Decomposition of generalized Lotka-Volterra systems in the microbiome

Eric W. Jones* and Jean M. Carlson
Department of Physics, University of California, Santa Barbara, California 93106, USA
(Dated: November 13, 2018)

In the gut microbiome, the successful administration of fecal microbiota transplantation (FMT) will convert a person's diseased microbial composition into a healthy one. We model these "healthy" and "diseased" microbial states as idealized ecological species, and characterize their behavior by homogenizing the properties of their constituitive microbial populations in a process called "steady state reduction" (SSR). This method decomposes a generalized Lotka-Volterra (gLV) system of many microbial species into two-dimensional (2D) subsystems that each span a pair of steady states of the high-dimensional model and obey gLV dynamics. We investigate an experimentally derived model of antibiotic-induced *C. difficile* infection (CDI), and study the clinically relevant transition between CDI-vulnerable (diseased) and CDI-resilient (healthy) states with the 2D subspace generated by SSR. Specifically, we investigate the ability of FMT, a type of bacteriotherapy, to convert a diseased state into a healthy state. This analysis is buoyed by the analytic calculation of the separatrix of a bistable 2D gLV system. In this model, we observe that FMT is most effective when administered immediately or nearly immediately after an antibiotic pulse, and that a delay in FMT administration may undermine its success.

I. INTRODUCTION

Biological and ecological models are intended to approximate real-world systems, and therefore a system's steady state outcome is often the key result of a simulation. For example, these models try to predict whether an invasive pest goes extinct [*], whether an epidemic ends [*], or whether human-manufactured chemicals alter an ecosystem's biodiversity [1]. In this paper, we explicitly account for this outcome-oriented perspective by reformulating a classic ecological model composed of ecological species into a model whose unit "species" are steady states of the classic model, in a process we call "steady state reduction" (SSR). We derive and apply this method in the context of the microbiome, a system that exhibits clear "diseased" states [2] that biomedical treatments attempt to heal.

In the microbiome, diseases ranging from inflammatory bowel disease to cancer having been found to be associated with particular "dysbiotic" microbial compositions in the gut [2], which we construe as diseased states. A promising frontier of medicine relies on the notion that the microbiome's composition can both influence and be influenced by disease, so that the deliberate alteration of the microbiome via bacteriotherapy might be a viable treatment option for a range of diseases [3, 4]. Fecal microbiota transplantation (FMT), which operates by engrafting gut microbes from a healthy donor into an infected patient, is a type of bacteriotherapy that has shown remarkable success in treating recurrent *C. difficile* infection (CDI) [5]. In order to better understand the mechanism of FMT, ecological frameworks are being actively developed that describe how bacteriotherapies can regulate the microbiome [6–8], experimental studies are leveraging these frameworks to design and test "optimal" FMT therapies [9], and clinical studies are identifying new diseases that FMT can successfully treat [*].

Within these FMT frameworks lie microbiome models. Thousands of microbial species populate the gut microbiome [10], but for modeling purposes it is common to coarse-grain at the genus or phylum level. In this paper we rely on one particular model by Stein et al. [7] that fit a genus-level ecological model of CDI, using bacterial abundance data from a mouse experiment [11]. Despite the fact that this model did not resolve individual species, it still captured the clinically- and experimentally-observed phenomenon of antibiotic-induced CDI, suggesting that the true microbiome's dimensionality could be approximated by an 11-dimensional model.

In this paper we simplify the dimensionality of the microbiome even further: instead of thousands of microbial species or eleven dominant genera, we idealize each steady state of the microbiome as a single ecological population with SSR. By considering two such steady states (interpreted as healthy and diseased), we analyze a two-state ecological system that demonstrates how FMT and antibiotic administration can induce a lasting shift in microbial composition.

We take advantage of the analytic tractability of this two-state model to investigate the role of FMT timing on its efficacy. While manipulating a treatment's time course has innovated treatment strategies in cancer immunotherapy [12] and HIV vaccination [13], this perspective has not been previously considered in the context of bacteriotherapy. We discover the clinically-relevant prediction that delaying the administration of FMT can lead to its failure, as in Fig. 3.

* ewj@physics.ucsb.edu

II. DECOMPOSITION OF GENERALIZED LOTKA-VOLTERRA SYSTEMS

We begin by demonstrating how to compress the ecological dynamics described by the generalized Lotka-Volterra (gLV) equations into a subspace spanned by a pair of fixed points in a process we call *steady state reduction* (SSR), which we depict schematically in Fig. 1. Specifically, we recast a pair of fixed points of an N-dimensional gLV model as idealized ecological species in a two-dimensional (2D) gLV model, and characterize the interactions between these two composite states by homogenizing the species interactions of the high-dimensional system. We prove that these reduced dynamics constitute the best possible 2D gLV approximation of the actual gLV dynamics in the subspace spanned by the two fixed points.

For N interacting species, the gLV equations model the populations of each ecological species y_i as

$$\frac{\mathrm{d}}{\mathrm{d}t}y_i(t) = y_i(t)\left(\rho_i + \sum_{j=1}^{11} K_{ij}y_j(t)\right) \tag{1}$$

for $i \in 1, \ldots, N$, so that we can write microbial dynamics in vector form as $\frac{\mathrm{d}\vec{y}}{\mathrm{d}t} = \sum_{i=1}^{N} \frac{\mathrm{d}y_i}{\mathrm{d}t}\hat{y}_i$. Here, the growth rate of species i is ρ_i, and the effect of species j on species i is given by the interaction term K_{ij}. In the following derivation, we assume that this model observes distinct stable fixed points $\vec{y}_a$ and $\vec{y}_b$.

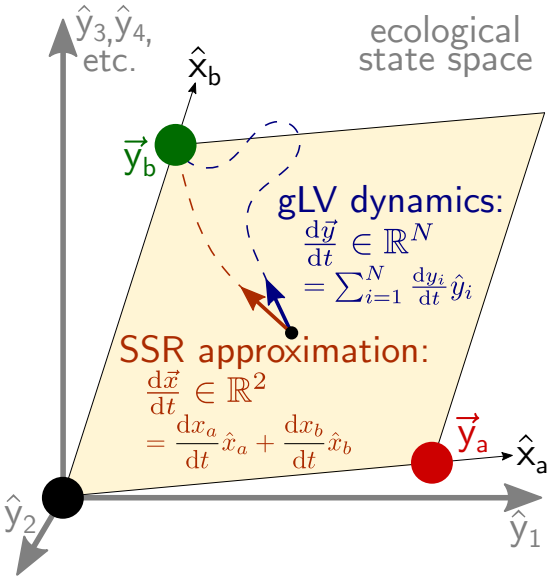

FIG. 1. Schematic of how steady state reduction (SSR) compresses a generalized Lotka-Volterra (gLV) system. We consider a gLV system of N ecological species whose dynamics are given in Eq. (1), and we assume that this model exhibits two steady states $\vec{y}_a$ and $\vec{y}_b$ that we characterize as diseased (red) and healthy (green). We consider the family of gLV dynamics that exist on the two-dimensional (2D) subspace spanned by these two steady states, given in Eq. (2), and through SSR identify the gLV parameters in Eq. (3) that minimize the deviation between the actual N-dimensional gLV dynamics and the embedded 2D SSR-reduced dynamics.

We consider new variables x_a and x_b in the direction of unit vectors $\hat{x}_a = \vec{y}_a/\|\vec{y}_a\|_2$ and $\hat{x}_b = \vec{y}_b/\|\vec{y}_b\|_2$, where $\|\cdot\|_k$ is the k-norm, and consider the general class of gLV dynamics on the subspace spanned by $\hat{x}_a$ and $\hat{x}_b$

$$\begin{aligned}\frac{\mathrm{d}x_a}{\mathrm{d}t} &= x_a(\mu_a + M_{aa}x_a + M_{ab}x_b)\\ \frac{\mathrm{d}x_b}{\mathrm{d}t} &= x_b(\mu_b + M_{ba}x_a + M_{bb}x_b),\end{aligned} \tag{2}$$

so that the *in-plane dynamics* on this subspace in vector form are $\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \frac{\mathrm{d}x_a}{\mathrm{d}t}\hat{x}_a + \frac{\mathrm{d}x_b}{\mathrm{d}t}\hat{x}_b$.

SSR links the parameters of the in-plane dynamics to the actual gLV dynamics by setting

$$\begin{aligned}\mu_\gamma &= \frac{\vec{\rho}\cdot\vec{y}_\gamma^{\,2}}{\|\vec{y}_\gamma\|_2^2} \qquad \text{for } \gamma \in a,\ b, \text{ and}\\ M_{\gamma\delta} &= \frac{(\vec{y}_\gamma^{\,2})^T K \vec{y}_\delta}{\|\vec{y}_\gamma\|_2^2\|\vec{y}_\delta\|_2}, \qquad \text{for } \gamma,\ \delta \in a,\ b.\end{aligned} \tag{3}$$

Here, we have assumed that $\vec{y}_a$ and $\vec{y}_b$ are orthogonal, and we have introduced the vector $\vec{y}_\gamma^{\,2}$ that is an element-wise square of steady state $\vec{y}_\gamma$ given by $\vec{y}_\gamma^{\,2} \equiv \mathrm{diag}(\vec{y}_\gamma)\vec{y}_\gamma$ for $\gamma \in a,\ b$. When $\vec{y}_a$ and $\vec{y}_b$ are not orthogonal, the cross-interaction terms M_{ab} and M_{ba} are more complicated, and are given in Eqs. (A.17) and (A.18) of the Appendix.

This choice of parameters minimizes the deviation between the in-plane and true gLV dynamics $\epsilon = \|\frac{\mathrm{d}\vec{y}}{\mathrm{d}t} - \frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\|_2$ for any point on the subspace spanned by $\hat{x}_a$ and $\hat{x}_b$. We prove this in Appendix * by showing that, when evaluated with the SSR-prescribed parameter values of Eq. (3), $\frac{\partial\epsilon}{\partial c_i} = 0$ for every coefficient $c_i \in \{\mu_a,\ \mu_b,\ M_{aa},\ M_{ab},\ M_{ba},\ M_{bb}\}$, and that $\frac{\partial^2\epsilon}{\partial c_i \partial c_j} > 0$ for every pair of coefficients c_i and c_j. Under this construction, the high-dimensional steady states $\vec{y}_a$ and $\vec{y}_b$ will have in-plane steady state counterparts at $(\|\vec{y}_a\|_2,\ 0)$ and $(0,\ \|\vec{y}_b\|_2)$, respectively. It is for this reason we call this method *steady state reduction*. Further, if $\vec{y}_a$ and $\vec{y}_b$ are stable, then the corresponding 2D steady states will be stable as well [14], which guarantees the existence of a separatrix in the reduced 2D system.

If the system's ecological dynamics lie entirely on the plane spanned by $\hat{x}_a$ and $\hat{x}_b$, the SSR approximation will be exact. In this case, the plane will contain a slow manifold on which the ecological dynamics evolve. Therefore, SSR describes the ecological dynamics when the slow manifold of steady states $\vec{y}_a$ and $\vec{y}_b$ lies on the plane spanned by $\hat{x}_a$ and $\hat{x}_b$, which is a linear approximation to the slow manifold of the system. Therefore, the dynamics generated by SSR result from a linear approximation of the slow manifold.

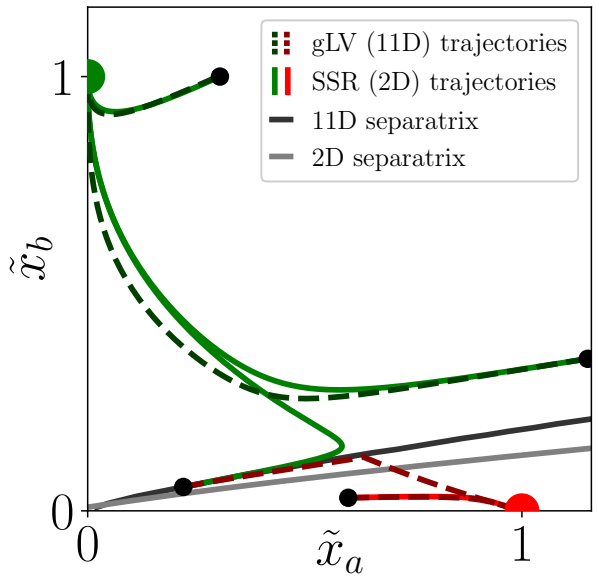

FIG. 2. Fidelity of steady state reduction, as applied to the experimentally-derived 11-dimensional (11D) gLV model of *C. difficile* infection of Stein et al. [7]. This CDI model exhibits steady states $\vec{y}_a$ and $\vec{y}_b$ that are respectively vulnerable (diseased) and resilient (healthy) to invasion by the pathogen *C. difficile* [14]. We consider 11D microbial trajectories whose initial conditions lie the plane spanned by these two steady states, and plot the in-plane projection of these trajectories (dashed lines). Alongside these trajectories, we plot the SSR-generated dynamics (solid lines), which largely approximate the 11D trajectories. Lastly, we demonstrate that basins of attraction are largely preserved under SSR by plotting the separatrix for both systems: as a proxy for the 11D separatrix (actually an 10-dimensional surface), we numerically generate and plot the in-plane separatix (dark grey); for the 2D separatrix, we use the exact expression given in Eq. (6) (light grey).

III. STEADY STATE REDUCTION APPLIED TO A MICROBIOME MODEL

SSR may be applied to any gLV system that exhibits two or more stable steady states. Recently, many experimentally derived gLV microbiome models have been constructed with tools such as MDSINE, a computational pipeline that infers gLV parameters from time-series microbial abundance data [15]. We consider one such experimentally derived model, constructed by Stein et al. [7], that studies CDI in the mouse gut microbiome. This model takes the same form as Eq. (1) but tracks the abundances of 10 different microbial genera and the pathogen *C. difficile* (CD), all of which can inhabit the mouse gut. The 11-dimensional (11D) parameters of this model were fit with data from an experimental mouse study [11], and the resulting model captures the clinically observed phenomenon of antibiotic-induced CDI vulnerability.

This CDI model exhibits five steady states that are reachable from experimentally measured initial conditions. Previously, we identified which of these steady states were vulnerable to invasion by CD, which we interpreted as CDI-susceptible states [8]. Based on this classification, here we interpret a CD-susceptible steady state $\vec{y}_a$ of the 11D model as "diseased," interpret a CD-resilient steady state $\vec{y}_b$ as "healthy," and provide explicit details about these two states in [14]). We use these two states for SSR.

Since the coarse-grained 11D model approximates features of the actual microbiome (which consists of thousands of species), testing how well SSR reduces the 11D model to 2D is therefore a proxy for how well SSR compresses the true high-dimensional microbiome into a 2D model. We generate the 2D parameters according to Eq. (3), and introduce new scaled variables $\tilde{x}_a = x_a/\|\vec{y}_a\|_2$ and $\tilde{x}_b = x_b/\|\vec{y}_b\|_2$ so that the 2D system exhibits steady states at $(1,\ 0)$ and $(0,\ 1)$.

The basins of attraction and microbial trajectories of the 11D gLV model are largely preserved through SSR. To compare the two models, in Fig. 2 we consider the trajectories of the initial conditions that lie on the embedded 2D subspace spanned by $\vec{y}_a$ and $\vec{y}_b$ alongside their corresponding SSR trajectories. By tracking the fates of these initial conditions, we numerically construct the in-plane 11D separatrix (dark grey). We compute the 2D separatrix (light grey) analytically with Eq. (*), which is derived in the next section. Since the 11D system has been compressed, naturally the low- and projected high-dimensional trajectories do not exactly match, but there is an obvious qualitative agreement between both the trajectories and the separatrices of the two models.

In Section VI, we compare the transient dynamics of the two models and find that they also largely agree. Before this, in Sections IV and V we analytically examine the steady state and transient dynamics of the 2D model, which we use to inform the dynamics of higher-dimensional systems by proxy of SSR.

IV. DYNAMICAL LANDSCAPES OF 2D GENERALIZED LOTKA-VOLTERRA SYSTEMS

Having demonstrated a method of linking a high-dimensional gLV system to a 2D system via SSR, here we provide novel analytic results for generic 2D gLV systems. In this section we diverge from the previous section's choice of SSR parameters, and instead consider a set of unconstrained gLV parameters. Through nondimensionalization we set $M_{aa} = M_{bb} = \mu_a = 1$, and to simplify the presentation of our results we assume $\mu_b = 1$. SSR-generated parameters satisfy $M_{aa} = \mu_a$ and $M_{bb} = \mu_b$, and therefore also satisfy the assumption $\mu_b = 1$ after nondimensionalization [14]. Figs. 3, 4, and 5 are generated with the randomly chosen yet representative parameter values $M_{ab} = 1.167$ and $M_{ba} = 1.093$. In the Supplementary Information [14] we verify that the results of this paper hold for $\mu_b \neq 1$, and provide the

explicit calculations supporting this paper's claims.

The dynamical system Eq. (7) exhibits three nontrivial steady states at (1, 0), (0, 1), and $(x_a^*,\ x_b^*) \equiv (\frac{1-M_{ab}\mu_b}{1-M_{ab}M_{ba}},\ \frac{\mu_b-M_{ba}}{1-M_{ab}M_{ba}})$. Since we are interested in the switching behavior between healthy and diseased states, we only consider systems in which the two homogeneous steady states are stable, so that $0 < x_a^* < 1$ and $0 < x_b^* < 1$, and so that the mixed steady state $(x_a^*,\ x_b^*)$ is hyperbolic [14].

The long-term dynamics of this system are dictated by the basins of attraction of the stable steady states, and these basins are delineated by a separatrix that, for topological reasons, must be the stable manifold of the hyperbolic fixed point $(x_a^*,\ x_b^*)$. In Fig. 3 these basins are depicted topographically via isoclines of the split Lyapunov function $V(x_a, x_b)$ (lightly shaded contours), which acts as a potential energy landscape [14, 16].

We may analytically compute this separatrix $h(x_a)$ in a power series expansion about the hyperbolic fixed point as a function of the gLV system parameters

$$h(x_a) = \sum_{n=0}^{\infty} \frac{c_n}{n!}(x_a - x_a^*)^n, \tag{4}$$

that as an invariant manifold must obey the relation [17]

$$\frac{\mathrm{d}h(x_a)}{\mathrm{d}x_a} = \frac{\mathrm{d}x_b}{\mathrm{d}t} \Big/ \frac{\mathrm{d}x_a}{\mathrm{d}t}, \tag{5}$$

resulting in the recursive relations

$$\begin{aligned}
c_0 &= x_b^*, \\
c_1 &= \frac{1}{2M_{ab}x_a^*}\Big[M_{bb}x_b^* - M_{aa}x_a^*\ , \\
&\qquad +\sqrt{(M_{bb}x_b^* - M_{aa}x_a^*)^2 + 4M_{ab}x_a^*M_{ba}x_b^*}\Big], \\
c_2 &= \frac{2c_1(M_{ba} + M_{bb}c_1 - M_{aa} - M_{ab}c_1)}{2x_a^*M_{aa} + 3x_a^*M_{ab}c_1 - M_{bb}x_b^*}, \qquad \text{and} \\
c_n &= \frac{1}{(nx_a^*M_{aa} + (n+1)x_a^*M_{ab}c_1 - M_{bb}x_b^*)} \\
&\quad \times \Bigg\{ nc_{n-1}(M_{ba} + M_{bb}c_1 - (n-1)(M_{aa} + M_{ab}c_1)) \\
&\quad + n!\sum_{\ell=2}^{n-1}\Bigg[\frac{c_\ell}{\ell!\ (n-\ell)!}\,(M_{bb}c_{n-\ell} - (n-\ell)M_{ab}c_{n-\ell} \\
&\qquad\qquad -x_a^*M_{ab}c_{n-\ell+1})\Bigg]\Bigg\}, \qquad \text{for } n > 2,
\end{aligned} \tag{6}$$

as derived in Eqs. (S*) [14]. Therefore, from this calculation we may *a priori* classify the fate of a given initial condition, without need for simulation. We note that this algebraic computation of the separatrix is considerably faster than numerical methods that rely on relatively costly quadrature computations. Further, in conjunction with SSR, this analytic form offers an efficient approximation to the in-plane separatrix of high-dimensional systems.

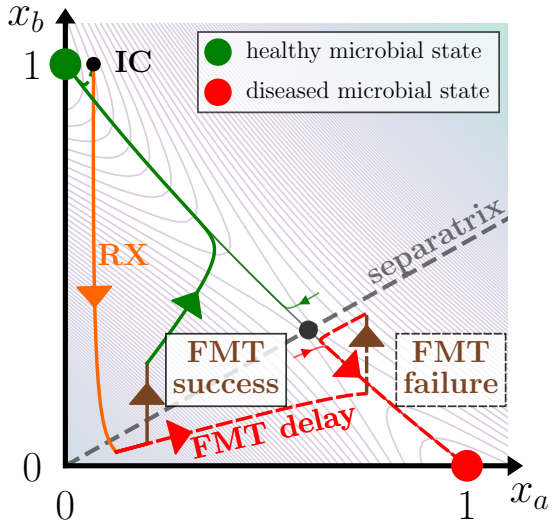

FIG. 3. The success or failure of fecal microbiota transplantation (FMT) depends on the timing of its administration. We consider the two-state ecological system of Eq. (7) in a clinically-inspired scenario. First, a health-prone initial condition (IC) is depleted by antibiotics (RX, orange). If FMT (brown) is administered shortly after the antibiotics, the treatment steers the composition to a healthy state (FMT success). If FMT administration is delayed, the microbial trajectory instead attains the diseased state (FMT failure). The basins of attraction of the healthy and diseased steady states are delineated by the separatrix, and isoclines of the system's potential energy landscape (light contours) are given by the split Lyapunov function Eq. (S21) [14].

V. TRANSIENT DYNAMICS OF 2D GLV SYSTEMS, WITH APPLICATION TO FMT

Next, we describe the transient dynamics of 2D gLV models. We are driven by the scenario of antibiotic-induced CDI, in which antibiotics deplete a health-prone initial condition, requiring administration of FMT in order to recover, as in Fig. 3. The efficacy of a FMT treatment depends on the microbial state of the system, which itself evolves in time. Therefore, the transient dynamics of this 2D ecological model, which resemble the dynamics of a more highly-resolved microbial system through SSR, may offer clinically pertinent predictions for the administration of FMT.

In previous work, we supplemented the 11D model of Stein et al. [7] with an in-silico FMT framework [8], which numerically demonstrated how FMT could successfully salvage a CDI-prone microbial state, mirroring clinical observations. Here we analytically identify how FMT operates in the 2D system, which through SSR pro-

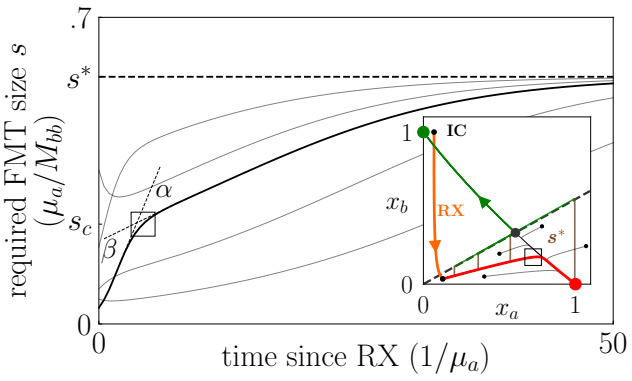

FIG. 4. The FMT transplant size needed to revert an antibiotic-depleted state back to health grows as FMT administration is delayed. We calculate the minimum FMT transplant size required to cure five disease-prone microbial trajectories that evolve according to Eq. (7). As trajectories attain the diseased steady state, the required transplant size approaches s^*. The required transplant size changes at two different rates, α and β, with the crossover point between these two rates indicated by a hollow square. The transplant size dynamics $\mathrm{d}s/\mathrm{d}t$ as well as the rates α and β are derived in Eq. (9) and the surrounding text.

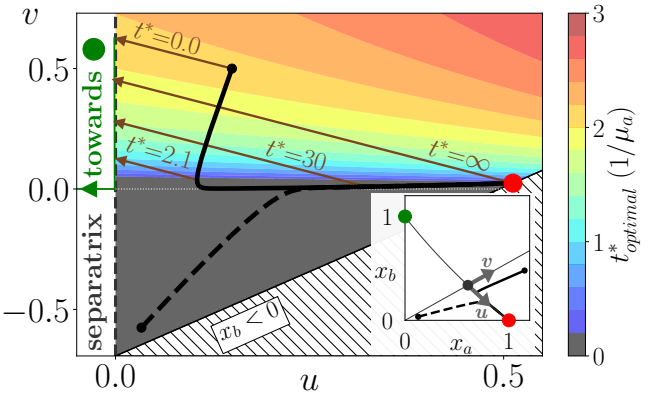

FIG. 5. The role of timing in FMT administration. For disease-prone initial conditions, FMT is most effective when administered immediately ($t^*_{opt} = 0$, grey) or nearly immediately ($t^*_{opt} > 0$, colored). We can calculate the optimal transplant time t^*_{opt} for any initial condition (u_0, v_0) (colorbar) according to Eq. (S62) of [14], which can reduce to Eq. (10). We show two microbial trajectories that are plotted in (u, v) (main panel) and (x_a, x_b) (inset) coordinates. For $v_0 > 0$ we show four possible FMT transplants, including the optimal one that occurs at $t^*_{opt} = 2.1$. For $v_0 < 0$ it is always best to administer FMT immediately.

vides a theoretical underpinning for the behavior of FMT in the 11D model. We find that following antibiotics, the efficacy of FMT declines as its administration is increasingly delayed, a result which we had previously identified in the 11D model.

To include the therapeutic interventions of antibiotics and FMT, we consider a modified form of the 2D gLV equations given by

$$\begin{aligned}\frac{\mathrm{d}x_a}{\mathrm{d}t} &= x_a(\mu_a - M_{aa}x_a - M_{ab}x_b)\\ &\quad + x_a\varepsilon_a u(t) + w_a\delta(t-t^*) \quad \text{and}\\ \frac{\mathrm{d}x_b}{\mathrm{d}t} &= x_b(\mu_b - M_{ba}x_a - M_{bb}x_b)\\ &\quad + x_b\varepsilon_b u(t) + w_b\delta(t-t^*),\end{aligned} \tag{7}$$

which includes optional antibiotic administration $u(t)$ operating with efficacy $\vec{\varepsilon}$, and optional FMT with transplant $\vec{w}$ administered instantaneously at time t^*. As before, the variable x_a corresponds to a diseased state, while x_b corresponds to a healthy state. We also reverse the sign of the interaction coefficients for notational convenience. Though the transplant $\vec{w}$ can consist of a mixture of the healthy and diseased steady states, for demonstrative purposes we assume that the transplant consists of exclusively healthy microbes so that $w_a = 0$.

Altering the fate of an initial condition requires crossing the separatrix by some external means, which we achieve through FMT. At each point along a microbial trajectory in the diseased basin of attraction, we may calculate the minimum FMT size s required to transfer the microbial state into the healthy basin of attraction. We use this metric to quantify our notions of "FMT efficacy." In Fig. 4 we present s as a function of time (main panel) for several trajectories that originate in the diseased basin of attraction (inset), including the main trajectory of Fig. 3.

Currently, in clinical practice FMT administration varies in transplant size, transplant composition, and how many transplants are performed. Further, it is unclear how these factors influence the success of FMT [18]. For the purposes of this paper, we consider a hypothetical FMT treatment of size s_t (i.e. a horizontal cut across Fig. 4), and study how its success depends on transplant time. We find that the minimum required FMT size increases with time— often dramatically— and that there are two discernible rates of increase, which we denote α and β in the figure. In the following text, we link these two rates to the fast and slow manifolds of the ecological system, and we derive an expression that governs the minimum required transplant size over time.

To reflect the importance of the separatrix in dictating the microbial dynamics, we change coordinates to the eigenvectors (u, v) of the hyperbolic steady state, shown in Fig. 5 (inset). In these coordinates the v-axis corresponds to the separatrix, and u is proportional to the minimum FMT size required for a successful transplant s, such that $s = u/(\hat{u}\cdot\hat{x}_b)$, where ($\hat{u}$, $\hat{v}$) and ($\hat{x}_a$, $\hat{x}_b$) are the unit vectors associated with their associated coordinates.

In this new (u, v) basis, the 2D gLV equations become

$$\begin{aligned}\frac{\mathrm{d}u}{\mathrm{d}t} &= A_{10}u - A_{11}uv - A_{20}u^2\\ \frac{\mathrm{d}v}{\mathrm{d}t} &= -B_{01}v - B_{02}v^2 + B_{20}u^2,\end{aligned} \tag{8}$$

where each coefficient is a positive algebraic function of the original gLV parameters that we have analytically determined in Eqs. (S36-S48) [14]. When $\mu_b \neq 1$, these equations contain additional quadratic terms described in [14] that account for the nonlinearity of the separatrix. Naturally, in the small u and small v limit this model reduces to the linearization about the hyperbolic fixed point. Near this fixed point there is a separation of time scales between u and v ($B_{01}/A_{10} > 1$ always, with median of 5.9 and IQR of [2.7, 9.1] over random parameter draws [14]), indicating that there are inherent fast and slow manifolds in this system.

This coordinate change also reveals the role of timing in FMT administration, since the minimum required transplant size s is precisely governed by Eq. (8), by proxy of u. To demonstrate this analytically, we consider an initial condition condition (u_0, v_0) that is located near the fast manifold in a system with clear separation of time scales, so that (i) $B_{20}u_0^2$ is negligible, (ii) $B_{01} >> A_{10}$, and (iii) $B_{02}v_0^2 << B_{01}v_0$ (though we relax this assumption in Eq. (S62) [14]). In this case, the dynamics in the fast $\hat{v}$ direction may be approximated as $v(t) \approx v_0 \mathrm{e}^{-B_{01}t}$, and so the required transplant size dynamics become

$$\frac{\mathrm{d}s}{\mathrm{d}t} = s\left(A_{10} - A_{20}(\hat{u}\cdot\hat{x}_b)s - A_{11}v_0\mathrm{e}^{-B_{01}t}\right). \quad (9)$$

With Eq. (9) we can interpret the rates of increase in required transplant size s of the main trajectory in Fig. 4, α and β [14]. Specifically, $\alpha = \frac{\mathrm{d}s(0)}{\mathrm{d}t}\big|_{s=s_c}$ and $\beta = \frac{\mathrm{d}s(\infty)}{\mathrm{d}t}\big|_{s=s_c}$ where s_c is the transplant size required at the crossover point between these rates (e.g. as shown in Fig. 4).

For an initial condition with $v_0 < 0$, which can occur when a nearly healthy state is depleted by antibiotics as in Fig. 3, we have $\alpha > \beta$. In this case the required transplant size monotonically increases until it attains s^* at the infected steady state, and so it is best to administer FMT as soon as possible. Alternatively, when $v_0 > 0$ we have $\alpha < \beta$. When $A_{11}v_0$ is sufficiently large α can become negative, which means that there is a nonzero transplant time at which the required transplant size is minimized (corresponding to $\frac{\mathrm{d}s}{\mathrm{d}t} = 0$). The concave-up trajectories in Fig. 4 exhibit this optimal transplant time. For $v_0 > 0$ and under the same conditions for which Eq. (9) was derived, we find that this optimal transplant time t^* is

$$t^*_{opt} = \frac{1}{B_{01}}\ln\left(\frac{A_{11}v_0}{A_{10} - A_{20}u_0}\right). \quad (10)$$

This nonzero transplant time reflects ecological pressures that temporarily drive the system closer to the separatrix, overpowering the slow unstable manifold. In Fig. 5 we show two trajectories that numerically recapitulate these two cases.

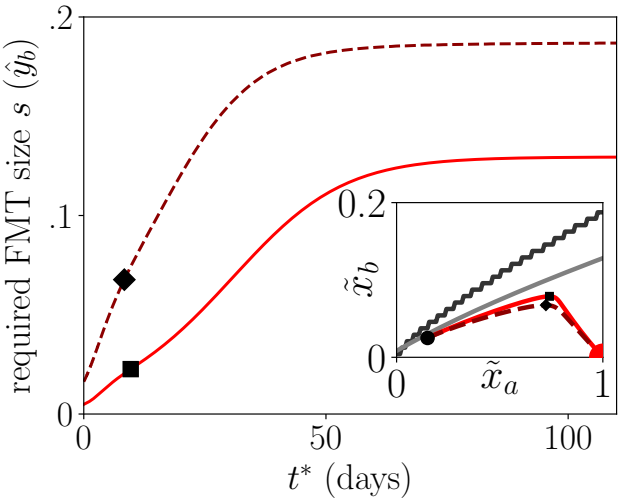

FIG. 6. Transient dynamics are laregly preserved under SSR. (inset) We display microbial trajectories of the CDI model (in-plane projection, dashed) and of the associated SSR-reduced model (solid). (main panel) At each time along these trajectories, we plot the minimum FMT size required to make the state health-prone, for a transplant made up of $\vec{y}_b$ (11D, dashed) or (0, 1) (2D, solid). We link phase space to the FMT size dynamics by indicating the time at which $\tilde{y}_b$ begins to decrease with a solid square (2D) or diamond (11D) in both the inset and main figure. Since the required FMT size s is the distance between a state and the separatrix, the similarity between the two time-courses of s indicate that SSR preserves transient dynamics.

VI. SSR PRESERVES FMT EFFICACY DEPENDENCE ON TIMING

The time courses of the required transplant size s qualitatively agree in both the high-dimensional CDI and 2D SSR systems, indicating that SSR largely preserves the transient microbial dynamics. In Fig. 6 (inset), as in Fig. 2, we consider the in-plane projection of the 11D (dashed) and corresponding SSR-reduced 2D (solid) microbial trajectories whose initial conditions lie on the plane spanned by ($\hat{y}_a$, $\hat{y}_b$) (11D) or ($\hat{x}_a$, $\hat{x}_b$) (2D). For each state along the two microbial trajectories, in Fig. 6 (main panel) we plot the transplant size s required to convert that state to health: for 11D (dashed) we consider a transplant made up of $\hat{y}_a$ and calculate s numerically with a bisection method, and for 2D (solid) we consider a transplant made up of $\hat{x}_a = (0,\ 1)$ and compute s analytically by using the exact form of the 2D separatrix given in Eq. (6).

In both systems, the microbial trajectories follow a fast stable manifold before switching to a slow manifold of some hyperbolic fixed point. As in the low-dimensional case, the flow along these fast and slow manifolds underpins the required transplant size s dynamics. In Fig. 2, the transition point between the fast and slow manifolds occurs at 8.4 days in 11D (solid diamond, main panel and

inset) and at 9.7 days in 2D (solid square). The similarity between the FMT timing results in 2D and in 11D suggests that the theoretical analysis of Section V may be applicable to more highly-resolved systems.

Some circumstantial evidence exists that supports these predictions of optimal FMT timing. Kang et al. [19] used a promising variant of FMT to induce seemingly long-term changes in the microbiome and symptoms of children with autism spectrum disorders. This FMT variant ("Microbiota Transfer Therapy," or MTT) first prescribed two weeks of vancomycin, then bowel cleaning, then a large dose of Standardized Human Gut Microbiota, and finally 2 months of daily maitenance FMT doses. In their study, they hoped that the efficacy of FMT would be improved by clearing the microbiome with antibiotics. This is consistent with our model's formulation, in which recently antibiotic-depleted microbiomes are especially malleable, during which time FMT is especially effective. The perceived success of Kang et al.'s clinical trial corroborates some of the results predicted from the 2D gLV equations, but future experiments are needed to explicitly test whether this timing property is a generic feature of microbial systems.

VII. DISCUSSION

SSR differs from other model reduction techniques [20, 21] by preserving core observable ecological features, since steady states and their stabilities are preserved under SSR. SSR could be directly applied to raw microbial data and by construction choose basis vectors that correspond to experimentally observed steady states [22]. The resulting models would consist of fewer variables and parameters that described interactions between steady states, with improved explanatory power.

Broadly, SSR realizes a natural progression towards the simplification of dynamical systems: while linearization approximates a dynamical system about a single steady state, SSR approximates a dynamical system about *two* steady states. We have shown that SSR produces the best possible in-plane 2D gLV approximation to high-dimensional gLV dynamics. Further, we have demonstrated the extent to which the 2D model captures the basins of attraction and transient dynamics of the 11D model. In all, these results suggest that SSR can act as a conduit between tractable low-dimensional analysis and the realistic behavior of high-dimensional systems. In addition to the computational efficiency of this technique, which employs the earlier analytic results rather than expensive simulations, SSR generates an intuition for the high-dimensional system with 2D cross-sections that link pairs of steady states.

With this tool we compressed an experimentally-derived model of CDI, uncovering the prediction that FMT is more effective when administered directly after antibiotics. By decomposing this complex and classic ecological model into analytically tractable ecological subspaces, we anchored a high-dimensional system to well-characterized two-dimensional systems. These techniques allow for previously inaccessible insight into high-dimensional dynamical landscapes, and offer a way to unravel the complicated landscape that accompanies complex systems and their behaviors.

ACKNOWLEDGMENTS

This material was based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1650114. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work was also supported by the David and Lucile Packard Foundation and the Institute for Collaborative Biotechnologies through contract no. W911NF-09-D-0001 from the U.S. Army Research Office. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix: Derivation of steady state reduction

We consider an N-dimensional gLV system given by Eq. (1) that exhibits steady states $\vec{y}_a$ and $\vec{y}_b$. As before, we consider new variables x_a and x_b in the direction of unit vectors $\hat{x}_a = \vec{y}_a/\|\vec{y}_a\|_2$ and $\hat{x}_b = \vec{y}_b/\|\vec{y}_b\|_2$, where $\|\cdot\|_k$. We consider the class of possible 2D gLV dynamics that exist on the plane spanned by $\hat{x}_a$ and $\hat{x}_b$, which we call in-plane dynamics. Here, we prove that the parameters prescribed by steady state reduction, given in Eq. (3), minimize the 2-norm of the deviation $\vec{\epsilon}$ between the actual and in-plane dynamics at every point on the plane.

We consider coefficients $\mathbf{c} = \{c_1, \ldots, c_6\}$ that parameterize the 2D gLV equations

$$\begin{aligned} \frac{\mathrm{d}x_a}{\mathrm{d}t} &= x_a\left(c_1 + c_2 x_a + c_3 x_b\right) \\ \frac{\mathrm{d}x_b}{\mathrm{d}t} &= x_b\left(c_4 + c_5 x_a + c_6 x_b\right), \end{aligned} \tag{A.1}$$

so that the in-plane dynamics are given by $\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \frac{\mathrm{d}x_a}{\mathrm{d}t}\hat{x}_a + \frac{\mathrm{d}x_b}{\mathrm{d}t}\hat{x}_b$. We note that on this plane, we may decompose any point $\vec{y} = \vec{y}_a x_a + \vec{y}_b x_b$.

Then, the deviation between the actual and in-plane dynamics $\vec{\epsilon}$ is

$$\vec{\epsilon}(x_a,\ x_b) = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} - \frac{\mathrm{d}\vec{y}}{\mathrm{d}t} \tag{A.2}$$

and is defined at every point on the plane $(x_a,\ x_b)$. We will show that the parameters precribed by SSR minimize the 2-norm of this deviation $\|\vec{\epsilon}\|_2$ at point on the plane.

The deviation $\vec{\epsilon}$ can be decomposed into the N-dimensional unit vectors $\hat{y}_i$, so that $\vec{\epsilon} = \sum_i \hat{y}_i \epsilon_i$, where the components ϵ_i are given by

$$\begin{aligned}\epsilon_i &= y_{ai}x_a\left((c_1-\rho_i)+\left(c_2-\sum_{j=1}^{N}K_{ij}y_{aj}\right)x_a+\left(c_3-\sum_{j=1}^{N}K_{ij}y_{bj}\right)x_b\right)\\ &\quad+y_{bi}x_b\left((c_4-\rho_i)+\left(c_5-\sum_{j=1}^{N}K_{ij}y_{aj}\right)x_a+\left(c_6-\sum_{j=1}^{N}K_{ij}y_{bj}\right)x_b\right)\\ &\equiv \epsilon_i^{10}x_a+\epsilon_i^{20}x_a^2+\epsilon_i^{11}x_ax_b+\epsilon_i^{01}x_b+\epsilon_i^{02}x_b^2,\end{aligned} \tag{A.3}$$

where we define components ϵ_i^{jk} that correspond to contributions by $x_a^j x_b^k$ terms. In the same way, we can decompose the deviation vector according to

$$\vec{\epsilon}=\vec{\epsilon}_{10}x_a+\vec{\epsilon}_{20}x_a^2+\vec{\epsilon}_{11}x_ax_b+\vec{\epsilon}_{01}x_b+\vec{\epsilon}_{02}x_b^2. \tag{A.4}$$

Since we wish to minimize this deviation at each point $(x_a,\ x_b)$, we must minimize each orthogonal contribution $\vec{\epsilon}_{jk}$. We note that each contribution is a function of one or two parameters ($\vec{\epsilon}_{10}(c_1)$, $\vec{\epsilon}_{20}(c_2)$, $\vec{\epsilon}_{01}(c_4)$, $\vec{\epsilon}_{02}(c_6)$, and $\vec{\epsilon}_{11}(c_3,\ c_5)$), which simplifies the minimization process. We therefore now find the set of optimal coefficients $\mathbf{c}^*=\{c_1^*,\ \ldots,\ c_6^*\}$ that minimize the 2-norm of each contribution $\|\vec{\epsilon}_{jk}\|_2$. For notational convenience, we note this is equivalent to minimizing the square of the 2-norm. Also, we define the vector quantity $\vec{y}_a{}^2$ to be an element-wise square of $\vec{y}_a$, given by $\vec{y}_a{}^2\equiv\sum_{i=1}^{N}\vec{y}_{ai}{}^2=\mathrm{diag}(\vec{y}_a)\vec{y}_a$ (and likewise for $\vec{y}_b{}^2$).

We first consider $\|\vec{\epsilon}_{10}\|_2^2$, given by

$$\|\vec{\epsilon}_{10}\|_2^2=\sum_{i=1}^{N}y_{ai}^2(c_1-\rho_i)^2. \tag{A.5}$$

To minimize this quantity with respect to c_1, we set

$$\frac{\mathrm{d}\|\vec{\epsilon}_{10}\|_2^2}{\mathrm{d}c_1}=\sum_{i=1}^{N}2y_{ai}^2(c_1-\rho_i)=0, \tag{A.6}$$

which is satified for

$$c_1^*=\frac{\sum_{i=1}^{N}y_{ai}^2\rho_i}{\sum_{i=1}^{N}y_{ai}^2}=\frac{\vec{y}_a{}^2\cdot\vec{\rho}}{\|\vec{y}_a\|_2^2}. \tag{A.7}$$

In a similar way, we minimize $\|\vec{\epsilon}_{20}\|_2^2$, $\|\vec{\epsilon}_{01}\|_2^2$, and $\|\vec{\epsilon}_{02}\|_2^2$ to find

$$c_2^*=\frac{\sum_{i=1}^{N}\left(y_{ai}^2\sum_{j=1}^{N}K_{ij}y_{aj}\right)}{\sum_{i=1}^{N}y_{ai}^2}=\frac{(\vec{y}_a{}^2)^TK\vec{y}_a}{\|\vec{y}_a\|_2^3}, \tag{A.8}$$

$$c_4^*=\frac{\sum_{i=1}^{N}y_{bi}^2\rho_i}{\sum_{i=1}^{N}y_{bi}^2}=\frac{\vec{y}_b{}^2\cdot\vec{\rho}}{\|\vec{y}_b\|_2^2}, \tag{A.9}$$

and

$$c_6^*=\frac{\sum_{i=1}^{N}\left(y_{bi}^2\sum_{j=1}^{N}K_{ij}y_{bj}\right)}{\sum_{i=1}^{N}y_{bi}^2}=\frac{(\vec{y}_b{}^2)^TK\vec{y}_b}{\|\vec{y}_b\|_2^3}. \tag{A.10}$$

Lastly, we consider the squared norm of the cross-term $\|\vec{\epsilon}_{11}\|_2$, given by

$$\begin{aligned}\|\vec{\epsilon}_{11}\|_2^2=\sum_{i=1}^{N}\Bigg[&y_{ai}\left(c_3-\sum_{j=1}^{N}K_{ij}y_{bj}\right)\\ &+y_{bi}\left(c_5-\sum_{j=1}^{N}K_{ij}y_{aj}\right)\Bigg]^2,\end{aligned} \tag{A.11}$$

and when minimized by c_3 and c_5 will satisfy

$$\begin{aligned}\frac{\mathrm{d}\|\vec{\epsilon}_{11}\|_2^2}{\mathrm{d}c_3}&=\sum_{i=1}^{N}2\left[y_{ai}^2\left(c_3-\sum_{j=1}^{N}K_{ij}y_{bj}\right)\right.\\ &\qquad\left.+y_{ai}y_{bi}\left(c_5-\sum_{j=1}^{N}K_{ij}y_{aj}\right)\right]\\ &=0\end{aligned} \tag{A.12}$$

and

$$\begin{aligned}\frac{\mathrm{d}\|\vec{\epsilon}_{11}\|_2^2}{\mathrm{d}c_5}&=\sum_{i=1}^{N}2\left[y_{ai}y_{bi}\left(c_3-\sum_{j=1}^{N}K_{ij}y_{bj}\right)\right.\\ &\qquad\left.+y_{bi}^2\left(c_5-\sum_{j=1}^{N}K_{ij}y_{aj}\right)\right]\\ &=0.\end{aligned} \tag{A.13}$$

After rearranging terms, these conditions read

$$\begin{aligned}c_3\sum_{i=1}^{N}y_{ai}^2+c_5\sum_{i=1}^{N}y_{ai}y_{bi}&=\sum_{i=1}^{N}\sum_{j=1}^{N}y_{ai}K_{ij}(y_{ai}y_{bj}+y_{bi}y_{aj})\\ c_3\sum_{i=1}^{N}y_{ai}y_{bi}+c_5\sum_{i=1}^{N}y_{bi}^2&=\sum_{i=1}^{N}\sum_{j=1}^{N}y_{bi}K_{ij}(y_{ai}y_{bj}+y_{bi}y_{aj})\end{aligned} \tag{A.14}$$

which we can solve for the coefficients c_3 and c_5 to find

$$c_3^*=\frac{\sum_{i,j=1}^{N}K_{ij}(y_{ai}y_{bj}+y_{bi}y_{aj})\left(\sum_{k=1}^{N}y_{ai}y_{bk}^2-y_{bi}y_{ak}y_{bk}\right)}{\left(\sum_{i=1}^{N}y_{ai}^2\right)\left(\sum_{i=1}^{N}y_{bi}^2\right)-\left(\sum_{i=1}^{N}y_{ai}y_{bi}\right)^2} \tag{A.15}$$

and

$$c_5^* = \frac{\sum_{i,j=1}^N K_{ij}(y_{ai}y_{bj} + y_{bi}y_{aj})\left(\sum_{k=1}^N y_{bi}y_{ak}^2 - y_{ai}y_{ak}y_{bk}\right)}{\left(\sum_{i=1}^N y_{ai}^2\right)\left(\sum_{i=1}^N y_{bi}^2\right) - \left(\sum_{i=1}^N y_{ai}y_{bi}\right)^2}. \tag{A.16}$$

However, when $\vec{y}_a$ and $\vec{y}_b$ are orthogonal, these coefficients simplify and become

$$c_3^* = \frac{\sum_{i=1}^N \left(y_{ai}^2 \sum_{j=1}^N K_{ij}y_{bj}\right)}{\sum_{i=1}^N y_{ai}^2} = \frac{(\vec{y}_a{}^2)^T K \vec{y}_b}{\|\vec{y}_a\|_2^2 \|\vec{y}_b\|_2}, \tag{A.17}$$

and

$$c_6^* = \frac{\sum_{i=1}^N \left(y_{bi}^2 \sum_{j=1}^N K_{ij}y_{aj}\right)}{\sum_{i=1}^N y_{bi}^2} = \frac{(\vec{y}_b{}^2)^T K \vec{y}_a}{\|\vec{y}_b\|_2^2 \|\vec{y}_a\|_2}. \tag{A.18}$$

Since the squared norms of the deviations $\|\epsilon_{jk}\|_2$ are convex, the coefficient set $\mathbf{c}^*$ that we have identified is a global minimum. Therefore, we have achieved our goal of identifing parameters that minimize the deviation between the in-plane and actual gLV dynamics for any point on the plane spanned by $\vec{y}_a$ and $\vec{y}_b$.

[1] J. R. Rohr, C. J. Salice, and R. M. Nisbet, Critical Reviews in Toxicology **46**, 756 (2016), https://doi.org/10.1080/10408444.2016.1190685.
[2] J. Lloyd-Price, G. Abu-Ali, and C. Huttenhower, Genome Medicine **8**, 51 (2016).
[3] V. B. Young, BMJ **356** (2017).
[4] J. E. Belizário and M. Napolitano, Frontiers in Microbiology **6**, 1050 (2015).
[5] D. Merenstein, N. El-Nachef, and S. V. Lynch, Journal of Pediatric Gastroenterology and Nutrition **59** (2014).
[6] J. Walter, M. X. Maldonado-Gómez, and I. Martínez, Current Opinion in Biotechnology **49**, 129 (2018).
[7] R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Rätsch, E. G. Pamer, C. Sander, and J. B. Xavier, PLoS Comput Biol **9**, 1 (2013).
[8] E. W. Jones and J. M. Carlson, PLOS Computational Biology **14**, 1 (2018).
[9] C. G. Buffie, V. Bucci, R. R. Stein, P. T. McKenney, L. Ling, A. Gobourne, D. No, H. Liu, M. Kinnebrew, A. Viale, E. Littmann, M. R. M. van den Brink, R. R. Jenq, Y. Taur, C. Sander, J. Cross, N. C. Toussaint, J. B. Xavier, and E. G. Pamer, Nature **517**, 205 (2015).
[10] J. L. Round and S. K. Mazmanian, Nat Rev Immunol **9**, 313 (2009).
[11] C. G. Buffie, I. Jarchum, M. Equinda, L. Lipuma, A. Gobourne, A. Viale, C. Ubeda, J. Xavier, and E. G. Pamer, Infect Immun **80**, 62 (2012).
[12] D. J. Messenheimer, S. M. Jensen, M. E. Afentoulis, K. W. Wegmann, Z. Feng, D. J. Friedman, M. J. Gough, W. J. Urba, and B. A. Fox, Clinical Cancer Research **23**, 6165 (2017).
[13] S. Wang, PLOS Computational Biology **13**, 1 (2017).
[14] Supplementary Information available at ****.
[15] V. Bucci, B. Tzen, N. Li, M. Simmons, T. Tanoue, E. Bogart, L. Deng, V. Yeliseyev, M. L. Delaney, Q. Liu, B. Olle, R. R. Stein, K. Honda, L. Bry, and G. K. Gerber, Genome Biology **17**, 121 (2016).
[16] Z. Hou, B. Lisena, M. Pireddu, F. Zanolin, S. Ahmad, and I. Stamova, *Lotka-Volterra and Related Systems: Recent Developments in Population Dynamics*, De Gruyter Series in Mathematics and Life Sciences (De Gruyter, 2013).
[17] S. Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*, Vol. 2 (Springer Science & Business Media, 2003).
[18] S. S. Hota and S. M. Poutanen, Open Forum Infect Dis **5**, ofy045 (2018), 29600252[pmid].
[19] D.-W. Kang, J. B. Adams, A. C. Gregory, T. Borody, L. Chittick, A. Fasano, A. Khoruts, E. Geis, J. Maldonado, S. McDonough-Means, E. L. Pollard, S. Roux, M. J. Sadowsky, K. S. Lipson, M. B. Sullivan, J. G. Caporaso, and R. Krajmalnik-Brown, Microbiome **5** (2017), 10.1186/s40168-016-0225-7.
[20] C. Gu, *Model order reduction of nonlinear dynamical systems*, Ph.D. thesis, UC Berkeley (2011).
[21] A. Goeke, S. Walcher, and E. Zerz, Physica D: Nonlinear Phenomena **345**, 11 (2017).
[22] P. I. Costea, F. Hildebrand, M. Arumugam, F. Bäckhed, M. J. Blaser, F. D. Bushman, W. M. de Vos, S. Ehrlich, C. M. Fraser, M. Hattori, C. Huttenhower, I. B. Jeffery, D. Knights, J. D. Lewis, R. E. Ley, H. Ochman, P. W. O'Toole, C. Quince, D. A. Relman, F. Shanahan, S. Sunagawa, J. Wang, G. M. Weinstock, G. D. Wu, G. Zeller, L. Zhao, J. Raes, R. Knight, and P. Bork, Nature Microbiology **3**, 8 (2018).