2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Stackelberg evolutionary games for cancer modeling and treatment

Chiara Romano, Alessandro Borri*, Maria Domenica Di Benedetto

Abstract— **Stackelberg Evolutionary Game (SEG) theory models the interacting dynamics between a rational leader and a population of evolving followers, merging classical and evolutionary game theory. Although both methods are well-developed individually, the potential of SEG itself has not been appropriately recognized. Thus, in this paper we propose a novel eco-evolutionary model with resource mutualism and competition and we introduce a control framework based on SEGs, to steer the eco-evolutionary dynamics of followers at will. As a case study, we consider the treatment of cancer, where tumour cells are the followers that evolve in response to changes in the tumour micro-environment (and thus on available resources) and to the medical therapy, where the physician is the leader. An interesting aspect of this approach is that the objective function can be tailored according to the goal, e.g. jointly balancing tumour size, developed resistance, and toxicity of therapy, to ensure maximum quality of life for the patient. Simulations confirm the effectiveness and the potential of the proposed approach.**

Index Terms— **Game theory; Stackelberg Evolutionary Games; Mathematical modeling; Cancer treatment**

I. INTRODUCTION

While classical control theory aims at steering the dynamics of a plant towards a desired objective, in presence of two or more entities each with its own goal, the resulting interaction can be formalized and solved as a game [1]. However, traditional game theory often assumes static players with fixed strategies. This does not apply to complex systems, such as biological systems, where strategies can evolve over time through adaptation and natural selection. In this type of games, called *evolutionary games*, players are evolving individuals, while inherited behavioral phenotypes (traits) act as strategies [2]. These players, if evolutionarily identical, can be grouped into species, and this allows to define a fitness generating function, or G-function [3]. The success of a species corresponds to a high reproduction rate or fitness. In these complex games, some strategies are more resilient than others, in the sense that once adopted by most individuals, cannot be easily replaced by any other strategy, leading to the definition of an evolutionarily stable strategy. The stability comes from the strategy effectiveness in the current environment, which ensures its success and spread [4]. Challenges arise when a rational agent disrupts this environment to manipulate the eco-evolutionary dynamics of the species, for its own purpose. It is precisely in this scenario that the Stackelberg evolutionary game (SEG) theory is of fundamental help. This theory comes from the application to evolutionary game theory (EGT) of the dynamic duopoly model introduced by Stackelberg in 1934 [5], in which a leader company moves first and a subordinate or follower company moves second, after observing the leader's action. SEGs provide a powerful tool for modeling the dynamics between a strategic leader aiming at influencing the adaptive responses of the followers toward a desired outcome, and the followers themselves evolving through the principles of natural selection [6].

In this paper, we extend the recent work [7] that reviews SEG theory and shows some of its possible applications, including cancer treatment, which looks highly innovative and is therefore the subject of our current investigation. In particular, with respect to the existing literature and [7], our contribution is three-fold: i) we introduce a novel, general, eco-evolutionary model with resource mutualism and competition; ii) we restate the SEG problem in a rigorous control-theoretical formalism, both in the transient and at the steady state; iii) we introduce a realistic model of tumour growth that takes into account the heterogeneity of cells, their interactions, and the influence of external factors, such as resources and treatment, on the evolution of their traits.

A game-theoretic approach to the problem of cancer treatment may lead to interesting practical results, as it considers the fact that cells are subject to natural laws, in which evolution is their strongest advantage. Instead, in most of the existing papers, the tumour is simply considered as a process to be governed, not much different from a vehicle or chemical plant. Examples in this regard include [8], [9], providing optimal control strategies to minimize the dosage of drugs used during treatment; the issue of robustness of the control efforts with respect to parametric variations of the model is addressed in [10]; linear optimal control based on extremal variation, H_∞ control and nonlinear optimal control are investigated e.g. in [11]; in [12], Lyapunov Redesign, Synergetic and Sliding Mode controllers are designed to reduce tumour below a certain threshold. More recently, deterministic and stochastic approaches to tumour modeling and treatment have been investigated in [13], exploiting the formalism of chemical reaction networks (CRNs).

The manuscript is organized as follows. Section II in-

* Corresponding author.

This work has been partially supported by the Center of Excellence for Research DEWS, University of L'Aquila, Italy, and by the project "Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care" (D^{3}4-HEALTH), funded by the Italian National Recovery and Resilience Plan (NRRP) complementary fund.

C. Romano and M. D. Di Benedetto are with the Department of Information Engineering, Computer Science, and Mathematics, Center of Excellence for Research DEWS, University of L'Aquila, Italy. E-mail addresses: chiara.romano1@graduate.univaq.it, mariadomenica.dibenedetto@univaq.it.

A. Borri is with the CNR-IASI Biomathematics Laboratory, National Research Council of Italy, Rome, Italy. E-mail address: alessandro.borri@iasi.cnr.it.

troduces a general eco-evolutionary model with resources mutualism and competition and formalize the SEG control-theoretical framework, later applied to the problem of tumour growth control in Section III. Numerical results are discussed in Section IV. Section V offers some concluding remarks.

II. Model and problem formulation

A. A general eco-evolutionary model with resources mutualism and competition

In this subsection, we propose a novel, general eco-evolutionary model with resources mutualism and competition, defined by the following system of ordinary differential equations (we omit time dependencies):

$$\dot{x}_i = H_i(x, U_i, z, m)x_i \qquad i = 1, ..., n \tag{1}$$

$$\dot{U}_{ij} = \sigma_{ij}\frac{\partial H_i(x, U_i, z, m)}{\partial U_{ij}} \quad i = 1, ..., n,\ j = 1, ..., p \tag{2}$$

$$\dot{z} = f_z(x, z, m) \tag{3}$$

where

- $x_i \in \mathbb{R}_{\geq 0}$ is the population size of species Y_i, with $i = 1, \ldots, n$, defining a population vector $x = [x_1 \cdots x_n]^T \in \mathbb{R}^n$;
- $z \in \mathbb{R}^r$ is a vector of resource availabilities;
- $m \in M \subseteq \mathbb{R}^q$ is an external input, representing the leader's strategy, where M is the set of admissible inputs;
- $U_i \in \mathbb{R}^p$ is the vector of genetic trait levels, representing the evolutionary strategies of population i, where U_{ij} is the value of trait E_j, with $j = 1, \ldots, p$, evolving with evolutionary speed $\sigma_{ij} \geq 0$; population trait vector levels U_i are stacked in a trait level vector $U = [U_1^T \cdots U_n^T]^T \in \mathbb{R}^{np}$;
- the function $H_i : (\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^r) \times \mathbb{R}^q \to \mathbb{R}$ (growth rate) is called fitness function of population Y_i, assumed to be affected by its own trait vector U_i and by all the other variables;
- $f_z : (\mathbb{R}^n \times \mathbb{R}^r) \times \mathbb{R}^q \to \mathbb{R}^r$ is a nonlinear function modeling the resource dynamics.

From the results in [14], [26], assuming the usual locally Lipschitz property of the functions in system (1)–(3), for any initial condition $(x(0), U(0), z(0))$ and for any Lebesgue measurable and locally essentially bounded control input function m, the evolution of the model is unique, thus there exists a unique locally absolutely continuous solution $(x(t), U(t), z(t))$, in a maximal interval $[0, b)$, with $0 < b \leq +\infty$.

Eq. (1) is a general ecological model where the dynamics of population i are influenced, through its fitness, by the size of other populations, resource availability, trait evolution, and external factors. Eq. (2) models Darwinian evolution which in absence of changes in the other variables, would imply that traits evolve so that the fitness of each population is non-decreasing in time:

$$\dot{H}_i(x, U_i, z, m)|_{\substack{x=constant\\ z=constant\\ m=constant}} = \frac{\partial H_i}{\partial U_i}\frac{\partial U_i}{\partial t} = \sum_{j=1}^{p}\frac{\partial H_i}{\partial U_{ij}}\dot{U}_{ij}$$
$$= \sum_{j=1}^{p}\sigma_{ij}\left(\frac{\partial H_i}{\partial U_{ij}}\right)^2 \geq 0. \tag{4}$$

The model differs from the one in [7] since it considers different fitness functions H_i for every population, and also the additional vector equation (3) that takes into account the resource consumption with respect to the eco-evolutionary model (1)–(2).

The equilibrium solution of the system of $n+r$ differential equations (1), (3), can be written as

$$\bar{x}(U, m) \tag{5}$$

$$\bar{z}(U, m) \tag{6}$$

and are called *ecological equilibrium* and *resource equilibrium*, respectively. Instead, the equilibrium solution of the whole system of $n + p + r$ differential equations (1)–(3), if it exists, is the joint *ecological*, *evolutionary* and *resource equilibrium*

$$\bar{x}(\bar{U}^*(m), m) \tag{7}$$

$$\bar{U}^*(m) \tag{8}$$

$$\bar{z}(\bar{U}^*(m), m) \tag{9}$$

expressed as a function solely of the leader's strategy.

Notice that, as a consequence of Eqs. (2) and (4), both evaluated at the steady state, each component $\bar{U}_i^*(m)$ of the evolutionary equilibrium (8) is a local maximizer of the individual fitness H_i with respect to the trait vector U_i. The vector $\bar{U}^*(m)$, aggregating the fitness maximizers for all populations, is usually called *Evolutionarily Stable Strategy* (ESS), highlighting the fact that, for living organisms, the utility of a strategy depends on the reproductive success associated with it. The interested reader is referred to [2], [4] for a general mathematical definition and deeper biological insights on the topic. The ESS is also referred to as the *followers' best response* to the leader's action.

B. Problem Formulation as a Stackelberg evolutionary game (SEG) and equilibrium solutions

In order to define a meaningful control problem on the system (1)–(3), we can formalize a Stackelberg evolutionary game (SEG) [7], [15] between a leader, which can decide the control input m to optimize an objective function

$$Q(x, U, m) \tag{10}$$

taking into account the dynamics of the n followers, (the populations in (1)), each of them evolving by trying to optimize the fitness function H_i of its own species. The combination of the strategies of the players leads to the equilibrium of the game.

It is possible to define the game in (at least) two different situations, which are explored in the following: (i) at the steady state, when the ecological, evolutionary and resource

equilibria are achieved, by means of constant strategies; (ii) during the transient, when neither ecological nor evolutionary equilibria are reached (yet) by the followers, by means of time-varying strategies.

SEG at the steady state: we analyze the best strategy when the steady state of the system is reached. This case can be solved exactly by means of algebraic equations, and it is possible to easily compare Stackelberg and Nash equilibrium strategies. In more detail, the Nash equilibrium [16] can be obtained by optimizing the leader's objective as a function of the followers' traits U:

$$\bar{m}^*(U) = \arg\max_{m \in M} Q(\bar{x}(U,m), U, m), \tag{11}$$

so that the Nash equilibrium $(\bar{U}_N^*, \bar{m}_N^*) = (\bar{U}^*(\bar{m}_N^*), \bar{m}^*(\bar{U}_N^*))$ is the solution of the algebraic system (8), (11).

In the Stackelberg game, the leader (acting first) is aware that the followers know and react (to their best) to its own action; so the leader will optimize its objective by taking into account the (known) followers' best response (8). This leads to changing the leader's optimization (11) into:

$$\bar{m}_S^* = \arg\max_{m \in M} Q(\bar{x}(\bar{U}^*(m), m), \bar{U}^*(m), m). \tag{12}$$

The Stackelberg equilibrium strategy $(\bar{U}_S^*, \bar{m}_S^*) = (\bar{U}^*(\bar{m}_S^*), \bar{m}_S^*)$, when applied to a leader-follower game (as the system evolving according to Eqs. (1)–(3)), leads to better results for the leader than those obtained with the Nash equilibrium strategy, which is more conservative since it ignores the followers' best response:

$$Q(x^*(\bar{U}_N^*, \bar{m}_N^*)), \bar{U}_N^*, \bar{m}_N^*) \leq Q(x^*(\bar{U}_S^*, \bar{m}_S^*), \bar{U}_S^*, \bar{m}_S^*). \tag{13}$$

SEG in the transient: let $T \in \mathbb{R}^+_{>0}$ be a finite time horizon of interest. Let $\mathcal{M}$ be the set of all admissible (for example, piecewise-continuous) functions in the form $m : [0,T] \longrightarrow M$. The leader can find the best strategy as a time-varying function in the interval $[0,T]$, starting from the initial condition $(x(0), U(0), z(0))$, by maximizing a weighted average of the quality of life (10) along the whole time horizon:

$$m_S^*(\cdot) = \arg \max_{\substack{m(\cdot) \in \mathcal{M} \\ \text{subject to (1),(2),(3)}}} \tilde{Q}(x(\cdot), u(\cdot), m(\cdot)), \tag{14}$$

where the functional $\tilde{Q}$ is defined as

$$\begin{aligned} \tilde{Q}(x(\cdot), u(\cdot), m(\cdot)) = w_T \; & Q(x(T), U(T), m(T)) \\ & + (1 - w_T) \; \frac{1}{T} \int_0^T Q(x(\tau), U(\tau), m(\tau)) d\tau \end{aligned} \tag{15}$$

and $w_T \in [0,1]$ is a design parameter, chosen to relatively weigh terminal and running costs as desired.

Notice that, with respect to what defined for the steady-state equilibrium, in the transient we more generally aim at optimizing a convex combination of the quality of life at the end of the interval $[0,T]$ (weighted by the coefficient w_T) and of the average quality of life in the same interval (weighted by $1 - w_T$). This results in a more complex dynamic game that is hard to solve exactly, but that can guarantee, even in an approximate solution (later shown in the case study), better quality of life than the one achieved by the exact steady-state optimization over the set of constant strategies, described in the previous paragraph.

The Stackelberg equilibrium strategy can be expressed, at each time $t \geq 0$, as the pair $(U^*(t), m_S^*(U^*(t)))$, where the followers' transient response $U^*(t)$ is the evolutionary part of the solution of the system (1)–(3) with the particular choice $m(t) = m_S^*(U^*(t))$, defined in (14), representing a feedback law operated by the leader from the traits evolution.

In the transient case, the computation of the Nash equilibrium still leads to a worse (more conservative) outcome for the leader with respect to Stackelberg strategy (14), and it is much harder to compute, since it would require the solution of the Hamilton-Jacobi-Bellman-Isaacs partial integro-differential equations [17]. This is left out of the present work for lack of space.

III. Tumour treatment as a Stackelberg game

A. A novel tumour model with mutualism, competition and Niche construction

In this section, we propose a model in the form (1)–(3) addressed in Section II for cancer treatment. The suggested model, compared with others in the literature, provides a unified view on the evolution and growth of cancer cells, addressing four main issues: heterogeneity [18], development of resistance to treatment [7], cellular cooperation [19], [20], and the *Niche* construction, that is the biunivocal relationship between cancer cells and the tumour environment [3]. The latter is crucial because the environment, particularly the resources within it, plays a key role in cellular evolution [21]. In detail, we identified oxygen as a key resource influencing cells. Indeed, low oxygen levels, a condition known as *hypoxia*, is often associated with tumour growth [22]. This led us to introduce a new trait: the ability to survive hypoxia, in addition to resistance to treatment. A deeper analysis of the evolutionary dynamics also led us to discover a relationship between the two traits, which has been integrated into an improved version of this eco-evolutionary model in [23]. Since hypoxia selects glycolysis as the predominant energy pathway and also promotes angiogenesis [21], we considered three types of cellular subpopulations: glycolytic cells (denoted as GLY), vascular overproducers (VOP) and defector cells (DEF). The latter species benefits from the *goods* (acidic environment and oxygen) provided by the former ones (GLY and VOP) [19]. Thus, the proposed model takes into account both aspects of mutualism (between the first two species) and competition (intra-species and between DEF and the others). We also distinguished between defector cells capable of developing resistance to treatment and sensitive ones, resulting in a total of 4 populations.

Therefore, our model in the form (1)–(3) considers $n = 4$ cell species, a single ($r = 1$) resource, a single therapy ($q = 1$), and $p = 2$ traits. The four species $Y_1 = G$, $Y_2 = V$, $Y_3 = D_r$ and $Y_4 = D_s$ stand for GLY, VOP, resistant DEF and sensitive DEF cells, respectively, the two traits $E_1 = H_r$ and $E_2 = T_r$, with levels U_1 and U_2, represent

the resistance to hypoxia and the resistance to treatment, respectively, and the resource $Z = O_x$ is the percentage of oxygen present in the environment. Consequently, the state variables are adimensional [#]. The therapy administration (in the unit of time) is assumed to be normalized over its maximum value, so $m \in M = [0, 1]$, where $m = 0$ and $m = 1$ correspond to no therapy and Maximum Tolerated Dose (MTD), respectively.

Denoting by $U_i \in \mathbb{R}^2_{\geq 0}$ the vector of trait levels of population i, with $i = 1, \ldots, 4$, we model the general fitness function for species $i = 1, \ldots, n$ in (1) as

$$H_i(x, U, z, m) = \bar{r}_i(z + c_i^T U_i)e^{-g_i^T U_i}\left(1 + \frac{\alpha_i^T x}{K_{max}}\right) - d - \frac{m}{k + b_i^T U_i}, \tag{16}$$

where

- $\bar{r}_i > 0$ (in units of [1/day]) is the nominal growth rate of population i;
- $c_i \in \mathbb{R}^2_{\geq 0}$ is a constant vector whose general entry c_i [#] is defined as

$$c_i = \begin{cases} [1 \quad 0]^T & i = 1, 2, \\ [0 \quad 0]^T & i = 3, 4, \end{cases}$$

  allowing the first two populations to have an increased growth rate for increasing values of the first trait;
- $g_i = [g_{i,H} \quad g_{i,T}]^T$, where $g_{i,H}$, $g_{i,T} \geq 0$ [#] represent population-dependent costs of resistance to low oxygen levels and treatment, respectively;
- $\alpha_i \in \mathbb{R}^4_{\geq 0}$ is a vector whose general entry α_{ij} [#] defines the cooperative ($\alpha_{ij} > 0$), competitive ($\alpha_{ij} < 0$) or null ($\alpha_{ij} = 0$) effect of a population j on population i, assuming $\alpha_{ii} = -1 \ \forall i$. Each $\alpha_{ij} = b_{ij} - c_{ij}$ is the difference between the benefit b_{ij} that species i gets from interacting with species j, and the associated cost c_{ij}. These interactions can be categorized as

$$\textit{Mutualism} \quad \text{if} \quad (\alpha_{ij} > 0 \wedge \alpha_{ji} > 0)$$

$$\textit{Competition} \quad \text{if} \quad (\alpha_{ij} < 0 \wedge \alpha_{ji} < 0)$$

$$\textit{Exploitation of species } j \quad \text{if} \quad (\alpha_{ij} > 0 \wedge \alpha_{ji} < 0)$$

  and are inferred from a 4-species extension of the 2-species Lotka-Volterra model proposed in [24];
- $K_{max} > 0$ [#] is the carrying capacity, namely the maximum population size reachable in absence of interactions ($\alpha_{ij} = 0$ for $i \neq j$);
- $d > 0$ [1/day] is the cell natural death rate;
- m [#] is the therapy administration rate, i.e. the external input;
- $k > 0$ [#] represents the cell innate resistance in absence of drug exposure;
- $b_i = [b_{i,H} \quad b_{i,T}]^T$, where $b_{i,H}$, $b_{i,T} \geq 0$ [#] represent the population-dependent advantages that cells gain from their ability to resist hypoxia and treatment, respectively.

Considering that H_i, for any i, is a locally Lipschitz function, by virtue of Rademacher's theorem [27], it is differentiable almost everywhere in $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^r$. This concept of differentiability asserts the existence of the coordinate-wise partial derivatives, thus (16) uniquely determines the following particular form of the dynamics (2), for all i and j:

$$\dot{U}_{ij} = \sigma_{ij}\bar{r}_i e^{-g_i^T U_i}\left(c_{ij} - g_{ij}(z + c_i^T U_i)\right)\left(1 + \frac{\alpha_i^T x}{K_{max}}\right) + \sigma_{ij}\frac{m b_{ij}}{(k + b_i^T U_i)^2}. \tag{17}$$

The resource equation in (3) is modeled as:

$$\dot{z} = f_z(x, z, m) = \left(P_{O_x}^T x\right)(1 - z) - \left(R_{O_x}^T x\right) z, \tag{18}$$

representing a logistic growth with linear clearance, where P_{O_x}, $R_{O_x} \in \mathbb{R}^4_{\geq 0}$ include the population-dependent *per capita* oxygen production and removal rates, respectively. In particular, we choose $P_{O_x} = [0 \quad P_V \quad 0 \quad 0]^T$, since only VOP are able to stimulate angiogenesis and thus increase oxygen availability, and $R_{O_x} = [R_G \quad R_A \quad R_A \quad R_A]^T$ where R_G and R_A represent the removal rates of glycolytic and aerobic cells (VOP and DEF), respectively, considering for the latter a higher rate ($R_A > R_G$), since these cells cannot rely on glycolysis.

B. Therapy optimization

For the optimization problem of the leader, namely the physician, it is necessary to define an objective function, as in (10). Our choice here is to generalize the objective of [7] by balancing tumour size, developed resistance, and toxicity of therapy to ensure maximum quality of life for the patient. We therefore set

$$Q(x, U, m) = \bar{Q} - c_x\left(\frac{\mathbf{1}^T x}{n K_{max}}\right)^2 - c_u\|U_{T_r}\|^2 - c_m m^2, \tag{19}$$

where

- $\bar{Q}$ is the maximum patient quality of life;
- $\mathbf{1}$ denotes the n-dimensional column vector of 1's (with $n = 4$ being the number of species), so that the term $\mathbf{1}^T x$ is the total tumour population; differently from [7], we divide $\mathbf{1}^T x$ by nK_{max} because mutualism among species results in increasing the maximum population by a factor of n with respect to the carrying capacity;
- the vector $U_{T_r} = [U_{12} \quad U_{22} \quad U_{32} \quad U_{42}]^T$ collects the treatment resistance trait levels for the 4 populations;
- the weights c_x, c_u, c_m represent the impact of tumour burden, resistance rate, and drug toxicity, respectively, on the quality of life.

As formalized in Section II, we can solve the Stackelberg game both at the steady state (i) and in the transient dynamics (ii), by obtaining a constant and a time-varying therapeutic strategy, respectively.

1) Constant tumour treatment (SEG at the steady-state): first, we computed some partial analytical expressions for the non-trivial (non-zero) ecological (7), evolutionary (8) and resource (9) equilibria as a function of the administered therapy. By imposing the fitness function H_i in (16) equal to zero for all i, and grouping the effects of cellular interactions, we obtain the algebraic linear system $A\bar{x} = \bar{b}$, where $A \in \mathbb{R}^{n\times n}$ is the matrix that captures all interaction effects. The vector $\bar{b} = [\bar{b}_1 \quad \bar{b}_2 \quad \bar{b}_3 \quad \bar{b}_4]^T$ has its general entries defined as $(i = 1, 2, 3, 4)$:

$$\bar{b}_i = \frac{K_{max} e^{g_i^T U_i}}{\bar{r}_i (z + c_i^T U_i)} \left(d + \frac{m}{k + b_i^T U_i} - \bar{r}_i (z + c_i^T U_i) e^{-g_i^T U_i} \right). \tag{20}$$

Considering that A is invertible for the parameters in Table I, an explicit form for the ecological equilibrium $\bar{x}$ can be readily obtained as a function of the other variables:

$$\bar{x} = A^{-1}\bar{b}. \tag{21}$$

For the resource equilibrium, setting Eq. (18) to zero, we obtain:

$$\bar{z}(x) = \frac{P_{O_x}^T x}{P_{O_x}^T x - R_{O_x}^T x}. \tag{22}$$

Unfortunately, the remaining Eq. (17) at the steady state cannot be easily solved analytically with respect to the trait level U_{ij}, to find the evolutionary stable strategy (ESS) $\bar{U}^*(m)$ in (8) in closed form, but the ESS can be readily obtained numerically, jointly with the analytical solutions (21), (22), so that a numerical evaluation of the equilibria (7)–(9), as a function of the therapy m, is finally available.

2) Time-varying tumour treatment (SEG during transient): since the optimal solution (14) in the transient would require the solution of an optimal control problem over a system of dimension $n + np + r = 13$, we propose here a sub-optimal piecewise-constant solution, computed by solving an approximate version of (14), in the spirit of Model Predictive Control (MPC) [25]. Let the interval $[0, T]$ be partitioned into N equal subintervals of length Δ, we aim at finding a suboptimal Stackelberg strategy $\tilde{m}_S^*(\cdot)$ defined as

$$\tilde{m}_S^*(t) = \tilde{m}_{S,k}^* \qquad t \in [k\Delta, (k+1)\Delta), \quad k = 0, 1, \ldots, N-1, \tag{23}$$

where the therapy value at each interval solves the following approximate version of the control problem in (14):

$$\tilde{m}_{S,k}^* = \arg \max_{\substack{m_k \in M \\ \text{subject to (1),(2),(3)}}} \tilde{Q}_k(x(\cdot), u(\cdot), m_k), \tag{24}$$

with

$$\tilde{Q}_k(x(\cdot), u(\cdot), m_k) = w_T \; Q(x(T), U(T), m(T)) + (1 - w_T) \; \frac{1}{(N-k)\Delta} \int_{k\Delta}^{T} Q(x(\tau), U(\tau), m_k) d\tau, \tag{25}$$

where $T = N\Delta$, starting from the initial condition $(x(k\Delta), U(k\Delta), z(k\Delta))$ of interval k. Notice that the value of $\tilde{m}_{S,0}^*$ can be close to the Stackelberg steady-state strategy $\bar{m}_S^*$ if T is large and $\omega_T = 1$, but the time-varying strategy is free to vary in the following intervals $k = 1, \ldots, N-1$, to possibly obtain a higher quality of life with respect to the one obtained by the constant Stackelberg steady-state strategy.

IV. Simulations

Several simulations were carried out in MATLAB® on the proposed tumour growth model (1)–(3), (16)–(18), for the choice of parameter values in Table I, which have been recalibrated with respect to the values in [7] in order to obtain more meaningful results in terms of tumour dimension and response/treatment time.

TABLE I
NUMERICAL VALUES OF THE MODEL PARAMETERS

Parameter	Value		Parameter	Value	
K_{max}	10^{11}		$\bar{c}_{ii,i=1,\ldots,4}$	1	
$\bar{r}_{i,i=1,\ldots,4}$	0.45	$[1/day]$	$\bar{c}_{1i,i\neq1}$	1/6	
$\bar{g}_{iH,i=1,\ldots,4}$	1		$\bar{c}_{2i,i\neq2}$	1/6	
$\bar{g}_{iT,i=1,\ldots,4}$	1		$\bar{c}_{3i,i\neq3}$	0	
σ_{ij}	0.1	$[1/day]$	$\bar{c}_{4i,i\neq4}$	0	
$\bar{b}_{i1,i\neq1}$	0.5		$b_{iH,i,i=1,\ldots,4}$	0	
$\bar{b}_{i2,i\neq2}$	0.25		$b_{iT,i,i=1,\ldots,4}$	10	
$\bar{b}_{i3,i\neq3}$	0		k	1	
$\bar{b}_{i4,i\neq4}$	0		d	0.01	$[1/day]$
$\bar{b}_{ii,i=1,\ldots,4}$	0		c_x	0.6	
P_V	0.1		c_u	0.32	
R_G	0.01		c_m	0.08	
R_A	0.03		$\bar{Q}$	1	

First, we numerically verified the uniqueness and the stability of the equilibrium (7)–(9), for non-zero initial populations, as a function of the therapy level m. In order to evaluate the effectiveness of the proposed strategy, both in steady state and during the transient, we considered as initial state of the simulation the equilibrium reached in absence of treatment: $x(0) = \bar{x}(\bar{U}^*(0), 0)$, $U(0) = \bar{U}^*(0)$, and $z(0) = \bar{z}(\bar{U}^*(0), 0)$, corresponding to an initial tumour size in the order of 10^{11} cells.

Regarding the behavior of the SEG at the steady state by means of constant treatments, in Fig. 1 we plot $\bar{Q}(m) := Q(\bar{x}(\bar{U}^*(m), m), \bar{U}^*(m), m)$, which confirms (as expected from Eq. (13)) that the Stackelberg equilibrium strategy (corresponding to the therapy $m = \bar{m}_S^* \simeq 0.5208$) guarantees the maximum steady-state quality of life, compared with other constant strategies, including the Nash equilibrium strategy ($m = \bar{m}_N^* \simeq 0.5498$) and the extreme cases of no therapy ($m = 0$) and MTD ($m = 1$).

The situation changes when considering the SEG during transient and time-varying therapies, which, in general, lead to an overall better quality of life, over the entire observation period of 700 days ($\simeq$ 2 years), with respect to constant therapies. In Fig. 2 we plot $\tilde{Q}(x(t), u(t), m(t))$ in (15) for different choices of the input function $m(\cdot)$, providing a comprehensive comparison of some investigated possibilities, including (i) the Stackelberg MPC-like solution (23)–(24) with $\omega_T = 0.5$ (equally weighing average and final quality of life), achieving the best outcome, (ii) a static proportional state-feedback controller $m(t) = \frac{\mathbf{1}^T}{\mathbf{1}^T x(0)} x(t)$, reducing the therapy proportionally to the tumor shrinking in time, and (iii) the extreme cases of constant strategies.

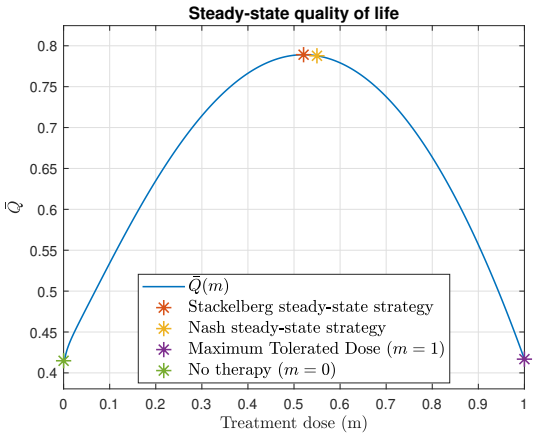

Fig. 1. Quality of life achieved at the steady state by means of constant strategies: comparison between Stackelberg equilibrium, Nash equilibrium, MTD ($m = 1$) and null therapy ($m = 0$).

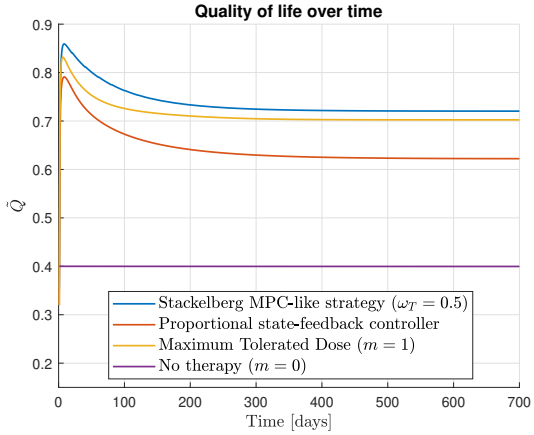

Fig. 2. Quality of life over a time span of 700 days: comparison among the Stackelberg MPC-like strategy (with $\omega_T = 0.5$), a proportional static state-feedback controller, and the cases of MTD and null therapy.

V. Conclusions

In this paper, we formalized a control-theoretical setting for Stackelberg evolutionary games (SEGs) together with a novel general eco-evolutionary modeling framework accounting for resource, mutualism and competition. The approach was applied to the case study of cancer treatment, which presents the interest of taking into account how cancer cells evolve in response to treatment, considering the now well-known fact that cancer is a Darwinian process. Compared with other models of tumour growth already in the literature, our model provides a more realistic view of tumour evolution, addressing four main issues: heterogeneity, development of treatment resistance, cellular cooperation, and the *Niche* construction. The potential of the approach is validated by means of numerical simulations. Future work will address a deeper comparison between the SEG approach and the optimal control methods available in the context of cancer treatment, further model refinements and an extensive *in silico* simulation campaign, as a first step in view of a future translation of the approach to the clinical practice.

References

[1] R. B. Myerson, Game Theory: Analysis of Conflict, *Harvard University Press*, 1991.

[2] J. Maynard Smith, Evolution and the Theory of Games, *Cambridge University Press*, 1982.

[3] T. L. Vincent, and J. S. Brown, Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics, *Cambridge Univ. Press*, 2005.

[4] J. Maynard Smith, and G. R. Price, The Logic of Animal Conflict, *Nature*, vol. 246, pp. 15-18, 1973.

[5] Heinrich von Stackelberg, Marktform und gleichgewicht, 1934.

[6] M. Kleshnina, S. Streipert, J. Brown, and K. Staňková, Game theory for managing evolving systems: challenges and opportunities of including vector-valued strategies and life-history traits, *Dyn Games Appl*, vol. 13, no. 4, pp. 1035-1048, 2023.

[7] A. Stein, M. Salvioli, H. Garjani, J. Dubbeldam, Y. Viossat, J.S. Brown, and K. Staňková, Stackelberg evolutionary game theory: how to manage evolving systems, in *Phil. Trans. R. Soc. B*, vol. 378, no. 1876, 2023.

[8] L. G. De Pillis, and A Radunskaya, The dynamics of an optimally controlled tumour model: A case study, *Mathematical and Computer Modelling*, vol. 37, pp. 1221-1244, 2003.

[9] L. G. de Pillis, W. Gu, K. R. Fister, T. Head, K. Maples, A. Murugan, T. Neal, and K. Yoshida, Chemotherapy for tumours: An analysis of the dynamics and a study of quadratic and linear optimal controls, *Mathematical Biosciences*, vol. 209, no. 1, pp. 292-315, 2007.

[10] E. A. Sarapata, and L. G. de Pillis, A Comparison and Catalog of Intrinsic tumour Growth Models, *Bull Math Biol*, vol. 76, no. 8, pp. 2010-2024, 2014.

[11] H. Moradi, M. Sharifi, and G. Vossoughi, Adaptive robust control of cancer chemotherapy in the presence of parametric uncertainties: a comparison between three hypotheses, *Comput Biol Med*, vol. 56, pp. 145-157, 2015.

[12] I. Lodhi, I. Ahmad, M. Uneeb, and M. Liaquat, Nonlinear Control for Growth of Cancerous tumour Cells, *IEEE Access*, vol. 7, pp. 177628-177636, 2019.

[13] F. Papa, A. Borri, and P. Palumbo, Tumour growth control: analysis of alternative approaches, *Journal of Theoretical Biology*, 2023.

[14] H. Khalil, Nonlinear Systems, 3rd Edition, *Prentice Hall, Upper Saddle River* , 2002.

[15] T. Başar, and G. J. Olsder, Dynamic noncooperative game theory, *Society for Industrial and Applied Mathematics*, 1998.

[16] J. F. Nash, Equilibrium points in n-person games, *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48-49, 1950.

[17] J. P. Hespanha, Noncooperative game theory: An introduction for engineers and computer scientists, *Princeton University Press*, 2017.

[18] M. Gluzman, J. G. Scott, and A. Vladimirsky, Optimizing adaptive cancer therapy: dynamic programming and evolutionary game theory, *Proceedings of the Royal Society B: Biological Sciences*, 2020.

[19] M. Archetti, and K. J. Pienta, Cooperation among cancer cells: applying game theory to cancer, *Nat Rev Cancer*, vol. 19, 2019.

[20] J.P. Capp, F. Thomas, A. Marusyk , M. A. Dujon, S. Tissot, R. Gatenby, J.S. Brown et al., The paradox of cooperation among selfish cancer cells, *Evol Appl*, vol. 16, no. 7, pp. 1239-1256, 2023.

[21] D. C. Belisario, J. Kopecka, M. Pasino, M. Akman, E. De Smaele, M. Donadelli, and C. Riganti, Hypoxia Dictates Metabolic Rewiring of tumours: Implications for Chemoresistance, *Cells*, vol. 9, 2020.

[22] J. Zhou, T. Schmid, S. Schnitzer, and B. Brüne, Tumour hypoxia and cancer progression, *Cancer Letters*, vol. 237, no. 1, pp. 10-21, 2006.

[23] C. Romano, A. Borri and M. D. Di Benedetto, Adaptive responses of cancer cells to tumor microenvironment changes: an evolutionary model perspective, *Proceedings of the 20th IEEE Conference on Automation Science and Engineering (CASE 2024)*, Bari, Italy, Aug. 28–Sept. 1, 2024 (*IEEE CASE 2024 Best Student Paper Award*).

[24] S. Mao, M. Zhu, X. Wang, and X. Xiao, Grey–Lotka–Volterra model for the competition and cooperation between third-party online payment systems and online banking in China, *Applied Soft Computing*, vol. 95, 2020.

[25] E. F. Camacho, C. Bordons, Model predictive control, *Springer*, London, 2007.

[26] I. Karafyllis, Z.P. Jiang, Stability and Stabilization of Nonlinear Systems, *Springer*, London, 2011.

[27] H. Rademacher, *Ü*ber partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und *ü*ber die Transformation der Doppelintegrale, *Math. Ann.*, vol. 79, pp. 340-359, 1919.