bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

Dynamic Flux Balance Analysis Game

G. Iyengar and M. Perry

IEOR Department, Columbia University

This version: November 2, 2021

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).

1 Introduction

The interest in modeling microbial communities continues to grow as their impact on human health [6, 18, 19, 31] and the environment [11, 12, 25, 32] becomes more evident. The relationships between species in a community can show2 both aspects of 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 steady state compositions of the microbial community, and understanding the dynamics that lead to these steady states have important implications for potentially controlling the community composition by introducing new species, or specific 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 [24]. Flux balance analysis (FBA) [27] and its many variants leverage genome-scale chemical reaction networks [33] 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. In this work, we propose a *non-cooperative* model where each species attempts to 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 the 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 can be efficiently computed by iteratively solving a sequence of quadratic programs. We establish a convergence guarantee for this method, and show that it scales to the problem sizes encountered using genome-scale metabolic models.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

(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 how to efficiently compute steady-state compositions of these dynamics 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.

(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. 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 [13, 28]. See Section 3 for more details.

The rest of this paper is organized as follows. In Section 1.1 we review related literature and provide background for our analysis. In Section 2 we introduce our model for predicting the composition of microbial communities, characterize properties of stable steady states, and show how to use these properties to efficiently compute the stable steady states. Section 3 applies the model to a small example community, and section 4 concludes with a discussion on how the model can be improved.

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 all 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 [27, 29]. 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 [8]. 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 [20, 21, 26].

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. 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 [4, 35]. The model discussed in this paper is a departure from this last point – in our model each species maximizes its own biomass without any regard for any community level objective function.

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. [7] and Goyal, Dubinkina, and Maslov [14] predict the composition of the microbiome by proposing that the equilibrium composition is a stable matching between species and

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

nutrients. Harcombe et. al. [15] 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. Chan et. al. [4] and Zomorrodi and Maranas [35] study the steady-state composition of a microbial community via a bi-level FBA optimization model where each species maximize its own growth rate while simultaneously cooperating with the other species to maximize the overall community growth rate. Chan et. al. [4] assume 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. However, the model by Zomorrodi and Maranas [35] results in a large bi-level non-convex problem. Zomorrodi and Segre [36] 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 that studying the values of fluxes for each reaction in the entire set of reactions.

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 [30]. We also discuss the stability of the associated GNEs.

1.1.3 Dynamic FBA

Typically, the quantity of interest when analyzing a microbial ecosystem is the relative proportion of each species. However, FBA, by itself, does not provide this information since FBA focuses on fluxes of the biomass reactions and not the biomasses. Dynamic FBA (DFBA) [22] 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 [34], and a number of efficient implementations have been developed [2, 17]. We consider the dynamics of general communities by assuming that the instantaneous growth rates of all species is given by the Nash equilibrium at each point in time. We show that we can directly compute steady-state community compositions, bypassing the need to simulate the solution for the Lotka-Volterra differential equation.

Our method is motivated by SteadyCom [4]. However, our model departs from SteadyCom in several important ways: we do not assume that at steady-state the species are collaborating to maximize the common growth rate, and instead assume that each species focuses just on maximizing its own steady-state biomass reaction rate. We also consider the stability of steady-states to perturbations in the biomasses as well as to invasion by other species.

1.1.4 Generalized Nash equilibria

We model the interaction between the species as a non-cooperative game where the actions for each player (i.e. species) can depend on the action of the other players. This feature is needed to model the fact that species in a community exchange metabolites, and the available actions, i.e. reaction rates, can be limited by the supply of metabolites that are secreted by or shared with other species. Nash equilibria for such games are called generalized Nash equilibria, 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. A large number of approaches to solving this type of problem have been developed, including methods based on solving the concatenated KKT conditions for all players [5], formulating the GNEP as the minimization of a Nikaido-Isoda function [16], and penalty methods that solve a sequence of simpler Nash equilibrium problems [9]. 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 [10]. This approach

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

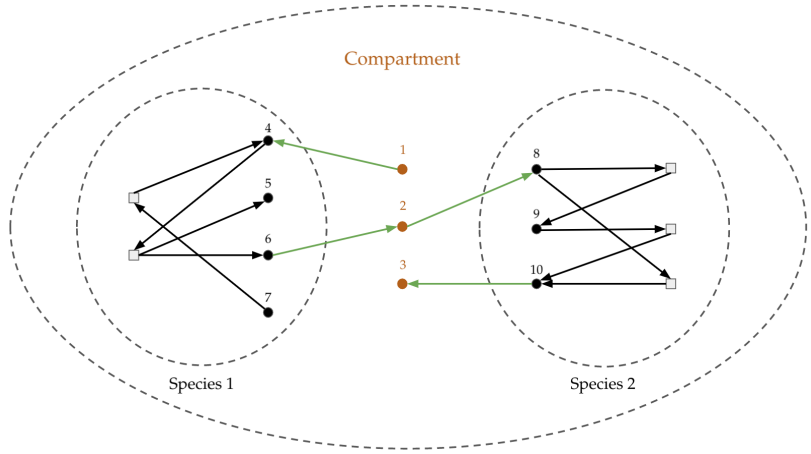


Figure 1: Schematic of microbial interactions, with grey squares representing reactions consuming and producing metabolites, represented by circles. Orange circles denote metabolites in the shared compartment, whereas black circles represent metabolites within a species. Arrows represent metabolites 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.

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 Model

We start by formulating a GNEP involving just reaction fluxes without considering biomasses. Adding in the biomasses for each species is straightforward, and does not change the details of the model or how to compute GNE.

2.1 Chemical reactions

Let $[K] = \{1, 2, ..., K\}$ denote the set of K microbial species present in a compartment or community. In Figure 1 we have $K = 2$ species in the compartment. Let $[I_k] = \{1, \ldots, I_k\}$ index the set of metabolites in species 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, $I_1 = 4$, $I_2 = 3$ and $I_c = 3$.

There are two sets of reactions associated with each species 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 $[J_k^{\text{ex}}] = \{1, \ldots, J_k^{\text{ex}}\}$ of reactions 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 concentration 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. $\boldsymbol{v}_k^{\text{ex}} \in \mathbb{R}^{J_k^{\text{ex}}}$) denote 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 metabolite concentration. The (i, j)-th element $R_k(i, j)$ of $\boldsymbol{R}_k$ denotes the change in the concentration of metabolite

[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 November 4, 2021. 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.

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 term equal to ± 1, where $R_{ex,k}(i,j) = +1$ (resp. $R_{ex,k}(i,j) = -1$) indicates that the reaction j is 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 of* $(\boldsymbol{v}_k, \boldsymbol{v}_k^{ex})$ *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 metabolite outflow from the compartment as a result of the exchange reactions associated with 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 there is an inflow $\boldsymbol{f}_c$ of metabolites into the compartment. Thus, we require that

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

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.

2.2 Game formulation

Let $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) = \{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) : k \in [K]\}$ denote the reaction rates for all the species. Then the set of feasible rate profiles for all species is given by the set

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

For $(\boldsymbol{v}, \boldsymbol{v}^{\text{ex}}) \in \chi$, 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. Let

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

denote the set of feasible reaction rates for species k given the rates of all other species $\ell \neq k$.

We call a reaction rate profile $(\boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ a generalized Nash equilibrium (GNE) if, for all $k \in [K]$,

$$(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}}) = \underset{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) \in \chi_{-k}((\boldsymbol{\nu}_\ell, \boldsymbol{\nu}_\ell^{\text{ex}})_{\ell \neq k})}{\text{argmax}} U_k((\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}), (\boldsymbol{\nu}_\ell, \boldsymbol{\nu}_\ell^{\text{ex}})_{\ell \neq k}) = \boldsymbol{g}_k^\top \boldsymbol{v}_k,$$

i.e. $(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}})$ is the best response if the reaction rate profile of all the other species is given by $(\boldsymbol{\nu}_\ell, \boldsymbol{\nu}_\ell^{\text{ex}})_{\ell \neq k}$. We propose that the reaction rates for a microbial community is given by a GNE. And, therefore, we are faced with the computational task of efficiently computing a GNE.

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

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 [1].

Proposition 12.5 in [10] 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 [10] 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

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

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.

2.3 Population dynamics and steady states

We now describe a model for computing the evolution of the biomass of each species over time. These dynamics are similar to those in [4], but with the biomass fluxes calculated using the Nash equilibrium approach described in Section 2.2 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, but that the biomasses are not necessarily in equilibrium.

Let x_k denote the biomass of species k and let $\boldsymbol{x} = [x_1, \ldots, x_K]$ denote the vector of quantities of all species. We define an $\boldsymbol{x}$-weighted game as follows. The set of "players" is given by $[K(\boldsymbol{x})] = \{k \in [K] : x_k > 0\}$, i.e. only species with positive quantity can play the game. The feasible set of rates $\{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) : k \in [K(\boldsymbol{x})]\}$ is now given by

$$\chi(\boldsymbol{x}) = \Big\{ (\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) : (\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}}) \text{ satisfies (2) and (3) for all } k \in [K(\boldsymbol{x})], \sum_{k \in [K(\boldsymbol{x})]} x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} \leq \boldsymbol{f}_c \Big\}, \tag{8}$$

i.e. the conservation equation for the compartment is now given by $\sum_{k \in [K(\boldsymbol{x})]} x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} \leq \boldsymbol{f}_c$. We are assuming that the external supply $\boldsymbol{f}_c$ remains constant and each cell of species k sets the same rates $(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})$, and therefore, the total outflow from the compartment is given by $\sum_{k \in [K(\boldsymbol{x})]} x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}^{\text{ex}}$. Thus, it follows that $(\boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ is a GNE for the $\boldsymbol{x}$-weighted game if, and only if, for all $k \in [K(\boldsymbol{x})]$

$$\begin{aligned} (\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}}) \ \in \underset{(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})}{\text{argmax}} \ & \boldsymbol{g}_k^{\top} \boldsymbol{v}_k, \\ \text{s.t. } & \boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = \boldsymbol{0}, \\ & x_k \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} + \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\nu}_\ell^{\text{ex}} \leq \boldsymbol{f}_c \\ & \boldsymbol{\ell}_k \leq \ \boldsymbol{v}_k \ \leq \boldsymbol{u}_k. \end{aligned} \tag{9}$$

Note that a species k is not a player if $x_k = 0$, and therefore, reaction rates $(\boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})$ are *not* defined for all k with $x_k = 0$.

We assume that the quantity of each species of bacteria evolves according to a Lotka-Volterra system of equations, with the growth rate of each species at each point in time being equal to its biomass reaction rate

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

at that point in time. Then the dynamics for the quantity $x_k(t)$ of bacterial species k are given by

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

where δ_k denotes the death rate for species k and $\boldsymbol{g}_k^\top \boldsymbol{\nu}_k$ denotes the growth rate for species k corresponding to a Nash equilibrium $(\boldsymbol{\nu}, \boldsymbol{\nu}_{\text{ex}})$ of the $\boldsymbol{x}(t)$-weighted Nash game. We are interested in efficiently computing steady states $\lim_{t\to\infty} \boldsymbol{x}(t)$ and analyzing their stability. Before moving on to how to do these tasks, we formalize the idea of a steady state. Suppose $\bar{\boldsymbol{x}} = \lim_{t\to\infty} \boldsymbol{x}(t)$ exists. Then the dynamics in (10) imply that either $\bar{x}_k = 0$ or $g_k^\top \nu_k(\bar{\boldsymbol{x}}) = \delta_k$ for all k.

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

Lemma 2.1. *A vector of quantities* $\boldsymbol{x}$ *is a steady state for the dynamics described by* (10) *if, and only if, there exist vectors* $(\boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ *such that* $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ *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 $\boldsymbol{g}_k^\top \boldsymbol{\nu}_k$ for some GNE equilibrium $(\boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ for an $\boldsymbol{x}$-weighted game. Thus, it follows $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ is a steady state GNE.

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

2.4 Stability of steady states

Suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ is a steady-state Nash equilibrium. We want to understand the stability of this steady state. 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, we're interested in whether or not the dynamics are able to dampen any such perturbation to keep 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 [3].

We discuss our approach informally here before providing a more formal analysis. Let $h_k(\boldsymbol{x}) = \boldsymbol{g}_k^\top \boldsymbol{\nu}_k(\boldsymbol{x}) = \delta_k$ denote the growth rate at the steady state $\boldsymbol{x}$, where we emphasize that the GNE $\boldsymbol{\nu}$ 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 that

$$h_k(\boldsymbol{x} + \boldsymbol{p}) \approx h_k(\boldsymbol{x}) + \sum_{\ell \in [K]} \frac{\partial h_k(\boldsymbol{x})}{\partial x_\ell} p_\ell,$$

and we can ignore quadratic terms in $\boldsymbol{p}$, 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 $M_{kj} = x_k \frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$. If the real part of all the eigenvalues of $\boldsymbol{M}$ are all less than 0, $\boldsymbol{x}$ is a stable steady state for (10). To define $\boldsymbol{M}$, we need to evaluate $\frac{\partial h_k(\boldsymbol{x})}{\partial x_j}$ for all k, j. We need some notation in order to explain the construction.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

(i) Let

$$U_k(\boldsymbol{x}) = \{i : \nu_{ki}(\boldsymbol{x}) = u_{ki}\} \tag{11}$$
$$L_k(\boldsymbol{x}) = \{i : \nu_{ki}(\boldsymbol{x}) = l_{ki}\} \tag{12}$$

denote the indices for which the optimal solution $\boldsymbol{\nu}_k(\boldsymbol{x})$ for the linear program (9) at biomasses $\boldsymbol{x}$ is equal to the reaction upper or lower bound. Let

$$F_k(\boldsymbol{x}) = \{i : i \notin U_k(\boldsymbol{x}) \cup L_k(\boldsymbol{x})\} \tag{13}$$

be the sets of free non-exchange reactions at the optimal solution for (9) 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

$$E(\boldsymbol{x}) = \left\{ j : x_k \boldsymbol{R}_k^{\mathrm{c}}[j] \boldsymbol{\nu}_k^{\mathrm{ex}}(\boldsymbol{x}) = f_c[j] - \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^{\mathrm{c}}[j] \boldsymbol{\nu}_\ell^{\mathrm{ex}}(\boldsymbol{x}) \right\} \tag{14}$$

denote the set of community exchange constraints that are active in (9). 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 2.2. *Suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ denotes any steady state GNE. Suppose the dual linear program of (9) 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. Then we have*

$$\frac{\partial h_k(\boldsymbol{x})}{\partial x_j} = (\boldsymbol{\lambda}_k)^\top \left(-\frac{1}{x_k} \left(\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_l \boldsymbol{R}_\ell^c \frac{\partial \boldsymbol{\nu}_l^{ex}(\boldsymbol{x})}{\partial x_j} + \delta_{jk} \boldsymbol{b}_k(\boldsymbol{x}) + (1 - \delta_{jk}) \boldsymbol{R}_j^c \boldsymbol{\nu}_j^{ex}(\boldsymbol{x}) \right) \right) \tag{15}$$

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 $k, j \in [K(\boldsymbol{x})]$.

$$\boldsymbol{R}_k^E \frac{\partial \boldsymbol{\nu}_k^{ex}}{\partial x_j} = -\frac{1}{x_k} \left(\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^E \frac{\partial \boldsymbol{\nu}_\ell^{ex}(\boldsymbol{x})}{\partial x_j} + \delta_{jk} \boldsymbol{b}_k(\boldsymbol{x}; E) + (1 - \delta_{jk}) \boldsymbol{R}_j^E \boldsymbol{\nu}_j^{ex}(\boldsymbol{x})) \right) \tag{16}$$

$$0 = \boldsymbol{R}_k^F \frac{\partial \boldsymbol{\nu}_k(F_k(\boldsymbol{x}))}{\partial x_j} + \boldsymbol{R}_k^{ex} \frac{\partial \boldsymbol{\nu}_k^{ex}(\boldsymbol{x})}{\partial x_j} \tag{17}$$

Proof. Recall the "utility" $h_k(\boldsymbol{x})$ for species k is given by the optimal value of the LP (9) at a steady state Nash equilibrium $(\boldsymbol{\nu}(\boldsymbol{x}), \boldsymbol{\nu}^{\mathrm{ex}}(\boldsymbol{x}))$ with the biomasses set to $\boldsymbol{x}$. Let

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

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 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,

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

with $\delta_{jk} = 1$ if $j = k$, $\delta_{jk} = 0$ otherwise. Thus, it follows that the optimal solution for $\boldsymbol{b}_k$ will have $\boldsymbol{\nu}_k[U_k] = \boldsymbol{u}_k[U_k(\boldsymbol{x})]$, $\boldsymbol{\nu}_k[L_k(\boldsymbol{x})] = \boldsymbol{\ell}_k[L_k(\boldsymbol{x})]$, and

$$\boldsymbol{R}_k^E \boldsymbol{\nu}_k^{\text{ex}} = \frac{1}{x_k}\left(\boldsymbol{f}_c(E_k(\boldsymbol{x})) - \sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^E \boldsymbol{\nu}_\ell^{\text{ex}}\right) \tag{18}$$

and the free variables are determined by the systems of equations

$$\boldsymbol{R}_k^U \boldsymbol{u}_k(U_k(\boldsymbol{x})) + \boldsymbol{R}_k^L \boldsymbol{\ell}_k(L_k(\boldsymbol{x})) + \boldsymbol{R}_k^F \boldsymbol{\nu}_k(F_k(\boldsymbol{x})) + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{\nu}_k^{\text{ex}} = 0 \tag{19}$$

We can differentiate (18) and (19) to solve for $\frac{\partial \boldsymbol{\nu}_k^{\text{ex}}(\boldsymbol{x})}{\partial x_j}$ and $\frac{\partial \boldsymbol{\nu}_k(F_k(\boldsymbol{x}))}{\partial x_j}$ for all $k, j \in [K(\boldsymbol{x})]$. □

Lemma 2.2 is concerned with species k such that $x_k > 0$, i.e. species that survive at steady state. We also want to understand the stability of a steady state to invasion by a new species. Let $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\text{ex}})$ 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 $\lim_{\epsilon \downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon) > \delta_{K+1}$. The next result shows $\lim_{\epsilon \downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon)$ can be computed by solving an optimization problem.

Lemma 2.3. *Suppose a steady-state GNE $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ satisfies the conditions in Lemma 2.2. Then, there exists matrices $\{\boldsymbol{B}_\ell : \ell \in K[\boldsymbol{x}]\}$ such that*

$$\begin{aligned} \lim_{\epsilon \downarrow 0} h_{K+1}(\boldsymbol{x}, \epsilon) = \max_{(\boldsymbol{v}, \boldsymbol{v}^{ex})} \; & \boldsymbol{g}_{K+1}^\top \boldsymbol{v}, \\ s.t. \; & \boldsymbol{R}_{K+1} \boldsymbol{v} + \boldsymbol{R}_{K+1}^{ex} \boldsymbol{v}^{ex} = \boldsymbol{0}, \\ & \boldsymbol{e}_j^\top \Big(\boldsymbol{R}_{K+1}^c + \sum_{\ell \in K[\boldsymbol{x}]} x_\ell \boldsymbol{R}_\ell^c \boldsymbol{B}_\ell\Big) \boldsymbol{v}^{ex} \leq 0, \; \forall j \in E(\boldsymbol{x}), \\ & \boldsymbol{\ell}_{K+1} \leq \; \boldsymbol{v} \; \leq \boldsymbol{u}_{K+1}, \end{aligned} \tag{20}$$

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 (9). Convert this LP to standard form by adding slacks $\boldsymbol{s}_k \geq 0$ 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 $\boldsymbol{v}_k^{\text{ex},\pm} \geq 0$. Let $F_k(\boldsymbol{x}) = \{i : i \notin U_k(\boldsymbol{x}) \cup L_k(\boldsymbol{x})\}$ denote components of the vector $\boldsymbol{\nu}$ 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 (9). Then, for all small enough perturbations $\Delta \boldsymbol{f}_c$, the change in the optimal solution $(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}})$ is given by the solution to the following linear equations:

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

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 $\Delta \boldsymbol{\nu}_k^{\text{ex}}[\overline{F}_k^{\text{ex}}(\boldsymbol{x})] = \boldsymbol{0}$, where $\overline{F}_k^{\text{ex}}(\boldsymbol{x})$ denotes the complement of the set $F_k^{\text{ex}}(\boldsymbol{x})$. Consequently,

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

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

$$\begin{aligned} \boldsymbol{\ell}_k[F_k(\boldsymbol{x})] \leq \boldsymbol{\nu}_k[F_k(\boldsymbol{x})] + \Delta \boldsymbol{\nu}[F_k(\boldsymbol{x})] \leq \boldsymbol{u}[F_s(\boldsymbol{x})], \\ x_k \boldsymbol{R}_k^c[\bar{E}(\boldsymbol{x}), :](\boldsymbol{\nu}_k^{\text{ex}} + \Delta \boldsymbol{\nu}_k^{\text{ex}}) + \textstyle\sum_{\ell \neq k, \ell \in [K(\boldsymbol{x})]} x_\ell \boldsymbol{R}_\ell^c[\bar{E}(\boldsymbol{x}), :] \boldsymbol{\nu}_\ell^{\text{ex}} \leq \boldsymbol{f}_c[\bar{E}(\boldsymbol{x})] + \Delta \boldsymbol{f}[\bar{E}(\boldsymbol{x})], \end{aligned} \tag{21}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

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

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

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

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

Thus, the perturbation can be written as $\Delta\boldsymbol{\nu}^{\mathrm{ex}}_k[F^{\mathrm{ex}}_k(\boldsymbol{x})] = \epsilon \boldsymbol{B}_k \boldsymbol{R}^{\mathrm{c}}_{K+1}[E(\boldsymbol{x}), :]\boldsymbol{v}^{\mathrm{ex}}_{K+1}$, and $\Delta\boldsymbol{\nu}_k[F_k(\boldsymbol{x})] = \epsilon \boldsymbol{C}_k \boldsymbol{R}^{\mathrm{c}}_{K+1}[E(\boldsymbol{x}), :]\boldsymbol{v}^{\mathrm{ex}}_{K+1}$,, 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 small enough ϵ, $(\Delta\boldsymbol{\nu}^{\mathrm{ex}}_k[F^{\mathrm{ex}}_k(\boldsymbol{x})], \Delta\boldsymbol{\nu}_k[F_k(\boldsymbol{x})])$ will satisfy (21) for all $k \in K[\boldsymbol{x}]$.

More formally, let $V_{K+1} = \big\{(\boldsymbol{v}, \boldsymbol{v}^{\mathrm{ex}}) : \boldsymbol{R}_{K+1}\boldsymbol{v} + \boldsymbol{R}^{\mathrm{ex}}_{K+1}\boldsymbol{v}^{\mathrm{ex}} = \boldsymbol{0}, \boldsymbol{\ell}_{K+1} \leq \ \boldsymbol{v} \ \leq \boldsymbol{u}_{K+1}\big\}$. By Assumption 1, we have that the set V_{K+1} is bounded. Define

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

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 $\boldsymbol{v}^{\mathrm{ex}} \in V_{K+1}$ for $\epsilon < \epsilon^*$. Consequently, it follows that

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

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

Before moving on, we address the Assumption in Lemma 2.2 and Lemma 2.3 that the optimal dual vector corresponding to the LP (9) is unique. Note that LPs with continuous parameters [23] 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.

2.5 Computing steady states

Definition 2.1 allows us to efficiently check whether a vector $\boldsymbol{x}$ is a steady-state GNE for the dynamics in (10). However, we do not yet know how to efficiently compute a steady-state GNE $\boldsymbol{x}$. In this section, we introduce a new quantity-weighted game which allows us to identify candidates for the steady-state GNE.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

The players in this new quantity-weighted game are still the species, and the actions for each player are a triple of quantities $(x_k, \boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{\text{ex}})$, the utility of the player k is $U_k(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{\text{ex}}) = x_k$ and the set of feasible actions is

$$\chi = \left\{ (\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{\text{ex}}) : \begin{array}{ll} R_k \boldsymbol{\omega}_k + R_k^{\text{ex}} \boldsymbol{\omega}_k^{\text{ex}} = 0, & k \in [K] \\ x_k \boldsymbol{\ell}_k \leq \boldsymbol{\omega}_k \leq x_k \boldsymbol{u}_k, & k \in [K] \\ \sum_{k \in [K]} \boldsymbol{R}_k^{\text{c}} \boldsymbol{\omega}_k^{\text{ex}} \leq \boldsymbol{f}_c. & \end{array} \right\} \tag{22}$$

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

$$\begin{aligned} (x_k, \boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{\text{ex}}) \in \underset{(\xi_k, \boldsymbol{v}_k, \boldsymbol{v}_k^{\text{ex}})}{\operatorname{argmax}} \; & \xi_k, \\ \text{s.t. } & \boldsymbol{R}_k \boldsymbol{v}_k + \boldsymbol{R}_k^{\text{ex}} \boldsymbol{v}_k^{\text{ex}} = \boldsymbol{0}, \\ & \boldsymbol{R}_k^{\text{c}} \boldsymbol{v}_k^{\text{ex}} + \sum_{\ell \neq k, \ell \in [K]} \boldsymbol{R}_\ell^{\text{c}} \boldsymbol{\omega}_\ell^{\text{ex}} \leq \boldsymbol{f}_c, \\ & \xi_k \delta_k \leq \boldsymbol{g}_k^T \boldsymbol{v}_k, \\ & \xi_k \boldsymbol{\ell}_k \leq \boldsymbol{v}_k \leq \xi_k \boldsymbol{u}_k. \end{aligned} \tag{23}$$

We next want to consider how well the set of Nash equilibria to (23) matches up with the set of steady state Nash equilibria. We show that a GNE for (23) is a steady state GNE for (9).

Lemma 2.4. *Suppose $(\boldsymbol{x}, \boldsymbol{\omega}, \boldsymbol{\omega}^{ex})$ is GNE for the quantity-weighted game. Define $(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{ex}) = \frac{1}{x_k}(\boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{ex})$ for $k \in [K(\boldsymbol{x})]$. Then $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ is a steady-state GNE.*

Proof. $(\boldsymbol{\nu}, \boldsymbol{\nu}_{\text{ex}})$ is clearly feasible for (9) 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 (9) such that $\boldsymbol{g}_k^\top \boldsymbol{\epsilon}_k > g_k^T \boldsymbol{\nu}_k = x_k^{-1} g_k^\top \boldsymbol{\omega}_k \geq \delta_k$.

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 (23). We have then that $(x_k, \boldsymbol{\omega}_k, \boldsymbol{\omega}_{\text{ex},k})$ is not optimal for (23). A contradiction. □

We would like the converse result to hold because this would allow us to use (23) to compute steady state GNEs. However, this is not the case. There exist steady-state GNE $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{\text{ex}})$ such that $(x_k, x_k \boldsymbol{\nu}_k, x_k \boldsymbol{\nu}_k^{\text{ex}})_{k \in [K(\boldsymbol{x}]}$ is not a GNE for (23). However, a *resource-constrained* steady state GNE is a GNE for (23).

Definition 2.2 (Resource-constrained steady-state GNE)**.** *A steady-state GNE $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ is* resource-constrained *if for all species k, the optimal value of the best response problem (9), when the biomasses x_ℓ and fluxes $\boldsymbol{\nu}_\ell^{ex}$ 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.*

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

Lemma 2.5. *Let $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ be a resource-constrained steady-state GNE. Suppose dual optimal solutions for each species' best response problem (9) at $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{ex})$ is unique. Let $U_k(\boldsymbol{x})$, $L_k(\boldsymbol{x})$, and $E(\boldsymbol{x})$ denote the active sets defined in* (11)-(14)*. Then for each species $k \in [K(\boldsymbol{x})]$, there exists an active constraint $j \in E(\boldsymbol{x})$ such that*

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

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

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

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

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

Thus, it follows that $(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\mathrm{ex}})$ is feasible for (9) 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 $\boldsymbol{\nu}_\ell^{\mathrm{ex}}$ 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, $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\mathrm{ex}})$ is not resource-constrained. □

Lemma 2.5 implies that a resource-constrained GNE is one where 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. 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 in [4]). We have the following correspondence between resource-constrained steady states and GNE for (23).

Lemma 2.6. *$(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{ex})$ is a resource-constrained steady-state GNE if and only if $(x_k, x_k\boldsymbol{\nu}_k, x_k\boldsymbol{\nu}_k^{ex})_{k\in[K]}$ is a GNE for (23).*

Proof. Suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\mathrm{ex}})$ is a resource-constrained steady-state GNE, but $(x_k, x_k\boldsymbol{\nu}_k, x_k\boldsymbol{\nu}_k^{\mathrm{ex}})_{k\in[K]}$ is not an equilibrium for (23). Then there exists a species k that can unilaterally improve its solution to (23) i.e. there exists biomass $z_k > x_k$ and fluxes $(\boldsymbol{\omega}_k, \boldsymbol{\omega}_k^{\mathrm{ex}}) \neq x_k(\boldsymbol{\nu}_k, \boldsymbol{\nu}^{\mathrm{ex}})$ that are feasible for (23). Thus, it follows that $z_k^{-1}(\boldsymbol{\omega}_k, \boldsymbol{\omega}^{\mathrm{ex}})$ is feasible (9) with an objective value at least δ_k. This contradicts that $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\mathrm{ex}})$ is resource-constrained.

Now suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{\mathrm{ex}})$ is not a resource-constrained steady state Nash equilibrium. If $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{\mathrm{ex}})$ is not a steady state GNE, it follows from Lemma 2.4 that $(x_k, x_k\boldsymbol{\nu}_k, x_k\boldsymbol{\nu}_k^{\mathrm{ex}})_{k\in[K(\boldsymbol{x})]}$ is not a Nash equilibrium for (23). Suppose instead $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\mathrm{ex}})$ 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 (9) with its biomass $z_k > x_k$ is at least δ_k, i.e. there exist feasible fluxes $(\boldsymbol{v}_k, \boldsymbol{v}_k^{\mathrm{ex}})$ for (9) such that $\boldsymbol{g}_k^\top \boldsymbol{v}_k \geq \delta_k$. Thus, it follows that $(z_k, z_k\boldsymbol{v}_k, z_k\boldsymbol{v}_k^{\mathrm{ex}})$ is feasible for (23) with an objective $z_k > z_k$. A contradiction. □

Lemma 2.7. *Suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{ex})$ is as steady state GNE. If $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{ex})$ is not resource-constrained, it is unstable to perturbation in the biomass $\boldsymbol{x}$.*

Proof. Suppose $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}_{\mathrm{ex}})$ is not resource-constrained. Then there exists a species k and $z_k > x_k$ such that the problem (9) with the biomasses x_ℓ fluxes $\boldsymbol{\nu}_\ell^{\mathrm{ex}}$ fixed and the biomass of species k set to z_k has an optimal solution $(\boldsymbol{h}_k, \boldsymbol{h}_k^{\mathrm{ex}})$ with a value at least δ_k.

Thus, we have that $(x_k, x_k\boldsymbol{\nu}_k, x_k\boldsymbol{\nu}_k^{\mathrm{ex}})$ and $(z_k, z_k\boldsymbol{h}_k, z_k\boldsymbol{h}_k^{\mathrm{ex}})$ are both feasible for (23). Since the feasible set for (23) is convex, it follows that $\epsilon(z_k, z_k\boldsymbol{h}_k, z_k\boldsymbol{h}_k^{\mathrm{ex}}) + (1-\epsilon)(x_k, x_k\boldsymbol{\nu}_k, x_k\boldsymbol{\nu}_k^{\mathrm{ex}})$ is also feasible for (23) for all $\epsilon \in [0, 1]$. Thus, it follows that the fluxes

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

are feasible for (9) 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 $(\boldsymbol{x}, \boldsymbol{\nu}, \boldsymbol{\nu}^{\mathrm{ex}})$ is unstable.

□

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

Since observing a microbial community at an unstable equilibrium is unlikely, Lemma (2.7) allows us to focus our attention on resource-constrained steady state equilibria, and therefore, justifies using (23) to compute steady states of the dynamics.

The results in this section imply the following GameCom algorithm for computing a steady state GNE.

(a) Sample an initial biomass level x_k for $k \in K$.

(b) Solve (9) to compute reaction rate vectors $\{(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}}) : k \in K\}$ that form a GNE with respect to $\boldsymbol{x}$. If (9) is infeasible for any k, or if the optimal growth rate $\boldsymbol{g}_k^\top \boldsymbol{\nu}_k \neq \delta_k$ for any k, go to the next step. Else stop. We have a candidate steady state GNE.

(c) Solve (23) to identify a stable steady state GNE.

3 Numerical Results

Recall that GameCom is motivated by SteadyCom [4], which is another method for computing steady-state community compositions using FBA. However, our computational model is different: SteadyCom assumes that all the species collectively maximize their common growth rate, whereas we assume each species maximizes its own growth rate. In order to compare and contrast the predictions of the two approaches, we also consider the same community of auxotrophic modified E. Coli. species considered in the SteadyCom paper. This community has both elements of competitio,sn and competition.

3.1 Auxotrophic E. Coli. community

The community was created by taking four copies of a metabolic reconstruction 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 [4]. 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%.

3.2 Analysis of SteadyCom steady state

The first numerical experiment was to check whether the steady state predicted by SteadyCom 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 $\delta \leq g_{sc}^* = 0.736\text{h}^{-1}$, 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 (9). We find that this re-scaled version of the output from SteadyCom is a steady-state GNE, but that it is *unstable* to perturbations in the biomasses.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

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

Next, we compute other steady state GNEs for the problem and analyze their stability. To do this, we first sample $2^{11} = 2048$ biomasses for the four species uniformly from $[0, 1]^4$ using a Sobol sequence generator, and solve (9) for each species to compute fluxes $(\boldsymbol{\nu}_k, \boldsymbol{\nu}_k^{\text{ex}})_{k\in[K]}$ that act as a starting point for computing fluxes that, paired with the sampled biomasses, form a steady state GNE. We throw away the sample if (9) is infeasible for any of the four species. After this screening step, we had 408 samples remaining. For the remaining samples, we computed the GNE for (9) with the biomasses $\boldsymbol{x}$ set equal to the sampled values. If the growth rate for each species k is equal to the death rate δ in the GNE, then the chosen biomasses $\boldsymbol{x}$ define a steady state GNE, and we stop further analysis for this sample. If the growth rate of any species is different from the death rate δ, we solve for the GNE of the quantity-weighted game with the best response dynamics given by (23) using the sampled biomasses and fluxes as a starting point. Checking for feasibility and computing the steady took approximately 3.5s, on average, for each of the 408 steady states. Analyzing all the 2048 biomass samples (i.e. computing steady state fluxes if the biomass sample is feasible, and throwing out the sample if it's not) took approximately 1.9s per sample. Finally, calculating the stability to perturbations for each of the 408 steady states took approximately 8.5s per steady state.

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, we find that 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 than that for stable steady states. The SteadyCom steady state corresponds to a steady state with the largest community biomass (indicated by the vertical dotted line in the left-hand plot).

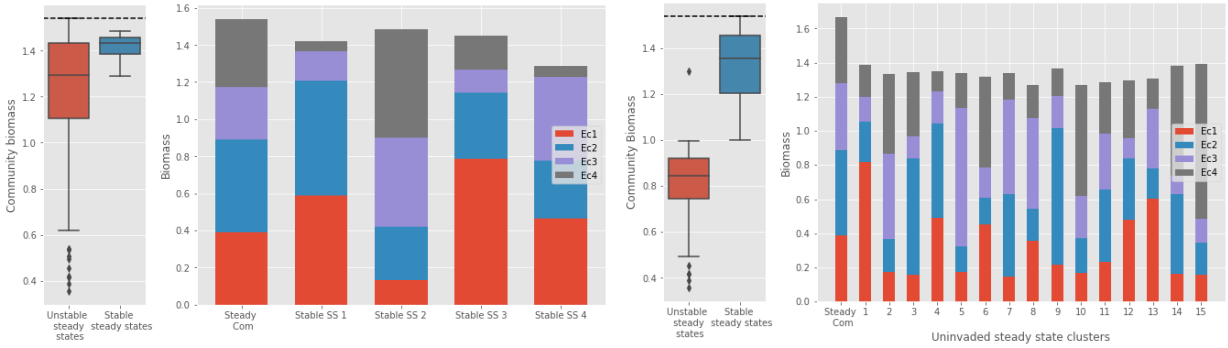
In the right panel in Figure (2) (a) we plot the (unstable) SteadyCom steady state, and the 4 stable steady states computed by GameCom. We see that the 4 steady states are very different from each other, and also from SteadyCom. Thus, we find that assuming perfect competition, as is done in SteadyCom,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

Next, we carefully examine the variation in the biomass distributions for the SteadyCom solution, SS1 and SS3. The goal here is to understand how two very distinct steady states can both be stable. 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 the SteadyCom steady state, as well as SS1 and SS3, all 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 SS1 and SS3, when compared to the SteadyCom steady state. Furthermore, we find that the very different steady states can be supported because even the 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, the key metabolites encoding this auxotrophic behavior are not severely growth limiting, and this allows for qualitatively different steady state GNEs.

3.4 Steady states robust to invasion

Next, we consider stability to invasion by a new species for each of the 408 steady states. In particular, we look at whether or not an E. Coli. 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 of already existing species, stability to invasion by this fifth E. Coli. species is much more common - 333 of the 408 steady state GNE cannot be invaded by new mutant. We summarize our 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

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

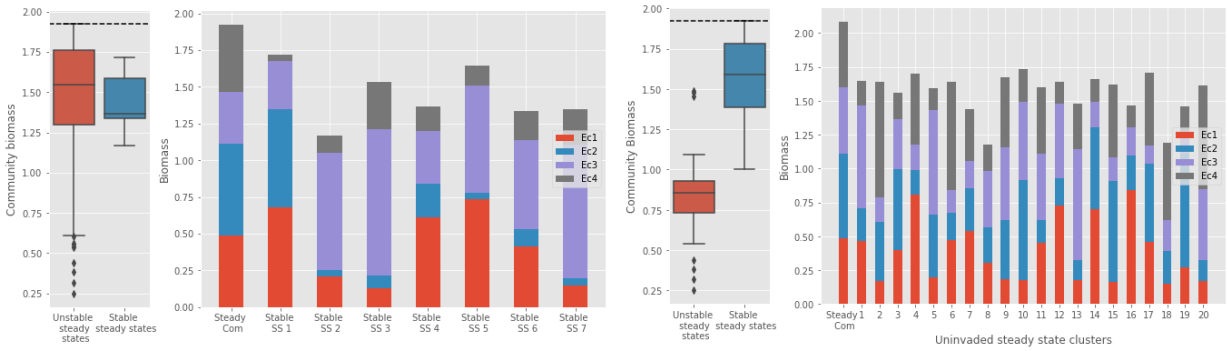


(a) (b)

Figure 2: (a) Results for the community of four E. Coli. mutants. The box plots on the left consider the sum of the biomasses for the four mutants, with equilibria split based on whether or not they're stable to perturbations in the biomasses. The bar charts on the right compares the different types of stable steady state GNEs sampled to the biomasses computed by SteadyCom. The height of the bar for each species is the biomass for that species in that steady state. (b) Similar to (a) but for stability to invasion. The bar charts on the right compare the different types of uninvaded steady state GNEs. The 333 uninvaded steady states are grouped into 15 clusters, with the height of the bar for each species being the mean biomass for that species in that cluster.

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 [4]. 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.



(a) (b)

Figure 3: Similar to figure 2, but for death rate $\delta = 0.40\ \text{h}^{-1}$.

4 Discussion

In this work, we propose a new dynamic model for predicting the composition of the microbiome. In our model the species in the community optimize their individual growth rate by competing for available

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

	Limited auxotrophy metabolites	Number of 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. $x_1\boldsymbol{R}_1^c[j]\boldsymbol{\nu}_1^{\text{ex}} + x_2\boldsymbol{R}_2^c[j]\boldsymbol{\nu}_2^{\text{ex}} + x_3\boldsymbol{R}_3^c[j]\boldsymbol{\nu}_3^{\text{ex}} + x_4\boldsymbol{R}_4^c[j]\boldsymbol{\nu}_4^{\text{ex}} = 0$, 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.

metabolites. Therefore, we avoid imposing a community-level objective function that all species attempt to optimize in a cooperative manner. We argue that our approach has potential to more flexibly predict the composition of microbial communities, since the appropriate community-level objective function is not clear, and imposing such a community wide objective may not even be appropriate. We show that the Nash equilibria corresponding to the non-cooperative game can be computed efficiently by solving a sequence of quadratic programs. Moreover, we provide a convergence guarantee for our solution algorithm. Depending on the chosen community-level objective function, computing Nash equilibria can be more efficient than solving the bi-level optimization problems encountered while computing the solution for a community level objective. For example, 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, computing the steady state becomes more complicated. In the case where the community-level objective function is more complicated than just maximizing the growth rate of the community, one has to solve a non-convex bi-level optimization problem with no convergence guarantees.

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 [13, 28].

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. In our numerical experiments, we see that stability is a very strong selection criterion – most steady states are unstable.

We are able to identify the set of stable steady states; however, we are not able to efficiently identify the particular steady state that is likely to be the convergence point for a given initial state. One would like a method more efficient than simply solving the ODE (10). Similarly, our model is able to predict whether a particular steady state is stable to invasion; however, it is not 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, whether through introducing new species or metabolites, changing the supply of metabolites, or changing the death rates, is a larger question that's 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.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

References

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

[2] Brunner, James D and Chia, Nicholas. "Minimizing the number of optimizations for efficient community dynamic flux balance analysis". In: *PLoS computational biology* 16.9 (2020), e1007786.

[3] Caputo, Michael R. "The envelope theorem and comparative statics of Nash equilibria". In: *Games and Economic Behavior* 13.2 (1996), pp. 201–224.

[4] Chan, Siu Hung Joshua, Simons, Margaret N, and Maranas, Costas D. "SteadyCom: Predicting microbial abundances while ensuring community stability". In: *PLoS computational biology* 13.5 (2017), e1005539.

[5] Dreves, Axel et al. "On the solution of the KKT conditions of generalized Nash equilibrium problems". In: *SIAM Journal on Optimization* 21.3 (2011), pp. 1082–1108.

[6] Du Toit, Andrea. "The gut microbiome and mental health". In: *Nature Reviews Microbiology* 17.4 (2019), pp. 196–196.

[7] Dubinkina, Veronika et al. "Multistability and regime shifts in microbial communities explained by competition for essential nutrients". In: *Elife* 8 (2019), e49720.

[8] Edwards, Jeremy S, Covert, Markus, and Palsson, Bernhard. "Metabolic modelling of microbes: the flux-balance approach". In: *Environmental microbiology* 4.3 (2002), pp. 133–140.

[9] Facchinei, Francisco and Kanzow, Christian. "Penalty methods for the solution of generalized Nash equilibrium problems". In: *SIAM Journal on Optimization* 20.5 (2010), pp. 2228–2253.

[10] Facchinei, Francisco and Pang, Jong-Shi. "Nash Equilibria: the variational approach". In: *Convex optimization in signal processing and communications* (2010), p. 443.

[11] Falkowski, Paul G, Fenchel, Tom, and Delong, Edward F. "The microbial engines that drive Earth's biogeochemical cycles". In: *science* 320.5879 (2008), pp. 1034–1039.

[12] Fuhrman, Jed A, Cram, Jacob A, and Needham, David M. "Marine microbial community dynamics and their ecological interpretation". In: *Nature Reviews Microbiology* 13.3 (2015), pp. 133–146.

[13] Gonze, Didier et al. "Multi-stability and the origin of microbial community types". In: *The ISME journal* 11.10 (2017), pp. 2159–2166.

[14] Goyal, Akshit, Dubinkina, Veronika, and Maslov, Sergei. "Microbial community structure predicted by the stable marriage problem". In: *bioRxiv* (2017), p. 235374.

[15] Harcombe, William R et al. "Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics". In: *Cell reports* 7.4 (2014), pp. 1104–1115.

[16] Heusinger, Anna von and Kanzow, Christian. "Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions". In: *Computational Optimization and Applications* 43.3 (2009), pp. 353–377.

[17] Höffner, Kai, Harwood, Stuart M, and Barton, Paul I. "A reliable simulator for dynamic flux balance analysis". In: *Biotechnology and bioengineering* 110.3 (2013), pp. 792–802.

[18] John, George Kunnackal and Mullin, Gerard E. "The gut microbiome and obesity". In: *Current oncology reports* 18.7 (2016), pp. 1–7.

[19] Kau, Andrew L et al. "Human nutrition, the gut microbiome and the immune system". In: *Nature* 474.7351 (2011), pp. 327–336.

[20] King, Zachary A et al. "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models". In: *Nucleic acids research* 44.D1 (2016), pp. D515–D522.

[21] Magnúsdóttir, S. et al. "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota". In: *Nature biotechnology* 35.1 (2017), p. 81.

[22] Mahadevan, Radhakrishnan, Edwards, Jeremy S, and Doyle III, Francis J. "Dynamic flux balance analysis of diauxic growth in Escherichia coli". In: *Biophysical journal* 83.3 (2002), pp. 1331–1340.

[23] May, Jerrold H and Smith, Robert L. "Random polytopes: their definition, generation and aggregate properties". In: *Mathematical programming* 24.1 (1982), pp. 39–54.

[24] Momeni, Babak, Xie, Li, and Shou, Wenying. "Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions". In: *Elife* 6 (2017), e25051.

[25] Nelson, Eric B. "The seed microbiome: origins, interactions, and impacts". In: *Plant and Soil* 422.1 (2018), pp. 7–34.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.02.466952; this version posted November 4, 2021. 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.

[26] Noronha, Alberto et al. "The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease". In: *Nucleic acids research* 47.D1 (2019), pp. D614–D624.

[27] Orth, Jeffrey D, Thiele, Ines, and Palsson, Bernhard Ø. "What is flux balance analysis?" In: *Nature biotechnology* 28.3 (2010), pp. 245–248.

[28] Pagaling, Eulyn et al. "Assembly of microbial communities in replicate nutrient-cycling model ecosystems follows divergent trajectories, leading to alternate stable states". In: *Environmental microbiology* 19.8 (2017), pp. 3374–3386.

[29] Raman, Karthik and Chandra, Nagasuma. "Flux balance analysis of biological systems: applications and challenges". In: *Briefings in bioinformatics* 10.4 (2009), pp. 435–449.

[30] Ruppin, Eytan et al. "Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks". In: *Current opinion in biotechnology* 21.4 (2010), pp. 502–510.

[31] Schwabe, Robert F and Jobin, Christian. "The microbiome and cancer". In: *Nature Reviews Cancer* 13.11 (2013), pp. 800–812.

[32] Souza, Rocheli de, Ambrosini, Adriana, and Passaglia, Luciane MP. "Plant growth-promoting bacteria as inoculants in agricultural soils". In: *Genetics and molecular biology* 38.4 (2015), pp. 401–419.

[33] Terzer, Marco et al. "Genome-scale metabolic networks". In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1.3 (2009), pp. 285–297.

[34] Zomorrodi, Ali R, Islam, Mohammad Mazharul, and Maranas, Costas D. "d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities". In: *ACS synthetic biology* 3.4 (2014), pp. 247–257.

[35] Zomorrodi, Ali R and Maranas, Costas D. "OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities". In: *PLoS Comput Biol* 8.2 (2012), e1002363.

[36] Zomorrodi, Ali R and Segrè, Daniel. "Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities". In: *Nature communications* 8.1 (2017), pp. 1–12.