Evolutionarily Stable Stackelberg Equilibrium

Sam Ganzfried
Ganzfried Research
sam.ganzfried@gmail.com

Abstract— **We present a new solution concept called evolutionarily stable Stackelberg equilibrium (SESS). We study the Stackelberg evolutionary game setting in which there is a single leading player and a symmetric population of followers. The leader selects an optimal mixed strategy, anticipating that the follower population plays an evolutionarily stable strategy (ESS) in the induced subgame and may satisfy additional ecological conditions. We consider both leader-optimal and follower-optimal selection among ESSs, which arise as special cases of our framework. Prior approaches to Stackelberg evolutionary games either define the follower response via evolutionary dynamics or assume rational best-response behavior, without explicitly enforcing stability against invasion by mutations. We present algorithms for computing SESS in discrete and continuous games, and validate the latter empirically. Our model applies naturally to biological settings; for example, in cancer treatment the leader represents the physician and the followers correspond to competing cancer cell phenotypes.**

I. Introduction

While Nash equilibrium has emerged as the standard solution concept in game theory, it is often criticized as being too weak: often games contain multiple Nash equilibria (sometimes even infinitely many), and we want to select one that satisfies other natural properties. For example, one popular concept that refines Nash equilibrium is evolutionarily stable strategy (ESS). A mixed strategy in a two-player symmetric game is an evolutionarily stable strategy if it is robust to being overtaken by a mutation strategy. Formally, $\mathbf{x}^*$ is an ESS if for every mixed strategy $\mathbf{x}$ that differs from $\mathbf{x}^*$, there exists $\epsilon_0 = \epsilon_0(\mathbf{x}) > 0$ such that, for all $\epsilon \in (0, \epsilon_0)$,

$$(1-\epsilon)u_1(\mathbf{x}, \mathbf{x}^*)+\epsilon u_1(\mathbf{x}, \mathbf{x}) < (1-\epsilon)u_1(\mathbf{x}^*, \mathbf{x}^*)+\epsilon u_1(\mathbf{x}^*, \mathbf{x}).$$

From a biological perspective, we can interpret $\mathbf{x}^*$ as a distribution among "normal" individuals within a population, and consider a mutation that makes use of strategy $\mathbf{x}$, assuming that the proportion of the mutation in the population is ϵ. In an ESS, the expected payoff of the mutation is smaller than the expected payoff of a normal individual, and hence the proportion of mutations will decrease and eventually disappear over time, with the composition of the population returning to being mostly $\mathbf{x}^*$. An ESS is therefore a mixed strategy that is immune to being overtaken by mutations. ESS was initially proposed by mathematical biologists motivated by applications such as population dynamics (e.g., maintaining robustness to mutations within a population of humans or animals) [21], [20]. A common example game is the 2x2 game where strategies correspond to an "aggressive" Hawk or a "peaceful" Dove strategy. A paper has recently proposed a similar game in which an aggressive malignant cell competes with a passive normal cell for biological energy, which has applications to cancer eradication [6].

While Nash equilibrium is defined for general multiplayer games, ESS is traditionally defined specifically for two-player symmetric games. ESS is a refinement of Nash equilibrium. In particular, if $\mathbf{x}^*$ is an ESS, then $(\mathbf{x}^*, \mathbf{x}^*)$ (i.e., the strategy profile where both players play $\mathbf{x}^*$) is a (symmetric) Nash equilibrium [19]. Of course the converse is not necessarily true (not every symmetric Nash equilibrium is an ESS), or else ESS would be a trivial refinement. In fact, ESS is not guaranteed to exist in games with more than two pure strategies per player (while Nash equilibrium is guaranteed to exist in all finite games). For example, while rock-paper-scissors has a mixed strategy Nash equilibrium (which puts equal weight on all three actions), it has no ESS [19] (that work considers a version where payoffs are 1 for a victory, 0 for loss, and $\frac{2}{3}$ for a tie).

There exists a polynomial-time algorithm for computing Nash equilibrium (NE) in two-player zero-sum games, while for two-player non-zero-sum and multiplayer games computing an NE is PPAD-complete and it is widely conjectured that no efficient (polynomial-time) algorithm exists. However, several algorithms have been devised that perform well in practice. The problem of computing whether a game has an ESS was shown to be both NP-hard and CO-NP hard and also to be contained in Σ_2^P (the class of decision problems that can be solved in nondeterministic polynomial time given access to an NP oracle) [7]. Subsequently it was shown that the exact complexity of this problem is that it is Σ_2^P-complete [5], and even more challenging for more than two players [2]. Note that this result is for determining whether an ESS exists (as discussed above there exist games which have no ESS), not for the complexity of computing an ESS in games for which one exists. Thus, computing an ESS is significantly more difficult than computing an NE. Several approaches have been proposed for computing ESS in two-player games [14], [1], [4], [3], [23].

A *normal-form game* consists of a finite set of players $N = \{1, \ldots, n\}$, a finite set of pure strategies S_i for each player i, and a real-valued utility for each player for each strategy vector (aka *strategy profile*), $u_i : \times_i S_i \to \mathbb{R}$. In a *symmetric normal-form game*, all strategy spaces S_i are equal and the utility functions satisfy the following symmetry condition: for every player $i \in N$, pure strategy profile $(s_1, \ldots, s_n) \in S^n$, and permutation π of the players,

$$u_i(s_1, \ldots, s_n) = u_{\pi(i)}(s_{\pi(1)}, \ldots, s_{\pi(n)}).$$

This allows us to remove the player index of the utility function and just write $u(s_1, \ldots, s_n)$, where it is implied that the utility is for player 1 (we can simply permute the players to obtain the utilities of the other players). We write u_i for notational convenience, but note that only a single utility function must be specified which applies to all players.

Let Σ_i denote the set of mixed strategies of player i (probability distributions over elements of S_i). If players follow mixed strategy profile $\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)})$, where $\mathbf{x}^{(i)} \in \Sigma_i$, the expected payoff to player i is

$$u_i(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}) = \sum_{s_1, \ldots, s_n \in S} x^{(1)}_{s_1} \cdots x^{(n)}_{s_n} u_i(s_1, \ldots, s_n).$$

We write $u_i(\mathbf{x}) = u_i(\mathbf{x}^{(i)}, \mathbf{x}^{(-i)})$, where $\mathbf{x}^{(-i)}$ denotes the vector of strategies of all players except i. If all players follow the same mixed strategy $\mathbf{x}$, then for all players i:

$$u_i(\mathbf{x}) = u_i(\mathbf{x}, \ldots, \mathbf{x}) = \sum_{s_1, \ldots, s_n \in S} x_{s_1} \cdots x_{s_n} u_i(s_1, \ldots, s_n).$$

Definition 1: A mixed strategy profile $\mathbf{x}^\star$ is a Nash equilibrium if for each player $i \in N$ and for each mixed strategy $\mathbf{x}^{(i)} \in \Sigma_i$: $u_i(\mathbf{x}^{*(i)}, \mathbf{x}^{*(-i)}) \geq u_i(\mathbf{x}^{(i)}, \mathbf{x}^{*(-i)})$.

Definition 2: A mixed strategy profile $\mathbf{x}^\star$ in a symmetric normal-form game in a symmetric Nash equilibrium if it is a Nash equilibrium and: $\mathbf{x}^{*(1)} = \mathbf{x}^{*(2)} = \ldots = \mathbf{x}^{*(n)}$.

Definition 3: A mixed strategy $\mathbf{x}^\star \in \Sigma_1$ is *evolutionarily stable* in a symmetric normal-form game if for each mixed strategy $\mathbf{x} \neq \mathbf{x}^\star$ exactly one of the following holds:

1) $u_1(\mathbf{x}^*, \mathbf{x}^*, \ldots, \mathbf{x}^*) > u_1(\mathbf{x}, \mathbf{x}^*, \ldots, \mathbf{x}^*)$,
2) $u_1(\mathbf{x}^*, \mathbf{x}^*, \ldots, \mathbf{x}^*) = u_1(\mathbf{x}, \mathbf{x}^*, \ldots, \mathbf{x}^*)$ *and* $u_1(\mathbf{x}^*, \mathbf{x}, \mathbf{x}^*, \ldots, \mathbf{x}^*) > u_1(\mathbf{x}, \mathbf{x}, \mathbf{x}^*, \ldots, \mathbf{x}^*)$.

It has been proven that every symmetric normal-form game has at least one symmetric Nash equilibrium [24]. It is clear from Definition 3 that every evolutionarily stable strategy in a symmetric normal-form game must be a symmetric Nash equilibrium (SNE). Thus, a natural approach for ESS computation is to first compute SNE and then perform subsequent procedures to determine whether they are ESS.

Now we present the Stackelberg evolutionary game setting. The leader's set of pure strategies is $\mathcal{M} = \{1, \ldots, m\}$ and the follower phenotypes are $\mathcal{P} = \{1, \ldots, n\}$. The leader's utility function is given by $\mathbf{A_L} \in \mathbb{R}^{m \times n}$, where $\mathbf{A_L}(\ell, i)$ gives the payoff to the leader when the leader plays pure strategy $\ell \in \mathcal{M}$ and the follower population state is phenotype $i \in \mathcal{P}$. The leader's expected utility when the leader plays mixed strategy $\boldsymbol{\sigma}$ and the follower plays mixed strategy $\mathbf{x}$ is $U_L(\boldsymbol{\sigma}, \mathbf{x}) = \sum_{\ell=1}^{m} \sum_{i=1}^{n} \sigma_\ell x_i \mathbf{A_L}(\ell, i)$. The follower payoff tensor is $\mathbf{A_F} \in \mathbb{R}^{m \times n \times n}$, where $\mathbf{A_F}(\ell, i, j)$ is the payoff (fitness) to a follower of phenotype $i \in \mathcal{P}$ when the leader plays pure strategy $\ell \in \mathcal{M}$ and the interacting follower has phenotype $j \in \mathcal{P}$. For each fixed $\ell \in \mathcal{M}$, $\mathbf{A_F}(\ell, \cdot, \cdot)$ is an $n \times n$ symmetric payoff matrix. After the leader commits to mixed strategy $\boldsymbol{\sigma} \in \Delta(\mathcal{M})$, the induced follower payoff matrix is $\mathbf{B}^{\boldsymbol{\sigma}} \in \mathbb{R}^{n \times n}$ with $\mathbf{B}^{\boldsymbol{\sigma}} = \sum_{\ell=1}^{m} \sigma_\ell \mathbf{A_F}(\ell, \cdot, \cdot)$. The follower population then plays a symmetric evolutionary game with payoff matrix $\mathbf{B}^{\boldsymbol{\sigma}}$. A population state $\mathbf{x} \in \Delta(\mathcal{P})$ is admissible for $\boldsymbol{\sigma}$ iff it is an ESS of $\mathbf{B}^{\boldsymbol{\sigma}}$. Let $\mathcal{E}(\boldsymbol{\sigma})$ denote the set of ESSs of $\mathbf{B}^{\boldsymbol{\sigma}}$, and let $g^{\boldsymbol{\sigma}} : \Delta(\mathcal{P}) \rightarrow \mathbb{R}$ denote the follower's *ESS selection function.* Finally let $\mathcal{G}(\boldsymbol{\sigma}) = \arg\max_{\mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma})} g^{\boldsymbol{\sigma}}(\mathbf{x})$.

We now present our general definition for *evolutionarily stable Stackelberg equilibrium* (SESS) in Definition 4. We also define two natural special cases. In the first case, the followers play the ESS that is most beneficial to the leader. We refer to this as an *optimistic evolutionarily stable Stackelberg equilibrium* (OSESS). In this case we have $g^{\boldsymbol{\sigma}}(\mathbf{x}) = U_L(\boldsymbol{\sigma}, \mathbf{x})$. We also consider *pessimistic evolutionarily stable Stackelberg equilibrium* (PSESS), in which the followers play the ESS that is worst for the leader. While the followers are not rationally trying to harm the leader, we can view PSESS as SESS that is robust to worst-case evolutionary stability. For PSESS we have $g^{\boldsymbol{\sigma}}(\mathbf{x}) = -U_L(\boldsymbol{\sigma}, \mathbf{x})$. OSESS and PSESS are defined below in Definitions 5–6. For OSESS we have $\mathcal{G}(\boldsymbol{\sigma}) = \arg\max_{\mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma})} U_L(\boldsymbol{\sigma}, \mathbf{x})$, and for PSESS $\mathcal{G}(\boldsymbol{\sigma}) = \arg\min_{\mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma})} U_L(\boldsymbol{\sigma}, \mathbf{x})$. SESS refines Stackelberg equilibrium by restricting the follower response to strategies that are evolutionarily stable in the induced population game.

Definition 4 (Evolutionarily Stable Stackelberg Equilibrium): Given a selection correspondence $\mathcal{G}(\boldsymbol{\sigma}) \subseteq \mathcal{E}(\boldsymbol{\sigma})$, $(\boldsymbol{\sigma}^*, \mathbf{x}^*)$ is an evolutionarily stable Stackelberg equilibrium if

$$\boldsymbol{\sigma}^* \in \operatorname*{arg\,max}_{\boldsymbol{\sigma} \in \Delta(\mathcal{M})} \max_{\mathbf{x} \in \mathcal{G}(\boldsymbol{\sigma})} U_L(\boldsymbol{\sigma}, \mathbf{x}) \quad \text{and} \quad \mathbf{x}^* \in \mathcal{G}(\boldsymbol{\sigma}^*).$$

Definition 5 (Optimistic SESS): A pair $(\boldsymbol{\sigma}^*, \mathbf{x}^*) \in \Delta(\mathcal{M}) \times \Delta(\mathcal{P})$ is an *optimistic evolutionarily stable Stackelberg equilibrium* if

$$(\boldsymbol{\sigma}^*, \mathbf{x}^*) \in \operatorname*{arg\,max}_{\boldsymbol{\sigma} \in \Delta(\mathcal{M}),\ \mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma})} U_L(\boldsymbol{\sigma}, \mathbf{x}).$$

Definition 6 (Pessimistic SESS): A pair $(\boldsymbol{\sigma}^*, \mathbf{x}^*) \in \Delta(\mathcal{M}) \times \Delta(\mathcal{P})$ is a *pessimistic evolutionarily stable Stackelberg equilibrium* if

$$\boldsymbol{\sigma}^* \in \operatorname*{arg\,max}_{\boldsymbol{\sigma} \in \Delta(\mathcal{M})} \min_{\mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma})} U_L(\boldsymbol{\sigma}, \mathbf{x}) \quad \text{and} \quad \mathbf{x}^* \in \operatorname*{arg\,min}_{\mathbf{x} \in \mathcal{E}(\boldsymbol{\sigma}^*)} U_L(\boldsymbol{\sigma}^*, \mathbf{x}).$$

II. Continuous Stackelberg–ESS

We now extend the Stackelberg–ESS framework to continuous Stackelberg evolutionary games, which arise naturally in biological and control applications such as cancer treatment. In this setting, the leader selects a continuous decision variable, while the followers correspond to an evolving population whose response is described by coupled ecological and evolutionary dynamics.

Let $\mathcal{M} \subseteq \mathbb{R}^d_+$ denote the leader decision set (e.g., treatment doses or control parameters). For a given leader decision $\mathbf{m} \in \mathcal{M}$, the followers are described by a collection of continuous state variables, including population sizes and traits. We denote the follower outcome by $\mathbf{z} = (\mathbf{x}, \mathbf{u}) \in \mathcal{Z}$, where $\mathbf{x} \in \mathbb{R}^n_+$ represents population abundances and $\mathbf{u} \in \mathcal{U} \subseteq \mathbb{R}^n$ represents trait or strategy variables. As prior work has done we will assume that $\mathcal{U} = [0, 1]^n$. In the context of cancer treatment d denotes the number of drugs and n denotes the number of cancer cell phenotypes. The leader objective is given by a function $Q : \mathcal{M} \times \mathcal{Z} \rightarrow \mathbb{R}$. For each $\mathbf{m} \in \mathcal{M}$, the leader induces an eco–evolutionary system governing the dynamics of $(\mathbf{x}, \mathbf{u})$.

For a fixed leader decision $\mathbf{m}$, we define the set of admissible follower responses $\mathcal{E}(\mathbf{m}) \subseteq \mathcal{Z}$ as the set of *eco–evolutionarily stable* outcomes under $\mathbf{m}$. Informally, a follower outcome $\mathbf{z}^* = (\mathbf{x}^*, \mathbf{u}^*)$ is admissible if it satisfies any required ecological equilibrium conditions and is stable against invasion by rare mutant phenotypes in the induced eco–evolutionary system. We assume throughout that $\mathcal{E}(\mathbf{m})$ is nonempty for the leader decisions under consideration. When multiple evolutionarily stable outcomes exist for the same leader decision, we model their selection via a correspondence $\mathcal{G}(\mathbf{m}) \subseteq \mathcal{E}(\mathbf{m})$.

We now define continuous analogues of Stackelberg–ESS.

Definition 7 (Continuous Stackelberg–ESS): Let $\mathcal{G}(\mathbf{m}) \subseteq \mathcal{E}(\mathbf{m})$ be a selection correspondence. A pair $(\mathbf{m}^*, \mathbf{z}^*) \in \mathcal{M} \times \mathcal{Z}$ is a *continuous Stackelberg–ESS* if

$$\mathbf{m}^* \in \arg\max_{\mathbf{m}\in\mathcal{M}} \max_{\mathbf{z}\in\mathcal{G}(\mathbf{m})} Q(\mathbf{m}, \mathbf{z}) \quad \text{and} \quad \mathbf{z}^* \in \mathcal{G}(\mathbf{m}^*).$$

Definition 8 (Optimistic continuous Stackelberg–ESS): A pair $(\mathbf{m}^*, \mathbf{z}^*) \in \mathcal{M} \times \mathcal{Z}$ is an *optimistic continuous Stackelberg–ESS* if

$$(\mathbf{m}^*, \mathbf{z}^*) \in \arg\max_{\mathbf{m}\in\mathcal{M},\ \mathbf{z}\in\mathcal{E}(\mathbf{m})} Q(\mathbf{m}, \mathbf{z}).$$

Definition 9 (Pessimistic continuous Stackelberg–ESS): A pair $(\mathbf{m}^*, \mathbf{z}^*) \in \mathcal{M} \times \mathcal{Z}$ is a *pessimistic continuous Stackelberg–ESS* if

$$\mathbf{m}^* \in \arg\max_{\mathbf{m}\in\mathcal{M}} \min_{\mathbf{z}\in\mathcal{E}(\mathbf{m})} Q(\mathbf{m}, \mathbf{z}) \quad \text{and} \quad \mathbf{z}^* \in \arg\min_{\mathbf{z}\in\mathcal{E}(\mathbf{m}^*)} Q(\mathbf{m}^*, \mathbf{z}).$$

First we present the Stackelberg evolutionary game formulation of cancer treatment from prior work [18]. In this formulation (Equation 1), the objective Q corresponds to the quality of life function that is maximized by the leader. For each cell type i, there is a fitness function $G_i(u_i, \mathbf{m}, \mathbf{x})$ that the follower is trying to maximize. We assume that the dynamics of the population $\mathbf{x}$ are governed by $\dot{\mathbf{x}} = \mathbf{G}(t)\mathbf{x}$. In order to ensure that we are in equilibrium of the ecological dynamics we must have that $\dot{x}_i = 0$ for all i. Note that in this formulation the follower is playing a best response by maximizing the fitness function. By contrast, in Stackelberg-ESS we assume that the behavior of the cancer cells is determined by evolutionary stability conditions as opposed to explicit utility maximization. An algorithm for solving this problem has been presented that is based on a nonconvex quadratic program formulation [8]. This algorithm has been demonstrated to quickly compute a global optimal solution on a realistic example problem proposed by prior work [18].

$$\begin{aligned}
\max_{\mathbf{m}^*,\mathbf{u}^*,\mathbf{x}^*} \quad & Q(\mathbf{m}^*, \mathbf{u}^*, \mathbf{x}^*) \\
\text{s.t.} \quad & \dot{x}_i^* = 0, \quad i = 1, \ldots, n, \\
& u_i^* \in \arg\max_{u_i\in[0,1]} G_i(u_i, \mathbf{m}^*, \mathbf{x}^*), \quad i = 1, \ldots, n, \\
& \mathbf{m}^* \geq \mathbf{0}, \quad \mathbf{x}^* \geq \mathbf{0}, \quad \mathbf{0} \leq \mathbf{u}^* \leq \mathbf{1}.
\end{aligned} \tag{1}$$

The formulation for optimistic continuous Stackelberg-ESS is given by Equation 2. This formulation is similar to the previous one for standard Stackelberg equilibrium except that now instead of maximizing the fitness functions the follower ensures that their strategy is an ESS in the induced subgame given the leader is playing m^*. We will focus our attention on the problem of computing OSESS under this formulation, since it is perhaps the easiest case of SESS to solve since both players are aligned in maximizing Q.

$$\begin{aligned}
\max_{\mathbf{m}^*,\mathbf{u}^*,\mathbf{x}^*} \quad & Q(\mathbf{m}^*, \mathbf{u}^*, \mathbf{x}^*) \\
\text{s.t.} \quad & \dot{x}_i^* = 0, \quad i = 1, \ldots, n, \\
& (\mathbf{x}^*, \mathbf{u}^*) \in \mathcal{E}(\mathbf{m}^*), \\
& \mathbf{m}^* \geq \mathbf{0}, \quad \mathbf{x}^* \geq \mathbf{0}, \quad \mathbf{0} \leq \mathbf{u}^* \leq \mathbf{1}.
\end{aligned} \tag{2}$$

Equation 3 provides an equivalent reformulation of Equation 2. Note that we use the continuous-trait definition of ESS for our setting [16], [25], [11] as opposed to the discrete ESS definition presented previously. Given fixed $\mathbf{m}$, a resident state $(\mathbf{x}^*, \mathbf{u}^*)$ is an ESS iff $\mathbf{x}^*$ is an ecological equilibrium (i.e., $\dot{\mathbf{x}}^* = \mathbf{0}$) and no rare mutant can grow, i.e., $\sup_{u_i} G_i(u_i, \mathbf{m}, \mathbf{x}^*) \leq 0 \ \forall i$. Recall that $\dot{\mathbf{x}} = \mathbf{G}\mathbf{x}$. So the constraint $\dot{x}_i^* = 0$ is equivalent to $G_i(\mathbf{u}_i^*, \mathbf{m}^*, \mathbf{x}^*)x_i^* = 0$. Recall that $(\mathbf{x}^*, \mathbf{u}^*) \in \mathcal{E}(\mathbf{m})$ iff $(\mathbf{x}^*, \mathbf{u}^*)$ is evolutionarily stable in the eco-evolutionary system induced by $\mathbf{m}$. This is true iff for all i and u_i a mutation to u_i does not lead to positive fitness (growth rate), i.e., $G_i(u_i, \mathbf{m}, \mathbf{x}^*) \leq 0$.

$$\begin{aligned}
\max_{\mathbf{m}^*,\mathbf{u}^*,\mathbf{x}^*} \quad & Q(\mathbf{m}^*, \mathbf{u}^*, \mathbf{x}^*) \\
\text{s.t.} \quad & x_i^* \, G_i(u_i^*, \mathbf{m}^*, \mathbf{x}^*) = 0, \quad i = 1, \ldots, n, \\
& G_i(u_i, \mathbf{m}^*, \mathbf{x}^*) \leq 0 \quad \forall u_i \in [0,1], \ i = 1, \ldots, n, \\
& \mathbf{m}^* \geq \mathbf{0}, \quad \mathbf{x}^* \geq \mathbf{0}, \quad \mathbf{0} \leq \mathbf{u}^* \leq \mathbf{1}.
\end{aligned} \tag{3}$$

III. Algorithms

We first present an algorithm for computing OSESS in normal-form games, followed by an algorithm for continuous-trait games. Pseudocode for the main algorithm in the normal-form setting is given in Algorithm 1. We compute an OSESS by enumerating candidate follower supports and for each support computing a Stackelberg equilibrium of the game where the follower is restricted to strategies with the given support. This is done by using the subroutine given in Algorithm 2. Given the Stackelberg equilibrium $(\boldsymbol{\sigma}, \boldsymbol{x})$, we next check whether $\boldsymbol{x}$ is an ESS in the subgame induced by the leader's strategy $\boldsymbol{\sigma}$. This is done in Algorithm 3. If the strategies do constitute an ESS, then we have found an evolutionarily stable Stackelberg equilibrium (SESS). We then calculate the objective $\boldsymbol{\sigma}^\top \mathbf{A}_L$, and return the SESS with highest value as an OSESS. The algorithm can be easily adapted to output the first SESS found, all SESSs found, or an SESS that optimizes a different objective. The values used for all numerical tolerance parameters are given in Table I.

Algorithm 2 computes a Stackelberg equilibrium strategy with given support T by solving a quadratically-constrained quadratic program (QCQP). The constraints $\boldsymbol{x}_i \geq \epsilon_s$ ensure that the pure strategies in the support are played with positive probability. The objective ensures that the leader's payoff is maximized, and the final two sets of constraints ensure that the follower is best-responding to the leader's strategy. Our algorithm is implicitly assuming that the game is *nondegenerate* and that there exists at least one Stackelberg equilibrium for each follower support. This is a common assumption in equilibrium-finding algorithms [22], [12], [15], and has been studied recently in the context of ESS computation [10], [9]. If the game is degenerate with an infinite continuum of equilibria our algorithm will fail to find all of them, and therefore may fail to find all SESSs; however it may still find one. Algorithm 3 also involves solving a QCQP. In practice we will be solving nonconvex QCQPs with Gurobi's solver, which can solve small problem instances quickly despite their computational complexity. The QCQP checks whether any mutant strategy $\mathbf{y}$ can successfully invade against a population following strategy $\mathbf{x}$ in the game induced by the leader's strategy σ. The parameter ϵ_p allows for a small amount of numerical imprecision, and the final constraint ensures that $\mathbf{y}$ is numerically distinct from $\mathbf{x}$ by using numerical separation parameter δ.

Algorithm 1 COMPUTEDISCRETEOSESS

```
Require: Follower payoff tensor $\mathbf{A}_F \in \mathbb{R}^{m\times n\times n}$
Require: Leader payoff matrix $\mathbf{A}_L \in \mathbb{R}^{m\times n}$
Require: Minimum support mass $\epsilon_s$, ESS tolerances $\epsilon_p, \delta$
Ensure: OSESS $(\boldsymbol{\sigma}^\star, \boldsymbol{x}^\star)$
  $v^\star \leftarrow -\infty$
  $(\boldsymbol{\sigma}^\star, \boldsymbol{x}^\star) \leftarrow (\varnothing, \varnothing)$
  for all nonempty supports $T \subseteq P$ do
    $(status, \boldsymbol{\sigma}, \boldsymbol{x}) \leftarrow$ SOLVESESUPPORT$(\mathbf{A}_F, \mathbf{A}_L, T, \epsilon_s)$
    if $status =$ FEASIBLE then
      if ISESS$(\mathbf{A}_F, \boldsymbol{\sigma}, \boldsymbol{x}, \epsilon_p, \delta)$ then
        $v \leftarrow \boldsymbol{\sigma}^\top \mathbf{A}_L \boldsymbol{x}$
        if $v > v^\star$ then
          $v^\star \leftarrow v$
          $(\boldsymbol{\sigma}^\star, \boldsymbol{x}^\star) \leftarrow (\boldsymbol{\sigma}, \boldsymbol{x})$
        end if
      end if
    end if
  end for
  return $(\boldsymbol{\sigma}^\star, \boldsymbol{x}^\star)$
```

Pseudocode for the main algorithm in continuous setting is given in Algorithm 4. The algorithm follows a generate-and-certify structure: first a candidate solution satisfying the KKT conditions of the invasion maximization problems is computed, and then a global optimization check verifies that no mutant phenotype can invade. The main subroutine Algorithm 5 generates a candidate OSESS by solving the optimization problem presented in Equation 3. The formulation finds a solution that satisfies the KKT optimality conditions, though it is not guaranteed to be a globally optimal solution. To test if this candidate solution is globally optimal we perform the *ex-post* certification procedure described in Algorithm 6, using a numerical tolerance of ϵ_{inv}. We set $\epsilon_{inv} = 10^{-3}$, which is biologically negligible but significantly larger than Gurobi's default tolerance of 10^{-6}. Note that if the

Algorithm 2 SOLVESESUPPORT

```
Require: Follower payoff tensor $\mathbf{A}_F \in \mathbb{R}^{m\times n\times n}$
Require: Leader payoff matrix $\mathbf{A}_L \in \mathbb{R}^{m\times n}$
Require: Support $T \subseteq P$, minimum support mass $\epsilon_s$
Ensure: $(status, \boldsymbol{\sigma}, \boldsymbol{x})$
  Solve the following optimization problem:
      $\max_{\boldsymbol{\sigma}, \boldsymbol{x}, v} \quad \boldsymbol{\sigma}^\top \mathbf{A}_L \boldsymbol{x}$
      s.t. $\boldsymbol{\sigma}_\ell \geq 0, \quad \forall \ell \in M,$
           $\sum_{\ell \in M} \boldsymbol{\sigma}_\ell = 1,$
           $\boldsymbol{x}_i \geq \epsilon_s, \quad \forall i \in T,$
           $\boldsymbol{x}_j = 0, \quad \forall j \notin T,$
           $\sum_{i \in P} \boldsymbol{x}_i = 1,$
           $\sum_{\ell \in M} \sum_{k \in P} \boldsymbol{\sigma}_\ell \mathbf{A}_F(\ell, i, k) \boldsymbol{x}_k = v, \quad \forall i \in T,$
           $\sum_{\ell \in M} \sum_{k \in P} \boldsymbol{\sigma}_\ell \mathbf{A}_F(\ell, j, k) \boldsymbol{x}_k \leq v, \quad \forall j \notin T.$
  if the problem is feasible then
    return $(\texttt{FEASIBLE}, \boldsymbol{\sigma}, \boldsymbol{x})$
  else
    return $(\texttt{INFEASIBLE}, \varnothing, \varnothing)$
  end if
```

Algorithm 3 ISESS

```
Require: Follower payoff tensor $\mathbf{A}_F \in \mathbb{R}^{m\times n\times n}$
Require: Leader mixed strategy $\boldsymbol{\sigma} \in \Delta(M)$
Require: Follower mixed strategy $\boldsymbol{x} \in \Delta(P)$
Require: Tolerances $\epsilon_p, \delta$
Ensure: TRUE if $\boldsymbol{x}$ is an ESS of the induced follower game;
    otherwise FALSE
  Define the induced follower payoff matrix $\mathbf{B} \in \mathbb{R}^{n\times n}$:
      $\mathbf{B}_{ij} \leftarrow \sum_{\ell \in M} \boldsymbol{\sigma}_\ell \mathbf{A}_F(\ell, i, j), \qquad \forall i, j \in P.$
  $v \leftarrow \boldsymbol{x}^\top \mathbf{B} \boldsymbol{x}$
  Solve the following optimization problem:
      $\max_{\boldsymbol{y}} \quad \boldsymbol{y}^\top \mathbf{B} \boldsymbol{y} - \boldsymbol{x}^\top \mathbf{B} \boldsymbol{y}$
      s.t. $\boldsymbol{y}_j \geq 0, \quad \forall j \in P,$
           $\sum_{j \in P} \boldsymbol{y}_j = 1,$
           $\left| \boldsymbol{y}^\top \mathbf{B} \boldsymbol{x} - v \right| \leq \epsilon_p,$
           $\| \boldsymbol{y} - \boldsymbol{x} \|_2^2 \geq \delta.$
  if the optimal objective value $\leq \epsilon_p$ then
    return TRUE
  else
    return FALSE
  end if
```

quality of life functions G_i are concave in u_i then the KKT conditions are also sufficient. In our experiments we will use previously-studied Q and G_i functions, and the algorithm will involve solving nonconvex QCQPs.

```
Algorithm 4 COMPUTECONTINUOUSOSESS
Require: Leader objective $Q(\mathbf{m}, \mathbf{u}, \mathbf{x})$
Require: Fitness functions $\{G_i(u_i, \mathbf{m}, \mathbf{x})\}_{i=1}^{n}$
Require: Feasible set $\mathcal{M}$, tolerance $\epsilon_{\text{inv}}$
Ensure: OSESS $(\mathbf{m}^\star, \mathbf{u}^\star, \mathbf{x}^\star)$ if one is found
  $(\mathbf{m}^\star, \mathbf{u}^\star, \mathbf{x}^\star) \leftarrow$ GENOSESS$(Q, \{G_i\}, \mathcal{M})$
  for $i = 1$ to $n$ do
      $G_i^{\max} \leftarrow$ CERTIFY$(i, \mathbf{m}^\star, \mathbf{x}^\star)$
  end for
  if $\max_i G_i^{\max} \leq \epsilon_{\text{inv}}$ then
      return $(\mathbf{m}^\star, \mathbf{u}^\star, \mathbf{x}^\star)$
  else
      return FAILURE
  end if
```

```
Algorithm 5 GENOSESS
Require: Leader objective $Q(\mathbf{m}, \mathbf{u}, \mathbf{x})$
Require: Fitness functions $\{G_i(u_i, \mathbf{m}, \mathbf{x})\}_{i=1}^{n}$
Require: Feasible set $\mathcal{M}$
Ensure: Candidate solution $(\mathbf{m}^\star, \mathbf{u}^\star, \mathbf{x}^\star)$
  Solve the following optimization problem:
      $\max_{\mathbf{m},\mathbf{u},\mathbf{x},\mathbf{u}',\boldsymbol{\lambda}^L,\boldsymbol{\lambda}^U} \quad Q(\mathbf{m}, \mathbf{u}, \mathbf{x})$
      s.t. $x_i G_i(u_i, \mathbf{m}, \mathbf{x}) = 0, \quad i = 1, \ldots, n,$
           $\mathbf{m} \in \mathcal{M},\ \mathbf{x} \geq 0,\ 0 \leq \mathbf{u} \leq 1,\ 0 \leq \mathbf{u}' \leq 1,$
           $G_i(u_i', \mathbf{m}, \mathbf{x}) \leq 0, \quad i = 1, \ldots, n,$
           $\frac{\partial G_i}{\partial u_i}(u_i', \mathbf{m}, \mathbf{x}) + \lambda_i^L - \lambda_i^U = 0,\ i = 1, \ldots, n,$
           $\lambda_i^L u_i' = 0, \quad i = 1, \ldots, n,$
           $\lambda_i^U (1 - u_i') = 0, \quad i = 1, \ldots, n,$
           $\lambda_i^L \geq 0,\ \lambda_i^U \geq 0, \quad i = 1, \ldots, n.$
  return optimal solution $(\mathbf{m}^\star, \mathbf{u}^\star, \mathbf{x}^\star)$
```

```
Algorithm 6 CERTIFY
Require: Phenotype index $i \in \{1, \ldots, n\}$
Require: Fixed $(\mathbf{m}^\star, \mathbf{x}^\star)$
Ensure: $G_i^{\max}$
  Computing globally optimal solution:
      $\max_{u} \quad G_i(u, \mathbf{m}^\star, \mathbf{x}^\star)$
      s.t. $\quad 0 \leq u \leq 1$
  return optimal objective $G_i^{\max}$
```

IV. EXPERIMENTS

We experiment with our continuous algorithm on an evolutionary cancer game model previously studied [18]. Note that prior work for this model has computed a Stackelberg equilibrium (SE), where the leader commits to a strategy that optimizes the quality of life function Q while the cancer cell phenotypes optimize their respective fitness functions G_i (subject to the equilibrium conditions of the ecological dynamics being satisfied) [18], [8]. In contrast, we compute an OSESS in this game, which guarantees that the follower cancer cells are following an ESS in the game induced by the leader's strategy and no mutant trait can have positive growth (while SE does not preclude existence of mutants with positive growth). This is fundamentally a different solution concept, so the players' payoffs may increase or decrease. The model is instantiated by the following functional forms for the fitness functions G_i and quality of life function Q. Note that the model presentation is slightly different between the paper [18] and the implementation in the code repository [17], and we will use the model from the code.

Parameter	Value
ϵ_s	10^{-4}
ϵ_p	10^{-5}
δ	10^{-2}
ϵ_{inv}	10^{-3}

TABLE I
TOLERANCE PARAMETER VALUES FOR THE ALGORITHMS.

$$G_0 = r_{\max}\left(1 - \frac{\alpha_{00}x_0 + \alpha_{01}x_1 + \alpha_{02}x_2}{K}\right) - d - \frac{m_1}{k_1} - \frac{m_2}{k_2}$$

$$G_1 = r_{\max}e^{-g_1 u_1}\left(1 - \frac{\alpha_{10}x_0 + \alpha_{11}x_1 + \alpha_{12}x_2}{K}\right) - d - \frac{m_1}{b_1 u_1 + k_1} - \frac{m_2}{k_2}$$

$$G_2 = r_{\max}e^{-g_2 u_2}\left(1 - \frac{\alpha_{20}x_0 + \alpha_{21}x_1 + \alpha_{22}x_2}{K}\right) - d - \frac{m_1}{k_1} - \frac{m_2}{b_2 u_2 + k_2}$$

$$Q = Q^{\max} - c\left(\frac{x_0 + x_1 + x_2}{K}\right)^2 - w_1 m_1^2 - w_2 m_2^2 - r_1 u_1^2 - r_2 u_2^2$$

The model has several parameters, whose interpretations are summarized in Table II. Note that in the code additional parameters a_0, a_1, a_2, a_3 are defined, with

$$\begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} \\ \alpha_{10} & \alpha_{11} & \alpha_{12} \\ \alpha_{20} & \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_1 \\ a_2 & a_0 & a_3 \\ a_2 & a_3 & a_0 \end{bmatrix}$$

We ran experiments comparing the algorithm for computing a Stackelberg equilibrium [8] to our new algorithm for computing OSESS on a problem instance using the same parameter values that have been previously used [17], which are provided in Table III. Both approaches involve solving QCQPs using Gurobi's nonconvex mixed-integer quadratically-constrained programming (MIQCP) solver version 11.0.3 build v11.0.3rc0 [13] with Java version 14.0.2. Gurobi's nonconvex MIQCP solver guarantees global optimality (up to a numerical tolerance). We used Gurobi's default numerical tolerance value of 10^{-6}. We used Gurobi's default values for all settings other than DualReductions which we set to 0 since we observed improved performance by disabling aggressive presolve reductions. We used an Intel Core i7-1065G7 processor with base clock speed of 1.30 GHz (and

Parameter	Interpretation
$r_{\max}$	Max cell growth rate
g_i	Cost of resistance strategy (cell type) i
α_{ij}	Interaction coefficient between cell types i and j
K	Carrying capacity
d	Natural death rate
k_i	Innate resistance that may be present before drug exposure
b_i	Benefit of the evolved resistance trait in reducing therapy efficacy
$Q^{\max}$	Quality of life of a healthy patient
w_i	Toxicity of drug i
r_i	Effect of resistance rate of cell type i
c	Weight for impact of tumor burden vs. drug toxicity/treatment-induced resistance rate

TABLE II

INTERPRETATIONS OF MODEL PARAMETERS

maximum turbo boost speed of up to 3.9 GHz) with 16 GB of RAM under 64-bit Windows 11 (8 logical cores/threads).

Parameter	Value
$r_{\max}$	0.45
g_1	0.5
g_2	0.5
a_0	1
a_1	0.15
a_2	0.9
a_3	0.9
K	10,000
d	0.01
k_1	5
k_2	5
b_1	10
b_2	10
$Q^{\max}$	1
w_1	0.5
w_2	0.2
r_1	0.4
r_2	0.4
c	0.5

TABLE III

PARAMETER VALUES USED IN EXPERIMENTS

The main results from our experiments are depicted in Table IV. Note that the optimal physician strategies $(\mathbf{m}^*)$ vary slightly between the two solutions, and the optimal quality of life Q^* is slightly higher for the OSESS solution than the SE solution. The SE algorithm ran significantly faster than the OSESS algorithm (around 2 seconds versus 2.2 minutes). Both algorithms are based on KKT necessary optimality conditions and require the application of *ex-post* certification procedures to ensure global optimality. For both algorithms we recompute the values of Q and G_i exactly by calculating them directly from the solutions output from Gurobi (since the values of Q and G_i in the optimizations themselves contain numerical error). For the OSESS algorithm the certification procedure obtains $G_1^{max} = 7.85 \times 10^{-5}$, $G_2^{max} = 1.41 \times 10^{-4}$, which are both significantly below $\epsilon_{inv} = 10^{-3}$, so we conclude that our solution is in fact an OSESS. As it turns out, running the same certification procedure on the Stackelberg equilibrium solution obtains $G_1^{max} = 8.00 \times 10^{-5}$, $G_2^{max} = 1.42 \times 10^{-4}$, which are also both significantly below $\epsilon_{inv} = 10^{-3}$. So for these particular game parameters it turns out that the SE solution obtained is also an OSESS, though this will not be the case in general.

	Stackelberg Equilibrium	OSESS
m_1^*	0.4105	0.4003
m_2^*	0.4680	0.4571
u_1^*	0.2139	0.1827
u_2^*	0.2856	0.2828
x_0^*	5731.0481	5823.7239
x_1^*	0.0087	9.5179
x_2^*	950.7623	946.4278
Q^*	0.5978	0.6029
Runtime (seconds)	2.2070	134.6910

TABLE IV

EXPERIMENTAL RESULTS FOR EVOLUTIONARY CANCER GAME

V. CONCLUSION

Stackelberg evolutionary games have emerged as a powerful model for studying many types of biological interactions, including cancer therapy, fishery management, and pest control [18]. In these games, a rational human leader first selects their strategy, then evolutionary followers select their strategies conditional on the leader's strategy (possibly subject to ecological population equilibrium constraints). Prior work has assumed that the follower players act to maximize fitness functions G_i. In the resulting Stackelberg equilibrium, the leader and followers are simultaneously maximizing their respective objective functions. However, this solution does not guarantee that the follower strategy is resistant to invasion by rare mutants. To ensure evolutionary stability we must impose the condition that the fitness functions are nonpositive (subject to numerical tolerance) for all possible mutations so that their populations do not continue to grow. We introduced the new solution concept *evolutionarily stable Stackelberg equilibrium* (SESS) in which the leader is a rational utility maximizer, while the followers employ an evolutionarily stable strategy (ESS) that is resistant to invasion. In an SESS, the leader commits to a strategy that maximizes their objective assuming the followers will play an ESS in the subgame induced by the leader's strategy. This imposes a stricter stability requirement on the follower strategy than standard Stackelberg equilibrium, which only requires best-response behavior without ensuring evolutionary stability.

We define SESS both for discrete normal-form games, as well continuous-trait games, which arise naturally in biological and control applications such as cancer treatment. In our general formulation, the followers play any ESS in the subgame induced by the leader's strategy. We also consider the special cases where the followers play the ESS that is most beneficial to the leader (OSESS) and the ESS that is worst for the leader (PSESS). OSESS is perhaps the easiest case of SESS to compute since in a sense all players are aligned in maximizing the leader's payoff (subject to the followers playing an ESS), which simplifies the optimization to a single outer maximization problem. We present algorithms for computing OSESS in both the discrete and continuous settings. If we view the follower strategies as population frequencies, then our model is applicable to settings with a large number of follower players; e.g., there may be a large number of cancer cells where each cell has one of a small number of possible phenotypes.

We implemented our OSESS algorithm on an evolutionary cancer game model that has been previously studied and demonstrated that its runtime is reasonable in practice. It is not surprising that OSESS computation takes longer than SE computation, since OSESS is essentially SE with additional stability constraints imposed and therefore involves solving a more challenging optimization problem. As it turns out, the Stackelberg equilibrium solution for this game ended up being an OSESS as well, though this is not guaranteed in general. This suggests that for larger games perhaps a good approach would be to first compute an SE, which is more tractable than SESS, and test whether it also satisfies evolutionary stability; if not then we can continue on to the more computationally expensive SESS algorithm.

References

[1] Andris Abakuks. Conditions for evolutionarily stable strategies. *Journal of Applied Probability*, 17(2):559–562, 1980.

[2] Manon Blanc and Kristoffer Arnsfelt Hansen. Computational complexity of multi-player evolutionarily stable strategies. In *International Computer Science Symposium in Russia (CSR 2021)*, volume 12730 of *Lecture Notes in Computer Science*, pages 44–58. Springer, 2021.

[3] I. M. Bomze. Detecting all evolutionarily stable strategies. *Journal of Optimization Theory and Applications*, 75:313–329, 11 1992.

[4] Mark Broom and Jan Rychtář. *Game-Theoretical Models in Biology*. CRC Press, 2013.

[5] Vincent Conitzer. The exact computational complexity of evolutionarily stable strategies. In *Conference on Web and Internet Economics (WINE-13)*, pages 96–108, 2013.

[6] David Dingli, Francisco A. C. C. Chalub, Francisco C. Santos, Sven Van Segbroeck, and Jorge M. Pacheco. Cancer phenotype as the outcome of an evolutionary game between normal and malignant cells. *British Journal of Cancer*, 101(7):1130–1136, 2009.

[7] Kousha Etessami and Andreas Lochbihler. The computational complexity of evolutionarily stable strategies. *International Journal of Game Theory*, 37(1):93–113, 2008.

[8] Sam Ganzfried. Computing Stackelberg equilibrium for cancer treatment. *Games*, 15(6), 2024.

[9] Sam Ganzfried. Computing evolutionarily stable strategies in imperfect-information games, 2025. arXiv:2512.10279 [cs.GT].

[10] Sam Ganzfried. Computing evolutionarily stable strategies in multiplayer games, 2025. arXiv:2511.20859 [cs.GT].

[11] Stefan A. H. Geritz, Éva Kisdi, Gábor Meszéna, and Johan A. J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12(1):35–57, 1998.

[12] Srihari Govindan and Robert Wilson. A global Newton method to compute Nash equilibria. *Journal of Economic Theory*, 110:65–86, 2003.

[13] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2026.

[14] John Haigh. Game theory and evolution. *Advances in Applied Probability*, 7(1):8–11, 1975.

[15] P. Jean-Jacques Herings and Ronald Peeters. Homotopy methods to compute equilibria in game theory. *Economic Theory*, 42(1):119–156, 2010.

[16] Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK, 1998.

[17] Maria Kleshnina. `https://github.com/kleshnina/SEGopinion`, 2023.

[18] Maria Kleshnina, Sabrina Streipert, Joel Brown, and Kateřina Staňková. Game theory for managing evolving systems: Challenges and opportunities of including vector-valued strategies and life-history traits. *Dynamic Games and Applications*, 13:1–26, 2023.

[19] Michael Maschler, Eilon Solan, and Shmuel Zamir. *Game Theory*. Cambridge University Press, 2013.

[20] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK, 1982.

[21] John Maynard Smith and George R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

[22] Richard D. McKelvey and Andrew McLennan. Computation of equilibria in finite games. In H. Amann, D. Kendrick, and J. Rust, editors, *Handbook of Computational Economics*, volume 1, pages 87–142. Elsevier, 1996.

[23] John McNamara, James N. Webb, E. J. Collins, Tamás Székely, and Alasdair I. Houston. A general technique for computing evolutionarily stable strategies based on errors in decision-making. *Journal of Theoretical Biology*, 189:211–225, 1997.

[24] John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[25] William H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, MA, 2010.