bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Steady state flux balance analysis game

Garud Iyengar and Mitch Perry*

IEOR Department, Columbia University

This version: August 6, 2022

Abstract

Flux balance analysis (FBA) for microbial communities often assumes a global objective function that all species cooperatively maximize in addition to maximizing their own growth. Combining community FBA with dynamic FBA to understand the time course and steady states of communities typically entails discretizing time and solving a community FBA model at each time point, a time-intensive process. We propose a dynamic community FBA model where species compete for metabolites to grow off of without needing to cooperate to maximize a community-level objective. An efficient method for computing steady state community compositions is provided, as well as methods for determining the stability of a steady state community to perturbations in biomass and invasion by species outside the community. The model and methods are applied to a model of four *E. coli* mutants with elements of competition (for shared metabolites) and cooperation (via mutants being auxotrophic for metabolites exported by other mutants), as well as a nine-species gut microbiome model.

1 Introduction

Interest in modeling microbial communities is growing as their impact on human health [11, 22, 23, 39] and the environment [16, 17, 30, 42] becomes more evident. The relationships between species in a community can show aspects of both cooperation, or dependence on the metabolites produced by other species, and competition for metabolites available in the environment. These interactions can lead to very different compositions of the microbial community, and understanding the underlying dynamics has important implications for potentially controlling the community composition by introducing new species or metabolites, or selectively eliminating certain species via antibiotics or immune action.

The simplest model for interaction between different species is the Lotka-Voltera model. This is a reduced-form model that summarises the interactions between species by an interaction matrix, and therefore is not able to represent interactions that are mediated by the exchange of metabolites [29]. Flux balance analysis (FBA) [32] and its many variants leverage genome-scale chemical reaction networks [43] to understand the interactions between microbial species at the level of metabolic interactions. FBA models typically assume that all species in the community cooperatively maximize a global community-wide objective function, possibly at the expense of lowering their own individual growth rates – notable exceptions being [6, 40] (see Section 1.1.2 for a detailed discussion). In this work, we propose a *non-cooperative* model where each species maximize its own grow rate, albeit taking into consideration the second-order effects of the metabolites it secretes into the environment. Our main contributions are as follows.

(a) We propose a game-theory based model for predicting the composition of a microbial community. In this model, the various species are players in a non-cooperative game where the goal of the players is to maximize their own utility, which we define as their growth rate. However, unlike most game theoretic models, the "actions", i.e. the reaction rates, feasible for a particular species depend on the actions taken by other species. We show that a generalized Nash equilibrium for this game exists, and can be efficiently computed by a globally convergent iterative algorithm that solves a convex quadratic

*Correspondence to: mop2107@columbia.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

program in each step. Our numerical experiments show that this algorithm scales to genome-scale metabolic models.

(b) We consider the dynamic setting where biomasses of the species evolve according to the instantaneous growth rate given by the generalized Nash equilibrium of the game corresponding to the current biomass composition. Computing the steady state of these dynamics by solving the ordinary differential equations involves discretizing time and solving a Nash equilibrium problem at each time point – a very time-intensive procedure. We show that a steady state always exists, and that it can be efficiently computed by solving a single generalized Nash equilibrium problem. We also show how to test the robustness of a steady state to perturbations in the biomasses of each species, as well as to invasion by a new microbial species not currently in the community. We develop an algorithm to identify a steady state that is stable to biomass perturbation if one exists.

(c) We discuss results of numerical experiments predicting the composition of a community of four *E. coli* mutants, with each mutant auxotrophic for a metabolite exported by one of the other mutants, and a nine-species gut microbiome model that can be directly compared to experimental data. We find that the steady state computed by a previously proposed community flux balance model that enforces cooperation is, in fact, unstable with respect to perturbations in the biomasses. We show that our approach implies that there are several qualitatively different steady states consistent with a given level of metabolite supply. This aligns well with studies showing that microbial communities can have a wide range of steady state behaviors even when the environmental conditions are the same [18, 33].

The rest of this paper is organized as follows. In Section 1.1 we review related literature and provide a background for our analysis. In Section 2.1 we introduce our game theory based model for predicting reaction rate profiles for each species in the community for a given biomass. In Section 2.2 we propose a model for the biomass evolution and show how to efficiently compute steady state biomass distributions. In Section 2.3 we characterize properties of stable steady states, and show how to use these properties to efficiently compute stable steady states. In Section 3 we use our procedure to predict biomass and reaction rate profiles for two model communities. Section 4 includes a discussion of our findings, a more in-depth analysis of how our model compares to two closely-related models [6, 8], and avenues for extending and improving our approach.

1.1 Related literature

Our work draws and builds on the following four different streams of literature.

1.1.1 Flux balance analysis (FBA)

FBA is an approach for studying metabolic networks, i.e. the collection of metabolites found in an organism as well as the genes that encode enzymes that catalyze metabolic reactions. FBA uses genome-scale metabolic models and tools from optimization to compute the steady-state consumption and production of metabolites in an organism in order to predict the growth rate of the organism [32, 34]. A major benefit of FBA is that it only requires the stoichiometric constants of the metabolites in each reaction, rather than kinetic rate constants needed for a traditional dynamical systems approach to metabolic modeling that, in practice, are difficult to experimentally estimate [13]. The FBA approach has been gaining popularity in the last decade because high-quality metabolic network reconstructions for a number of different organisms have become available [24, 25, 31].

FBA imposes mass balance constraints that ensure the amount of each metabolite excreted and absorbed across all species are equal, as well as lower and upper bounds on the rate of each reaction. Reaction rates are calculated, subject to these constraints, so as to maximize an objective function. In single species models, this objective function is typically the rate of some "biomass" reaction that represents the consumption of metabolites needed for members of the species to reproduce. On the other hand, canonical multi-species FBA models are bi-level optimization problems wherein a community-level objective function is optimized in the outer problem, and each species maximizes its biomass reaction in the sub-problems [8, 47]. Our model is a departure from this last point – in our model each species maximizes its own biomass without any regard

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

for any community level objective function, i.e. the species interact in a Nash equilibrium-like fashion. The importance of accounting for the behavior of selfish agents in community metabolism and biological systems in general has been increasingly recognized [6, 38], and requires game theoretical extensions of classic community FBA models.

1.1.2 Microbial communities as a game

We model the steady-state composition of a community of microbial species as the equilibrium outcome of the competition between the various bacterial species for nutrients (i.e. metabolites). While not an FBA-based approach, Dubinkina et. al. [12] and Goyal, Dubinkina, and Maslov [19] predict the composition of the microbiome by proposing that the equilibrium composition is a stable matching between species and nutrients. Harcombe et. al. [20] consider competition between microbial species on a lattice, where space is the limiting resource and each species is described by an FBA model.

Combining FBA with game-theoretic/multi-agent system models has potential to be a useful approach to modeling the interaction between multiple bacterial species. Zomorrodi and Maranas [47] study the composition of a microbial community via a bi-level optimization model, referred to as OptCom, where each species maximizes its own growth rate while simultaneously cooperating with the other species to maximize the overall community growth rate. Our model can be considered an extension of OptCom that considers the Nash equilibrium behavior of the community compositions without consideration for a community-level objective function, accounts for and explicitly calculates the biomasses of each species in the community, and considers the dynamic stability of computed solutions. The solution of OptCom also entails solving a large bi-level non-convex problem, whereas our model only requires solving a sequence of single-level quadratic programs.

The SteadyCom model presented by Chan et. al. [8] is another modification of OptCom that specifically addresses the issues of simultaneously computing species biomasses and computing solutions in a more tractable manner. SteadyCom assumes that, at steady-state, each species has the same biomass-weighted reaction flux and that the different species cooperate to maximize their common biomass reaction rate. With these assumptions the bi-level problem reduces to solving a sequence of linear programs. Unlike our model, SteadyCom does not consider Nash equilibrium behavior from the species, nor does it consider dynamic stability of its solutions. The model in [40] is also a multi-level model that first identifies a species that can grow off the supplied nutrients and solves the corresponding FBA problem; next, it works with species that can grow off the metabolic by products of the first set of species, until the total community biomass is optimized.

Zomorrodi and Segre [48] used FBA to compute the payoff matrix for a producer-consumer game amongst n bacterial species. While Nash equilibria for this game are efficiently computed for moderate-size payoff matrices via an integer linear programming formulation, the size of the payoff matrix grows exponentially in the number of bacterial species as well as the number of metabolites considered. The model also focuses on the presence or absence of reactions producing metabolites of interest, rather than studying the values of fluxes for each reaction in the entire set of reactions.

NECom [6] considers a game theoretic model very similar to the one proposed in this paper when the biomasses are fixed. There are subtle modeling differences between the two approaches that have important methodological consequences, in particular, in our approach a generalized Nash equilibrium is guaranteed to exist, whereas this might not be the case in NECom. Furthermore, [6] does not discuss biomass evolution dynamics. See Section 2.1.2 for a more detailed discussion of the two approaches.

We model the interaction between the species as a non-cooperative game in which each species attempts to maximize its growth rate while being constrained by FBA constraints that can depend on the fluxes of other species. We assume that the reaction rates are given by generalized Nash equilibria (GNE). This approach has the advantage of not imposing a community-level objective function on the individual species, allowing us to study the more realistic case where the microbial community can operate sub-optimally at the community level [36]. We also discuss the stability of the associated GNEs to perturbation in biomasses and invasion by new species.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1.1.3 Dynamic FBA

Typically, the quantity of interest when analyzing a microbial ecosystem is the relative proportion of each species. However, FBA does not provide this information, since it focuses on fluxes of the biomass reactions and not the biomasses. Dynamic FBA (DFBA) [26] models the dynamics of the biomasses by a Lotka-Volterra model with the growth rates given by the rates of the biomass reactions as calculated by FBA at that point in time. DFBA has been extended to model cooperative community dynamics [46], and a number of efficient implementations have been developed [5, 21]. We model the dynamics of microbial communities by assuming that the instantaneous growth rates for all species are given by a generalized Nash equilibrium for the FBA problem. However, we do not explicitly simulate the full dynamics. In fact, one of our main results is that one can compute steady states by solving for the generalized Nash equilibrium of a related game. This is more efficient that simulating the full dynamics, and similar in spirit to how SteadyCom calculates community biomass composition.

1.1.4 Generalized Nash equilibria

The available actions, i.e. reaction rates, for any player (species) in our model are limited by the supply of metabolites that are secreted by or shared with other species. Consequently, the actions for each player depend on the actions of the other players. Nash equilibria for such games are called generalized Nash equilibria (GNE), and the associated games are called generalized Nash equilibrium problems (GNEP). We consider GNEPs where the objective and the constraints are described by linear functions. Several different approaches exists for solving such games: solving the concatenated KKT conditions for all players [10], formulating the GNEP as the minimization of a Nikaido-Isoda function [45], and penalty methods that solve a sequence of simpler Nash equilibrium problems [14]. We use a method that iteratively solves, in parallel, a quadratically regularized version of the best response problem of each player with the other players' strategies fixed [15]. This approach is guaranteed to converge to a GNE, and we find in our numerical experiments that the method scales well as the number of species increases.

2 Methods

2.1 Model

We start by formulating a GNEP assuming that the biomass for each species is fixed. We consider the dynamics of the biomasses in Section 2.2.

2.1.1 Chemical reactions

Let $[K] = \{1, 2, ..., K\}$ denote the set of K microbial species present in a compartment or community. Let $\boldsymbol{x} = [x_k]_{k \in K}$ denote the biomasses of the K species. In this section, we assume that $\boldsymbol{x}$ is fixed. Let $[I_k] = \{1, \ldots, I_k\}$ index the set of metabolites in species $k \in [K]$. Let $[I_c] = \{1, \ldots, I_c\}$ index the set of metabolites found in the common space shared by the microbes[1]. In Figure 1, $K = 2$, $I_1 = 4$, $I_2 = 3$ and $I_c = 3$.

There are two sets of reactions associated with each species $k \in [K]$. The first set of reactions $[J_k] = \{1, \ldots, J_k\}$ change the concentration of metabolites within an individual cell of species k. The second set of reactions $[J_k^{\text{ex}}] = \{1, \ldots, J_k^{\text{ex}}\}$ are exchange reactions, i.e. these ingest or excrete metabolites from the cell to the shared compartment. In Figure 1, black lines schematically denote the internal set of reactions $[J_k]$ and the green lines indicate the exchange reactions $[J_k^{\text{ex}}]$.

Let $\boldsymbol{c}_k(t) \in \mathbb{R}_+^{I_k}$ denote the metabolite concentrations in species k at time t. Then

$$\dot{\boldsymbol{c}_k} = \boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}}, \tag{1}$$

where $\boldsymbol{v}_k \in \mathbb{R}^{J_k}$ (resp. $\boldsymbol{v}_k^{\text{ex}} \in \mathbb{R}^{J_k^{\text{ex}}}$) denotes the rates of the internal (resp. exchange) reactions, and the stochiometric matrices $\boldsymbol{R}_k \in \mathbb{R}^{I_k \times J_k}$ and $\boldsymbol{R}_k^{\text{ex}} \in \mathbb{R}^{I_k \times J_k^{\text{ex}}}$ detail the impact of the reactions on the

[1] Note that the same metabolite may be indexed differently in two different species and the compartment.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

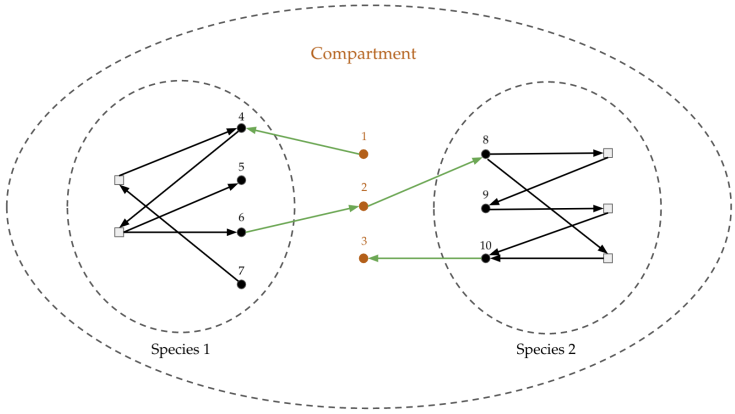


Figure 1: Schematic of microbial interactions, with grey squares representing reactions consuming and producing metabolites, represented by circles. Orange circles denote metabolites in the shared compartment, whereas black circles represent metabolites within a species. Arrows represent metabolites taking place in a reaction, with an arrow from (towards) a metabolite representing consumption (production) of that metabolite by the reaction. Green arrows represent exchange reactions that move metabolites between the shared compartment and a species, whereas black arrows represent reactions within a species. Note that the same type of metabolite is tracked as multiple different metabolites i.e. metabolite 2 in the shared compartment is the same type of metabolite as metabolite 6 in species 1 and metabolite 8 in species 2.

metabolite concentrations. The (i,j)-th element $R_k(i,j)$ of $\boldsymbol{R}_k$ denotes the change in the concentration of metabolite i when reaction j proceeds at a unit rate: $R_k(i,j) > 0$ if the metabolite i is produced by reaction j and $R_k(i,j) < 0$ if metabolite i is consumed by reaction j. The stochiometric matrices $\boldsymbol{R}_k^{\text{ex}}$ for the exchange reactions are special in the sense that every row has exactly one element equal to ± 1, where $R_{ex,k}(i,j) = +1$ (resp. $R_{ex,k}(i,j) = -1$) indicates that the reaction j ingests (resp. excretes) metabolite i from the compartment, i.e. increases (resp. decreases) the concentration of the metabolite. At steady-state, each species must ensure that $\dot{\boldsymbol{c}}_k = 0$ i.e.

$$\boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = \boldsymbol{0} \tag{2}$$

The fluxes within the species are further restricted by upper and lower bounds $\boldsymbol{u}_k$ and $\boldsymbol{\ell}_k$, respectively, that define maximum and minimum allowable fluxes, determined by the chemistry of the reactions

$$\boldsymbol{\ell}_k \leq \boldsymbol{v}_k \ \leq \boldsymbol{u}_k. \tag{3}$$

Assumption 1 (Bounded fluxes)**.** *The bounds* (3) *and the flux balance constraints* (2) *imply that the set of feasible values* $(\boldsymbol{v}_k, \boldsymbol{v}_k^{ex})$ *are bounded for all* $k \in [K]$.

The exchange reactions result in a change in concentration of the metabolites in the compartment. Let $\boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}}$, for $\boldsymbol{R}_k^{\text{c}} \in \mathbb{R}^{I_c \times J_k^{\text{ex}}}$, denote the metabolites ingested by each cell of species k. Note that $\boldsymbol{R}_k^{\text{c}} \neq \boldsymbol{R}_k^{\text{ex}}$ since the metabolites might be indexed differently in species k and the compartment. We assume that the inflow rate of metabolites into the compartment is fixed and is given by $\boldsymbol{f}_c \in \mathbb{R}_+^{I_c}$. Thus, we require that

$$\sum_{k \in [K]} x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} \leq \boldsymbol{f}_c, \tag{4}$$

where x_k denotes the biomass for species k. By allowing an inequality above, we are implicitly assuming that any unused metabolite diffuses away. The constraints (2)-(3) are species specific constraints, whereas (4) is a constraint *across* species.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

2.1.2 Nash equilibrium

The goal of the model is to predict the reaction rate profiles $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ that are likely to be observed in experiments. To this end, we posit that $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ corresponds to a Nash equilibrium of an appropriately defined game where the "players" are the species. This approach is similar to that in [6]. We will comment on the differences between the two approaches later in this section.

For given biomasses $\boldsymbol{x}$, the set of feasible reaction rate profiles $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ for all species is given by the set

$$\chi(\boldsymbol{x}) = \{(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) : (\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) \text{ satisfies (2) and (3) for all } k \in [K], \boldsymbol{v}^{\text{ex}} \text{ satisfies (4)}\} \tag{5}$$

We assume that the payoff for species k is given by

$$U_k(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) = \boldsymbol{g}_k^\top \boldsymbol{v}_k, \tag{6}$$

where $\boldsymbol{g}_k$ denotes a combination of reactions that lead to biomass growth, and the set of feasible reaction rates for species k given the rates of all other species $\ell \neq k$ is given by the set

$$\chi_k((\boldsymbol{v}_\ell, \boldsymbol{v}_\ell^{ex})_{\ell \neq k}; \boldsymbol{x}) = \left\{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) : ((\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}), (\boldsymbol{v}_\ell, \boldsymbol{v}_\ell^{ex})_{\ell \neq k}) \in \chi(\boldsymbol{x})\right\} \tag{7}$$

Let $[K(\boldsymbol{x})] = \{k : x_k > 0\}$ denote the set of species with positive biomasses. A reaction rate profile $(\boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ is a GNE for the game defined by (6) and (7) if there is no incentive for any species $k \in [K(\boldsymbol{x})]$ to deviate from $(\boldsymbol{\pi}_k, \boldsymbol{\pi}_k^{\text{ex}})$, i.e. for all $k \in K[\boldsymbol{x}]$,

$$\begin{aligned}
(\boldsymbol{\pi}_k, \boldsymbol{\pi}_k^{\text{ex}}) \in \underset{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})}{\text{argmax}}\ & \boldsymbol{g}_k^\top \boldsymbol{v}_k, \\
\text{s.t. } & \boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = \boldsymbol{0}, \\
& x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} + \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\pi}_\ell^{\text{ex}} \leq \boldsymbol{f}_c \\
& \boldsymbol{\ell}_k \leq \ \boldsymbol{v}_k \ \leq \boldsymbol{u}_k.
\end{aligned} \tag{8}$$

Our first result is that such a GNE exists.

Lemma 2.1 (Theorem 1 in [35])**.** *A GNE exists for the game* (8) *for any* $\boldsymbol{x}$ *and any chemical reaction parameters satisfying Assumption 1.*

We describe an algorithm for computing a GNE in Section A.1. Our game formulation is identical to the NECom algorithm proposed in [6], with one notable difference. In NECom, the constraints on the exchange reactions in the best response problem (8) are

$$x_k \boldsymbol{R}_k^c v_k^{\text{ex}} \leq \max\left\{\boldsymbol{0}, \boldsymbol{f}_c - \sum_{\ell \neq k} x_\ell \boldsymbol{R}_\ell^c \boldsymbol{\pi}_\ell^{\text{ex}}\right\}. \tag{9}$$

The rationale for introducing (9) is that it does not force altruism between the species i.e. no species is forced to produce metabolites for another species when doing so does not provide an immediate personal utility gain. However, introducing (9) results in serious methodological consequences. The best-response problem in NECom is non-convex, and therefore, Lemma 2.1 does not guarantee that a GNE exists. Furthermore, even if a GNE were to exist, the non-differentiability of (9) will make establishing stability of the GNE nearly impossible. It is also not clear whether imposing (9) is appropriate from a modeling perspective. Prohibiting forced altruism can cause the best-responses to be too myopic, and, as a consequence, rule out reaction profiles that are experimentally observed. For example, NECom when directly applied to auxotrophic communities, e.g. the four *E. coli* community analyzed in Section 3.1, predicts that the only possible community composition is the one where each species has *zero* biomass. In [6], this issue is handled in an ad hoc manner by introducing the additional objective of maximizing the total community biomass; it also employs a dFBA-like model with an NECom model solved at each time step and an initial supply of the auxotrophic metabolites. In contrast, our model is able to calculate non-zero steady states for this auxotrophic community without adding a community objective function, using dFBA, or having an external supply of the auxotrophic metabolites. Recall that our goal is to design a model that reliably predicts observed reaction rate profiles. In this context, the enforced altruism implicit in (8) is a better model since it is able reproduce observed behavior without any additional corrections.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

2.2 Population dynamics and steady states

In this section, we describe a model for biomass evolution. These dynamics are similar to those in SteadyCom [8], but with the biomass fluxes calculated using the Nash equilibrium approach described in Section 2.1.2 i.e. we assume that at each point in time the vector of reaction rates is a GNE with respect to a biomass-weighted FBA game. Note that the model in this section is an NECom model where the biomasses are not fixed parameters, but rather evolve over time.

We assume that the quantity of each species of bacteria evolves according to the following Lotka-Volterra system of equations

$$\frac{dx_k}{dt} = (\boldsymbol{g}_k^\top \boldsymbol{\pi}_k - \delta_k) x_k(t), \tag{10}$$

where δ_k denotes the death rate for species k and $\boldsymbol{g}_k^\top \boldsymbol{\pi}_k$ denotes the growth rate for species k corresponding to a GNE $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ of the game where the species biomasses are given by $\boldsymbol{x}(t)$. We are interested in efficiently computing steady states $\lim_{t\to\infty} \boldsymbol{x}(t)$ and analyzing their stability. Although our focus is on the steady state behavior of (10), temporal dynamics of (10) can be easily computed by discretizing time, and computing a GNE for (8) at every time step.

Next, we show how to compute the steady states directly, i.e. without solving (10). We start with the following definition.

Definition 2.1. $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *is a* steady state GNE *if* $(\boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *is a GNE for the* $\boldsymbol{x}$*-weighted game, and* $\boldsymbol{g}_k^\top \boldsymbol{\pi}_k = \delta_k$ *for all* $k \in K[\boldsymbol{x}]$.

The following result characterizes the set of steady states for (10).

Lemma 2.2. $\boldsymbol{x}$ *is a steady state for* (10) *if, and only if, there exist* $(\boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *such that* $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *is a steady state GNE as defined in Definition 2.1.*

Lemma 2.2 allows us to efficiently check whether a given $\boldsymbol{x}$ is a steady-state for the dynamics in (10). However, we do not yet know how to efficiently compute a steady-state $\boldsymbol{x}$. Next, we introduce a game that allows us to identify candidates for the steady-state. The players in this new quantity-weighted game are still the species, and the actions for each player are a triple of quantities $(x_k, \boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})$, the utility of the player k is $U_k(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) = x_k$, and the set of feasible actions is

$$\chi = \left\{ (\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) : \begin{array}{ll} R_k \boldsymbol{v}_k + R_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = 0, & k \in [K] \\ x_k \boldsymbol{\ell}_k \le \boldsymbol{v}_k \le x_k \boldsymbol{u}_k, & k \in [K] \\ \sum_{k\in[K]} \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} \le \boldsymbol{f}_c. & \end{array} \right\} \tag{11}$$

Thus, $(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{\text{ex}})$ is a GNE for this new quantity weighted game if, and only if,

$$\begin{aligned} (x_k, \boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{\text{ex}}) \in \underset{(\xi_k, \boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})}{\operatorname{argmax}} \ & \xi_k, \\ \text{s.t. } & \boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = \boldsymbol{0}, \\ & \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} + \sum_{\ell \neq k, \ell \in [K]} \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\omega}_\ell^{\text{ex}} \le \boldsymbol{f}_c, \\ & \xi_k \delta_k \le \boldsymbol{g}_k^T \boldsymbol{v}_k, \\ & \xi_k \boldsymbol{\ell}_k \le \boldsymbol{v}_k \le \xi_k \boldsymbol{u}_k. \end{aligned} \tag{12}$$

The next result establishes that a GNE $(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{\text{ex}})$ for (12) is a steady state GNE for (8), and therefore, $\boldsymbol{x}$ is a steady state for (10).

Lemma 2.3. *Let* $(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{ex})$ *be a GNE for the game* (12)*. Define* $(\boldsymbol{\pi}_k, \boldsymbol{\pi}_k^{ex}) = \frac{1}{x_k}(\boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{ex})$ *for* $k \in [K(\boldsymbol{x})]$*. Then* $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *is a steady-state GNE.*

Since all constraints and objective functions for (12) are linear, a simple extension of Lemma 2.1 implies the following result.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Corollary 2.1. *Suppose there exists* $(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{v}^{ex}) \in \chi$ *with* $x_k\delta_k \geq \boldsymbol{g}_k^\top \boldsymbol{v}_k$ *for all* k*. Then a GNE for the game* (12) *exists.*

A GNE for (12) can be efficiently computed by solving a sequence of quadratic programs, see Section A.1 in the appendix for more details. Ideally, one would like the set of GNEs for (12) to coincide with the set of steady state GNEs for (8). This is not necessarily the case. However, we are able to establish that the set of GNEs of (12) is a super set of steady state GNEs that are *stable* with respect to biomass fluctuations.

2.3 Stability to perturbation in biomass and invasion

In Section 2.2 we showed how to compute one possible steady state for the Lotka-Voltera dynamics (10). However, this steady state may be unstable to fluctuation in the biomasses. Such an unstable steady state is unlikely to be experimentally observed, and therefore, we need a methodology for identifying stable steady states. In the next section, we first describe how to check stability of a given steady state, and then describe an algorithm that can identify if one exists. In Section 2.3.2 we show how to check whether a given steady state is stable under invasion by another species.

2.3.1 Stability of steady state to fluctuations in biomass

Let $\boldsymbol{x}$ denote any steady state for 10. Then, there exists $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ such that $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ is a steady-state GNE for (8). We want to understand the stability of this steady state for fluctuations in the biomass $\boldsymbol{x}$, since stability is a crucial property of microbial ecosystems [9]. Suppose the steady state $\boldsymbol{x}$ is perturbed to $\boldsymbol{x}+\boldsymbol{p}(0)$, where $\boldsymbol{p}(0) \approx \boldsymbol{0}$. We are interested in understanding the dynamics $\boldsymbol{p}(t)$ for $t \geq 0$. In particular, whether $\boldsymbol{p}(t) \to 0$, and thereby keeps the community composition stable. Since the dynamics (10) are defined by Nash equilibria, we are effectively interested in understanding the comparative statics of the Nash equilibria [7].

We first discuss our approach informally and then state the formal mathematical result. Let $h_k(\boldsymbol{x}) = \boldsymbol{g}_k^\top \boldsymbol{\pi}_k(\boldsymbol{x}) = \delta_k$ denote the growth rate at the steady state $\boldsymbol{x}$, where we emphasize that the GNE $\boldsymbol{\pi}$ is a function of $\boldsymbol{x}$. Then

$$\begin{aligned}
\frac{dp_k}{dt} &= (x_k + p_k)(h_k(\boldsymbol{x}+\boldsymbol{p}) - \delta_k), \\
&\approx (x_k + p_k)\left(h_k(\boldsymbol{x}) + \sum_{\ell\in[K]} \frac{\partial h_k(\boldsymbol{x})}{\partial x_\ell} p_\ell - \delta_k\right) \\
&\approx x_k \sum_{\ell\in[K]} \frac{\partial h_k(\boldsymbol{x})}{\partial x_\ell} p_\ell,
\end{aligned}$$

where we assume that the perturbation is small enough so that quadratic terms in $\boldsymbol{p}$ can be ignored, and use the fact that $x_k\big(h_k(\boldsymbol{x}) - \delta_k\big) = 0$ for all k. Thus, we have that

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{M}\boldsymbol{p},$$

where $\boldsymbol{M} \in \mathbb{R}^{K\times K}$ with the (k, j) element $M_{kj} = x_k \frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$.

Lemma 2.4 (Theorem 9.3.2 in [3])**.** *The steady state* $\boldsymbol{x}$ *is stable to fluctuations in the biomass if the real part of all the eigenvalues of* $\boldsymbol{M}$ *are all less than* 0*.*

In Lemma A.1 we show how to compute $\frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$ for all k, j, in order to define $\boldsymbol{M}$.

Next, we establish a partial converse for Lemma 2.3. As a first step, we define a subset of steady state GNEs.

Definition 2.2 (Resource-constrained steady-state GNE)**.** *A steady-state GNE* $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *is* resource-constrained *if for all species* k*, the optimal value of the best response problem (8), when the biomasses* x_ℓ *and fluxes* $\boldsymbol{\pi}_\ell^{ex}$ *for all species* $\ell \neq k$ *are fixed, and the biomass of species* k *is set to* $z_k > x_k$*, is less than* δ_k*.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Lemma A.2 establishes that at a resource-constrained GNE each species' growth is limited by the consumption of at least one external metabolite. However, this is not a sufficient condition for a GNE to be resource-constrained. For that, we must have that increasing the biomass to $z_k > x_k$ forces a decrease in the uptake of a metabolite that ultimately results in a decrease in the rate of biomass producing reactions. A similar concept was formalized for the computation of steady states in SteadyCom (see the proofs of Theorems 1 and 2 in the Appendix of [8]).

Lemma 2.5. *$(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ is a resource-constrained steady-state GNE if and only if $(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\pi}_k^{ex})_{k\in[K]}$ is a GNE for (12).*

Thus, we have a partial converse for Lemma 2.3. Next, we show that stable steady state GNEs are a subset of resource-constrained steady states, and therefore, a subset of GNEs for (12).

Lemma 2.6. *Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{ex})$ is a steady state GNE, but it is not resource-constrained. Then it is unstable to perturbation in the biomass $\boldsymbol{x}$.*

Since observing a microbial community at an unstable equilibrium is unlikely, Lemma (2.6) allows us to focus our attention on resource-constrained steady state equilibria, and therefore, justifies using (12) to compute steady states. Recall that Corollary 2.1 establishes that the game (12) has a GNE. However, there are no general results that guarantee the existence of a stable steady state. In Appendix A.4 we develop an algorithm that exhaustively lists all possible candidates for steady state GNEs.

Lemma 2.7. *There exists an exhaustive search algorithm to check for the existence of stable steady state GNE that terminates in finite time.*

The search algorithm is not scalable to very large communities, since it involve enumerating the vertices of high-dimensional polytopes. However, it is highly parallelizable, and therefore, the existence of stable steady state GNE can be verified given sufficient computational resources. See Appendix A.4 for more details.

2.3.2 Stability to invasion

In the previous section, we were concerned with stability with respect to fluctuations in the biomass $\boldsymbol{x}$. Here, we consider stability of a steady state to invasion by a new species. Let $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ be a steady-state GNE for the K species, and index this new species by $K+1$. Let $h_{K+1}(\boldsymbol{x}, \epsilon)$ denote the growth rate of the invading species when its biomass $x_{K+1} = \epsilon$. Then species $K+1$ successfully invades if $\lim_{\epsilon\downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon) > \delta_{K+1}$. Lemma A.4 establishes that $\lim_{\epsilon\downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon)$ can be computed by solving an optimization problem.

3 Results

In this section, we use our proposed methodology for computing the steady state biomass and reaction profiles for two different bacterial communities: the four auxotrophic modified *E. coli* species and the nine species gut microbiome model studied in [8].

The results in Section 2.2 establish that a steady state for (10) can be computed by solving for a GNE of the game (12), and Lemma 2.4 allows us to check whether the steady state is stable. Furthermore, the exhaustive algorithm in Appendix A.4 is able to search for all possible stable steady states, if any, for (10). However, the running time of this algorithm was prohibitive even for the 4 species community. Instead, we used a sampling based algorithm to identify steady states. We sample N biomass vectors $\{\boldsymbol{x}^{(s)} : s = 1, \ldots, N\}$ and then call Algorithm 1 on each of the samples. We found that this procedure efficiently identified a stable steady state with $N = 2^{11}$ for the *E. coli* model and $N = 100$ for the nine species gut model. Therefore, we did not have to use the exhaustive search algorithm described in Appendix A.4.

SteadyCom [8] and NECom [6] are two other algorithms available in the literature to compute reaction rate profiles. In contrast to our approach, SteadyCom assumes that all the species collectively maximize their common growth rate. NECom computes reaction rate profiles by solving for a GNE for given biomasses $\boldsymbol{x}$; however, with a different best-response optimization – see the discussion below (8) in Section 2.1.2. The fluxes computed by NECom for the 4 bacteria *E. coli* community are different enough from our results that

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Algorithm 1 GameCom

1: Find feasible solution $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ of (5) with additional constraints $\boldsymbol{g}_k^\top \boldsymbol{v}_k \geq \delta_k$ for all species $k \in [K(\boldsymbol{x})]$
2: **if** (5) with additional constraints is infeasible **then**
3: **return** Infeasible
4: **else if** $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ is a GNE of (8) **then**
5: **return** $\boldsymbol{x}$
6: **else**
7: Solve (12) starting from $[x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\pi}_k^{\text{ex}}]_{k\in[K(\boldsymbol{x})]}$ to identify a steady state GNE.
8: **end if**

we have ommitted a direct comparison to NECom in the results section. We comment on the differences between NECom and GameCom in Section 4.

3.1 Auxotrophic *E. coli* community

The community was created by taking four copies of *E. coli* metabolism and inhibiting the export and production of specific metabolites in each copy so that each of the four modified models can only grow in the presence of the other three. More specifically:

(1) EC1 is able to synthesize arginine and phenylalanine and export arginine.

(2) EC2 is able to synthesize lysine and methionine and export lysine.

(3) EC3 is able to synthesize lysine and methionine and export methionine.

(4) EC4 is able to synthesize arginine and phenylalanine and export phenylalanine.

Note that, in spite of the interdependence, there is also a degree of competition, e.g. EC1 and EC4 compete for lysine and methionine, and all four species compete for common resources, e.g. glucose. Metabolites necessary for the community to grow are externally supplied; we use the "western diet" provided in [8]. SteadyCom computes a steady-state for this model with a common growth rate of $g_{sc}^* = 0.736\text{h}^{-1}$. In this steady-state, the proportion of EC1 is 25.3%, EC2 is 32.5%, EC3 is 18.5%, and EC4 is 23.7%.

3.1.1 Analysis of SteadyCom steady state SC

The first numerical experiment was to check whether the SteadyCom steady state (referred to as "SC") is a steady state in our computational framework, and if so, whether it is stable.

SteadyCom assumes that the death rate for each species is identical, and predicts the proportion of each species; therefore, the death rate is not required for the calculations. We predict the absolute biomasses and account for the slow down in growth as the biomasses increase but the nutrient supply remains constant. Consequently, we need to explicitly specify the death rate. In order to compare our predictions, we also assume that the death rate of each of the species is identically equal to δ. In order to ensure that there is a non-zero steady state, we ensure that $\delta \leq g_{sc}^* = 0.736\text{h}^{-1}$, the maximum SteadyCom growth rate. We found that our results were robust to the specific value of δ. For our numerical experiments we set $\delta = 0.5\text{h}^{-1}$, and then rescaled the SteadyCom biomass proportions and reaction fluxes to obtain a set of biomasses and fluxes that are feasible for (8). We find that this re-scaled SC is a steady-state GNE, but it is *unstable* to perturbations in the biomasses.

3.1.2 Steady states robust to perturbation in biomass $\boldsymbol{x}$

Next, we compute other steady state GNEs for the problem and analyze their stability. We sampled $N = 2^{11} = 2048$ biomasses for the four species uniformly from $[0, 1]^4$ using a Sobol sequence generator [41]. Of these, 408 samples remain after Step 2 in Algorithm 1. Each of these samples yielded a steady state GNE either in Step 4 or Step 7 of Algorithm 1. The screening step on all of the $N = 2^{11}$ biomass samples took approximately 1.9s per sample. Computing the steady state GNE for each of these 408 samples took approximately 3.5s on average, and checking for stability took 8.5s on average.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Very few stable steady states Our first result is that, although there are many steady state GNEs, very few of them are actually stable – only 4 of the 408 steady state GNE are stable to biomass perturbations. Thus, stability is a very strong selection criterion, and the analysis of the results of the model reduces to only a handful of steady states. In the left panel of Figure (2)(a), we compare the total community biomasses for stable and unstable steady states. We find the range for the total biomass is much larger for the unstable steady states when compared to stable steady states. As expected, the steady state SC has the largest community biomass (see vertical dotted line in the left-hand plot).

In the right panel in Figure (2)(a) we plot the (unstable) steady state SC, and the 4 stable steady states computed by GameCom. We see that the 4 steady states are very different from each other, and also from SC. Thus, we find that assuming perfect cooperation, as in SteadyCom, does not predict the range of behavior that is possible in the microbial community.

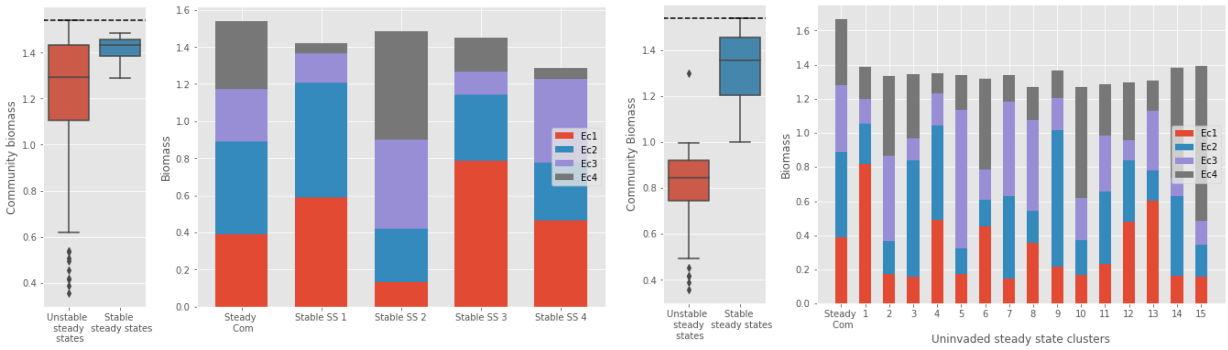
Biomass distribution across stable steady states Next, we carefully examine the variation in the biomass distributions for SC, SS2 and SS3. The goal here is to understand how two very distinct steady states can both be stable, and how they compare to SC. In Table 1 we list the externally supplied metabolites that are limiting, i.e. the corresponding compartment constraint is tight, and also the number of reactions that are equal to their upper or lower bounds, i.e. the bounds are active. We find that in all three steady states the four metabolites that are part of the auxotrophic interdependence between the species are fully utilized, and, among the externally supplied metabolites, only glucose is fully utilized. We also find that many more of the reaction bounds are active in SS2 and SS3, as compared to SC. Furthermore, we find that the very different steady states can be supported because even a small amount of key metabolites (e.g. arginine EC1, or lysine EC2, etc.) can sustain a relatively large quantity of the other species dependent on the key exported metabolite. Therefore, while the existence of each species is essential for the existence of the other species, they are not severely growth limiting, allowing for qualitatively different steady state GNE.

Instability of SC To understand why SC is unstable, we compare its fluxes to those in SS3. As noted earlier, in both SC and SS3, the four auxotrophic metabolites and glucose are fully utilized. Therefore, no species can simultaneously increase the rate at which it takes up one of these metabolites, and keep its biomass the same. As a rule of thumb, this will prohibit perturbations with increases in the biomasses to drive the biomass away from the steady state. In contrast to SC a number of the SS3 fluxes are at their upper or lower bound. For each *E. coli* mutant, a handful of these reactions are involved in carbon metabolism. For example, in EC1 the rate for the reaction catalyzed by glycogen branching enzyme is at its maximum possible value; consequently, in EC1 the rate of conversion of glucose to glycogen is limited even if more glucose becomes available. This prevents perturbations where species biomasses are decreased from destabilizing the steady state. As a simple example of this phenomenon, suppose the biomass of EC1 at SS3 is decreased while the biomasses of the other three species are held constant. The decrease in the biomass of EC1 results in more carbon available for the other three species; however, these species cannot utilize the available carbon because the rate of some reaction involved in glucose conversion is at its upper bound. The unused carbon is completely available for EC1 to grow back to its original biomass level. None of the carbon metabolic pathways are at their maximum fluxes in SC; therefore, such a perturbation can potentially be unstable. In a sense, SC is unstable because it is too efficient – in maximizing the community growth rate, the species' reactions had to be perfectly balanced, and as a result the community is capable of utilizing additional resources that become available due to perturbations.

3.1.3 Steady states stable to invasion

Next, we consider stability to invasion by a new species for each of the 408 steady states. In particular, we explore whether an *E. coli* mutant that can produce arginine, methionine, lysine, and phenylalanine, but not export any of these four metabolites, can invade. Computing the stability for each of the 408 steady states took approximately 25s per steady state. In contrast to stability to perturbations in the biomasses, stability to invasion by this fifth *E. coli* species is much more common - 333 of the 408 steady state GNEs cannot be invaded by new mutant. Note that the four stable steady states found in the previous section, as

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



(a) (b)

Figure 2: (a) Results for the community of four *E. coli* mutants. The box plots on the left the total community biomass for the four mutants. The bar charts on the right compare four stable steady states with the (unstable) SteadyCom steady state SC. (b) Compares steady states stable to invasion. The bar charts on the right compare steady states stable/unstable to invasion. The 333 steady states stable to invasion are grouped into 15 clusters.

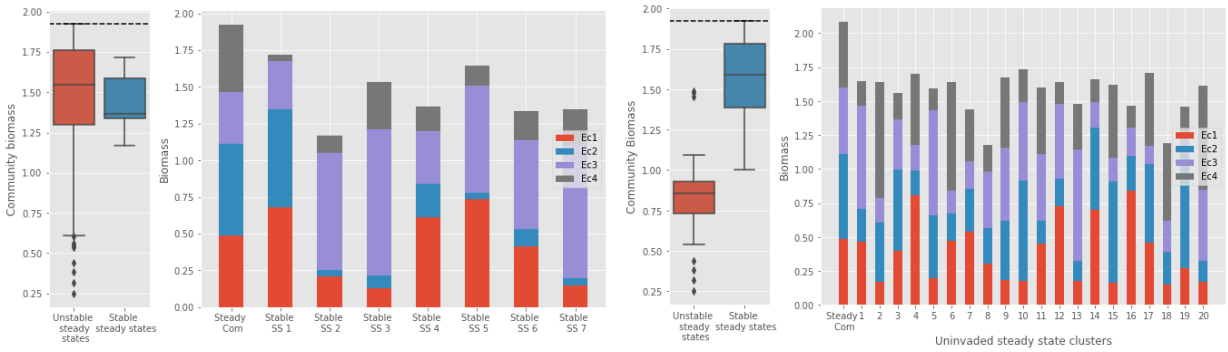
	Limited auxotrophy metabolites	Limited external resources	Number of active lower bounds	Number of active upper bounds
Stable steady state 2	Arg, Met, Lys, Phe	Glucose	Ec1: 12, Ec2: 12 Ec3: 11, Ec4: 15	Ec1: 10, Ec2: 12 Ec3: 11, Ec4: 15
Stable steady state 3	Arg, Met, Lys, Phe	Glucose	Ec1: 12, Ec2: 13 Ec3: 12, Ec4: 15	Ec1: 8, Ec2: 13 Ec3: 14, Ec4: 17
SteadyCom steady state	Arg, Met, Lys, Phe	Glucose	Ec1: 0, Ec2: 2 Ec3: 2, Ec4: 0	Ec1: 0, Ec2: 3 Ec3: 3, Ec4: 1

Table 1: Comparison of metabolite utilization and internal reaction fluxes for the second stable steady state in the right-hand part of figure (2), the third stable steady state from figure (2), and the SteadyCom steady state. The first column lists which of the four metabolites that the species are auxotrophic for are fully utilized i.e. $x_1\boldsymbol{R}_1^c[j]\boldsymbol{\pi}_1^{\text{ex}} + x_2\boldsymbol{R}_2^c[j]\boldsymbol{\pi}_2^{\text{ex}} + x_3\boldsymbol{R}_3^c[j]\boldsymbol{\pi}_3^{\text{ex}} + x_4\boldsymbol{R}_4^c[j]\boldsymbol{\pi}_4^{\text{ex}} = 0$, j being corresponding to the fully utilized metabolite. The second column lists the externally supplied metabolites that are fully utilized i.e. metabolites j where $\boldsymbol{f}_j > 0$. The third and fourth columns list, for each species, the number internal reactions that are at their lower bound and upper bounds, respectively.

well as SC, are all stable to invasion. The results in this section are summarized in Figure 2 (b). On the left panel, we plot the biomass distribution for steady states that are stable/unstable to invasion. On the right panel we plot the steady states resistant to invasion clustered into 15 groups. As was the case with stability to biomass perturbations, steady states resistant to invasion have a larger community biomass. In contrast to stability to biomass perturbations, the stability to invasion is determined by the availability of externally supplied metabolites: stable steady states have at least one fully exhausted externally supplied metabolite whereas unstable steady states do not. The exhausted nutrient prevents the growth of the invading species.

Figure (3) plots the results for the death rate $\delta = 0.4\text{h}^{-1}$. We note that the results are qualitatively similar to those in Figure 2 that corresponds to a death rate $\delta = 0.50\text{h}^{-1}$, in the sense that there are multiple stable steady states with a large variation in the biomass distribution. There are more steady states when the death rate is lower: 753 steady states, 7 of which are stable to perturbations in the biomass $\boldsymbol{x}$. This increase is in line with the flux variability analysis summarized in Figure 2 of [8]. Increasing the death rate to 0.60h^{-1} resulted in only 120 steady states, and none of them were stable. Hence, establishing that it is more difficult to sustain all four species as the death rate increases.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



(a) (b)

Figure 3: Results for death rate $\delta = 0.40$ h^{-1}. See footnote for Figure 2 for explanation of the plots.

3.2 Nine species gut microbiome model

Next, we apply our procedure to identify stable steady states for the gut microbiome model analyzed in SteadyCom [8]. The model consists of nine species comprising four phyla, as described in Table 2. We use the same model files for each species as used in SteadyCom, and use the same "Western diet" as the auxotrophic *E. coli* community in the previous section. In addition to being a more realistic community, the results of this model can be directly compared to experimental data from [28, 44].

We sampled $N = 100$ biomass vectors close to the experimentally observed data in [28, 44] as candidate steady states. The biomass x_k for each species k was chosen uniformly from an interval of width 0.1 centered at the average value for that species' proportion in [28] and [44], with a minimum value of 0.001. Even though the values reported in [28, 44] are proportions, we treated them as absolute biomass weights. This is justified as long as the death rate is set close to the maximum growth rate computed in SteadyCom, since SteadyCom normalizes the biomasses to 1 and finds a growth rate that's compatible with the absolute biomass values adding up to 1. The analysis in [8] shows a maximum possible community growth rate for this model is approximately 0.08h^{-1}, therefore, we set the death rate to $\delta = 0.06$h^{-1} for each species. From these 100 sampled biomass distributions, we were able to compute 20 steady states, and found that 2 of them were stable. See Figure 4 for the plots of the proportions of the different phyla for the two stable steady states. Also plotted are the phyla proportions as computed by SteadyCom and experimentally determined proportions from [28, 44].

The predictions from our model match up well with the experimental results in [28, 44], with the exceptions of having too much proteobacteria (10% to 15% in our model compared to 1% to 2% in experiments), and the firmicutes being comprised of a larger fraction of bacili compared to clostridia.

Note that we did note take any additional ad hoc steps to improve the match between our results and the experimental data. SteadyCom requires ad hoc manipulations of the model to get similar results. The grey bars in Figure 4 plot the proportions computed by SteadyCom using the model and diet as is – the entire community is comprised almost entirely of firmicutes. This is because firmicutes grow most efficiently off of fiber, the main carbon source in the diet. SteadyCom is able to overcome this issue in one of two ways: by requiring that the proportions of all species are constrained to be sufficiently above zero, and/or imposing upper bounds on the rate of fiber uptake for each species. However, since there are no experimental measurements for these species-specific fiber uptake upper bounds, SteadyCom resorts to sampling many different fiber uptake bounds, and computes average species composition across the samples. The results of this are the purple bars in Figure 4 – putting a cap on the amount of fiber that any given species can uptake allows for multiple species to uptake the fiber needed for their growth. In contrast, our competition-based approach is able to obtain reasonable community proportions without artificially constraining uptake reactions with unknown upper bounds.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Species	Phyla
B. theta	Bacteroidete
E. rectale	Firmicute
F. prausnitzii	Firmicute
E. faecalis	Firmicute
L. casei	Firmicute
S. thermophilus	Firmicute
B. adolescentis	Actinobacteria
E. coli	Proteobacteria
K. pneumoniae	Proteobacteria

Table 2: Species used in gut microbiome model

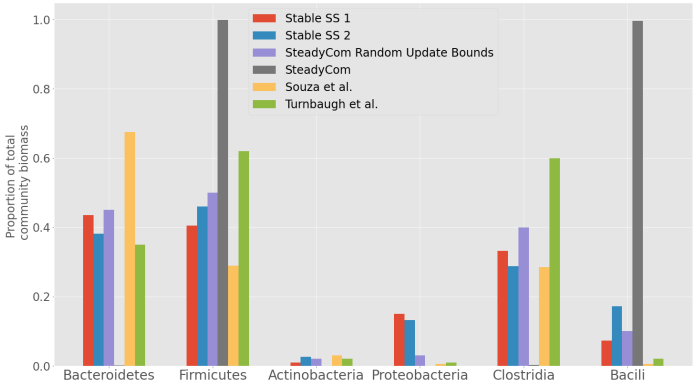


Figure 4: Stable steady state phyla proportions computed by our method (red and blue bar plots). Phyla proportions computed using SteadyCom without modifications to the model (grey bars) and with restrictions on the maximum fiber uptake rate (purple bars). Experimental data from [28] (yellow bars) and [44] (green bars) included for comparison. Two classes of firmicutes, clostridia and bacili, are also plotted since the majority of the firmicute population should be comprised of clostridia.

4 Discussion

We propose a new FBA model for predicting the composition of microbial communities. In our model the species in the community optimize their individual growth rate by competing for available metabolites. Therefore, we avoid imposing a community-level objective function that all species attempt to optimize in a cooperative manner. We show that a Generalized Nash Equilibrium (GNE) for the non-cooperative game always exists and can be computed by a globally convergent iterative algorithm that solves a quadratic program in each iteration. Depending on the chosen community-level objective function, computing a GNE can be significantly more efficient than solving the bi-level optimization problems encountered while computing the solution for a community level objective. SteadyCom simplifies to solving a small number of linear programs when the death rate for all species in the community is identical; however, when the death rates are species dependent one has to solve a non-convex bi-level optimization problem with no convergence guarantees.

Stability of a steady state to perturbations in the biomass composition and invasion from other microbial species is an important consideration for predicting community composition. This is because unstable steady states are unlikely to survive in biological environments that are inherently noisy. We propose a methodology for checking the stability of such steady states to both perturbation in biomasses and the introduction of new species. We also present an algorithm that can exhaustively list all possible candidates for steady states stable to biomass perturbation, and therefore, one can check whether a community has a such a stable steady state.

The numerical results reported in Section 3 confirm that removing the requirement that the members of the community must cooperate to optimize a shared objective allows a richer variety of steady state communities. This aligns well with studies showing that microbial communities can have a wide range of steady state behaviors even when the environmental conditions are the same [18, 33]. In our numerical experiments, we see that stability is a very strong selection criterion – most steady states are unstable.

GameCom shares a number of similarities to NECom, which have already been discussed in Sections 1.1.2 and 2.1.2. One of the differences between GameCom and NECom is that the shared metabolite constraints in NECom do not force the export of metabolites needed by other members of the community. Prohibiting forced

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

altruism can cause the best-responses to be too myopic, and, as a consequence, rule out reaction profiles that are experimentally observed. For example, NECom when directly applied to auxotrophic communities, e.g. the four *E. coli* community analyzed in Section 3.1, predicts that the only possible community composition is the one where each species has *zero* biomass. In [6], this issue is handled in an ad hoc manner by introducing an additional objective of maximizing total community biomass; and employing a dFBA-like model with a NECom model solved at each time step and an initial supply of the auxotrophic metabolites. In contrast, our model is able to calculate non-zero steady states for this auxotrophic community without adding a community objective function, using dFBA, or having an external supply of the auxotrophic metabolites. Recall that our goal is to design a model that reliably predicts observed reaction rate profiles. In this context, the enforced altruism implicit in (8) is a better model since it is able reproduce observed behavior without any additional corrections.

While we are able to identify the set of steady states to biomass perturbations, we are not able to efficiently identify the particular steady state that is likely to be the convergence point for a given initial state, other than simply solving the ODE (10). We are able to predict whether a particular steady state is stable or unstable to invasion. However, we are not able to list all possible candidates for steady states that are immune to invasion, nor are we able to efficiently predict the new steady state, or whether or not the invading species survives. Understanding how to control the composition of a microbial community, by introducing new species or metabolites, changing the supply of metabolites, or changing the death rates, is a larger question that is both theoretically interesting and practically important. Analyzing the impact of the death rates in particular could be an interesting way of understanding the interaction between the microbiome and the immune system (e.g. intestinal toll-like receptors). We believe GameCom is a first step towards answering these questions.

References

[1] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and Z. J. Kolter. Differentiable convex optimization layers. *Advances in neural information processing systems*, 32, 2019.

[2] A. Agrawal, S. Barratt, S. Boyd, E. Busseti, and W. M. Moursi. Differentiating through a cone program. *arXiv preprint arXiv:1904.09043*, 2019.

[3] W. E. Boyce, R. C. DiPrima, and D. B. Meade. *Elementary differential equations and boundary value problems*. John Wiley & Sons, 2021.

[4] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[5] J. D. Brunner and N. Chia. Minimizing the number of optimizations for efficient community dynamic flux balance analysis. *PLoS computational biology*, 16(9):e1007786, 2020.

[6] J. Cai, T. Tan, and S. H. J. Chan. Predicting nash equilibria for microbial metabolic interactions. *Bioinformatics*, 36(24):5649–5655, 2020.

[7] M. R. Caputo. The envelope theorem and comparative statics of nash equilibria. *Games and Economic Behavior*, 13(2):201–224, 1996.

[8] J. S. H. Chan, M. N. Simons, and C. D. Maranas. Steadycom: Predicting microbial abundances while ensuring community stability. *PLoS computational biology*, 13(5):e1005539, 2017.

[9] K. Z. Coyte, J. Schluter, and K. R. Foster. The ecology of the microbiome: networks, competition, and stability. *Science*, 350(6261):663–666, 2015.

[10] A. Dreves, F. Facchinei, C. Kanzow, and S. Sagratella. On the solution of the kkt conditions of generalized nash equilibrium problems. *SIAM Journal on Optimization*, 21(3):1082–1108, 2011.

[11] A. Du Toit. The gut microbiome and mental health. *Nature Reviews Microbiology*, 17(4):196–196, 2019.

[12] V. Dubinkina, Y. Fridman, P. P. Pandey, and S. Maslov. Multistability and regime shifts in microbial communities explained by competition for essential nutrients. *Elife*, 8:e49720, 2019.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[13] J. S. Edwards, M. Covert, and B. Palsson. Metabolic modelling of microbes: the flux-balance approach. *Environmental microbiology*, 4(3):133–140, 2002.

[14] F. Facchinei and C. Kanzow. Penalty methods for the solution of generalized nash equilibrium problems. *SIAM Journal on Optimization*, 20(5):2228–2253, 2010.

[15] F. Facchinei and J-S. Pang. Nash Equilibria: the variational approach. *Convex optimization in signal processing and communications*, page 443, 2010.

[16] P. G. Falkowski, T. Fenchel, and E. F. Delong. The microbial engines that drive earth's biogeochemical cycles. *science*, 320(5879):1034–1039, 2008.

[17] J. A. Fuhrman, J. A. Cram, and D. M. Needham. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*, 13(3):133–146, 2015.

[18] D. Gonze, L. Lahti, J. Raes, and K. Faust. Multi-stability and the origin of microbial community types. *The ISME journal*, 11(10):2159–2166, 2017.

[19] A. Goyal, V. Dubinkina, and S. Maslov. Microbial community structure predicted by the stable marriage problem. *bioRxiv*, page 235374, 2017.

[20] W. R. Harcombe, W. J. Riehl, I. Dukovski, B. R. Granger, A. Betts, A. H. Lang, G. Bonilla, A. Kar, N. Leiby, P. Mehta, et al. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell reports*, 7(4):1104–1115, 2014.

[21] K. Höffner, S. M. Harwood, and P. I. Barton. A reliable simulator for dynamic flux balance analysis. *Biotechnology and bioengineering*, 110(3):792–802, 2013.

[22] G. K. John and G. E. Mullin. The gut microbiome and obesity. *Current oncology reports*, 18(7):1–7, 2016.

[23] A. L. Kau, P. P. Ahern, N. W. Griffin, A. L. Goodman, and J. I. Gordon. Human nutrition, the gut microbiome and the immune system. *Nature*, 474(7351):327–336, 2011.

[24] Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson, and N. E. Lewis. Bigg models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic acids research*, 44(D1):D515–D522, 2016.

[25] S. Magnúsdóttir, A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jäger, J. Baginska, P. Wilmes, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature biotechnology*, 35(1):81, 2017.

[26] R. Mahadevan, J. S. Edwards, and F. J. Doyle III. Dynamic flux balance analysis of diauxic growth in escherichia coli. *Biophysical journal*, 83(3):1331–1340, 2002.

[27] J. H. May and R. L. Smith. Random polytopes: their definition, generation and aggregate properties. *Mathematical programming*, 24(1):39–54, 1982.

[28] M. Mitreva, Human Microbiome Project Consortium, et al. Structure, function and diversity of the healthy human microbiome. *nature*, 486(7402):207–214, 2012.

[29] B. Momeni, L. Xie, and W. Shou. Lotka-volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *Elife*, 6:e25051, 2017.

[30] E. B. Nelson. The seed microbiome: origins, interactions, and impacts. *Plant and Soil*, 422(1):7–34, 2018.

[31] A. Noronha, J. Modamio, Y. Jarosz, E. Guerard, N. Sompairac, G. Preciat, A. D. Daníelsdóttir, M. Krecke, D. Merten, H. S. Haraldsdóttir, et al. The virtual metabolic human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic acids research*, 47(D1):D614–D624, 2019.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[32] J. D. Orth, I. Thiele, and B. O. Palsson. What is flux balance analysis? *Nature biotechnology*, 28(3):245–248, 2010.

[33] E. Pagaling, K. Vassileva, C. G. Mills, T. Bush, R. A. Blythe, J. Schwarz-Linek, F. Strathdee, R. J. Allen, and A. Free. Assembly of microbial communities in replicate nutrient-cycling model ecosystems follows divergent trajectories, leading to alternate stable states. *Environmental microbiology*, 19(8):3374–3386, 2017.

[34] K. Raman and N. Chandra. Flux balance analysis of biological systems: applications and challenges. *Briefings in bioinformatics*, 10(4):435–449, 2009.

[35] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

[36] E. Ruppin, J. A. Papin, L. F. De Figueiredo, and S. Schuster. Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. *Current opinion in biotechnology*, 21(4):502–510, 2010.

[37] S. Schuster and C. Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems*, 2(02):165–182, 1994.

[38] S. Schuster, J-U. Kreft, A. Schroeter, and T. Pfeiffer. Use of game-theoretical methods in biochemistry and biophysics. *Journal of biological physics*, 34(1):1–17, 2008.

[39] R. F. Schwabe and C. Jobin. The microbiome and cancer. *Nature Reviews Cancer*, 13(11):800–812, 2013.

[40] S. Shoaie, P. Ghaffari, P. Kovatcheva-Datchary, A. Mardinoglu, P. Sen, E. Pujos-Guillot, T. De Wouters, C. Juste, S. Rizkalla, J. Chilloux, et al. Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell metabolism*, 22(2):320–331, 2015.

[41] I. M. Sobol, D. Asotsky, A. Kreinin, and S. Kucherenko. Construction and comparison of high-dimensional sobol generators. *Wilmott*, 2011(56):64–79, 2011.

[42] R. Souza, A. Ambrosini, and L. M. P. Passaglia. Plant growth-promoting bacteria as inoculants in agricultural soils. *Genetics and molecular biology*, 38(4):401–419, 2015.

[43] M. Terzer, N. D. Maynard, M. W. Covert, and J. Stelling. Genome-scale metabolic networks. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 1(3):285–297, 2009.

[44] P. J. Turnbaugh, M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, W. J. Jones, B. A. Roe, J. P. Affourtit, et al. A core gut microbiome in obese and lean twins. *nature*, 457(7228):480–484, 2009.

[45] A. von Heusinger and C. Kanzow. Optimization reformulations of the generalized nash equilibrium problem using nikaido-isoda-type functions. *Computational Optimization and Applications*, 43(3):353–377, 2009.

[46] A. R. Zomorrodi, M. M. Islam, and C. D. Maranas. d-optcom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS synthetic biology*, 3(4):247–257, 2014.

[47] A. R. Zomorrodi and C. D. Maranas. Optcom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput Biol*, 8(2):e1002363, 2012.

[48] A. R. Zomorrodi and D. Segrè. Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nature communications*, 8(1):1–12, 2017.

A Supplementary Materials

In this section, we include proofs for results presented in the main body of the paper, as well as additional results not included in the main body.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A.1 Algorithm for computing GNE

In this section, we describe an algorithm for computing a GNE for the game (8); computing GNE for (12) is similar.

Define the operator $\boldsymbol{T}(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) = [T_k(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})]_{k\in[K]}$ as follows:

$$T_k\left(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}\right) = \underset{(\boldsymbol{y},\boldsymbol{y}^{\text{ex}})\in\chi_{-k}((\boldsymbol{v}_\ell,\boldsymbol{v}_{\text{ex},\ell})_{\ell\neq k};\,\boldsymbol{x})}{\operatorname{argmax}} \left\{\boldsymbol{g}_k^\top \boldsymbol{y} + \frac{1}{2}||\boldsymbol{y}-\boldsymbol{v}_k||_2^2 + \frac{1}{2}||\boldsymbol{y}_k^{\text{ex}} - \boldsymbol{v}_k^{\text{ex}}||_2^2\right\}.$$

Note that the optimization problem implicit in the definition of T has a strongly convex quadratic objective function and linear constraints. Optimization problems of this form can be solved very efficiently [4].

Proposition 12.5 in [15] establishes that $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ is a GNE if, and only if, $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ is a fixed point of the operator $\boldsymbol{T}$. Proposition 12.17 in [15] applied to the problem defined here shows that $\boldsymbol{T}$ is non-expansive and, therefore, a fixed point of $\boldsymbol{T}$ can be computing using an averaging scheme of the form

$$(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) \leftarrow (\boldsymbol{v}, \boldsymbol{v}_k^{\text{ex}}) + \tau_j\left(\boldsymbol{T}(\boldsymbol{v}, \boldsymbol{v}_k^{\text{ex}}) - (\boldsymbol{v}, \boldsymbol{v}_k^{\text{ex}})\right)$$

for a step length $\tau_j \searrow 0$ such that $\sum_{j=0}^{\infty} \tau_j = +\infty$.

In practice, using such an averaging scheme to find a GNE is slow because the regularization term and the step length τ_j limit the change in (v, v^{ex}) at each iteration to be small. To counteract this, we define a new operator $\boldsymbol{T}^\alpha(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) = [T_k^\alpha(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})]_{k\in[K]}$ where

$$T_k^\alpha\left(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}\right) = \underset{(\boldsymbol{y},\boldsymbol{y}^{\text{ex}})\in\chi_{-k}((\boldsymbol{v}_\ell,\boldsymbol{v}_{\text{ex},\ell})_{\ell\neq k};\,\boldsymbol{x})}{\operatorname{argmax}} \left\{\boldsymbol{g}_k^\top \boldsymbol{y} + \frac{\alpha}{2}||\boldsymbol{y}-\boldsymbol{v}_k||_2^2 + \frac{\alpha}{2}||\boldsymbol{y}_k^{\text{ex}} - \boldsymbol{v}_k^{\text{ex}}||_2^2\right\}.$$

We compute a GNE by performing the averaging scheme with $\boldsymbol{T}^\alpha(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})$ where $\alpha = \alpha_j$ in iteration j, and it is a non-decreasing function of j that saturates at 1. By setting α_j small for j small, we allow the reaction rates to take large steps, and hence, speed up the convergence to a fixed point. By ensuring that α_j saturates at 1, we are guaranteed that the iterates converge to a fixed point of $\boldsymbol{T}$. Empirically, we found that we got very good performance by setting $\alpha_j = 10^{-3} + (1-10^{-3})\sigma(\frac{j}{50} - 6)$, where σ denotes the sigmoid function.

A.2 Calculating steady states

Lemma 2.2. *$\boldsymbol{x}$ is a steady state for* (10) *if, and only if, there exist $(\boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ such that $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ is a steady state GNE as defined in Definition 2.1.*

Proof. Suppose $\boldsymbol{x}$ is a steady state for the dynamics in (10). Then, for each k, either $x_k = 0$, or $x_k > 0$ and the growth rate of the species is equal to the death rate δ_k. By definition, the growth rate for a species k is given by $\boldsymbol{g}_k^\top \boldsymbol{\pi}_k$ for some GNE equilibrium $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ for an $\boldsymbol{x}$-weighted game. Thus, it follows $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ is a steady state GNE.

Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ is a steady state GNE. Then we have that $(\boldsymbol{\pi}, \boldsymbol{\pi}^{\text{ex}})$ is a GNE for the $\boldsymbol{x}$-weighted game, and $\boldsymbol{g}_k^\top \boldsymbol{\pi}_k = \delta_k$ for $x_k > 0$. Thus, it follows that $\boldsymbol{x}$ is a steady state for the dynamics in (10). □

Lemma 2.3. *Let $(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{ex})$ be a GNE for the game* (12)*. Define $(\boldsymbol{\pi}_k, \boldsymbol{\pi}_k^{ex}) = \frac{1}{x_k}(\boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{ex})$ for $k \in [K(\boldsymbol{x})]$. Then $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ is a steady-state GNE.*

Proof. $(\boldsymbol{\pi}, \boldsymbol{\pi}_{\text{ex}})$ is clearly feasible for (8) with weights $\boldsymbol{x}$. Suppose however that it is not a steady-state GNE i.e. for some k there exists $(\boldsymbol{\epsilon}_k, \boldsymbol{\epsilon}_k^{\text{ex}})$ feasible for (8) such that $\boldsymbol{g}_k^\top \boldsymbol{\epsilon}_k > g_k^T \boldsymbol{\pi}_k = x_k^{-1} g_k^\top \boldsymbol{\omega}_k \geq \delta_k$.

Define $(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) = x_k(\boldsymbol{\epsilon}_k, \boldsymbol{\epsilon}_k^{\text{ex}})$, and $\xi_k = x_k \boldsymbol{g}_k^\top \boldsymbol{\epsilon}_k / \delta_k > x_k$. Since $\boldsymbol{\ell}_k \leq 0$ and $\boldsymbol{u}_k \geq 0$, it follows that $(\xi_k, \boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})$ is feasible for (12). We have then that $(x_k, \boldsymbol{\omega}_k, \boldsymbol{\omega}_{\text{ex},k})$ is not optimal for (12). A contradiction. □

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A.3 Stability to perturbation in biomass

We need some notation in order to explain the construction of the system of equations for solving for $\frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$.

(i) Let

$$U_k(\boldsymbol{x}) = \{i : \nu_{ki}(\boldsymbol{x}) = u_{ki}\} \tag{13}$$
$$L_k(\boldsymbol{x}) = \{i : \nu_{ki}(\boldsymbol{x}) = l_{ki}\} \tag{14}$$

denote the indices for which the optimal solution $\boldsymbol{\pi}_k(\boldsymbol{x})$ for the linear program (8) at biomasses $\boldsymbol{x}$ is equal to the reaction upper or lower bound. Let

$$F_k(\boldsymbol{x}) = \{i : i \notin U_k(\boldsymbol{x}) \cup L_k(\boldsymbol{x})\} \tag{15}$$

be the sets of free non-exchange reactions at the optimal solution for (8) at biomasses $\boldsymbol{x}$.

(ii) Let $\boldsymbol{R}_k^{\mathrm{c}}[j]$ denote the j-th row of the matrix $\boldsymbol{R}_k^{\mathrm{c}}$. Let

$$E(\boldsymbol{x}) = \Big\{ j : x_k \boldsymbol{R}_k^{\mathrm{c}}[j] \boldsymbol{\pi}_k^{\mathrm{ex}}(\boldsymbol{x}) = f_c[j] - \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\mathrm{c}}[j] \boldsymbol{\pi}_\ell^{\mathrm{ex}}(\boldsymbol{x}) \Big\} \tag{16}$$

denote the set of community exchange constraints that are active in (8). Note that the set of active exchange constraints is independent of species $k \in K[\boldsymbol{x}]$.

The following characterizes the values of $\frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$ needed to construct $\boldsymbol{M}$.

Lemma A.1. *Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ denotes any steady state GNE. Suppose the dual linear program of (8) has a unique solution for all $k \in [K]$. In particular, let $\boldsymbol{\lambda}_k$ be the dual optimal solution corresponding to the metabolite exchange constraints. Also define $\mathbb{1}_{jk} = 1$ if $j = k$, $\mathbb{1}_{jk} = 0$ otherwise. Then we have*

$$\frac{\partial h_k(\boldsymbol{x})}{\partial x_j} = (\boldsymbol{\lambda}_k)^\top \left(-\frac{1}{x_k} \left(\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_l \boldsymbol{R}_\ell^c \frac{\partial \boldsymbol{\pi}_l^{ex}(\boldsymbol{x})}{\partial x_j} + \mathbb{1}_{jk} \boldsymbol{b}_k(\boldsymbol{x}) + (1 - \mathbb{1}_{jk}) \boldsymbol{R}_j^c \boldsymbol{\pi}_j^{ex}(\boldsymbol{x}) \right) \right) \tag{17}$$

where the partial derivatives of the fluxes with respect to changes in the biomasses are determined by solving the following system of equations over all $k, j \in [K(\boldsymbol{x})]$

$$\boldsymbol{R}_k^E \frac{\partial \boldsymbol{\pi}_k^{ex}}{\partial x_j} = -\frac{1}{x_k} \left(\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^E \frac{\partial \boldsymbol{\pi}_\ell^{ex}(\boldsymbol{x})}{\partial x_j} + \mathbb{1}_{jk} \boldsymbol{b}_k(\boldsymbol{x}; E) + (1 - \mathbb{1}_{jk}) \boldsymbol{R}_j^E \boldsymbol{\pi}_j^{ex}(\boldsymbol{x})) \right) \tag{18}$$
$$0 = \boldsymbol{R}_k^F \frac{\partial \boldsymbol{\pi}_k(F_k(\boldsymbol{x}))}{\partial x_j} + \boldsymbol{R}_k^{ex} \frac{\partial \boldsymbol{\pi}_k^{ex}(\boldsymbol{x})}{\partial x_j} \tag{19}$$

Proof. Recall the "utility" $h_k(\boldsymbol{x})$ for species k is given by the optimal value of the LP (8) at a steady state Nash equilibrium $(\boldsymbol{\pi}(\boldsymbol{x}), \boldsymbol{\pi}^{\mathrm{ex}}(\boldsymbol{x}))$ with the biomasses set to $\boldsymbol{x}$. Let

$$\boldsymbol{b}_k(\boldsymbol{x}) = \frac{1}{x_k} \left(\boldsymbol{f}_c - \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_{c,l} \boldsymbol{\pi}_\ell^{\mathrm{ex}} \right)$$

denote the right-hand side of the metabolite exchange constraints for the problem corresponding to species k. Since the dual solution is assumed to be unique, we have for all $\boldsymbol{b}_k$ sufficiently close to $\boldsymbol{b}_k(\boldsymbol{x})$ that the optimal

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

values $h_k(\boldsymbol{b}_k) = h_k(\boldsymbol{b}_k(\boldsymbol{x})) + (\boldsymbol{\lambda}_k)^\top(\boldsymbol{b}_k - \boldsymbol{b}_k(\boldsymbol{x}))$. Therefore,

$$\begin{aligned}\frac{\partial h_k(\boldsymbol{x})}{\partial x_j} &= (\boldsymbol{\lambda}_k)^\top \frac{\partial \boldsymbol{b}_k(\boldsymbol{x})}{\partial x_j} \\ &= (\boldsymbol{\lambda}_k)^\top \frac{\partial}{\partial x_j}\left(\frac{1}{x_k}\left(\boldsymbol{f}_c - \sum_{\ell\neq k, \ell\in[K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^c \boldsymbol{\pi}_\ell^{\text{ex}}(\boldsymbol{x})\right)\right) \\ &= (\boldsymbol{\lambda}_k)^\top \left(-\frac{1}{x_k}\left(\sum_{\ell\neq k, \ell\in[K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^c \frac{\partial \boldsymbol{\pi}_l^{\text{ex}}(\boldsymbol{x})}{\partial x_j} + \mathbb{1}_{jk}\boldsymbol{b}_k(\boldsymbol{x}) + (1-\mathbb{1}_{jk})\boldsymbol{R}_j^c \boldsymbol{\pi}_j^{\text{ex}}(\boldsymbol{x})\right)\right)\end{aligned}$$

Thus, it follows that the optimal solution for $\boldsymbol{b}_k$ will have $\boldsymbol{\pi}_k[U_k] = \boldsymbol{u}_k[U_k(\boldsymbol{x})]$, $\boldsymbol{\pi}_k[L_k(\boldsymbol{x})] = \boldsymbol{\ell}_k[L_k(\boldsymbol{x})]$, and

$$\boldsymbol{R}_k^E \boldsymbol{\pi}_k^{\text{ex}} = \frac{1}{x_k}\left(\boldsymbol{f}_c(E_k(\boldsymbol{x})) - \sum_{\ell\neq k, \ell\in[K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^E \boldsymbol{\pi}_\ell^{\text{ex}}\right) \tag{20}$$

and the free variables are determined by the systems of equations

$$\boldsymbol{R}_k^U \boldsymbol{u}_k(U_k(\boldsymbol{x})) + \boldsymbol{R}_k^L \boldsymbol{\ell}_k(L_k(\boldsymbol{x})) + \boldsymbol{R}_k^F \boldsymbol{\pi}_k(F_k(\boldsymbol{x})) + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{\pi}_k^{\text{ex}} = 0 \tag{21}$$

We can differentiate (20) and (21) to solve for $\frac{\partial \boldsymbol{\pi}_k^{\text{ex}}(\boldsymbol{x})}{\partial x_j}$ and $\frac{\partial \boldsymbol{\pi}_k(F_k(\boldsymbol{x}))}{\partial x_j}$ for all $k, j \in [K(\boldsymbol{x})]$. □

See Section A.5 for a discussion of the assumption that there exists a unique solution to the dual LP.

A.4 Existence of stable steady state equilibria

The following result gives more context to the resource-constrained steady state GNE defined in Section 2.3.1.

Lemma A.2. *Let $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{ex})$ be a resource-constrained steady-state GNE. Suppose dual optimal solutions for each species' best response problem (8) at $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{ex})$ is unique. Let $U_k(\boldsymbol{x})$, $L_k(\boldsymbol{x})$, and $E(\boldsymbol{x})$ denote the active sets defined in* (13)-(16)*. Then for each species $k \in [K(\boldsymbol{x})]$, there exists an active constraint $j \in E(\boldsymbol{x})$ such that*

$$f_{c,j} - \sum_{\ell\neq k} x_\ell \boldsymbol{R}_\ell^c[j]\boldsymbol{\nu}_\ell^{ex} > 0.$$

Proof. By contradiction suppose there exists a species k such that for all $j \in E(\boldsymbol{x})$, we have that

$$f_{c,j} - \sum_{\ell\neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}} \leq 0.$$

Recall the compartment exchange constraints imply that

$$x_k \boldsymbol{e}_j^\top \boldsymbol{R}_k^{\text{c}} \boldsymbol{\nu}_k^{\text{ex}} \leq \boldsymbol{e}_j^\top \left(\boldsymbol{f} - \sum_{\ell\neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}}\right) \leq 0.$$

Since $x_k > 0$, it follows that $\boldsymbol{R}_k^{\text{c}}[j]\boldsymbol{\nu}_k^{\text{ex}} \leq 0$ for all $j \in E(\boldsymbol{x})$. Thus, for all $z_k \geq x_k$

$$z_k \boldsymbol{e}_j^\top \boldsymbol{R}_k^{\text{c}} \boldsymbol{\nu}_k^{\text{ex}} \leq x_k \boldsymbol{e}_j^\top \boldsymbol{R}_k^{\text{c}} \boldsymbol{\nu}_k^{\text{ex}} \leq \boldsymbol{e}_j^\top \left(\boldsymbol{f}_c - \sum_{\ell\neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}}\right) \leq 0, \quad \forall j \in E(\boldsymbol{x}).$$

Since the constraint $j \notin E(\boldsymbol{x})$ are slack, it follows that there exists $z_k > x_k$ such that

$$z_k \boldsymbol{e}_j^\top \boldsymbol{R}_k^{\text{c}} \boldsymbol{\nu}_k^{\text{ex}} \leq x_k \boldsymbol{e}_j^\top \boldsymbol{R}_k^{\text{c}} \boldsymbol{\nu}_k^{\text{ex}} \leq \boldsymbol{e}_j^\top \left(\boldsymbol{f}_c - \sum_{\ell\neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}}\right) \leq 0, \quad \forall j \notin E(\boldsymbol{x}).$$

Thus, it follows that $(\boldsymbol{\pi}_k, \boldsymbol{\nu}_k^{\text{ex}})$ is feasible for (8) when the biomass of the k-th species is set $z_k > x_k$. Therefore, the optimal value when the biomasses x_ℓ and fluxes $\boldsymbol{\nu}_\ell^{\text{ex}}$ for all species $\ell \neq k$ are fixed, and the biomass of species k is set to $z_k > x_k$, is at least δ_k. Hence, $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{\text{ex}})$ is not resource-constrained. □

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Lemmas 2.5 and 2.6 established the relationship between resource-constrained steady states, GNE for (12), and stability of steady states.

Lemma 2.5. *$(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ is a resource-constrained steady-state GNE if and only if $(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\pi}_k^{ex})_{k\in[K]}$ is a GNE for (12).*

Proof. Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{\text{ex}})$ is a resource-constrained steady-state GNE, but $(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\pi}_k^{\text{ex}})_{k\in[K]}$ is not an equilibrium for (12). Then there exists a species k that can unilaterally improve its solution to (12) i.e. there exists biomass $z_k > x_k$ and fluxes $(\boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{\text{ex}}) \neq x_k(\boldsymbol{\pi}_k, \boldsymbol{\nu}^{\text{ex}})$ that are feasible for (12). Thus, it follows that $z_k^{-1}(\boldsymbol{\omega}_k, \boldsymbol{\omega}^{\text{ex}})$ is feasible (8) with an objective value at least δ_k. This contradicts that $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{\text{ex}})$ is resource-constrained.

Now suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{\text{ex}})$ is not a resource-constrained steady state Nash equilibrium. If $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{\text{ex}})$ is not a steady state GNE, it follows from Lemma 2.3 that $(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\nu}_k^{\text{ex}})_{k\in[K(\boldsymbol{x})]}$ is not a Nash equilibrium for (12). Suppose instead $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{\text{ex}})$ is a steady state GNE but it is not resource-constrained. Then there exists a species k such that the optimal value of the best response problem (8) with its biomass $z_k > x_k$ is at least δ_k, i.e. there exist feasible fluxes $(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})$ for (8) such that $\boldsymbol{g}_k^\top \boldsymbol{v}_k \geq \delta_k$. Thus, it follows that $(z_k, z_k\boldsymbol{v}_k, z_k\boldsymbol{v}_k^{\text{ex}})$ is feasible for (12) with an objective $z_k > z_k$. A contradiction. □

Lemma 2.6. *Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{ex})$ is a steady state GNE, but it is not resource-constrained. Then it is unstable to perturbation in the biomass $\boldsymbol{x}$.*

Proof. Suppose $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}_{\text{ex}})$ is not resource-constrained. Then there exists a species k and $z_k > x_k$ such that the problem (8) with the biomasses x_ℓ fluxes $\boldsymbol{\nu}_\ell^{\text{ex}}$ fixed and the biomass of species k set to z_k has an optimal solution $(\boldsymbol{h}_k, \boldsymbol{h}_k^{\text{ex}})$ with a value at least δ_k.

Thus, we have that $(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\nu}_k^{\text{ex}})$ and $(z_k, z_k\boldsymbol{h}_k, z_k\boldsymbol{h}_k^{\text{ex}})$ are both feasible for (12). Since the feasible set for (12) is convex, it follows that $\epsilon(z_k, z_k\boldsymbol{h}_k, z_k\boldsymbol{h}_k^{\text{ex}}) + (1-\epsilon)(x_k, x_k\boldsymbol{\pi}_k, x_k\boldsymbol{\nu}_k^{\text{ex}})$ is also feasible for (12) for all $\epsilon \in [0, 1]$. Thus, it follows that the fluxes

$$\boldsymbol{v}_k(\epsilon) = \frac{\epsilon z_k}{\epsilon z_k + (1-\epsilon)x_k}\,\boldsymbol{h} + \frac{(1-\epsilon)x_k}{\epsilon z_k + (1-\epsilon)x_k}\,\boldsymbol{\pi}_k$$

are feasible for (8) with species k biomass set to $\epsilon z_k + (1-\epsilon)x_k > x_k$ with an objective value at least δ_k. Therefore, the optimal value is at least δ_k

Consider perturbing the biomasses to $\hat{\boldsymbol{x}}$ with $\hat{x}_k = x_k + \epsilon(z_k - x_k)$ and $x_j = x_j$ for $j \neq k$. Then the result above shows the perturbation persists. Thus, it follows that the steady state GNE $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\nu}^{\text{ex}})$ is unstable. □

Now we describe the exhaustive search procedure that implies Lemma 2.7. The first step in the search is to write the FBA constraints in terms of the Elementary Flux Modes [37]. Assume that the elementary modes for species k are given by

$$\begin{bmatrix} \boldsymbol{v}_k \\ \boldsymbol{v}_k^{ex} \end{bmatrix} = \begin{bmatrix} \boldsymbol{N}_k \\ \boldsymbol{N}_k^{ex} \end{bmatrix} \boldsymbol{\lambda}_k,$$

where the mode vector $\boldsymbol{\lambda}_k$ satisfies $\mathbf{1}^\top\boldsymbol{\lambda}_k = 1$ and $\boldsymbol{\lambda}_k \geq \mathbf{0}$. Let $\{(x_k, \boldsymbol{\lambda}_k) : k \in [K]\}$ denote a Nash equilibrium defined in terms of the Elementary Flux modes. Then the analog of (12) in terms of $\boldsymbol{\lambda}_k$ is given by

$$\begin{aligned}
(x_k, \boldsymbol{\lambda}_k) \;\in\; & \operatorname*{argmax}_{(\xi_k, \boldsymbol{\mu}_k)} \xi_k, \\
& \text{s.t. } \boldsymbol{R}_k^c \boldsymbol{N}_k^{ex}\boldsymbol{\mu}_k + \sum_{\ell\neq k, \ell\in[K]} \boldsymbol{R}_\ell^c \boldsymbol{N}_\ell^{ex}\boldsymbol{\lambda}_\ell \leq \boldsymbol{f}_c, \\
& \xi_k\delta_k \leq \boldsymbol{g}_k^T \boldsymbol{N}_k\boldsymbol{\mu}_k, \\
& \xi_k\boldsymbol{\ell}_k \leq \boldsymbol{N}_k\boldsymbol{\mu}_k \leq \xi_k\boldsymbol{u}_k, \\
& \mathbf{1}^\top\boldsymbol{\mu}_k = \xi_k, \\
& \xi_k \geq 0, \quad \boldsymbol{\mu}_k \geq \mathbf{0}.
\end{aligned} \tag{22}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Define the per-species feasible set

$$\begin{aligned}\mathcal{C}_k &= \left\{\begin{bmatrix}\xi\\ \boldsymbol{\mu}\end{bmatrix} : \xi\delta_k \le \boldsymbol{g}_k^T\boldsymbol{N}_k\boldsymbol{\mu},\ \xi\boldsymbol{\ell}_k \le \boldsymbol{N}_k\boldsymbol{\mu} \le \xi\boldsymbol{u}_k,\ \mathbf{1}^\top\boldsymbol{\mu} = \xi,\ \xi \ge 0,\ \boldsymbol{\mu} \ge 0\right\},\\ &= \left\{\begin{bmatrix}\xi\\ \boldsymbol{\mu}\end{bmatrix} : \begin{bmatrix}\boldsymbol{c}_k & \boldsymbol{C}_k\end{bmatrix}\begin{bmatrix}\xi\\ \boldsymbol{\mu}\end{bmatrix} \ge \mathbf{0}, \xi \ge 0, \boldsymbol{\mu} \ge \mathbf{0}\right\}\end{aligned}$$

Note that the feasible set $\mathcal{C}_k$ is a cone, and the dual cone $\mathcal{C}_k^*$ is given by

$$\mathcal{C}_k^* = \left\{\begin{bmatrix}y\\ \bar{\boldsymbol{y}}\end{bmatrix} : \begin{bmatrix}y\\ \bar{\boldsymbol{y}}\end{bmatrix} \ge \begin{bmatrix}\boldsymbol{c}_k^\top\\ \boldsymbol{C}_k^\top\end{bmatrix}\boldsymbol{\gamma}, \boldsymbol{\gamma} \ge 0\right\}.$$

Let $\boldsymbol{A}_k = \boldsymbol{R}_k^c\boldsymbol{N}_k^{ex}$ for $k \in [K]$. Then (22) can be reformulated as follows:

$$\begin{aligned}(x_k, \boldsymbol{\lambda}_k) &\in \operatorname*{argmax}_{(\xi_k, \boldsymbol{\mu}_k)} \xi_k,\\ &\text{s.t. } \boldsymbol{A}_k\boldsymbol{\mu}_k + \sum_{\ell \ne k, \ell \in [K]} \boldsymbol{A}_\ell\boldsymbol{\lambda}_\ell \le \boldsymbol{f}_c,\\ &(\xi_k, \boldsymbol{\mu}_k) \in \mathcal{C}_k.\end{aligned} \tag{23}$$

Using duality one can show that $(\boldsymbol{x}, \boldsymbol{\lambda})$ is a steady state GNE if, and only if, there exists $\boldsymbol{z} = (\boldsymbol{z}_k)_{k\in[K]}$ and $\boldsymbol{w}$ such that $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{z}, \boldsymbol{w})$ is a solution to

$$\begin{aligned}\boldsymbol{A}_k^\top\boldsymbol{z}_k - \boldsymbol{e}_1 \in \mathcal{C}_k^*,\ \ \boldsymbol{z}_k \ge 0 \quad \forall k \in [K]\\ (x_k, \boldsymbol{\lambda}_k) \in \mathcal{C}_k \quad \forall k \in [K]\\ \sum_{\ell\in[K]} \boldsymbol{A}_\ell\boldsymbol{\lambda}_\ell + \boldsymbol{w} = \boldsymbol{f}_c,\ \ \boldsymbol{w} \ge \mathbf{0}\\ \boldsymbol{w}^\top\boldsymbol{z}_k = 0 \quad \forall k \in [K]\end{aligned} \tag{24}$$

Let m denote the dimension of the shared constraints. Then we can enumerate steady state GNE by computing the vertices of the polytopes given by (24) with $\boldsymbol{w}[J] = \mathbf{0}$ and $\boldsymbol{z}_k[J^c] = \mathbf{0}$ for all $k \in [K]$, and doing so for each $J \subseteq \{1, ..., m\}$. The following result shows that it is sufficient to only check the stability of such steady states.

Lemma A.3. *Fix $J \subset \{1, ..., m\}$. Let $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{z}, \boldsymbol{w})$ denote a feasible point of (24) corresponding to J that is not a vertex. Then it is unstable.*

Proof. Let $\mathcal{P}$ denote the polytope corresponding to J. Since $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{z}, \boldsymbol{w})$ is not a vertex of $\mathcal{P}$,

$$(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{z}, \boldsymbol{w}) = \gamma(\boldsymbol{x}^1, \boldsymbol{\lambda}^1, \boldsymbol{z}^1, \boldsymbol{w}^1) + (1-\gamma)(\boldsymbol{x}^2, \boldsymbol{\lambda}^2, \boldsymbol{z}^2, \boldsymbol{w}^2)$$

where $(\boldsymbol{x}^i, \boldsymbol{\lambda}^i, \boldsymbol{z}^i, \boldsymbol{w}^i) \in \mathcal{P}$, $i = 1, 2$, and $\gamma \in (0, 1)$. Thus, for every species i it is the case that either $x_i^1 > x_i$ or $x_i^2 > x_i$.

Fix a species k. Without loss of generality, assume $x_k^1 > x_k$. Define

$$(\boldsymbol{x}^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{z}^\alpha, \boldsymbol{w}^\alpha) = \big((1-\alpha)\gamma + \alpha\big)(\boldsymbol{x}^1, \boldsymbol{\lambda}^1, \boldsymbol{z}^1, \boldsymbol{w}^1) + (1-\alpha)(1-\gamma)(\boldsymbol{x}^2, \boldsymbol{\lambda}^2, \boldsymbol{z}^2, \boldsymbol{w}^2).$$

Note that $x_k^\alpha > x_k$ for all $\alpha > 0$, and $\boldsymbol{x}^\alpha = \boldsymbol{x} + \alpha(\boldsymbol{x}^1 - \boldsymbol{x})$, and therefore, for all α sufficiently small, $\boldsymbol{x}^\alpha$ is a small perturbation of $\boldsymbol{x}$. Since $(\boldsymbol{x}^\alpha, \boldsymbol{\lambda}^\alpha, \boldsymbol{z}^\alpha, \boldsymbol{w}^\alpha) \in \mathcal{P}$, it follows that $(\boldsymbol{x}^\alpha, \boldsymbol{\lambda}^\alpha)$ is a GNE for (12), and therefore, by Lemma 2.3, $(\boldsymbol{x}^\alpha, \boldsymbol{\pi}^\alpha)$ is a steady state GNE for (8), where $\boldsymbol{\pi}^\alpha$ is constructed using $\boldsymbol{\lambda}^\alpha$. Recall that $x_k^\alpha > 0$, therefore, for $\boldsymbol{x}^\alpha$ to be a steady state GNE, we must have that the growth $\delta_k = \boldsymbol{g}_k^\top\boldsymbol{\pi}_k^\alpha$. Thus, there exists a GNE such that the state $\boldsymbol{x}^\alpha$ does *not* revert back to $\boldsymbol{x}$ following a perturbation $\boldsymbol{p} = \alpha(\boldsymbol{x}^1 - \boldsymbol{x})$. □

The vertices of the polytopes given by (24) corresponding to J can be enumerated by generating random cost vectors $\boldsymbol{c}$ and solving LPs maximizing $(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{z}, \boldsymbol{w}) \cdot \boldsymbol{c}$ subject to (24) and the active sets J. Also, in practice one could just search over sets J corresponding to metabolites that are externally supplied and/or metabolites that are involved in cross-feeding reactions, rather than all possible subsets of $\{1, ..., m\}$ which will involve many community metabolites that are neither supplied nor involved in cross-feeding.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A.5 Stability to invasion

Lemma A.4. *Suppose a steady-state GNE* $(\boldsymbol{x}, \boldsymbol{\pi}, \boldsymbol{\pi}^{ex})$ *satisfies the conditions in Lemma A.1. Then, there exists matrices* $\{\boldsymbol{B}_\ell : \ell \in K[\boldsymbol{x}]\}$ *such that*

$$\begin{aligned} \lim_{\epsilon \downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon) = \max_{(\boldsymbol{v}, \boldsymbol{v}^{ex})} \; & \boldsymbol{g}_{K+1}^\top \boldsymbol{v}, \\ s.t. \; & \boldsymbol{R}_{K+1}\boldsymbol{v} + \boldsymbol{R}_{K+1}^{ex}\boldsymbol{v}^{ex} = \boldsymbol{0}, \\ & \boldsymbol{e}_j^\top \Big(\boldsymbol{R}_{K+1}^c + \sum_{\ell \in K[\boldsymbol{x}]} x_\ell \boldsymbol{R}_\ell^c \boldsymbol{B}_\ell \Big) \boldsymbol{v}^{ex} \le 0, \; \forall j \in E(\boldsymbol{x}), \\ & \boldsymbol{\ell}_{K+1} \le \; \boldsymbol{v} \; \le \boldsymbol{u}_{K+1}, \end{aligned} \tag{25}$$

where $E(\boldsymbol{x}) = \{j : \boldsymbol{e}_j^\top (\boldsymbol{f}_c - \sum_{\ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^c \boldsymbol{\nu}_\ell^{ex}) = 0\}$, $\boldsymbol{e}_j$ *denotes the* j*-th standard basis vector, and the matrices* $\{\boldsymbol{B}_\ell : \ell \in [K]\}$ *are defined in the proof below.*

Proof. Consider the LP (8). Convert this LP to standard form by adding slacks $\boldsymbol{s}_k \ge 0$ to the exchange constraints, and splitting the free variables $\boldsymbol{v}_k^{\text{ex}}$ into the difference $\boldsymbol{v}_k^{\text{ex}} = \boldsymbol{v}_k^{\text{ex},+} - \boldsymbol{v}_k^{\text{ex},-}$, of non-negative variables $\boldsymbol{v}_k^{\text{ex},\pm} \ge 0$. Let $F_k(\boldsymbol{x}) = \{i : i \notin U_k(\boldsymbol{x}) \cup L_k(\boldsymbol{x})\}$ denote components of the vector $\boldsymbol{\pi}$ that are *not* fixed to their upper or lower bounds, $F_k^{\text{ex}}(\boldsymbol{x}) = \{i : \nu_{k,i}^{\text{ex}}(\boldsymbol{x}) \neq 0\}$ the set of non-zero exchange reactions, and $E(\boldsymbol{x})$ the set of active exchange reaction constraints in (8). Then, for all small enough perturbations $\Delta \boldsymbol{f}_c$, the change in the optimal solution $(\boldsymbol{\pi}_k, \boldsymbol{\nu}_k^{\text{ex}})$ is given by the solution to the following linear equations:

$$\underbrace{\begin{bmatrix} \boldsymbol{R}_k^F & \boldsymbol{R}_k^{\text{ex}}[:, F_k^{ex}(\boldsymbol{x})] \\ \boldsymbol{0} & x_k \boldsymbol{R}_k^c[E(\boldsymbol{x}), F_k^{ex}(\boldsymbol{x})] \end{bmatrix}}_{\boldsymbol{\Omega}_k} \begin{bmatrix} \Delta \boldsymbol{\pi}_k[F_k(\boldsymbol{x})] \\ \Delta \boldsymbol{\pi}_k^{\text{ex}}[F_k^{\text{ex}}(\boldsymbol{x})] \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \Delta \boldsymbol{f}_c[E(\boldsymbol{x})] \end{bmatrix},$$

where $\boldsymbol{R}_k^{\text{ex}}[:, F_k^{ex}(\boldsymbol{x})]$ denotes the submatrix obtained by keeping all the rows and only the columns in $F_k^{ex}(\boldsymbol{x})$, $\boldsymbol{R}_k^c[E(\boldsymbol{x}), F_k^{ex}(\boldsymbol{x})]$ denote the submatrix obtained by taking the rows in $E(\boldsymbol{x})$ and the columns in $F_k^{ex}(\boldsymbol{x})$, and $\Delta \boldsymbol{\pi}_k^{\text{ex}}[\overline{F}_k^{\text{ex}}(\boldsymbol{x})] = \boldsymbol{0}$, where $\overline{F}_k^{\text{ex}}(\boldsymbol{x})$ denotes the complement of the set $F_k^{\text{ex}}(\boldsymbol{x})$. Consequently,

$$\Delta \boldsymbol{\nu}_k^{\text{ex}}[F_k^{\text{ex}}(\boldsymbol{x})] = \boldsymbol{\Gamma}_k \Delta \boldsymbol{f}_c[E(\boldsymbol{x})], \qquad \Delta \boldsymbol{\pi}[F_k(\boldsymbol{x})] = \boldsymbol{\Psi}_k \Delta \boldsymbol{f}_c[E(\boldsymbol{x})],$$

where the matrices $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\Psi}_k$ are appropriately defined submatrices of $\boldsymbol{\Omega}^{-1}$. This relationship is valid only if the following constraints hold

$$\begin{aligned} \boldsymbol{\ell}_k[F_k(\boldsymbol{x})] \le \boldsymbol{\pi}_k[F_k(\boldsymbol{x})] + \Delta \boldsymbol{\pi}[F_k(\boldsymbol{x})] \le \boldsymbol{u}[F_s(\boldsymbol{x})], \\ x_k \boldsymbol{R}_k^{\text{c}}[\bar{E}(\boldsymbol{x}), :](\boldsymbol{\nu}_k^{\text{ex}} + \Delta \boldsymbol{\nu}_k^{\text{ex}}) + \textstyle\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\text{c}}[\bar{E}(\boldsymbol{x}), :] \boldsymbol{\pi}_\ell^{\text{ex}} \le \boldsymbol{f}_c[\bar{E}(\boldsymbol{x})] + \Delta \boldsymbol{f}[\bar{E}(\boldsymbol{x})], \end{aligned} \tag{26}$$

where $\bar{E}[\boldsymbol{x}]$ denotes the complement of the set $E(\boldsymbol{x})$. Note that these constraints hold as long as the norm of the disturbance Δf_c is sufficiently small.

The effective change in the RHS for species k is given by

$$\Delta \boldsymbol{f}_c[E(\boldsymbol{x})] \;\; = \;\; \sum_{\ell \neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}}[E(\boldsymbol{x}), F_\ell^{\text{ex}}(\boldsymbol{x})] \Delta \boldsymbol{\nu}_\ell^{\text{ex}}[F_\ell^{\text{ex}}(\boldsymbol{x})] + \epsilon \boldsymbol{R}_{K+1}^{\text{c}}[E(\boldsymbol{x}), :] \boldsymbol{v}_{K+1}^{\text{ex}}.$$

Thus, for all $k \in K[\boldsymbol{x}]$, we have that

$$\begin{aligned} \Delta \boldsymbol{\nu}_k^{\text{ex}}[F_k^{\text{ex}}(\boldsymbol{x})] \;\; &= \;\; \boldsymbol{\Gamma}_k \Big(\sum_{\ell \neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}}[E(\boldsymbol{x}), F_\ell^{\text{ex}}(\boldsymbol{x})] \Delta \boldsymbol{\nu}_\ell^{\text{ex}}[F_\ell^{\text{ex}}(\boldsymbol{x})] + \epsilon \boldsymbol{R}_{K+1}^{\text{c}}[E(\boldsymbol{x}), :] \boldsymbol{v}_{K+1}^{\text{ex}} \Big), \\ \Delta \boldsymbol{\pi}_k[F_k(\boldsymbol{x})] \;\; &= \;\; \boldsymbol{\Psi}_k \Big(\sum_{\ell \neq k} x_\ell \boldsymbol{R}_\ell^{\text{c}}[E(\boldsymbol{x}), F_\ell^{\text{ex}}(\boldsymbol{x})] \Delta \boldsymbol{\nu}_\ell^{\text{ex}}[F_\ell^{\text{ex}}(\boldsymbol{x})] + \epsilon \boldsymbol{R}_{K+1}^{\text{c}}[E(\boldsymbol{x}), :] \boldsymbol{v}_{K+1}^{\text{ex}} \Big). \end{aligned}$$

Thus, the perturbation can be written as $\Delta \boldsymbol{\nu}_k^{\text{ex}}[F_k^{\text{ex}}(\boldsymbol{x})] = \epsilon \boldsymbol{B}_k \boldsymbol{R}_{K+1}^{\text{c}}[E(\boldsymbol{x}), :] \boldsymbol{v}_{K+1}^{\text{ex}}$, and $\Delta \boldsymbol{\pi}_k[F_k(\boldsymbol{x})] = \epsilon \boldsymbol{C}_k \boldsymbol{R}_{K+1}^{\text{c}}[E(\boldsymbol{x}), :] \boldsymbol{v}_{K+1}^{\text{ex}}$,, where $(\boldsymbol{B}_k, \boldsymbol{C}_k)$, $k \in K[\boldsymbol{x}]$, are a function of $(\boldsymbol{\Gamma}_k, \boldsymbol{\Psi}_k)$, $k \in K[\boldsymbol{x}]$. Thus, for all

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted August 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

small enough ϵ, $(\Delta\boldsymbol{\nu}_k^{\text{ex}}[F_k^{\text{ex}}(\boldsymbol{x})], \Delta\boldsymbol{\pi}_k[F_k(\boldsymbol{x})])$ will satisfy (26) for all $k \in K[\boldsymbol{x}]$. More formally, let $V_{K+1} = \left\{(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) : \boldsymbol{R}_{K+1}\boldsymbol{v} + \boldsymbol{R}_{K+1}^{\text{ex}}\boldsymbol{v}^{\text{ex}} = \boldsymbol{0}, \boldsymbol{\ell}_{K+1} \leq \ \boldsymbol{v} \ \leq \boldsymbol{u}_{K+1}\right\}$. By Assumption 1, we have that the set V_{K+1} is bounded. Define

$$
\begin{aligned}
\epsilon^* \quad = \quad \min \quad & \epsilon, \\
\text{s.t.} \quad & \max_{\boldsymbol{v}^{\text{ex}} \in V_{K+1}} \left\{\boldsymbol{e}_j^\top \left(\boldsymbol{R}_{K+1}^{\text{c}} + \textstyle\sum_{\ell \in K[\boldsymbol{x}]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{B}_\ell\right)\boldsymbol{v}^{\text{ex}}\right\} \\
& \qquad\qquad \leq \tfrac{1}{\epsilon}\boldsymbol{e}_j^\top \left(\boldsymbol{f}_c - \textstyle\sum_{\ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}}\right), \qquad \forall j \notin E(\boldsymbol{x})], \\
& \max_{\boldsymbol{v}^{\text{ex}} \in V_{K+1}} \left\{\boldsymbol{e}_j^\top \boldsymbol{C}_k R_{K+1}^{\text{c}}[F^{\text{ex}}(\boldsymbol{x}), :]\boldsymbol{v}^{\text{ex}}\right\} \leq \tfrac{1}{\epsilon}\boldsymbol{e}_j^\top (\boldsymbol{u}_k - \boldsymbol{\pi}_k), \quad \forall j \in F_k(\boldsymbol{x}), \forall k \in K[\boldsymbol{x}], \\
& \min_{\boldsymbol{v}^{\text{ex}} \in V_{K+1}} \left\{\boldsymbol{e}_j^\top \boldsymbol{C}_k R_{K+1}^{\text{c}}[F^{\text{ex}}(\boldsymbol{x}), :]\boldsymbol{v}^{\text{ex}}\right\} \geq \tfrac{1}{\epsilon}\boldsymbol{e}_j^\top (\boldsymbol{\ell}_k - \boldsymbol{\pi}_k), \quad \forall j \in F_k(\boldsymbol{x}), \forall k \in K[\boldsymbol{x}].
\end{aligned}
$$

From the definition of the sets $E(\boldsymbol{x})$ and $F_k(\boldsymbol{x})$, and Assumption 1, it follows that $\epsilon^* > 0$. The contraints corresponding to $j \notin E(\boldsymbol{x})$ are slack for all $\boldsymbol{v}^{\text{ex}} \in V_{K+1}$ for $\epsilon < \epsilon^*$. Consequently, it follows that

$$
\begin{aligned}
h_{K+1}(\boldsymbol{x}, \epsilon) = \max_{(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}})} \ & \boldsymbol{g}_{K+1}^\top \boldsymbol{v}, \\
\text{s.t.} \ & \boldsymbol{R}_{K+1}\boldsymbol{v} + \boldsymbol{R}_{K+1}^{\text{ex}}\boldsymbol{v}^{\text{ex}} = \boldsymbol{0}, \\
& \left(\boldsymbol{R}_{K+1}^{\text{c}} + \sum_{\ell \in K[\boldsymbol{x}]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{B}_\ell\right)\boldsymbol{v}^{\text{ex}} \leq \frac{1}{\epsilon}\left(\boldsymbol{f}_c - \sum_{\ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}}\right), \\
& \boldsymbol{\ell}_{K+1} \leq \ \boldsymbol{v} \ \leq \boldsymbol{u}_{K+1}.
\end{aligned}
$$

for all $\epsilon < \epsilon^*$. The result follows by taking the limit $\epsilon \searrow 0$. □

We now address the assumption in Lemma A.1 and Lemma A.4 that the optimal dual vector corresponding to the LP (8) is unique. Note that LPs with continuous parameters [27] have unique dual solutions with very high probability; therefore, it is unlikely that we will encounter problems where the optimal dual solution is non-unique. We can further increase the probability of an optimal dual solution by removing redundant constraints and variables. And, if degeneracy is still encountered, the metabolite supplies $\boldsymbol{f}_c$ and reaction bounds $\boldsymbol{\ell}_k$, $\boldsymbol{u}_k$ can be slightly perturbed to generate a new problem with a non-degenerate optimal solution, and such a perturbation is justified since reaction rates and metabolite supply are inherently noisy. Assuming uniqueness of the dual solutions is also standard in work related to differentiating through optimization problems [1, 2], which is exactly what we have done here and in Section A.3.