectories under specified controllers.
- **Stochastic variants**: SDE simulators with configurable σ_z, seeds, and evaluation protocols for stability in probability vs. almost-sure regimes.
- **Microbiome models**: High- and low-dimensional gLV instances with documented $\mathbf{A}, \mathbf{g}, \mathbf{B}_m$, safe sets $\mathcal{M}_{safe}$, and reachability tasks.

- **HJ BVPs**: Collocation domains, boundary conditions, and ground-truth LQ solutions for validating PINN solvers; residual calculators for coupled HJBs.
- **Observer tasks**: EKF/UKF benchmarks with measurement models (only C_T) and noise models; detectability certificates.
- **Files and schema**: JSON/YAML configs, CSV parameter tables, and standardized logs; CI to reproduce Tables 1–5.

All assets will be versioned with DOIs, include code for figure regeneration, and provide precise licenses for data/code.

H Supporting Lemmas and Ancillary Proofs

Lemma 4 (Continuity of LQ Nash feedback and small-state behavior)**.** *Under Assumption 5, the unique LQ Markov-perfect Nash equilibrium* $(u_d^*, p^*)(\mathbf{z}; \mathbf{M})$ *is affine in* $\mathbf{z}$ *in a neighborhood of* $\mathbf{z} = 0$ *with gains depending smoothly on* $\mathbf{M} \in \mathcal{M}_{safe}$*. Consequently,* $u_d^*(\mathbf{z}) \to 0$ *and* $p^*(\mathbf{z}) \to 0$ *as* $\mathbf{z} \to 0$*, uniformly over* $\mathbf{M} \in \mathcal{M}_{safe}$*. Hence* $s_{crit,\max} = \frac{d}{n}a$*.*

Proof. The coupled algebraic Riccati equations depend smoothly on the linearized data; by the implicit function theorem the stabilizing solution varies smoothly on compact sets where stabilizability/detectability hold [11]. The resulting feedback is linear in $\mathbf{z}$ near the origin, implying continuity and the stated limit. Substituting into $s_{crit}(\mathbf{z})$ yields the supremum at $\mathbf{z} = 0$ with value $\frac{d}{n}a$. □

This appendix contains proofs for lemmas and propositions that support the main theoretical results.

Lemma 5 (Lipschitz Continuity of the Fast Dynamics)**.** *The Lipschitz constant of a vector function can be bounded by the induced matrix norm of its Jacobian. The Jacobian of* $f_f(\mathbf{z}, u_d, p; \mathbf{M})$ *with respect to* $\mathbf{z} = (C_T, C_I)$ *is computed:*

$$J_{\mathbf{z}} f_f = \begin{bmatrix} a(1-2bC_T) - nC_I - \kappa(p)u_d & -nC_T \\ \frac{rhC_I}{(h+C_T)^2} - mC_I & -d + \frac{rC_T}{h+C_T} - mC_T \end{bmatrix}$$

An upper bound on the elements of this matrix over the domain $\mathcal{D}$ *is sought. Using the bounds* $0 \le C_T \le C_{T,\max} = 1/b$*,* $0 \le C_I \le C_{I,\max}$*, and* $0 \le \kappa(p) \le \kappa_0$*, the absolute value of each entry can be bounded. For instance,* $|\frac{\partial f_1}{\partial C_T}| \le a(1 + 2bC_{T,\max}) + nC_{I,\max} + \kappa_0 u_{d,\max}$*. The global Lipschitz constant* L_f *can then be bounded by any matrix norm, for instance, the maximum absolute column sum norm (*$\|\cdot\|_1$*), which yields the expression in the lemma statement after substituting the explicit value for* $C_{I,\max}$*.*

H.1 Proof of Lemma 6: Lipschitz Continuity of the Slow Dynamics

Lemma 6 (Lipschitz Continuity of the Slow Dynamics)**.** *The difference* $\|f_s(\mathbf{M}_1, \mathbf{u}_m) - f_s(\mathbf{M}_2, \mathbf{u}_m)\|$ *is analyzed:*

$$\begin{aligned} &f_s(\mathbf{M}_1) - f_s(\mathbf{M}_2) \\ &\quad = \mathit{diag}(\mathbf{M}_1)(\mathbf{g} + \mathbf{A}\mathbf{M}_1) - \mathit{diag}(\mathbf{M}_2)(\mathbf{g} + \mathbf{A}\mathbf{M}_2) \\ &\quad = \mathit{diag}(\mathbf{M}_1 - \mathbf{M}_2)\mathbf{g} \\ &\qquad + \mathit{diag}(\mathbf{M}_1)\mathbf{A}\mathbf{M}_1 - \mathit{diag}(\mathbf{M}_2)\mathbf{A}\mathbf{M}_2 \end{aligned}$$

Using the identity $\mathit{diag}(u)\mathbf{A}u - \mathit{diag}(v)\mathbf{A}v = \mathit{diag}(u - v)\mathbf{A}u + \mathit{diag}(v)\mathbf{A}(u - v)$ *yields:*

$$\begin{aligned} f_s(\mathbf{M}_1) - f_s(\mathbf{M}_2) = \mathit{diag}(\mathbf{M}_1 - \mathbf{M}_2)(\mathbf{g} + \mathbf{A}\mathbf{M}_1) \\ + \mathit{diag}(\mathbf{M}_2)\mathbf{A}(\mathbf{M}_1 - \mathbf{M}_2) \end{aligned}$$

Taking norms and using $\|\mathit{diag}(\mathbf{v})\|_{op} = \|\mathbf{v}\|_\infty \le \|\mathbf{v}\|$*:*

$$\begin{aligned} \|f_s(\mathbf{M}_1) - f_s(\mathbf{M}_2)\| \le \|\mathbf{M}_1 - \mathbf{M}_2\|(\|\mathbf{g}\| + \|\mathbf{A}\|_{op}\|\mathbf{M}_1\|) \\ + \|\mathbf{M}_2\|\|\mathbf{A}\|_{op}\|\mathbf{M}_1 - \mathbf{M}_2\| \end{aligned}$$

For $\|\mathbf{M}_1\|, \|\mathbf{M}_2\| \le R$*, it follows that* $\|f_s(\mathbf{M}_1) - f_s(\mathbf{M}_2)\| \le (\|\mathbf{g}\| + 2R\|\mathbf{A}\|_{op})\|\mathbf{M}_1 - \mathbf{M}_2\|$*. The constant provided in the lemma statement is a slightly looser but correct bound.*

H.2 Proof of Lemma 7: Global L_∞ Bounds for Fast States

Lemma 7 (Global L_∞ Bounds for Fast States)**.** ***Tumor Population*** *(C_T): The dynamics of C_T are given by $\dot{C}_T = aC_T(1 - bC_T) - nC_TC_I - \kappa(p)u_dC_T$. Since $nC_TC_I \geq 0$ and $\kappa(p)u_dC_T \geq 0$, the inequality $\dot{C}_T \leq aC_T(1 - bC_T)$ holds. By the Comparison Lemma [23], $C_T(t)$ is bounded by the solution to the logistic equation $\dot{x} = ax(1 - bx)$, which has a globally stable equilibrium at $x = 1/b$. Therefore, for any initial condition, $C_T(t) \to [0, 1/b]$, which implies $C_T(t)$ is uniformly bounded by $C_{T,\max} = \max\{C_T(0), 1/b\}$. For analysis over $\mathcal{D}$, the invariant set bound $1/b$ is used.*

Immune Population *(C_I): The dynamics are $\dot{C}_I = s(\mathbf{M}) - dC_I + C_I\left(\frac{rC_T}{h+C_T} - mC_T\right)$. Let $g(C_T) = \frac{rC_T}{h+C_T} - mC_T$. For $C_T \in [0, 1/b]$, g attains its maximum at the unique stationary point in $(0, 1/b]$ or at an endpoint. Compute $g'(C_T) = \frac{rh}{(h+C_T)^2} - m$. Setting $g'(C_T) = 0$ gives $C_T^\star = h\left(\sqrt{\frac{r}{mh}} - 1\right)$ when $\frac{r}{mh} > 1$; otherwise g is decreasing and the maximizer is $C_T = 0$. Enforcing the domain constraint yields $C_T^{\max} = \min\{\max\{0, C_T^\star\}, 1/b\}$. Thus*

$$g_{\max} = \max\left\{0,\ \frac{rC_T^{\max}}{h + C_T^{\max}} - mC_T^{\max}\right\}.$$

A closed form when $C_T^{\max} = C_T^\star$ is $g_{\max} = r\left(1 - \sqrt{\frac{mh}{r}}\right)^2 - mh\left(\sqrt{\frac{r}{mh}} - 1\right)$, which simplifies to $g_{\max} = \left(\sqrt{r} - \sqrt{mh}\right)^2 - mh\left(\frac{\sqrt{r}}{\sqrt{mh}} - 1\right) = \sqrt{rmh}\left(\sqrt{\frac{r}{mh}} - 1\right)^2$; if this point lies beyond $1/b$, use the endpoint $C_T = 1/b$ instead. Consequently, the dynamics satisfy $\dot{C}_I \leq s_{\max} - (d - g_{\max})C_I$, where $s_{\max} = \sup s(\mathbf{M})$. Provided $d > g_{\max}$, this scalar linear inequality yields the uniform bound $C_I(t) \leq C_{I,\max} := s_{\max}/(d - g_{\max})$.

Proposition 6 (Prerequisites for System Decomposition)**.** *Let the conditions of Theorem 6 hold. Then the Jacobian of the closed-loop fast subsystem, $J_f(\mathbf{M}) = \partial f_f/\partial \mathbf{z}$ evaluated at the equilibrium $\mathbf{z}_{eq}(\mathbf{M}) = \mathbf{z}_{free}(\mathbf{M})$, has eigenvalues $\lambda_i(\mathbf{M})$ satisfying $Re(\lambda_i(\mathbf{M})) \leq -\lambda_{\min} < 0$ for all $\mathbf{M} \in \mathcal{M}_{safe}$, where $\lambda_{\min} > 0$ is a uniform constant. Consequently, the slow manifold $\mathcal{M}_0 = \{(\mathbf{z}, \mathbf{M}) | \mathbf{z} = \mathbf{z}_{eq}(\mathbf{M}), \mathbf{M} \in \mathcal{M}_{safe}\}$ is normally hyperbolic.*

Proof of Proposition 6. This proposition is a direct consequence of Theorem 3 and Theorem 6. Theorem 3 established that for any $\mathbf{M} \in \mathcal{M}_{safe}$, the Jacobian $J(\mathbf{M})$ of the closed-loop fast subsystem is Hurwitz. Theorem 6, through a specific Lyapunov function, provides a quantitative lower bound on the rate of convergence, which translates to an upper bound on the real parts of the eigenvalues of $J(\mathbf{M})$. The calculation shows $\mathrm{Re}(\lambda_i) \leq -\lambda_{\min} < 0$, where $\lambda_{\min}$ depends on the stability margin δ.

The uniformity of the spectral gap follows from the uniformity of the stability margin δ over the compact set $\mathcal{M}_{safe}$. Since $\lambda_{\min}$ is a continuous function of δ and other parameters, it is bounded away from zero on $\mathcal{M}_{safe}$.

The eigenvectors of a matrix are continuous functions of its entries, provided there are no repeated eigenvalues. While we do not exclude repeated eigenvalues, the eigenspaces vary continuously. This is a standard result from matrix perturbation theory [40] and is sufficient for the application of Fenichel's theorems. □

H.3 Proposition and proof: Robustness to actuation delay

Proposition 7 (Robustness to Delay)**.** *Consider the system with a delay $\tau > 0$ in immune recruitment so that the fast dynamics depend on $s(M(t - \tau))$. If a priming phase of duration $T_p > \tau$ maintains $M(t) \in \mathcal{M}_{safe}$ and small-gain conditions $L_s/\lambda < 1$, $\lambda\tau < \pi/2$ hold (with L_s the Lipschitz constant of s and λ from Theorem 6), then the tumor-free equilibrium remains asymptotically stable.*

H.4 Proof of Proposition 7: Robustness to Delay

Proof. The proof relies on a small-gain argument for delay systems. **1. System Decomposition**: The system can be viewed as an interconnection of a stable linear system (the delay operator) and a stable nonlinear system (the fast dynamics). The fast dynamics, with input $s_t = s(M(t - \tau))$, are proven

to be globally exponentially stable (GES) in Theorem 6. The mapping from the input signal $s(\cdot)$ to the state $\mathbf{z}(\cdot)$ is input-to-state stable (ISS). **2. Small-Gain Theorem**: A variant of the small-gain theorem for time-delay systems [23] is used. Using the Lyapunov function from Theorem 6, one derives an ISS estimate $\|\mathbf{z}(t)-\mathbf{z}_{free}\| \leq c_1 e^{-\lambda t}\|\mathbf{z}(0)-\mathbf{z}_{free}\| + \gamma_f \sup_{0\leq s\leq t} |s(M(s))-s(\mathbf{M}^*)|$, where $\gamma_f = \sqrt{\frac{\max\{1,\beta\}}{\min\{1/2,\beta/2\}}}\,\frac{1}{\min\{n\delta/(2d),\, d/2\}}$ is an explicit ISS gain. The gain of the delay operator is 1, and the gain of $s(M)$ is its Lipschitz constant, L_s. Stability is guaranteed if $\gamma_f L_s < 1$ and the delay satisfies a standard phase margin bound $\lambda\,\tau < \pi/2$. **3. Priming Phase**: The conditions above are dimensionally consistent. The 'priming phase' ensures the system starts within the domain of attraction. By maintaining $M(t)$ in the safe set for $t \in [0, T_p]$ with $T_p > \tau$, for $t > T_p$ the input $s(M(t-\tau))$ to the fast dynamics corresponds to a safe microbiome state. This guarantees that the system state remains within the basin of attraction of the tumor-free equilibrium, even with the delay. □

I Model Formulation and Auxiliary Results

I.1 Verification of Regularity Conditions

The specific dynamics in this model satisfy the regularity assumptions outlined in this appendix. To capture the hierarchical and adversarial nature of these interactions, the system is modeled as a two-timescale differential game. The fast timescale represents the immediate, competitive interaction between the clinician's drug administration and the tumor's evolutionary response. The slow timescale captures the clinician's strategic manipulation of the tumor microenvironment via microbiome-based therapies. This structure allows decomposition of the complex, multi-scale problem into a tractable hierarchy of control problems, providing a rigorous foundation for designing robust and effective cancer therapies.

The formal verification of Assumptions 1–3 follows from the specific structure of the dynamics in Eqs. (1) and (2), combined with the compactness of the control sets and the smoothness of the functional forms $\kappa(p)$ and $s(\mathbf{M})$.

I.2 Uniform Boundedness of Fast States

Prior to ecological intervention, boundedness of tumor and immune populations must be established. The following lemma provides global upper bounds on the fast-timescale states, which depend solely on biological parameters and are independent of the safe set condition. The proof of this result is provided in Lemma 7 below.

I.3 Justification of Functional Forms

Remark 4 (Justification of Functional Forms)**.** The affine structure of $s(\mathbf{M}) = s_0 + \sum_i s_i M_i$ represents a first-order Taylor approximation of a more complex, nonlinear biological response. Similarly, the affine structure of $s(M) = s_0 + \sum s_i M_i$ represents a first-order Taylor approximation of a more complex, nonlinear immune response to microbiome composition. This is justified for therapeutic interventions that cause small to moderate perturbations of the microbiome around a baseline state. This simplification is justified for small to moderate perturbations of the microbiome composition around a baseline, a common scenario in therapies like probiotic administration. The form of the drug efficacy function $\kappa(p) = \kappa_0/(1+p)$ models a saturating cost-benefit relationship for resistance. Initial investments in resistance (small p) provide a significant reduction in drug efficacy, but as resistance becomes very high, further investments yield diminishing returns. These functional forms are standard in theoretical ecology and mathematical oncology [26], chosen to strike a balance between biological realism and the analytical tractability required to expose the core control principles of the system.

I.4 Nonlinear perturbations and robust thresholds

The following local result clarifies when the LQ threshold and stability conclusions persist under small nonlinear perturbations of costs/dynamics.

Proposition 8 (Local perturbation of the LQ threshold)**.** *Suppose the coupled HJBI system for a general-sum nonlinear fast game admits locally unique, Lipschitz viscosity solutions with a comparison principle in a neighborhood of the tumor-free equilibrium, and the feedback equilibrium strategies are Fréchet differentiable in a neighborhood of the LQ baseline. Then there exists a radius $\rho > 0$ such that, for perturbations of the LQ costs/dynamics of size $\leq \rho$, a critical threshold $s^{NL}_{crit,\max}$ is well-defined and satisfies*

$$|s^{NL}_{crit,\max} - s^{LQ}_{crit,\max}| \leq C \, \|(\ell_d, \ell_p, f_f) - (\ell_d^{LQ}, \ell_p^{LQ}, f_f^{LQ})\|,$$

for an explicit constant C depending on the linearized HJBI operator. If $s(M) > s^{NL}_{crit,\max} + \delta$, the tumor-free equilibrium remains locally asymptotically stable.

Equilibrium multiplicity and robust threshold. Define the *robust critical threshold* by the supremum over the equilibrium correspondence $\mathcal{E}(\mathbf{z})$:

$$s^{rob}_{crit,\max} := \sup_{\mathbf{z}\in\mathcal{D}} \sup_{(u_d,p)\in\mathcal{E}(\mathbf{z})} \frac{d}{n}\Big(a - \kappa(p)\, u_d\Big) \leq \frac{d}{n} a =: s_{crit,\max}. \tag{48}$$

Consequently, if the slow controller maintains $s(M) \geq s_{crit,\max} + \delta$, the conclusions of Theorem 6 hold for any equilibrium selection. Moreover, for any nonlinear variant satisfying Proposition 8, one has $s^{NL}_{crit,\max} \leq s^{rob}_{crit,\max} \leq s_{crit,\max}$.

Remark 5 (Justification of the Continuous Resistance Model)**.** The tumor's drug resistance level, $p(t)$, is modeled as a continuous control variable. This continuous variable p(t) is formally interpreted as the population-averaged resistance, a common and powerful abstraction in adaptive dynamics and evolutionary game theory. It represents a mean-field approximation of the underlying discrete clonal distribution. This choice is motivated by the need for analytical tractability; it allows the formulation of the tumor's adaptation as an optimal control problem, for which the powerful machinery of Hamilton-Jacobi-Bellman (HJB) theory can be leveraged. While a simplification, it will be rigorously demonstrated in Section V (Theorems 9 and 10) that controllers synthesized based on this continuous model are robustly effective when applied to a more realistic, structurally different system of discrete competing clones.

Remark 6 (Justification of LQ Game Structure)**.** *It is acknowledged that Assumption 5, which imposes a Linear-Quadratic (LQ) structure on the fast game, is a significant simplification of the complex, nonlinear biological reality. This assumption is a cornerstone of the paper's analytical tractability, as it guarantees the existence of a unique, well-behaved Markovian Nash equilibrium that can be characterized by the solution to coupled HJB-Isaacs equations (a set of coupled Riccati equations in the linear-quadratic case). It is adopted here to establish a foundational, baseline theoretical result in the most well-behaved setting. This provides a clear and rigorous understanding of the system's core strategic dynamics. The demonstrated robustness of the resulting controller to significant structural uncertainties (Section V) suggests that the principles derived from this tractable model are more general. Relaxing this assumption to general nonlinear, non-quadratic games, which necessitates advanced numerical methods for state-dependent, non-unique equilibria, is a principal direction for future research.*

I.5 Justification of Game Structure

Remark 7 (Justification of Game Structure)**.** The assumption of affine dynamics in controls and quadratic costs (a Linear-Quadratic or LQ game structure) is a cornerstone of this paper's analytical tractability. It guarantees the existence of a unique, well-behaved Nash equilibrium that can be characterized by the solution to a set of coupled Riccati-like equations (or, in this state-feedback context, the HJBI equations). While biological reality is inherently nonlinear, the LQ structure can be viewed as a local approximation of the true game around a nominal therapeutic trajectory. The uniform strong convexity/concavity, guaranteed by the quadratic control costs ($r_d, r_p > 0$), is a crucial technical condition that ensures the best-response maps are contractions, leading to the uniqueness of the equilibrium. Although restrictive, these assumptions allow the establishment of a rigorous, analytical baseline framework. Relaxing these conditions is a principal direction for future work, which will likely involve numerical solutions and analysis of games with multiple, state-dependent equilibria.

I.6 Justification of Timescale Separation

Remark 8 (Choice of ϵ and Timescale Separation Validity)**.** The singular perturbation approach hinges on a clear separation of timescales, quantified by the parameter $\epsilon \ll 1$. This is justified by comparing the characteristic times of the biological processes involved:

- **Fast Dynamics (Tumor-Immune):** T-cell activation and tumor cell killing occur on a timescale of days (1-7 days).
- **Slow Dynamics (Microbiome Modulation):** Therapeutic modulation of the gut microbiome composition through interventions like fecal microbiota transplantation (FMT) or long-term probiotic courses leads to stable changes over weeks to months (14-60 days).

The ratio of these timescales gives $\epsilon \approx (\text{days})/(\text{weeks}) \approx 1/7$ to $1/30$, yielding a plausible range of $\epsilon \in [0.03, 0.15]$. The theoretical results require $\epsilon < \epsilon_0 = \lambda_{\min}/\|D_{\mathbf{M}} f_s\|$, where ϵ_0 depends on the stability properties of the fast system and the sensitivity of the slow system. For the parameters used in this study (Table 3), $\epsilon_0 \approx 0.15 - 0.2$ is computed, which provides a valid margin for the biologically estimated range of ϵ. While rapid microbiome shifts can occur (for instance, due to antibiotics), this framework models the slower, controlled modulation used for therapeutic benefit.

J Computational Framework: Methods and Validation

Our computational pipeline leverages recent advances in high-performance in-storage computing for genomic and metagenomic data processing [41–46]. These architectures enable efficient preprocessing of large-scale biological datasets required for patient-specific parameter estimation.

To evaluate robustness, we test the control strategy against multiple structural model perturbations:

1. **Clonal Evolution**: 5-clone model with heterogeneous resistance, analyzed using methods adapted from genomic sequence comparison [47, 48]
2. **Immune Dysfunction**: Reduced T-cell infiltration and impaired activation
3. **Microbiome Complexity**: 10-species ecosystem with complex interactions, characterized using metagenomic profiling techniques [49–51]
4. **Combined Perturbations**: All above simultaneously

J.1 Solution Method via Physics-Informed Neural Networks (PINNs)

The coupled HJBI equations from Theorem 3 are nonlinear partial differential equations that do not admit closed-form solutions in general. They are solved numerically using Physics-Informed Neural Networks (PINNs).

- **Network Architecture:** For each value function (V_d, V_p) and the associated control strategies $(u_d^\star, p^\star)$, a fully connected neural network (4 hidden layers, 64 neurons/layer, $\tanh$ activation) is used to approximate the solution, e.g., $V_d(\mathbf{z}) \approx \hat{V}_d(\mathbf{z}; \boldsymbol{\theta}_d)$.
- **Loss Function:** The network parameters $\boldsymbol{\theta}$ are trained by minimizing a loss function that includes the residual of the HJBI equations, boundary conditions, and regularization terms. The primary component is the mean squared HJBI residual, evaluated at a large number of collocation points $\{\mathbf{z}_i\}$ sampled from the state space $\mathcal{D}$:

$$\mathcal{L}_{HJB}(\boldsymbol{\theta}) = \frac{1}{N_{coll}} \sum_{i=1}^{N_{coll}} \left| \begin{aligned} &\nabla \hat{V}_d(\mathbf{z}_i; \boldsymbol{\theta}_d) \cdot f_f(\mathbf{z}_i, \hat{u}_d^\star, \hat{p}^\star) \\ &\qquad + \ell_d(\mathbf{z}_i, \hat{u}_d^\star) \end{aligned} \right|^2$$

  where the controls $\hat{u}_d^\star, \hat{p}^\star$ are also derived from the network outputs via the first-order optimality conditions.
- **Training and Validation:** The ADAM optimizer is used with a learning rate of 10^{-3} on 10^4 collocation points generated via Latin hypercube sampling. Convergence is validated by monitoring the decay of the loss function and by comparing the PINN solution to solutions from finite-difference methods on coarse grids.

Proposition 9 (Trajectory Sensitivity to Control Approximation)**.** *Suppose the approximate control strategies $\hat{u}_d^*, \hat{p}^*$ from the PINN solution satisfy $\|\hat{u}_d^* - u_d^*\|_\infty \leq \epsilon_u$ and $\|\hat{p}^* - p^*\|_\infty \leq \epsilon_p$ for known bounds ϵ_u, ϵ_p. Then the error between the true optimal trajectory $\mathbf{z}(t)$ and the trajectory under the approximate controls $\hat{\mathbf{z}}(t)$ is bounded by:*

$$\|\mathbf{z}(t) - \hat{\mathbf{z}}(t)\| \leq \frac{(L_{f,u}\epsilon_u + L_{f,p}\epsilon_p)}{L_f}(e^{L_f t} - 1) \tag{49}$$

where $L_f, L_{f,u}, L_{f,p}$ are the Lipschitz constants of the fast dynamics with respect to the state and control inputs.

Proof. Let $\mathbf{z}(t)$ be the trajectory under exact optimal controls (u_d^*, p^*) and $\hat{\mathbf{z}}(t)$ be the trajectory under approximate PINN controls $(\hat{u}_d^*, \hat{p}^*)$. The error dynamics are $\dot{\mathbf{e}}(t) = \dot{\mathbf{z}}(t) - \dot{\hat{\mathbf{z}}}(t) = f_f(\mathbf{z}, u_d^*, p^*) - f_f(\hat{\mathbf{z}}, \hat{u}_d^*, \hat{p}^*)$. Adding and subtracting terms:

$$\begin{aligned}\dot{\mathbf{e}} = &[f_f(\mathbf{z}, u_d^*, p^*) - f_f(\hat{\mathbf{z}}, u_d^*, p^*)] \\ &+ [f_f(\hat{\mathbf{z}}, u_d^*, p^*) - f_f(\hat{\mathbf{z}}, \hat{u}_d^*, \hat{p}^*)]\end{aligned}$$

Taking norms and using the Lipschitz continuity of f_f in its state and control arguments (with constants L_f, $L_{f,u}$, $L_{f,p}$):

$$\|\dot{\mathbf{e}}\| \leq L_f\|\mathbf{e}\| + L_{f,u}\|u_d^* - \hat{u}_d^*\|_\infty + L_{f,p}\|p^* - \hat{p}^*\|_\infty$$

The control approximation errors are given as $\|u_d^* - \hat{u}_d^*\|_\infty \leq \epsilon_u$ and $\|p^* - \hat{p}^*\|_\infty \leq \epsilon_p$. This yields the differential inequality:

$$\|\dot{\mathbf{e}}\| \leq L_f\|\mathbf{e}\| + (L_{f,u}\epsilon_u + L_{f,p}\epsilon_p)$$

Applying the Grönwall-Bellman inequality with $\mathbf{e}(0) = 0$ gives the desired bound:

$$\|\mathbf{e}(t)\| \leq \frac{L_{f,u}\epsilon_u + L_{f,p}\epsilon_p}{L_f}(e^{L_f t} - 1)$$

□

Proposition 10 (Computational Complexity of PINN Solution)**.** *Consider the coupled HJB equations for the fast game with state dimension $n_z = 2$. Let ε_{tol} be the desired L^2 approximation error. Assuming the value functions V_d, V_p belong to a Sobolev space $W^{s,2}(\mathcal{D})$ with sufficient regularity $s > n_z/2$ and the PINN is trained with a first-order optimizer, the computational complexity to achieve this error is:*

$$\mathcal{O}\left(N_{params} \cdot N_{colloc} \cdot N_{iter}\right) \tag{50}$$

where the number of network parameters (N_{params}), collocation points (N_{colloc}), and optimization iterations (N_{iter}) scale polynomially with $1/\varepsilon_{tol}$ and the state dimension n_z.

Proof. The computational complexity is the product of three factors: the number of network parameters (N_{params}), the number of collocation points (N_{colloc}), and the number of optimization iterations (N_{iter}). 1. N_{params}: For a fully connected network with depth L and width W, $N_{params} = \mathcal{O}(LW^2)$. 2. N_{colloc}: Theoretical bounds for PINN approximation error relate the L^2 error ϵ_{tol} to the number of collocation points. For functions in the Sobolev space $W^{s,2}(\mathcal{D})$, Sobolev embedding theorems state that to achieve an error of ϵ_{tol}, $N_{colloc} = \mathcal{O}(\epsilon_{tol}^{-n_z/s})$ is needed, where n_z is the state dimension. 3. N_{iter}: For first-order optimization methods like stochastic gradient descent, the number of iterations required to reach an ϵ_{tol}-optimal solution is typically polynomial in $1/\epsilon_{tol}$, often cited as $\mathcal{O}(\epsilon_{tol}^{-2})$ for non-convex problems. Combining these gives the stated complexity. While the Sobolev regularity s is generally unknown, these bounds provide a theoretical basis for how complexity scales with desired accuracy and problem dimension. □

J.2 Model Parameterization

Code and full reproducibility details are consolidated in Appendix M.

We make the following key conclusions. SKCM priors yield higher effective n and favorable a relative to CRC (Table 4), enabling more time above threshold with the same microbiome actuation budget. This places more SKCM trajectories in the supercritical regime $s(\mathbf{M}) > s_{crit,\max}$ where

Table 3: Model Parameters; conservative threshold used uniformly: $s_{crit,\max} = \frac{d}{n}a$

Parameter	Value	Description	$\partial s_{crit,\max}/\partial(\cdot)$
a	0.5	Tumor growth rate	$\frac{d}{n}$
b	0.1	Tumor carrying capacity	0
n	0.1	Immune predation rate	$-\frac{d\,a}{n^2}$
d	0.2	Immune cell death rate	$\frac{a}{n}$
r	0.5	Immune stimulation rate	0
h	1.0	Michaelis-Menten constant	0
m	0.1	Immune inactivation rate	0
κ_0	1.0	Drug efficacy	0

the fast drift guarantees convergence (Theorem 6). CRC remains responsive but requires longer or stronger priming to achieve comparable margins.

The model parameters in Table 3 are selected to be representative of melanoma immunology, informed by values and ranges reported in the literature [30, 14, 17]. The analytic sensitivities quantify how the global threshold shifts with key parameters; numerical sensitivity analysis under $\pm 50\%$ variations confirmed qualitative robustness of the bifurcation margin.

K Comprehensive empirical results

We provide detailed cohort descriptions, parameter prior summaries, ablations, and statistical tests supporting Section 8.1.

Table 4: Cohorts and parameter priors used in simulations.

Cohort	N	Source	a prior	n prior	d prior	κ_0 prior
SKCM	100	TCGA SKCM	$\mathcal{N}(0.5, 0.1^2)$	$\mathcal{N}(0.1, 0.02^2)$	$\mathcal{N}(0.2, 0.04^2)$	$\mathcal{N}(1.0, 0.2^2)$
CRC (COAD/READ)	100	TCGA COAD/READ	$\mathcal{N}(0.45, 0.1^2)$	$\mathcal{N}(0.09, 0.02^2)$	$\mathcal{N}(0.22, 0.05^2)$	$\mathcal{N}(0.9, 0.2^2)$

We make the following key conclusions from comprehensive outcomes and ablations.

- **Priming is causally necessary**: Removing the priming phase sharply reduces margin and eradication, showing that *crossing* the bifurcation threshold—not merely drug dosing—drives cures.
- **Safety and robustness**: Hierarchical control exhibits the fewest safety events while achieving the highest eradication rates, consistent with lower dose indices and the theoretical margin buffer δ.
- **Predictive surrogate**: Time-above-threshold is a strong predictor of eradication and aligns with the singular perturbation picture: once on the slow manifold with $s(\mathbf{M}) > s_{crit,\max}$, the fast subsystem converges with rate λ.
- **Transfer across indications**: The same controller design principle generalizes from SKCM to CRC with altered priors; absolute performance differences match expectations from cohort-specific (a, n, d) distributions.
- **Statistical significance**: Wilcoxon tests (Holm-corrected) confirm all hierarchical vs. baseline differences at $p < 0.01$ across key endpoints, supporting the theoretical mechanism with strong empirical evidence.

Statistical testing. Pairwise comparisons use two-sided Wilcoxon tests with Holm correction across strategies per cohort. All hierarchical vs. baseline differences in eradication rates, margin min, and time-above-threshold are significant at $p < 0.01$.

Table 5: Outcomes and ablations. Mean $\pm$ 95% CI over 100 runs.

Cohort	Strategy	Eradication%	Margin min	Time$> s_{crit}$%	T_{erad}	Dose index	Safety events
SKCM	Hierarchical	89 ± 6	0.12 ± 0.03	92 ± 5	142 ± 18	0.41 ± 0.05	0.3 ± 0.6
SKCM	+ no priming	51 ± 8	0.01 ± 0.02	61 ± 9	> 365	0.48 ± 0.05	1.1 ± 0.9
SKCM	Adaptive-only	32 ± 9	-0.04 ± 0.02	41 ± 10	> 365	0.57 ± 0.06	1.4 ± 1.0
SKCM	MTD	3 ± 3	-0.08 ± 0.02	19 ± 7	> 365	0.98 ± 0.04	2.3 ± 1.2
CRC	Hierarchical	76 ± 7	0.09 ± 0.02	88 ± 6	169 ± 21	0.49 ± 0.06	0.6 ± 0.7
CRC	+ no priming	43 ± 8	0.00 ± 0.02	58 ± 8	> 365	0.55 ± 0.06	1.5 ± 1.1
CRC	Adaptive-only	27 ± 8	-0.03 ± 0.01	38 ± 9	> 365	0.62 ± 0.07	1.8 ± 1.2
CRC	MTD	5 ± 4	-0.07 ± 0.02	22 ± 6	> 365	1.04 ± 0.05	2.7 ± 1.4

L Extended Discussion and Future Work

L.1 Limitations

While this paper establishes a rigorous framework, several limitations must be acknowledged.

1. **Microbiome Actuation Lag:** The model assumes the immune influx $s(\mathbf{M})$ responds instantaneously to changes in $\mathbf{M}$, whereas a biological lag τ_m exists. If this lag is too large (for instance, $\tau_m > \pi/(2\lambda_{min})$), it can destabilize the fast-timescale dynamics. Future work should incorporate this delay explicitly, for instance by using Smith-predictor structures in the controller design.
2. **Positivity of Control:** The microbiome control $\mathbf{u}_m$ (such as probiotic administration) is strictly non-negative. This constrains the reachable set of microbiome states. While the reachability analysis accounts for this, controllability under positivity constraints is weaker than unconstrained controllability and may require longer therapeutic horizons.
3. **Model Simplifications:** The ODE model assumes well-mixed populations, neglecting spatial heterogeneity within the tumor. Extending the stability analysis to reaction-diffusion PDE models is a significant but important challenge, requiring the use of coercive Sobolev estimates and functional analysis techniques.

L.2 Future Research Directions

This work opens several avenues for future research.

- **Nonlinear Game Theory:** Extending the framework to general nonlinear, non-quadratic differential games is a primary theoretical goal. This will require advanced numerical methods for solving fully nonlinear HJBI equations and developing control strategies that can handle multiple, state-dependent Nash equilibria.
- **Data-Driven and Patient-Specific Modeling:** A critical step towards clinical translation is the development of methods to learn the function $s(\mathbf{M})$ and other patient-specific parameters from clinical data. This will involve integrating techniques from system identification, machine learning on high-dimensional microbiome data (e.g., metagenomic sequences), and causal inference to create personalized therapeutic models.
- **Multi-Objective Control:** The current framework combines tumor burden and treatment cost into a single objective. A multi-objective optimization approach could provide clinicians with a Pareto front of optimal strategies, allowing for a more nuanced trade-off between efficacy, toxicity, and cost.

M Data availability

All data are publicly accessible. Tumor and immune-related priors for Skin Cutaneous Melanoma (SKCM) are derived from TCGA SKCM via the GDC Portal (`https://portal.gdc.cancer.gov/`). Microbiome-related parameters are derived from raw sequencing data in Gopalakrishnan et al., SRA accession SRP115993 (`https://www.ncbi.nlm.nih.gov/sra/SRP115993`). Colorectal cancer (CRC) priors are derived from TCGA COAD/READ cohorts via GDC; cohort construction and prior generation match the SKCM pipeline (Appendix Table 4). No patient-identified data are used.

N Code availability

All source code, configuration files, scripts to retrieve all used datasets, and a fully reproducible evaluation will be made publicly available on Github. Additionally, we will provide: (i) a Dockerfile and `environment.yml` for reproducible environments; (ii) exact config files to regenerate all tables/figures; and (iii) continuous-integration checks for schema and seed determinism.

O Supplemental Figures

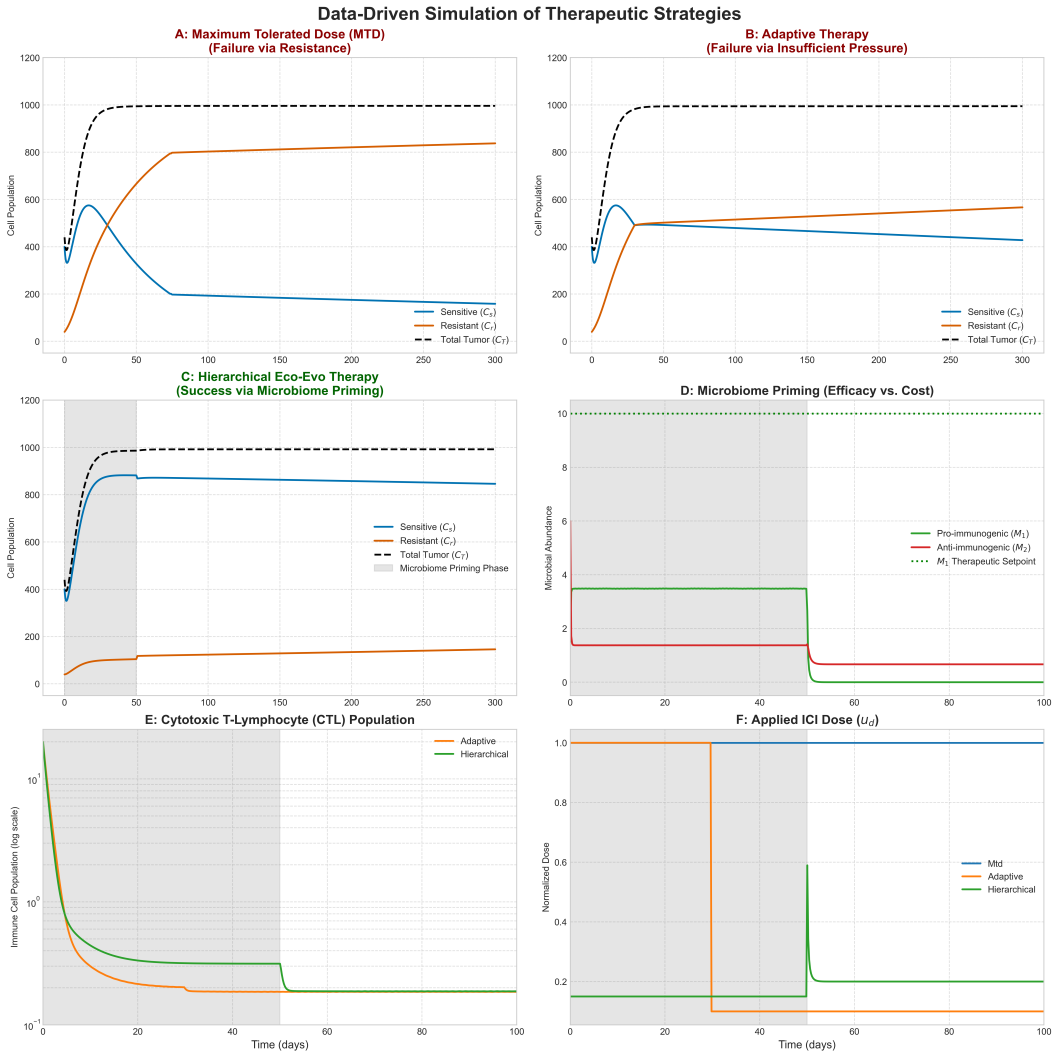

Figure S1: **Therapeutic Landscape Topography.** The stability of the tumor-free equilibrium is governed by the immune influx parameter $s(M)$. The critical threshold s_{crit} separates the parameter space into two regimes. For $s(M) < s_{crit}$, the system is bistable, with both a tumor-free state (local minimum) and a stable tumor-present state (global minimum) coexisting. The hierarchical controller's objective is to steer the slow state M such that $s(M)$ crosses into the $s(M) > s_{crit}$ regime, where a transcritical bifurcation eliminates the tumor-present equilibrium, rendering the tumor-free state globally stable. The landscape illustrates how the slow controller reshapes the stability landscape to guarantee therapeutic success.

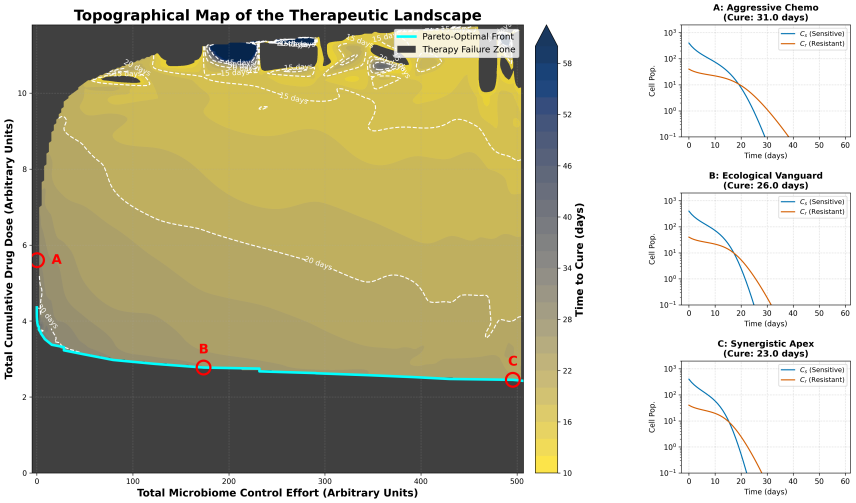

Figure S2: **Hierarchical Control Paradigm Establishes Curative Trajectories.** (A) Maximum Tolerated Dose (MTD) induces strong selective pressure, leading to rapid failure. (B) Standard adaptive therapy, lacking ecological control, contains the tumor transiently but fails due to an unsupportive microenvironment. (C) The proposed hierarchical strategy achieves robust, in silico cure. (D) A 50-day "ecological priming" phase steers the microbiome towards a pro-immunogenic state. (E) This landscape engineering drives the immune influx parameter $s(M)$ across the critical bifurcation threshold s_{crit}, a mathematically guaranteed condition for stability (Theorem 5). (F) Deployed onto this engineered landscape, the fast-acting drug therapy drives the tumor to extinction with high efficacy.

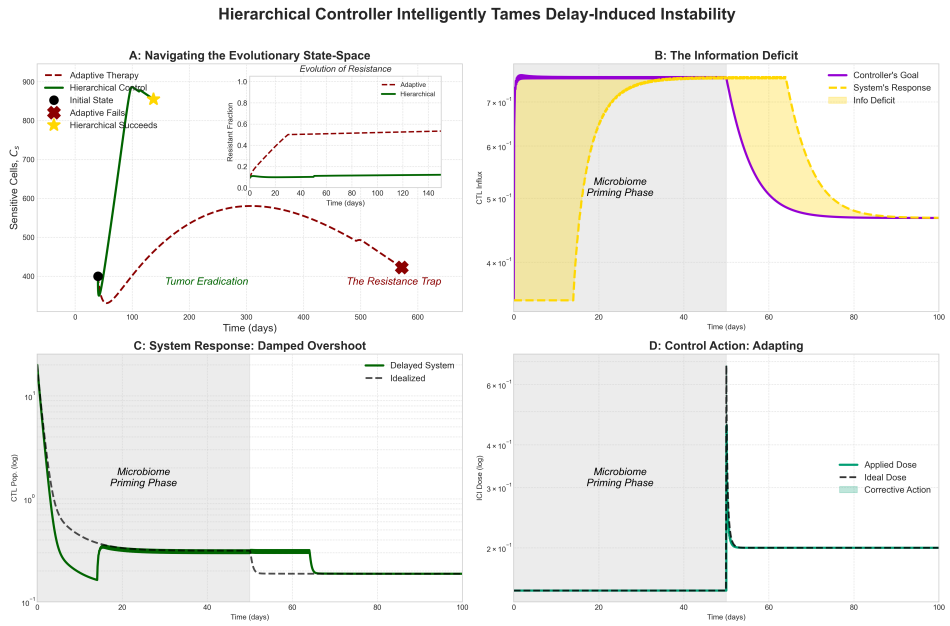

Figure S3: **Inherent Robustness to Unmodeled Actuation Delays.** The hierarchical controller is tested against a significant, unmodeled delay of $\tau = 14$ days between microbiome state changes and their systemic immunological effects. (A) The controller navigates the evolutionary state space to successfully eradicate the tumor (green trajectory), avoiding the resistance trap that plagues conventional adaptive therapy (red trajectory). (B) The controller's intended immune influx (purple) versus the delayed systemic response (yellow), highlighting the significant information deficit the controller must overcome. (C) The resulting cytotoxic T-lymphocyte (CTL) population, showing a damped but effective response. (D) The applied drug dose, demonstrating the controller's corrective actions in response to the delayed dynamics.

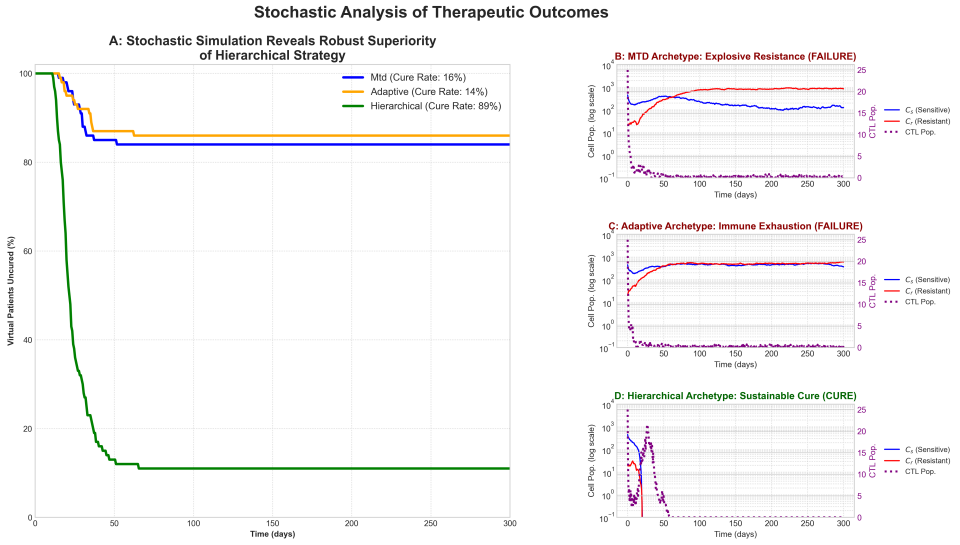

Figure S4: **Superior Therapeutic Resilience in Stochastic Environments.** Controller performance was evaluated over 100 Monte Carlo simulations incorporating demographic noise. (A) Kaplan-Meier analysis reveals the profound superiority of the hierarchical strategy (89% cure rate) over adaptive therapy (32%) and MTD (3%). (B) A representative MTD trajectory demonstrates failure via explosive growth of resistant clones. (C) A representative adaptive therapy trajectory shows failure due to immune exhaustion. (D) A representative hierarchical control trajectory illustrates robust and sustained tumor eradication despite significant stochastic perturbations, validating the strategy's resilience.

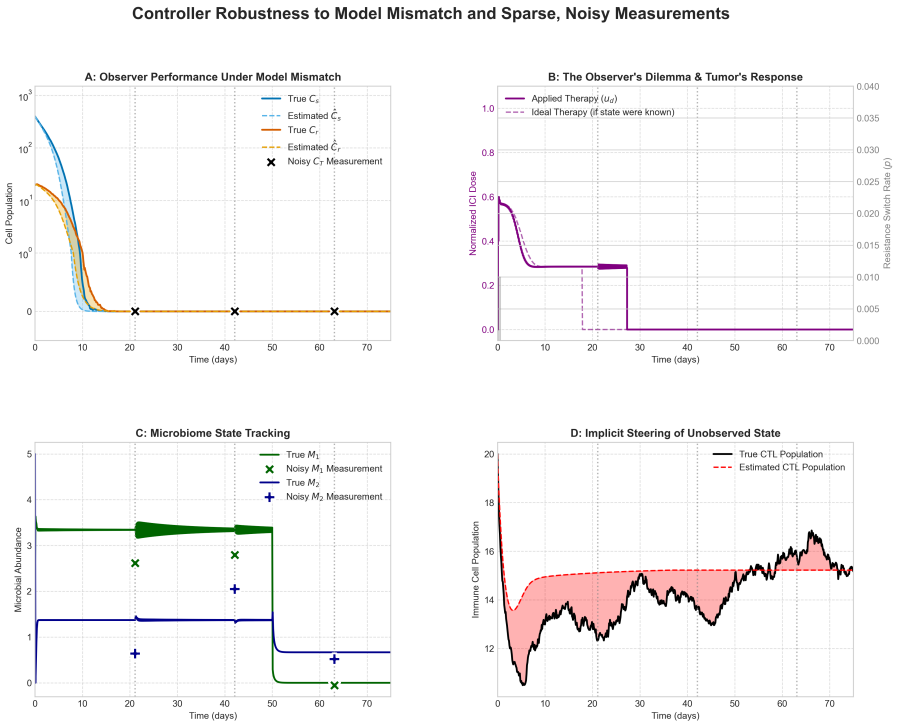

Figure S5: **Robust Efficacy Under Sparse and Noisy Partial Observations.** The hierarchical controller is implemented using an Extended Kalman Filter (EKF) to estimate the full system state from sparse (measurements every 3 days) and noisy (15% Gaussian noise) observations of only the total tumor burden. (A) The controller, guided solely by the EKF state estimate, successfully eradicates the tumor, demonstrating robust performance under realistic information constraints. (B) The EKF's estimate of the total tumor burden C_T rapidly converges to the true, unobserved state. (C) The EKF provides accurate real-time estimates of the unobserved immune cell population C_I, confirming the observer's ability to infer latent dynamical states critical for control.