SCIENCE CHINA
Information Sciences

• REVIEW •

April 2024, Vol. 67, Iss. 4, 141201:1–141201:28
https://doi.org/10.1007/s11432-022-3777-y

A survey of decision making in adversarial games

Xiuxian LI[1,2], Min MENG[1,2*], Yiguang HONG[1,2] & Jie CHEN[1,2]

[1]*College of Electronics and Information Engineering, Tongji University, Shanghai* 201804*, China;*
[2]*Shanghai Research Institute for Intelligent Autonomous Systems, Shanghai* 201210*, China*

Received 28 September 2022/Revised 21 January 2023/Accepted 21 March 2023/Published online 4 February 2024

Abstract In many practical applications, such as poker, chess, drug interdiction, cybersecurity, and national defense, players often have adversarial stances, i.e., the selfish actions of each player inevitably or intentionally inflict loss or wreak havoc on other players. Therefore, adversarial games are important in real-world applications. However, only special adversarial games, such as Bayesian games, are reviewed in the literature. In this respect, this study aims to provide a systematic survey of three main game models widely employed in adversarial games, i.e., zero-sum normal-form and extensive-form games, Stackelberg (security) games, and zero-sum differential games, from an array of perspectives, including basic knowledge of game models, (approximate) equilibrium concepts, problem classifications, research frontiers, (approximate) optimal strategy-seeking techniques, prevailing algorithms, and practical applications. Finally, promising future research directions are also discussed for relevant adversarial games.

Keywords adversarial games, zero-sum games, Stackelberg games, differential games, Nash equilibrium, correlated equilibrium, regret

1 Introduction

Game theory has long been a powerful and conventional paradigm for modeling complex and intelligent interactions among a group of players and improving decision making for selfish players since the seminal works of John von Neumann, John Nash, and others [1–3]. Hitherto, game theory has a vast range of real-world applications in a variety of domains, including economics, biology, finance, computer science, and politics, where each player is only concerned with his/her own interest [4–7]. Game theory played an extremely important role even during the Cold War in the 1960s and has been employed by many national institutions in defense, such as United States agencies for security control [8].

Adversarial games are a class of particularly important game models where players deliberately compete against each other while simultaneously minimizing their losses. To date, adversarial games have been an orthodox framework for ensuring highly efficient decision making in numerous realistic applications, such as poker, chess, evader pursuit, drug interdiction, coast guard, cybersecurity, and national defense. For example, in Texas Hold'em poker, which has been one of the primary competitions as a benchmark for testing researchers' proposed algorithms in game theory and artificial intelligence (AI) held by well-known international conferences, such as AAAI, multiple players compete against each other to win the game by seeking sophisticated strategies and techniques [9]. In general, adversarial games have the following main features: (1) hardness of efficient and fast algorithms' design with limited computing resources and/or samples [5]; (2) imperfect information on many practical problems [10], i.e., some information is private to one or more players, which, however, is hidden from other players, such as the card game of poker; (3) large models [11], including large action spaces and information sets (infosets), e.g., the adversarial space in the road network security problem is of the order 10^{18} [12]; (4) incomplete information on a multitude of real-world applications [10], i.e., one or more agents do not know what game is being played (e.g., the number of players, the strategies available to each player, and the payoff for each strategy), and in this case, the game being played is generally presented with players' uncertainties, such as uncertain payoff functions with uncertain parameters [10]; and (5) possible dynamic trait [13], i.e., the played game

* Corresponding author (email: mengmin@tongji.edu.cn)

Table 1 Relevant recent existing survey studies

Year	Refs.	Content summary	Remark
2018	[11]	It focuses on technical advances in SSG by 2018	Each existing literature provides an overview only on some special games, for example, SSG in [11], and no literature provides a thorough overview of adversarial games, which, however, is exactly the aim of this paper
2019	[13]	It focuses on dynamic game analysis of cyber-physical security problems	
2020	[19]	It reviews a combined use of game theory and optimization algorithms, along with the introduction of only basic knowledge of game theory, e.g., NE	
2020	[23]	It reviews multi-agent perimeter defense games, where a group of intruders try to reach the target region while another group of defenders try to intercept the intruders to protect the target region	
2020	[25]	It reviews Bayesian methods to deal with decision problems in reliability	
2020	[17]	It reviews the econometrics of static games with complete and incomplete information	
2021	[18]	It reviews counterfactual regret minimization methods for imperfect-information extensive-form games	
2021	[24]	It focuses on defensive deception research using game theory and machine learning, along with the introduction of basic concepts of game theory	
2022	[20]	It reviews optimization, Markov games (multi-agent reinforcement learning) Bayesian games, and mean-field games	
2022	[21]	It reviews distributed online optimization and online games (continuous games with time-varying cost functions)	
2022	[22]	It provides a succinct review of the literature using game theory to model decision making scenarios in defense applications	

is sometimes time-varying, instead of static, e.g., a poacher may have different poaching strategies in a wildlife park as the environment varies with seasons [13]. It is worth pointing out that incomplete information is distinct from imperfect information here, as distinguished by some researchers, although they are interchangeably used in some literature. Furthermore, other possible characteristics include bounded rationality [11], where players may not be fully rational, such as arbitrarily random lone wolf attacks by terrorists. However, it is noteworthy that not all adversarial games have imperfect and/or incomplete information, e.g., the game of Go has both perfect and complete information because it has explicit game rules and the positions of all chess pieces, as well as the actions of the opponent, are visible to both players at all times, which has been well solved by well-known AI agents, such as AlphaGo and AlphaZero [14–16].

As the competitive feature is ubiquitous in a large number of real-world applications, adversarial games have been extensively investigated, e.g., [11, 13, 17–25]. For example, in [11], the authors provided a broad survey of technical advances in Stackelberg security games (SSGs), and in [18], the authors reviewed some of the main Nash equilibrium (NE) computing algorithms for extensive-form games with imperfect information based on counterfactual regret minimization (CFR) methods. More details can be found in Table 1. Nonetheless, a thorough overview of adversarial games from the perspectives of the basic knowledge of game models, equilibrium concepts, optimal strategy seeking techniques, research frontiers, and prevailing algorithms is still lacking.

Motivated by the aforementioned facts, this survey aims to provide a systematic review of adversarial games from several dimensions, including the three main models frequently employed in adversarial games (i.e., zero-sum normal-form and extensive-form games, Stackelberg (security) games, and zero-sum differential games (DGs)), (approximate) optimal strategy concepts (i.e., NE, correlated equilibrium, coarse correlated equilibrium, strong Stackelberg equilibrium (SSE), team-maxmin equilibrium (TME), and their corresponding approximate equilibria), (approximate) optimal strategy computing techniques (e.g., CFR and AI methods), state-of-the-art results, prevailing algorithms, potential applications, and promising future research directions. To the best of our knowledge, this survey is the first systematic overview of adversarial games, generally providing an orthogonal and complementary component for the aforementioned survey papers, which may aid researchers and practitioners in relevant domains. Please note that the three main game models are not mutually exclusive but may overlap for the same game from different viewpoints. For example, the Stackelberg games (SGs) and DGs can also be zero-sum games. Furthermore, other models leveraged for adversarial games, such as Bayesian games, Markov games (or stochastic games), signaling games, behavioral game theory, and evolutionary game theory, exist. However, we do not attempt to review all of them in this survey because each of them is of independent interest and abundant in existing diverse materials. Finally, interested readers can refer to [17, 25] for related surveys of Bayesian games and [20] for an overview of Markov games.

This survey paper is organized as follows: the detailed game models and solution concepts are introduced in Section 2. The existing literature, along with state-of-the-art results, is reviewed in Section 3.

Table 2 Summary of primary abbreviations and notations in this survey

Abbreviation	Full name	Notation	Explanation
NFG	Normal-form game	$[n]$	$:= \{1, 2, \dots, n\}$ for integer $n > 0$
II-EFG	Imperfect-information extensive-form game	$\mathbb{R}$	The set of real numbers
NE	Nash equilibrium	$\mathbb{R}^d$	d-dimensional real vector set
CE	Correlated equilibrium	$\mathbb{R}^d_+$	$\mathbb{R}^d$ with nonnegative elements
CCE	Coarse correlated equilibrium	$\Delta(S)$	$= \Delta^m := \{x \in \mathbb{R}^m_+ : \sum_{i=1}^m x_i = 1\}$ for a finite set S with m elements
TG	Team game		
TME	Team-maxmin equilibrium	$\lvert S \rvert$	The cardinality of a set S
SG	Stackelberg game	$\mathbb{P}$	The mathematical probability
GSG	General Stackelberg game	$\mathbb{E}$	The mathematical expectation
SSG	Stackelberg security game	$x^\top$	The transpose of x
SSE	Strong Stackelberg equilibrium	$\langle \cdot, \cdot \rangle$	The inner product
WSE	Weak Stackelberg equilibrium	$\mathbf{0}$	Vectors/matrices of all entries 0, fit dimension in the context (sometimes given as a subscript)
DG	Differential game		
TP-ZS-DG	Two-player zero-sum differential game		
ZSG	Zero-sum game	$\mathbf{1}$	Vectors/matrices of all entries 1, fit dimension in the context (sometimes given as a subscript)
SPP	Saddle point problem		
MP-ZSG	Multi-player zero-sum game		

Some prevailing algorithms are expounded in Section 4. An array of applications are presented in Section 5. Promising future research directions are discussed in Section 6. Finally, the conclusion is drawn in Section 7. In addition, the primary abbreviations and notations in this survey can be found in Table 2.

2 Models of adversarial games

This section presents the three main models of adversarial games, i.e., zero-sum normal-form and extensive-form games, Stackelberg (security) games, and DGs, along with the solution concepts for these game models.

2.1 Zero-sum normal-form and extensive-form games

Normal-form and extensive-form games are two widely employed game models that account for simultaneous or sequential actions committed by the players in a game.

Normal-form games (NFGs). A normal-form (or strategic-form) game is defined as a tuple (N, A, u) [4], where $N := [n]$ is a finite set of players. In the meantime, $A := A_1 \times \cdots \times A_n$ is the action profile set for all players, where A_i is the set of pure actions or strategies available to player $i \in [n]$ and $a = (a_1, \dots, a_n) \in A$ is a joint action profile. Moreover, $u := (u_1, \dots, u_n)$, where $u_i : A_i \to \mathbb{R}$ is a real-valued utility (or payoff) function for player i. A mixed strategy/policy for player i is a probability distribution over its action set A_i, defined as $\pi_i \in \Delta(A_i)$, and $\pi_i(a_i)$ is the probability of player i to commit an action $a_i \in A_i$. The expected utility $u_i(\pi_i, \pi_{-i})$ of player i can be expressed as $\mathbb{E}_{a \sim \pi}(u_i(a))$, where $\pi := (\pi_1, \dots, \pi_n)$ is the joint (mixed) action profile and π_{-i} is the joint action profile of all players, except player i. Similarly, a_{-i} is the joint (pure) action profile of all players, except player i, and u_i is expressed as $u_i(a_i, a_{-i})$ to indicate the dependency of a joint pure action profile. Social welfare is defined as $\mathrm{SW}(a) := \sum_{i=1}^n u_i(a)$ for the pure action profile $a \in A$ whose mixed strategy correspondence is expressed as $\mathrm{SW}(\pi) := \mathbb{E}_{a \sim \pi} \mathrm{SW}(a)$. Furthermore, the game is called constant-sum if for any action profile $a \in A$, it holds that $\sum_{i=1}^n u_i(a) = c_s$ for a constant c_s, and called zero-sum if $c_s = 0$, as illustrated in Figure 1.

Note that for the case with continuous action sets that are generally assumed closed and convex, the games are usually called continuous games.

In what follows, extensive-form games with imperfect information are introduced, which become extensive-form games with perfect information when the infoset of each player is a singleton [5].

Imperfect-information extensive-form games (II-EFGs). An II-EFG is defined as a tuple $(N, H, Z, A, P, \mu, u, I)$, where $N = [n]$ is a finite set of n players, H is a set of histories (i.e., nodes) representing the possible sequence of actions, and $Z \subseteq H$ denotes the set of terminal nodes that have no

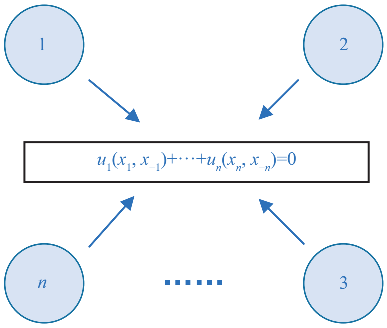

Figure 1 (Color online) A schematic illustration of zero-sum games with n players.

further actions and award a value to each player. Outside of N, a different "player" exists, defined as c, representing chance decisions. Moreover, the empty sequence $\emptyset$ is included in H, representing a unique root node. At a nonterminal node $h \in H$, $A(h) := \{a : (h, a) \in H\}$ is the action function assigning a set of available actions to h (here, $A(h)$ is different from A in NFGs, which should be clear from the context), and $P(h)$ is the player function assigning a player to node h, who takes an action at that node with $P(h) = c$ if chance determines the action at h. And $h \sqsubset h'$ means that h is led to h' by a sequence of actions; i.e., h is a prefix of h'. $u = (u_1, \ldots, u_n)$ is a set of utility functions, where $u_i : Z \to \mathbb{R}$ is the utility function of player i. If there is a constant c_s such that $\sum_{i=1}^{n} u_i(z) = c_s$ for all $z \in Z$, then the game is called a constant-sum game. If $c_s = 0$, then the game is called a zero-sum game.

The main feature, "imperfect information", is represented by the infosets of all players [5]. Specifically, $I = (I_1, \ldots, I_n)$ is a set of infosets, where I_i is a partition of $H_i := \{h : P(h) = i, h \in H\}$ satisfying $A(h_1) = A(h_2)$ and $P(h_1) = P(h_2)$ for any $h_1, h_2 \in I_{i,j}$ for some $I_{i,j} \in I_i$. That is, all nodes in the same infoset of I_i are indistinguishable to player i. Note that each node $h \in H$ is only in one infoset for each player. When all players can remember all historical information, it is called perfect recall. Formally, let h, h', g, g' be histories such that $h \sqsubset h', g \sqsubset g'$; then, perfect recall means that if g and h do not share an infoset and each is not a prefix of the other; then h' and g' also do not share an infoset.

A normal-form plan (or pure strategy) of player i is a tuple $a_i \in \Xi_i := \times_{I_{i,j} \in I_i} A(I_{i,j})$, which assigns an action to each infoset of player i. A normal-form strategy x_i means a probability distribution over Ξ_i, i.e., $x_i \in \Delta(\Xi_i)$. A behavioral strategy π_i (or simply, strategy) is a probability distribution over $A(I_{i,j})$ for each infoset of player i. A joint strategy profile π is composed of all players' strategies $\pi_i, i \in [n]$, i.e., $\pi = (\pi_1, \ldots, \pi_n)$, with π_{-i} representing all the strategies except π_i. Let $p(I_{i,j}, a)$ (or $p(h, a)$) denote the probability of a specific action a at infoset $I_{i,j}$ and $p^{\pi}(h)$ denote the reach probability of history h if all of the players select their actions according to π. For a strategy profile, player i has the total expected payoff of $u_i(\pi) = \sum_{h \in Z} p^{\pi}(h) u_i(h)$. Let Σ_i denote the set of all possible strategies for player i.

The best response of player i to π_{-i} is strategy $\mathrm{BR}(\pi_{-i}) := \arg\max_{\pi_i'} u_i(\pi_i', \pi_{-i})$. In a two-player zero-sum game, the exploitability $e(\pi_i)$ of strategy π_i, defined as $e(\pi_i) := u_i(\pi_i^*, \pi_{-i}^*) - u_i(\pi_i, \mathrm{BR}(\pi_i))$, where (π_i^*, π_{-i}^*) is an NE, which will be defined subsequently. In multiplayer games, the total exploitability (or NashConv) of strategy profile π is defined as [26] $e(\pi_i) := \sum_{i \in [n]} u_i(\mathrm{BR}(\pi_{-i}), \pi_{-i}) - u_i(\pi_i, \pi_{-i})$, and the average exploitability (or exploitability) is defined as $e(\pi)/|N|$, which is leveraged to measure how much can be gained by unilaterally deviating to their best response, generally interpreted as the distance from an NE.

Note that in addition to the above normal-form and extensive-form games, other classes of games may also be conducive in adversarial games, such as Markov games (or stochastic games) [27], where the game state changes according to a transition probability based on the current game state and actions of the players, and Bayesian games [28], which model game uncertainties with incomplete information.

In what follows, some solution concepts for related games are introduced.

The NE is the most widely adopted notion in [2].

Definition 1 (ϵ-Nash equilibrium (ϵ-NE))**.** For both normal-form and extensive-form games, a strategy $\pi = (\pi_1^*, \ldots, \pi_n^*)$ is called an ϵ-NE for a constant $\epsilon \geqslant 0$ if

$$u_i(\pi_i^*, \pi_{-i}^*) \geqslant u_i(\pi_i, \pi_{-i}^*) - \epsilon, \quad \forall \pi_i,\ i \in N, \tag{1}$$

that is, the gain is at most ϵ if any player changes his/her own strategy only. Moreover, it is called an NE when $\epsilon = 0$, that is, π_i^* is the best response of π_{-i}^* for any player $i \in [n]$, i.e., $\pi_i^* = \mathrm{BR}(\pi_{-i}^*), \forall i \in [n]$.

It is well known that at least one NE exists in mixed strategies for games with a finite number of players and a finite number of pure strategies for each player [2].

Even though NE may exist for many games and it is computationally efficient for two-player zero-sum games, it is well known by complexity theory that approximating an NE in κ-player ($\kappa \geqslant 3$) zero-sum games and even two-player nonzero-sum games are computationally hard, that is, it is PPAD-complete for general games [29–31]. As an alternative, (coarse) correlated equilibrium is often considered for NFGs in the literature, which is efficiently computable in all NFGs, as defined in the subsequent definition [32].

Definition 2 (ϵ-correlated equilibrium (ϵ-CE))**.** For the NFG (N, A, u), an ϵ-CE is a probability distribution μ over $\times_{i\in[n]} A_i$ if for each player $i \in [n]$ and any swap function $\phi_i : A_i \to A_i$ (usually called strategy modification),

$$\mathbb{E}_{a\sim\mu}[u_i(a_i, a_{-i})] \geqslant \mathbb{E}_{a\sim\mu}[u_i(\phi_i(a_i), a_{-i})] - \epsilon. \tag{2}$$

That is, no player can gain more payoff by unilaterally deviating its action that is privately informed by a coordinator, who samples a joint action $a = (a_1, \ldots, a_n)$ from that distribution. Furthermore, another relevant notion is defined below [33].

Definition 3 (ϵ-coarse correlated equilibrium (ϵ-CCE))**.** For the NFG (N, A, u), an ϵ-CCE is a probability distribution μ over $\times_{i\in[n]} A_i$ if for each player $i \in [n]$ and all actions $a'_i \in A_i$,

$$\mathbb{E}_{a\sim\mu}[u_i(a_i, a_{-i})] \geqslant \mathbb{E}_{a\sim\mu}[u_i(a'_i, a_{-i})] - \epsilon. \tag{3}$$

Except for the removal of the conditioning on the action a_i, the above condition is nearly the same as that for ϵ-CE by arbitrarily selecting an action a'_i on their own, instead of following the action a_i advised by the coordinator. For NE, CE, and CCE, it is known that they are payoff equivalent to each other in two-player zero-sum games by the minimax theorem [34]. Recently, the notions of CE and CCE have been extended to extensive-form games in [35,36], which, however, have been less investigated until now.

In an II-EFG, let us consider the case where all of the players in $\mathtt{T} := \{1, \ldots, n-1\}$ are cooperative, thus forming a team, who take actions independently and play against an adversary n, and $u_i = u_j, \forall i, j \in \mathtt{T}$ and $u_n = -u_{\mathtt{T}} = -\sum_{i\in\mathtt{T}} u_i$, called a zero-sum single-team single-adversary extensive-form team game (TG, or simply zero-sum TG) [37]. Before introducing the notion of TME, it is necessary to first prepare some essentials. Let S_i denote the set of action sequences of player i, where an action sequence of player i, defined by a node $h \in H$, is the ordered set of actions of player i that are on the path from the root to h. Let $\emptyset$ be the dummy sequence to the root. A realization plan $r_i : S_i \to [0, 1]$ is a function mapping each action sequence to a probability, satisfying the following expression:

$$\begin{aligned} & r_i(\emptyset) = 1, \\ & \sum_{a\in A(I_{i,j})} r_i(s_i, a) = r_i(s_i), \quad \forall I_{i,j} \in I_i, \ s_i = \mathrm{seq}_i(I_{i,j}), \\ & r_i(s'_i) \geqslant 0, \quad \forall s'_i \in S_i, \end{aligned} \tag{4}$$

where $\mathrm{seq}_i(I_{i,j})$ denotes the action sequence leading to $I_{i,j}$.

With the aforementioned preparations, the TME, first introduced in [38], is defined as follows [37].

Definition 4 (Team-maxmin equilibrium)**.** A TME is defined as follows:

$$\arg \max_{r_1,\ldots,r_{n-1}} \min_{r_n} \sum_{s=\times_{i\in N} s_i, s\in S} U_{\mathtt{T}}(s) \prod_{i=1}^{n} r_i(s_i), \tag{5}$$

where $U_{\mathtt{T}}$ is the team's utility defined as $U_{\mathtt{T}}(s) := \sum_{l\in Z'} u_{\mathtt{T}}(l)c(l)$ if at least one terminal node is achieved by the joint plan s (i.e., $Z' \subseteq Z$ is nonempty) with the chance $c(l)$ determined by chance nodes, and $U_{\mathtt{T}}(s) = 0$ otherwise.

The TME is generally unique, and it is an NE that maximizes the team's utility. In addition, the concept of ϵ-TME can be similarly defined, at which both the team and the adversary can gain at most ϵ if any player unilaterally changes its strategy.

Besides the aforementioned optimal strategy concepts, it is worth noting that there exist other notions, such as subgame perfect NE [4] and α-rank [39], which, however, are beyond this survey.

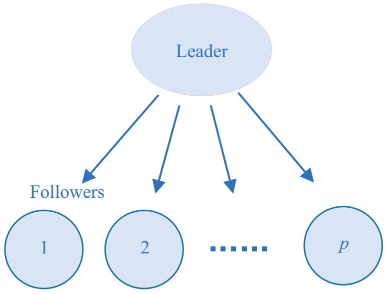

Figure 2 (Color online) Schematic illustration of GSGs, where directed edges mean that the leader first commits an action and the followers then play actions in response to the leader's action.

2.2 Stackelberg games

SGs (or leader-follower games) can date back to the Stackelberg competition introduced in [40] to model a strategic game between two firms, i.e., the leader and the follower, where the leader can take action first. SGs, as games with sequential actions and asymmetric information, have many practical applications, e.g., PROTECT, a system that the United States Coast Guard utilizes to assign patrols in Boston, New York, and Los Angeles [41], and ARMOR, an assistant deployed in the Los Angeles International Airport in 2007 to randomly schedule checkpoints on the roadways entering the airport. In what follows, general SGs and SSGs [42] are introduced, where SSGs are an important special case of general SGs.

General Stackelberg game (GSG). A GSG consists of a leader, who commits an action first, and p followers who observe and learn the leader's strategy and take actions in response to the leader's strategy (see Figure 2). Let F, A_l, and A_f be the sets of p followers, the leader's pure strategies, and each follower's pure strategies, respectively. The leader knows the probability of facing follower $k \in F$, denoted by $\varpi^k \in [0,1]$. Let $x \in \Delta(A_l)$ denote the mixed strategy of the leader, where the i-th component x_i represents the probability that the leader chooses the i-th pure strategy. Let $q_j^k \in \{0,1\}$ denote the decision of follower $k \in F$ to take a pure strategy $j \in A_f$ such that $\sum_{j\in A_f} q_j^k = 1$ for all $k \in F$. Note that only considering the pure strategies is enough for rational followers [43]. For the leader and each follower $k \in F$, the utilities (or payoffs/rewards) of the leader and follower are captured by a pair of matrices (R^k, C^k), where $R^k \in \mathbb{R}^{|A_l|\times|A_f|}$ is the utility matrix of the leader when facing follower k and $C^k \in \mathbb{R}^{|A_l|\times|A_f|}$ is the utility matrix of follower $k \in F$. Then, the expected utilities of the leader and follower k can be, respectively, expressed as

$$U_l(x,q) = \sum_{i\in A_l}\sum_{j\in A_f}\sum_{k\in F} \varpi^k x_i q_j^k R_{ij}^k, \tag{6}$$

$$U_f^k(x,q^k) = \sum_{i\in A_l}\sum_{j\in A_f} x_i q_j^k C_{ij}^k, \tag{7}$$

where $q := (q^1,\ldots,q^{|F|})$ and $q^k := (q_1^k,\ldots,q_{|A_f|}^k)$ for each $k \in F$.

SSG. In SSG, as a specific case of GSG, the leader and followers are viewed as the defender and attackers, respectively, where the defender aims to schedule a limited number of m security resources to protect (or cover) a subset of n targets from the attackers' attacks, with $m < n$. The definitions of the notations $F, A_l, A_f, \varpi^k, x, q_j^k$ are the same as those in the GSG. Note that $|A_f| = n$ in this case, the leader's pure strategy set A_l is now composed of all possible subsets of at most m targets that can be safeguarded simultaneously, and $q_j^k \in \{0,1\}$ indicates whether attacker $k \in F$ attacks target $j \in [n]$. Let $c_j \in [0,1]$ be the probability of coverage of target $j \in [n]$ such that $c_j = \sum_{i\in A_l, j\in i} x_i$, where $j \in i$ connotes that target j is covered by pure strategy i. When facing attacker $k \in F$, who attacks target $j \in [n]$, the defender's utility is $D_c^k(j)$ if the target is covered or protected, or $D_u^k(j)$ if the target is uncovered or unprotected. The utility of attacker $k \in F$ is $A_c^k(j)$ when attacking target j that is covered, or $A_u^k(j)$ when attacking target j that is uncovered. It is generally assumed that $D_c^k(j) \geqslant D_u^k(j)$ and $A_u^k(j) \geqslant A_c^k(j)$, are in line with common sense. The expected utilities for the defender and attacker $k \in F$ are, respectively, expressed as

$$U_d(x,q) = \sum_{j\in A_f}\sum_{k\in F} \varpi^k q_j^k [c_j D_c^k(j) + (1-c_j) D_u^k(j)], \tag{8}$$

$$U_a^k(x, q^k) = \sum_{j \in A_f} q_j^k [c_j A_c^k(j) + (1 - c_j) A_u^k(j)]. \tag{9}$$

The most widely adopted solution for GSG and SSG is the so-called SSE, which always exists in all SGs [42, 44]. Recall that it is enough for each follower to play pure strategies.

Definition 5 (Strong Stackelberg equilibrium). The strategy profile $(x^*, \{q^{k*}\}_{k \in F})$ for a GSG forms an SSE, if

(1) x^* is optimal for the leader,

$$(x^*)^\top R^k q^{k*} \geqslant x^\top R^k \mathcal{R}^k(x), \ \forall x \in \Delta(A_l), \ \forall k \in F,$$

where $\mathcal{R}^k(x)$ denotes attacker k's best response against x;

(2) Each follower k always plays a best-response, i.e.,

$$(x^*)^\top C^k q^{k*} \geqslant (x^*)^\top C^k q^k, \quad \forall \text{ feasible } q^k;$$

(3) Each follower k breaks ties in favor of the leader:

$$(x^*)^\top R^k q^{k*} \geqslant (x^*)^\top R^k \mathcal{R}^k(x^*), \quad \forall k \in F.$$

The tie-breaking rule is reasonable in cases of indifference because the leader can often induce favorable equilibrium by choosing a strategy arbitrarily close to the equilibrium that makes the follower prefer the desired strategy [45]. When the tie-breaking rule is in favor of the followers, the equilibrium is called weak Stackelberg equilibrium (WSE), which, however, does not always exist [46]. Moreover, the concept of SSE can be similarly defined for SSGs.

2.3 Zero-sum differential games

DGs, also known as dynamic games [46], are a natural extension of sequential games to continuous-time scenarios, which are expressed as differential equations and first introduced by Isaacs [47]. DGs can be regarded as an extension of optimal control [48], which usually has a single decision maker with a single objective function, whereas multiple players are involved in a DG with noncooperative objectives. Because this survey is concerned with adversarial games, zero-sum DGs (mostly involving two players in the literature) are considered here, although many other types of DGs emerge in the literature, including nonzero-sum DGs, mean-field games, differential graphical games, and Dynkin games [49, 50].

A two-player zero-sum differential game (TP-ZS-DG) is described as a dynamical system and expressed as follows:

$$\begin{aligned} &\dot{z}(t) = f(t, z(t), u(t), v(t)), \quad t \in [t_0, T], \\ &z(t_0) = z_0, \quad u(t) \in U, \quad v(t) \in V, \end{aligned} \tag{10}$$

where $z(t) \in \mathbb{R}^d$ is the state vector at time t, t_0 is the initial time, z_0 is the initial state, $U \subseteq \mathbb{R}^{m_1}$, $V \subseteq \mathbb{R}^{m_2}$ are control constraints for players 1 and 2, respectively, $u(t)$ and $v(t)$ are control actions (or signals) for players 1 and 2, respectively, and $f : [0, T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}^d$ is the dynamics, as illustrated in Figure 3.

For different setups in the literature, distinct cost functions are generally employed, most of which, however, are either based on or variants of an essential cost function, as given below:

$$J(u(\cdot), v(\cdot)) = \int_{t_0}^{T} f_0(t, z(t), u(t), v(t)) \mathrm{d}t + \phi(z(T)), \tag{11}$$

where $f_0 : [0, T] \times \mathbb{R}^d \times U \times V \to \mathbb{R}$ is the running cost (or stage cost) and $\phi : \mathbb{R}^d \to \mathbb{R}$ is the terminal cost (or final cost).

With (11), the goal of a DG (10) is for player 1 to minimize cost J and for player 2 to maximize cost J, i.e.,

$$\min_{u(\cdot) \in U} \max_{v(\cdot) \in V} J(u(\cdot), v(\cdot)). \tag{12}$$

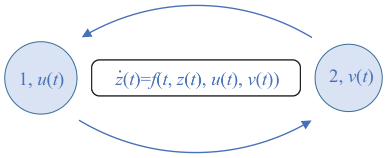

Figure 3 (Color online) Schematic illustration of two-player differential games.

Table 3 Summary of the solution concepts and complexity of adversarial games

Game models	Main solution concepts	Complexity of solutions
Zero-sum NFG Zero-sum EFG	NE	Polynomial-time (two-player ZSG); PPAD-complete (general games) [29–31]
	CE, CCE	Polynomial-time [55]
Zero-sum TG	TME	FNP-hard [37]
GSG and SSG	SSE	See Table 5
Zero-sum DG	NE	Computationally intractable in general [46]

Table 4 Main features of adversarial games

Game models	Player numbers	Action order	Information	Dynamics
Zero-sum NFG	$\geqslant 2$	Mostly simultaneous	Symmetric	Discrete-time continuous-time
Zero-sum EFG	$\geqslant 2$ (perfect information) mostly 2 (imperfect information)	Sequential	Symmetric	Mostly discrete-time
Zero-sum TG	$\geqslant 2$	Simultaneous, sequential	Symmetric	Discrete-time
GSG and SSG	Mostly one-leader p-follower	Sequential	Asymmetric	Mostly discrete-time
Zero-sum DG	Mostly 2	Mostly simultaneous	Mostly symmetric	Continuous-time

For (12), the optimal cost of J is called the value of the game, expressed as value function $\psi(t,z)$. Moreover, the solution notion is still the NE as in zero-sum normal-form and extensive-form games, also called minimax equilibrium (or minimax/saddle point) in the literature because the studied problem is, in fact, a saddle point game (or saddle point problem/optimization).

Note that for multi-player DGs [51, 52], say N players, the system (10) can be generally written as $\dot{z}(t) = f(t, z(t), u_1(t), \ldots, u_N(t))$, where u_i is the control action of player $i \in [N]$. Meanwhile, player i's cost function is of the form $J_i(u_i(\cdot), u_{-i}(\cdot)) = \int_{t_0}^{T} f_i(t, z(t), u_i(t), u_{-i}(t))\mathrm{d}t + \phi_i(z(T))$, where $u_{-i} := (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N)$ (i.e., all control actions, except u_i). The game is called zero-sum if $\sum_{i=1}^{N} J_i(u_i, u_{-i}) \equiv 0$ for all control actions. Moreover, it can be observed that dynamics (10) is deterministic. In the meantime, stochastic DGs have also been addressed in the literature (e.g., [53, 54]) and expressed as stochastic differential equations with the standard Brownian motion [49]. It is also noteworthy that the above DGs are usually studied under a set of assumptions, such as the compactness of U, V and the Lipschitz continuity of f, f_0, ϕ, among others [50].

Finally, the solution concepts and their complexity are epitomized in Table 3 [46, 55], and the main features of the aforementioned games are summarized in Table 4.

3 Research classification and frontiers

This section aims to succinctly summarize the relevant literature for zero-sum games, GSGs, SSGs, and TP-ZS-DGs along with the emerging state-of-the-art research.

3.1 Zero-sum games (ZSGs)

Both normal-form and extensive-form ZSGs investigated in the literature can be generally categorized into the following main aspects: bilinear games, SPPs, multi-player ZSGs, TGs, and imperfect-information ZSGs (II-ZSGs), as discussed below.

(1) Bilinear games. Bilinear games are simple models for delineating two-player games, generally in normal-form as [56] minimizing utilities $x^\top Ay$ and $x^\top By$ for players 1 and 2, respectively, where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ are payoff matrices, subject to strategy sets $x \in X := \{x \in \mathbb{R}^m : R_1x = r_1, x \in \mathbb{R}_+^m\}$ and $y \in Y := \{y \in \mathbb{R}^n : R_2y = r_2, y \in \mathbb{R}_+^n\}$ with some $R_1 \in \mathbb{R}^{k_1\times m}, R_2 \in \mathbb{R}^{k_2\times n}$,

and $r_1 \in \mathbb{R}^{k_1}, r_2 \in \mathbb{R}^{k_2}$. A bilinear game is usually defined as the payoff matrix pair (A, B), which is zero-sum when $B = -A$, and as an important notion, the rank of game (A, B) is defined as the rank of matrix $A + B$. Several interesting games can be viewed as special cases of bilinear games, such as bimatrix games [57–59], where $R_1 = \mathbf{1}_m^\top, R_2 = \mathbf{1}_n^\top$, and $r_1 = r_2 = 1$, imitation games (a special case of bimatrix games with $B = I$) [60], and the Colonel Blotto game (i.e., two colonels simultaneously allocate their troops across different battlefields) [61]. In addition, multi-player polymatrix games [62] can be equivalently transformed into bilinear games [56]. Generally speaking, the existing literature mainly focuses on the computational complexity and polynomial-time algorithm design for approximating the NE of bilinear games [63], bimatrix games [64], polymatrix games [65], and the Colonel Blotto game [66]. Recently, it is shown that the computation of the NE in two-player nonzero-sum games with rank $\geqslant 2$ is PPAD-hard [67, 68]. And computing a $1/n_s^c$-approximate NE is PPAD-hard even for imitation games for any $c > 0$ [60], where n_s is the number of moves available to the players, and a polynomial-time algorithm was developed for finding an approximate NE in [60]. Also, computing an NE in a tree polymatrix game with 20 actions per player is PPAD-hard [65], and a polynomial-time algorithm for 1/3-approximate NE in bimatrix games was proposed in [64], which is the state-of-the-art in the literature. For the Colonel Blotto game, efficient and simple algorithms have been recently provided in [69–71], and meanwhile, various scenarios have been extended for this game, including the dynamic Colonel Blotto game [72], generalized Colonel Blotto and generalized lottery Blotto games [73], and multi-player cases [71, 74]. Furthermore, bilinear games are generalized to hidden bilinear games in [75], where the inputs controlled by players are processed by a smooth function, i.e., a hidden layer, before coming into the conventional bilinear games.

(2) Saddle point problems (SPPs). SPPs are also called saddle point optimization, min-max/minimax games, or min-max/minimax optimization in the literature. The formulation of a general SPP [76] is given as $\min_{x \in X} \max_{y \in Y} f(x, y)$, where $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$ are closed and convex, possibly the entire Euclidean spaces or their compact subsets. For general SPPs, besides zero-sum bilinear games, two other types, i.e., non-bilinear and bilinear SPPs, have been extensively considered. A non-bilinear SSP [77, 78] is expressed as $\min_{x \in X} \max_{y \in Y} f(x) + \Theta(x, y) - h(y)$, where Θ is a general coupling function, and as a special case, when $\Theta(x, y) = x^\top C y$ with $C \in \mathbb{R}^{m \times n}$, the game is called a bilinear SPP [79–81] due to the bilinear coupling. The existing research mainly centers on equilibrium existence, computational and sampling complexity, and efficient algorithm design, for instance, as done in the aforementioned recent studies. Meanwhile, various scenarios have been investigated in the literature, including projection-free methods by applying the Frank-Wolfe algorithm [82, 83], nonconvex-nonconcave general SPPs [84, 85], linear last-iterate convergence [86], SSPs with adversarial bandits and delays [87], periodic zero-sum bimatrix games with continuous strategy spaces [88], compositional SSPs [89], decentralized setup [90], functional-form games [91], and hidden general SPPs [92], where the controlled inputs are first fed into smooth functions whose outputs are then treated as inputs for the traditional general SPPs. Finally, it is noteworthy that the general SPPs with sequential actions have also been studied, called min-max Stackelberg games, e.g., the recent work presented in [93] with dependent feasible sets.

(3) Multi-player zero-sum games (MP-ZSGs). The previously discussed games usually involve two players. It is well known that approximating an NE in multi-player zero-sum games and even two-player nonzero-sum games is PPAD-complete [29–31]. Moreover, it is known that multi-player symmetric zero-sum games might have only asymmetric equilibria, which is consistent with that of two-player and multi-player symmetric nonzero-sum games, but in contrast with the case in two-player symmetric zero-sum games that always have symmetric equilibria (if equilibria exist) [94]. In the literature, most of the studies focus on multi-player zero-sum polymatrix games (also called network matrix games in some studies), where the utility of each player is composed of the sum of utilities gained by playing with its neighbors in an undirected graph [62]. In [95], the authors generalized von Neumann's minimax theorem to multi-player zero-sum polymatrix games, thus, implying convexity of equilibria, polynomial-time tractability, and convergence of no-regret learning algorithms to NEs, and last-iterate convergence was studied in [96] for multi-player polymatrix zero-sum games. $O(1/T)$ time-average convergence was established using alternating gradient descent in [97], where T is the time horizon. Moreover, it is shown that for continuous-time algorithms, time-average convergence may fail even in a simple periodic multi-player zero-sum polymatrix game or replicator dynamics, but is Poincaré recurrent in [98, 99]. Furthermore, it is realized that mutual cooperations among players may benefit more than pursuing selfish exploitability, and in this case, team/alliance formation is also studied in the literature, e.g., [100], where it was demonstrated that team formation may be seen as a social dilemma. Additionally, other

pertinent research encompasses multi-player general-sum games [101–103] and machine learning based studies [104].

(4) Team games. Generically, TGs refer to those games where at least one team exists with the cooperation of team members with communications either before the play, or during the play, or simultaneously before and during the play, or without any communications [37]. In general, TGs in the literature can be classified from two perspectives. One perspective depends on the team number, i.e., one-team games (or adversarial TGs) [105], where the players in the team enjoying the same utility function play against an adversary independently, and two-team games [106] consisting of two teams in a game. The other perspective is on perfect-information and imperfect-information games. For TGs, TME is an important solution concept, for which it is known that computing a TME is FNP-hard and inapproximable in the additive sense [107, 108]. Even though, efficient algorithms for computing a TME in perfect-information zero-sum NFGs have been developed until now, e.g., [105]. Meanwhile, a class of zero-sum two-team games in perfect-information normal-form was studied in [106], where finding an NE is shown to be CLS-hard, i.e., unlikely to have a polynomial-time NE computing algorithm. Moreover, as two-team games, two-network zero-sum games are also addressed, where each network is thought of as a team [109–111]. For imperfect-information zero-sum TGs, the researchers have investigated a variety of scenarios centering around computational complexity and efficient algorithms, such as one-team games [37, 112, 113], one-team games with two members in the team [114], and the computation of team correlated equilibrium in two-team games [115].

(5) Imperfect-information ZSGs. Unlike perfect-information games, such as Chess, Go, and backgammon, II-ZSGs, involving individual players' private information that is hidden to other players, are more challenging due to information privacy and uncertainty, especially for large games with large action spaces and/or infosets. For example, the game of heads-up (i.e., two-player) limit Texas Hold'em poker, with over 10^{14} infosets [116], has been a challenging problem for AI for over 10 years, before being essentially solved by Cepheus [117], the first computer program for handling imperfect information games that are played competitively by humans. Also, the game of no-limit Texas Hold'em poker has more than 10^{161} infosets [116], for which DeepStack [118] and Libratus [119] are the first line of AI agents/algorithms to defeat professional humans in heads-up no-limit Texas Hold'em poker. As such, most of the research focuses on the computation of NEs in two-player II-ZSGs in [120, 121], aiming to develop efficient superhuman AI agents in the face of the challenges of imperfect information, large models, and uncertainties. To handle large games with imperfect information, several techniques have been successively proposed, e.g., pruning, abstraction, and search [116, 122, 123]. Roughly speaking, pruning aims to avoid traversing the entire game tree for an algorithm while simultaneously ensuring the same convergence, including regret-based pruning, dynamic thresholding, and best-response pruning [124]. Abstraction aims to generate a smaller version of the original game by bucketing similar infosets or actions, while maintaining as much as possible the strategic features of the original game [125], mainly including information abstraction and action abstraction. Meanwhile, search aims to improve the (approximate) solution of a game abstraction, which may be far from the true solution of the original game, by seeking a more precise equilibrium solution for a faced subgame, such as depth-limited search [123, 126]. Moreover, it has been shown recently that some two-player poker games can be represented as perfect-information sequential Bayesian extensive games with efficient implementation [127]. Recently, in [128], the authors bridged several standing gaps between NFG and EFG learning by directly transferring desirable properties in NFGs to EFGs, simultaneously guaranteeing last-iterate convergence, lower dependence on the game size, and constant regret in games. Furthermore, bandit feedback is of practical importance in real-world applications of II-ZSGs [129, 130], where only the interactive trajectory and payoff of the reached terminal node can be observed without prior knowledge of the game, such as the tree structure, observation/state space, and transition probabilities (for Markov games) [131]. On the other hand, multi-player II-ZSGs are more challenging and thus have been less researched except for a handful of studies, e.g., Pluribus [132], the first multi-player poker agent, which has defeated top humans in six-player no-limit Texas Hold'em poker (the most prevalent poker in the world) [133], and other endeavors [130, 134–137]. Aside from deterministic methods, AI approaches have achieved great success in II-ZSGs based on reinforcement learning, deep neural networks, and so on [118, 131, 138–149], for instance, DeepStack [118] and Pluribus [132], to name a few. More details can refer to a recent survey of AI in games [150]. Note that other closely related studies subsume imperfect-information general-sum games with full and bandit feedback [151–153], two-player zero-sum Markov games [154], and multi-player general-sum Markov games [155].

It should be noted that incomplete information is also important in adversarial games, mainly com-

Table 5 Complexity results of a game with a single attacker [174]

SSGs	Size of schedule		
	1	2	$\geqslant 3$
Homogeneous resources	P	P	NP-hard
Heterogeneous resources	P	NP-hard	NP-hard

prising Bayesian games [25, 28]. Finally, it is worth pointing out that the theory of zero-sum games helps model or resolve a variety of deep learning problems and is potential to improve the results of deep learning models [156], e.g., polymatrix games in semi-supervised learning [157]. For applications of game theory in deep learning, interested readers can refer to a recent survey [156].

3.2 Stackelberg games

SGs are roughly summarized from four perspectives, i.e., GSGs, SSGs, continuous SGs, and incomplete-information SGs.

(1) GSGs. The research on GSGs mainly lies in three aspects, i.e., computational complexity, solution methods, and their applications. For computational complexity, when only having one follower in GSGs, it is known that the problem can be solved in polynomial time, while it is NP-hard in the case of multiple followers [43]. Regarding solution methods, there are an array of proposed methods in the literature, but primarily depending upon approaches for coping with linear programming (LP) and mixed integer linear programming (MILP), including cutting plane methods, enumerative methods, and hybrid methods, among others [158, 159]. Note that both GSGs and SSGs can be formulated as bilevel optimization problems [158, 159], where bilevel optimization has a hierarchical structure with two level optimizations, i.e., one lower-level optimization (follower) nested in another upper-level optimization (leader) as constraints, which is an active research area unto itself [160]. As for practical applications, a multitude of real-world problems have been tackled using SGs, such as economics [161], smart grid [162, 163], wireless networks [164], dynamic inspection problems [165], and industrial Internet of Things [166]. It should be noted that other relevant cases have also been studied in the literature, such as multi-leader cases [167–171], cases with bounded rationality [172], and general-sum games [173].

(2) SSGs. In general, SSGs can be classified based on the functionality of security resources. Specifically, when every resource is capable of protecting every target, it is called homogeneous resources, and when resources are restricted to protecting only some subset of targets, it is called heterogeneous resources [174]. Meanwhile, resources can also be distinguished by how many targets they are able to cover simultaneously, and in this case, a notion, called schedule, is assigned to a resource with the size of the schedule being defined as the number of targets that can be simultaneously covered by the resource [174], including the case with size 1 [175] and greater than 1 [176]. For these scenarios, the computational complexity was addressed in [174] when having a single attacker, as shown in Table 5. With regard to the solution methods, similar methods for solving GSGs can be applied to handle SSGs. Moreover, the practical applications of SSGs encompass wildlife protection [177], passenger screening at airports [178], crime prevention [179], cybersecurity [180], information security [181], and border patrol [182, 183]. In the meantime, there are other scenarios addressed in the literature, such as multi-defender cases [184, 185], Bayesian generalizations [186], and cases with bounded rationality [187] and ambiguous information [188].

(3) Continuous SGs. This game is an SG with continuous strategy spaces. In general, continuous SGs have two players, i.e., a leader and a follower, who have cost functions $f_1 : \Omega \to \mathbb{R}$ and $f_2 : \Omega \to \mathbb{R}$ with $\Omega := X \times Y$, respectively, where $X \subseteq \mathbb{R}^{d_1}$ and $Y \subseteq \mathbb{R}^{d_2}$ are closed convex and possibly compact strategy sets for the leader and follower, respectively. Then, the problem can be formally written as

$$\min_{x\in X} \left\{ f_1(x,y) : y \in \arg\min_{y\in Y} f_2(x,y) \right\}, \tag{13}$$

where the follower still takes action in response to the leader after the leader makes a decision. In this case, strategy $x^* \in X$ of the leader is called a Stackelberg equilibrium strategy [189] if

$$\sup_{y\in \mathrm{BR}(x^*)} f_1(x^*,y) \leqslant \sup_{y\in \mathrm{BR}(x)} f_1(x,y), \ \forall x \in X_1, \tag{14}$$

where $\mathrm{BR}(x) = \{y \in Y : f_2(x,y) \leqslant f_2(x,y'), \forall y' \in Y\}$ is the best response of the follower to x. Along this line, a hierarchical Stackelberg v/s Stackelberg game was studied in [190], where the first general existence result for the games' equilibria is established without positing the single-valuedness assumption on the equilibrium of the follower-level game. Furthermore, the connections between NE and Stackelberg equilibrium were addressed in [189], where convergent learning dynamics is also proposed using Stackelberg gradient dynamics that can be regarded as a sequential variant of the conventional gradient descent algorithm, and both zero-sum and general-sum games are considered therein. Additionally, as a special case of the aforementioned game (13), min-max Stackelberg games are also considered, where the problem is of the form $\min_{x\in X}\max_{y\in Y} f(x,y)$ with $f:\Omega\to\mathbb{R}$ being the cost function. This problem has been investigated in the literature, especially for the case with dependent strategy sets [93,191]; i.e., inequality constraints $g(x,y)\geqslant \mathbf{0}$ are imposed for the follower for some function $g:\Omega\to\mathbb{R}$, for which the prominent minimax theorem [34] does not hold anymore.

(4) Incomplete-information SGs. Incomplete information means that the leader can only access partial information or cannot access any information about the followers' utility functions, moves, or behavior [10]. This is in contrast to the traditional SGs, where the followers' information is available to the leader [42]. Motivated by practical applications, this weak scenario has been extensively considered in recent years. For example, the authors in [192] studied situations in which only partial information on the attacker behavior can be observed by the leader. And a single-leader-multiple-followers SSG was considered in [193] with two types of misinformed information, i.e., misperception and deception, for which a stability criterion is provided for both strategic stability and cognitive stability of equilibria based on hyper NE. Additionally, one of the interesting directions is information deception of the follower, that is, the follower is inclined to deceive the leader by sending misinformation, such as fake payoffs, to benefit himself/herself as much as possible, while, at the same time, the leader needs to distinguish deceptive information to minimize its loss incurred by the deception. Recently, an interesting result on the nexus between the follower's deception and the leader's maximin utility is obtained to optimally deceive the leader in [194], that is, through deception, almost any (fake) Stackelberg equilibrium can be induced by the follower if and only if the leader procures at least their maximin utility at this equilibrium.

Finally, it is worthwhile to note that SGs play an important role in AI and machine learning. For example, SSGs are leveraged to deal with the security issue in AI [195], incomplete-information SSGs are employed to design an incentive mechanism in distributed machine learning [196], and robust/multi-agent reinforcement learning (MARL) is addressed by resorting to GSGs [197–200].

3.3 Zero-sum differential games

According to the existing literature, zero-sum DGs are categorized into five main dimensions, which, however, are not mutually exclusive, but from different angles of studied problems, i.e., linear-quadratic DGs, DGs with nonlinear dynamical systems, Stackelberg DGs, stochastic DGs, and terminal time and state constraints.

(1) Linear-quadratic DGs. This relatively simple model has been widely studied for DGs, where dynamical systems are linear differential equations and cost functions are quadratic [201,202]. In general, linear-quadratic DGs are analytically and numerically solvable, which can find a variety of real-world applications, such as pursuit-evasion problems [203,204]. Recently, singular linear-quadratic DGs were studied in [205], which cannot be handled either using the Isaacs MinMax principle or the Bellman-Isaacs equation approach. To solve this problem, an interception DG was introduced with the appropriate regularized cost function and dual representation. The authors in [206] studied a linear-quadratic-Gaussian asset defending DG where the state information of the attacker and defender is inaccessible to each other, but the trajectory of a moving asset is known by them. Meanwhile, a two-player linear-quadratic-Gaussian pursuit-evasion DG was investigated in [207] with partial information and selected observations, where the state of one player can be observed any time preferred by the other player and the cost function of each player consists of the direct cost of observing and the implicit cost of exposing his/her state. Two-player mean-field linear-quadratic stochastic DGs in an infinite horizon were investigated in [208], where the existence of both open-loop and closed-loop saddle points is analyzed by resorting to coupled generalized algebraic Riccati equations. A linear-quadratic DG with two defenders and two attackers against a stationary target was considered in [209], and multi-player zero-sum DGs were addressed using the neural-network-based synchronous iteration learning approach [210].

(2) Nonlinear DGs. DGs with nonlinear state dynamics have also been considered in the literature, given that many practical applications cannot be dealt with by linear-quadratic DGs. For example, the authors in [211] considered a class of nonlinear TP-ZS-DGs by employing an adaptive dynamic

programming. TP-ZS-DGs were addressed in [212] by proposing an approximate optimal critic learning algorithm based on policy iteration of a single neural network. Nonlinear DGs were also considered with time delays [213–215] and fractional-order systems [216] and investigated in [217] with the dynamical system depending on the system's distribution and the random initial condition. Aside from TP-ZS-DGs, multi-player zero-sum DGs with uncertain nonlinear dynamics were also considered and tackled using a new iterative adaptive dynamic programming algorithm in [218]. Multi-player nonlinear general-sum DGs were studied by considering efficient iterative linear-quadratic approximations [219]. And multi-player DGs were also considered for pursuit-evasion problems [220, 221] and reach-avoid problems [222, 223].

(3) Stackelberg DGs. Motivated by the sequential actions in some practical applications, such as SGs, DGs with sequential actions, called Stackelberg DGs, have been broadly addressed in the literature. For instance, a linear-quadratic Stackelberg DG was considered in [224] with mixed deterministic and stochastic controls, where the follower can select adapted random processes as its controller. The Stackelberg DG was employed to fight terrorism in [225]. Then, the authors in [226] investigated two classes of state-constrained Stackelberg DGs with a nonzero running cost and state constraint, for which the Hamilton-Jacobi equations are established.

(4) Stochastic DGs. In many realistic problems, the dynamics of a concerned system may not be completely modeled, but may undergo some uncertainties and/or noises; thereby, stochastic differential equations have been leveraged to model the system dynamics in stochastic DGs [227, 228]. In this respect, the authors in [229] considered two-person zero-sum stochastic linear-quadratic DGs, along with the investigation of the open-loop saddle point and the open-loop lower and upper values. A class of stochastic DGs with ergodic payoff was studied in [230], where it is unnecessary for the diffusion system to be non-degenerate. In addition, linear-quadratic stochastic Stackelberg DGs were taken into consideration in [231] with asymmetric roles for players, [232] for jump-diffusion systems, [233] without the solvability assumption on the associated Riccati equations, and [234] with model uncertainty. And a Stackelberg stochastic DG with nonlinear dynamics and asymmetric noisy observation was addressed in [235].

(5) Terminal time and state constraints. A basic classification of zero-sum DGs can be made based on terminal time and state constraints, i.e., whether the terminal time is finite (including two cases, i.e., a fixed constant or a variable to be specified) or infinite and whether the system state is unconstrained or constrained. Along this line, the case with fixed terminal time and unconstrained state was addressed [236], and the state-constrained case with fixed terminal time was also studied [237]. Meanwhile, the case with the terminal time being a variable was investigated in the literature, such as [238] without state constraints and [239, 240] with state constraints but with zero running-cost. Recently, the case with nonzero state constraint and underdetermined terminal time was investigated in [226]. In addition to the above finite horizon cases, the infinite horizon case has also been considered in the literature, e.g., [208, 241].

Finally, it is worth pointing out that other possible forms of zero-sum DGs exist in the literature, such as the case with continuous and/or impulse controls [241], mean-field DGs [242, 243], and risk-sensitive zero-sum DGs [228].

4 Prevailing algorithms and approaches

This section aims at encapsulating some of the main efficient algorithms and approaches for handling the reviewed adversarial games, as discussed in Section 2.

4.1 Zero-sum normal-form and extensive-form games

The bundle of algorithms can be roughly divided into two parts according to their applicability to NFGs or II-EFGs.

For NFGs, a large number of algorithms have so far been proposed, e.g., regret matching (RM, first proposed by Hart and Mas-Colell in 2000 [244]), RM+ [245], fictitious play [246, 247], double oracle [248], and online double oracle [59], among others. Wherein, the most prevalent algorithms are based on regret learning, usually called no-regret (or sublinear) learning algorithms, depending on external and internal regrets in general, defined as follows.

The external and internal regrets [249] of each player $i \in [n]$ are, respectively, defined as

$$R_i^E := \max_{a_i \in A_i} \sum_{t=1}^{T} \big[u_i\big(a_i, \pi_{-i}^t\big) - u_i\big(\pi^t\big)\big], \tag{15}$$

$$R_i^I := \max_{a_i', a_i \in A_i} \sum_{t=1}^{T} \mathbf{1}_{a_i^t = a_i^A} \big[u_i\big(a_i', a_{-i}^t\big) - u_i\big(a^t\big)\big], \tag{16}$$

where the superscript t stands for the iteration number, T is the time horizon, and $\mathbf{1}_E$ is the indicator function for event E. In general, the external regret measures the greatest regret for not playing action a_i, and the internal regret indicates the greatest regret for not swapping to action a_i' each time the player actually plays action a_i^A. The weighted external and internal regrets are also defined by adding a weight at each time t [250]. Other regrets, including swap regret [102] and several dynamic/static NE-based regrets [21, 251–254], are also considered in the literature.

With regrets at hand, two of the most widely employed algorithms, i.e., optimistic (or predictive) follow the regularized leader (Optimistic FTRL) and optimistic mirror descent (OMD) [96], are, respectively, given as

$$x^{t+1} = \arg\max_{x \in \mathcal{X}} \left\{ \alpha \left\langle x, m^t + \sum_{\tau=1}^{t} g^t \right\rangle - R(x) \right\}, \tag{17}$$

and

$$\begin{aligned} x^{t+1} &= \arg\max_{x \in \mathcal{X}} \big\{ \alpha \langle x, m^t \rangle - D_R\big(x, \hat{x}^t\big) \big\}, \\ \hat{x}^{t+1} &= \arg\max_{\hat{x} \in \mathcal{X}} \big\{ \alpha \langle \hat{x}, g^t \rangle - D_R\big(\hat{x}, \hat{x}^t\big) \big\}, \end{aligned} \tag{18}$$

where $\mathcal{X}$ is a generic closed convex constraint set, $\alpha > 0$ is the stepsize, g^t is a subgradient of function f^t returned by the environment after the player commits an action at time t, m^t is a subgradient prediction, which is often assumed to be $m^t = g^t$ in the literature, and $R(x)$ is a strongly convex function, serving as the base function for defining the Bregman divergence $D_R(x, y) := R(x) - R(y) - \langle \nabla R(y), x - y \rangle$ for any $x, y \in \mathbb{R}^d$.

Note that many widely employed algorithms, such as optimistic gradient descent ascent (OGDA) [86] and optimistic multiplicative weights update (OMWU, or optimistic hedge) [255], are special cases or variants of Optimistic FTRL and OMD. Other different efficient algorithms, such as optimistic dual averaging (OptDA) [256] and greedy weights [250], also exist.

For imperfect-information games, the most well-known algorithm is counterfactual regret minimization (CFR) [257], whose details are introduced as follows, with the same notations as in EFGs of Subsection 2.1.

Recall that $p^\pi(h)$ denotes the reach probability of history h with strategy profile π. For an infoset $J \in I$, let $p^\pi(J)$ denote the probability of reaching the infoset J via all possible histories in J, i.e., $p^\pi(J) = \sum_{h \in J} p^\pi(h)$. And let $p_i^\pi(J)$ denote the reach probability of infoset J for player i according to the strategy π, i.e., $p_i^\pi(J) = \Pi_{J' \cdot a' \sqsubseteq J, P(J) = i} p(J', a')$, and $p_{-i}^\pi(J)$ denote the counterfactual reach probability of infoset J, i.e., the probability of reaching J with strategy profile π except that the probability of reaching J is treated as 1 by the current actions of player i, i.e., without the contribution of player i to reach J. Meanwhile, $p^\pi(h, z)$ denotes the probability of going from history h to a nonterminal node $z \in Z$. Then, for player $i \in [n]$, the counterfactual value at a nonterminal history h is defined as

$$\nu_i^\pi(h) := \sum_{z \in Z, h \sqsubset z} p_{-i}^\pi(h) p^\pi(h, z) u_i(z), \tag{19}$$

the counterfactual value of an infoset J is defined as

$$\nu_i^\pi(J) := \sum_{h \in J} \nu_i^\pi(h), \tag{20}$$

and the counterfactual value of an action a is defined as

$$\nu_i^\pi(J, a) := \sum_{h \in J} \left[p_{-i}^\pi(h) \sum_{z \in Z} p^\pi(h \cdot a, z) u_i(z) \right]. \tag{21}$$

The instantaneous regret at iteration t and counterfactual regret at iteration T for action a in infoset J are, respectively, defined as

$$r_i^t(J,a) := \nu_i^{\pi^t}(J,a) - \nu_i^{\pi^t}(J), \tag{22}$$

$$R_i^T(J,a) := \sum_{t=1}^{T} r_i^t(J,a), \tag{23}$$

where π^t is the joint strategy profile leveraged at iteration t.

By defining $R_i^{T,+}(J,a) := \max\{R_i^T(J,a), 0\}$, applying RM can generate the strategy update as

$$\pi_i^{T+1}(J,a) = \begin{cases} \frac{R_i^{T,+}(J,a)}{\zeta_i^T(J,a)}, & \text{if } \zeta_i^T(J,a) > 0, \\ \frac{1}{|A(J)|}, & \text{otherwise}, \end{cases} \tag{24}$$

with $\zeta_i^T(J,a) := \sum_{a\in A(J)} R_i^{T,+}(J,a)$, and Eq. (24) is the essential CFR method for player i's strategy selection. Moreover, it is known that the CFR method can guarantee the convergence to NEs for the average strategy of players, i.e.,

$$\bar{\pi}_i^T(J,a) := \frac{\sum_{t=1}^T p_i^{\pi^t}(J)\pi_i^t(J,a)}{\sum_{t=1}^T p_i^{\pi^t}(J)}, \quad \forall i \in [n]. \tag{25}$$

Hitherto, various famous variants of CFR with superior performance, including CFR+ [245, 258], discounted CFR (DCFR) [259], linear CFR (LCFR) [260], exponential CFR (ECFR) [261], and AutoCFR [262], have been developed. More details can be found in [18, 116, 263].

Meanwhile, many AI methods have been proposed in the literature [104], such as policy space response oracles (PSRO) [26, 264], neural fictitious self-play [138], deep CFR [260], single deep CFR [265], unified deep equilibrium finding (UDEF) [147], player of games (PoG) [144], and neural auto-curricula (NAC) [148]. Among these methods, PSRO, which unifies fictitious play and double oracle algorithms, has been an effective approach in recent years. Meanwhile, UDEF provides a unified framework for PSRO and CFR, which are generally considered independently with their own advantages. Thus, UDEF is superior to both PSRO and CFR, as demonstrated by experiments on Leduc poker [147]. The recently developed PoG algorithm has unified several previous approaches by integrating guided search, self-play learning, and game-theoretic reasoning. It has been demonstrated theoretically and experimentally the achievement of strong empirical performance in large perfect-information and imperfect-information games, which outperforms the state-of-the-art in heads-up no-limit Texas Hold'em poker (Slumbot) [144]. Moreover, NAC, as a meta-learning algorithm proposed recently in [148], provides a potential future direction to develop general MARL algorithms solely from data because it can learn its own objective solely from its interactions with the environment without the need for human-designed knowledge about game-theoretic principles, and it can decide by itself what the meta-solution, i.e., who to compete with, should be during training. Furthermore, it is shown that NAC is comparable or even superior to state-of-the-art population-based game solvers, such as PSRO, on a series of games, such as Games of Skill, differentiable Lotto, non-transitive Mixture Games, Iterated Matching Pennies, and Kuhn poker [148]. It should be noted that although the above AI methods have been developed to compute the solutions for various games, partial AI methods are inversely beneficial from the theory of some game models. For instance, game-theoretic reasoning, resulting from the computation of (approximate) minimax-optimal strategies, is one of the key ingredients in the design of PoG [144]. And the idea of minimizing each player's exploitability is crucial in developing MARL algorithms in the framework of NAC [148].

Finally, it is worth pointing out that the CFR methods can guarantee the convergence to NEs in terms of the empirical distribution (i.e., time-average) of play, but generally fail to converge for the day-to-day play (i.e., the last-iterate convergence) [266, 267], although it enables last-iterate convergence in two-player zero-sum games [96]. In this respect, last-iterate convergence is of also importance to be explored as demonstrated in economics, and so on [86, 96, 268–270].

4.2 Stackelberg games

GSGs and SSGs can be expressed as bilevel linear programming (BLP) or mixed integer linear programming (MILP), which can be further transformed or relaxed as linear programming (LP) [158]. As discussed in Subsection 3.2, solving GSGs and SSGs is generally NP-hard, and most existing solution methods are variants of solution approaches for MILP and LP, including cutting plane methods, enumerative methods, and hybrid methods [159]. Some of the most widely used approaches in the literature are as follows.

(1) Multiple LP approach. This approach was proposed in [43] and is most widely employed for those easy problems that can be solved in polynomial time, including the case with a single follower for GSGs [43]. It was further improved in [271] by merging LPs into a single MILP. This approach was also improved in [174] to deal with SSGs and was generally efficient in the case with size 1 of the schedule and the case with size 2 of the schedule but for homogeneous resources, as shown in Table 5.

(2) Benders decomposition. Benders decomposition method, which can effectively handle the general MILP problems, was developed in [272]. The crux of this method is to divide the original problem into two other problems, i.e., the master problem, by relaxing some constraints, and the subproblem, along with a separation problem that is the dual of the subproblem. Then, the solution seeking procedure involves solving the master problem first, followed by solving the separation problem, and finally, checking the feasibility and optimality conditions for the subproblem with different contingent operations. Moreover, this approach can be improved by combining with other techniques, such as Farkas' lemma [273] and normalized cut [274], leading to an efficient algorithm, called normalized Benders decomposition [159].

(3) Branch and cut. The branch & cut method, as a hybrid method, combines the cutting plane method [275] with the branch and bound method [276]. This approach is pretty effective for resolving various (mixed) integer programming problems while still ensuring optimality. In general, the branch and cut algorithm is in the same spirit of the branch and bound scheme, but appends new constraints when necessary in each node by resorting to cutting plane approaches [159].

(4) Cut and branch. This method is similar to the branch and cut approach, and the difference lies in that the extra cuts are only added to the root node. Meanwhile, only the branching constraints are added to the other nodes. It is found in [159] that with variables in $\mathbb{R}$ in master problem and stabilization, the cut and branch method is superior to other methods to some extent.

(5) Gradient descent ascent. Gradient descent ascent, i.e., the classical gradient descent and ascent algorithm [277], is the most well-known algorithm for solving continuous SGs, where the descent and ascent operations are performed for the leader and follower, respectively, but in sequential order, and most of the other methods are based on this algorithm [93, 189]. For example, the max-oracle gradient-descent algorithm [93] is a variant of gradient descent ascent, where the ascent operation for the follower is directly replaced with an approximate best response provided by the max-oracle.

Finally, it is worth pointing out that AI methods have also been leveraged to cope with SGs, e.g., [278] and a survey [279] for reference.

4.3 Zero-sum differential games

Among the methods for solving zero-sum DGs, the viscosity solution approach is the most widely exploited one, for which it is known that a value function is the solution of the Hamilton-Jacobi-Isaacs (HJI) equation. In the sequel, this approach is introduced for DGs (10) and (11), and other detailed cases can be found in [50, 280].

For DGs (10) and (11), the Hamiltonian is defined as

$$H(t,x,\omega)=\min_{u\in U}\max_{v\in V}\{\langle f(t,x,u,v),\omega\rangle+f_0(t,x,u,v)\},\quad t\in[t_0,T],\quad x,\omega\in\mathbb{R}^d, \tag{26}$$

and the HJI equation is given as

$$\begin{aligned}&\partial_t\psi(t,z)+H(t,z,\partial_z\psi(t,z))=0,\\&\psi(T,z)=\phi(z),\quad t\in[t_0,T),\ z\in\mathbb{R}^d,\end{aligned} \tag{27}$$

where the second condition is called the terminal condition, $\psi:[t_0,T]\times\mathbb{R}^d\to\mathbb{R}$ is a function, and ∂_t,∂_z represent the subgradients with respect to t,z, respectively.

Let Ψ denote the set of functions $\psi:[t_0,T]\times\mathbb{R}^d\to\mathbb{R}$ satisfying the continuity condition in t and the Lipschitz condition on every bounded subset of $\mathbb{R}^d$ in the second argument. From [214], it is known that if a function $\psi\in\Psi$ is coinvariantly differentiable at each point $(t,z)\in[t_0,T]\times\mathbb{R}^d$ and satisfies HJI

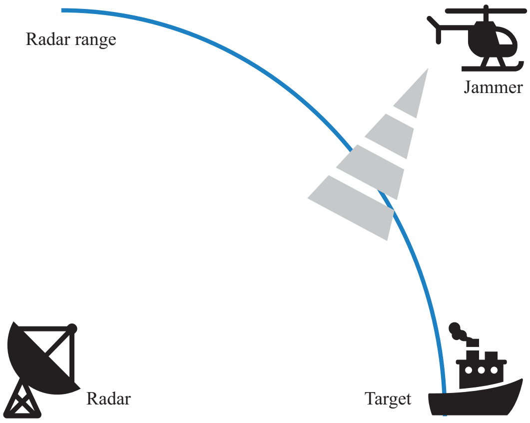

Figure 4 (Color online) Schematic illustration of radar jamming.

equation (27) and $\partial_t \psi, \partial_z \psi \in \Psi$, then ψ is the value function of DG (10) and (11). The optimal control strategies for two players are given as

$$\begin{aligned} u^*(t,z) &\in \arg \min_{u \in U} \max_{v \in V} \chi(t,z,u,v), \\ v^*(t,z) &\in \arg \max_{v \in V} \min_{u \in U} \chi(t,z,u,v), \end{aligned} \tag{28}$$

where

$$\chi(t,z,u,v) := \langle f(t,z,u,v), \partial_z \psi(t,z) \rangle + f_0(t,z,u,v). \tag{29}$$

Moreover, it should be noted that AI methods have also been applied to solve DGs, e.g., reinforcement learning was employed to deal with multi-player nonlinear DGs [281], where a novel two-level value iteration-based integral reinforcement learning algorithm, which only depends on partial information of system dynamics, was proposed.

5 Applications

This section presents some practical applications of adversarial games. Adversarial games have been leveraged to solve a large volume of realistic problems in the literature, including poker [144], StarCraft [282], politics [283], infrastructure security [11], pursuit-evasion problems [204], border defense [23, 182, 284], national defense [22], communication scheduling [285], autonomous driving [286], and homeland security [287]. In what follows, we provide three well-known examples to illustrate these applications.

Example 1 (Radar jamming)**.** Radar jamming is one of the widely studied applications of zero-sum games in modern electronic warfare [288, 289]. Radar jamming involves two players, i.e., one radar, which aims to detect a target with a probability that is as high as possible, and one jammer, which aims at minimizing the radar's detection by jamming it. Therefore, the two players are diametrically opposed, and the scenario forms a two-player zero-sum game (Figure 4 for a schematic illustration). Usually, according to the type of target (e.g., Swerling Type II target [290]), some kinds of utility functions can be constructed in distinct scenarios of jamming, and some constraints can be described mathematically relying on physical limitations, such as jammer power, spatial extent of jamming for the jammer, and threshold parameter and reference window size for the radar. For example, signal models of the radar and jammer in [289] are modeled as follows. For the radar, the transmitted signal is $s(t) = \sum_{m=1}^{M} s_m(t-(m-1)T)$, where M is the number of pulses, $s_m(t)$ is the transmitted signal at pulse m, and T is the pulse repetition time. Moreover, every $s_m(t)$ is composed of K subpulses expressed as follows:

$$s_m(t) = a(t) \sum_{k=1}^{K} \mathrm{rect}((t-kT_c)/T_c) \exp(\jmath 2\pi f_k^m t), \tag{30}$$

where $\jmath$ is the imaginary unit, $a(t)$ denotes the complex envelope, $T_c \in (0,T)$ represents the duration time of a subpulse, $\mathrm{rect}(t)$ is the rectangle function, i.e., $\mathrm{rect}(t) = 1$ if $t \in [0,1]$ and $\mathrm{rect}(t) = 0$ otherwise, and

$f_k^m \in \mathcal{F} := \{\bar{f}_1, \ldots, \bar{f}_N\}$ means the carrier frequency. On the other hand, for the jammer, the transmitted signal is $u(t) = \sum_{m=1}^{M} u_m(t - (m-1)T)$ with $u_m(t)$ being the jamming signal for the mth pulse. $u_m(t)$ is of the form

$$u_m(t) = \mathrm{rect}(t/T_J) v_m(t) \exp(\jmath 2\pi f_m t), \tag{31}$$

where $T_J > 0$ denotes the jamming signal's duration time, $v_m(t)$ represents the jamming signal envelope, and $f_m \in \mathcal{F}$ stands for the jamming signal's carrier frequency. In this problem, the actions of the radar and jammer are the selection of carrier frequency. In [289], the problem is studied by modeling it as an II-EFG, where the utility function is chosen as the probability of detection. The solution concept is the NE, which is computed using the CFR method as introduced in Subsection 4.1.

Example 2 (Border patrols)**.** Securing the national borders to avoid illicit behavior, such as drugs, contraband, and stowaways, is an important task. In this respect, border patrols are introduced here as one application of SSGs, as proposed by Carabineros de Chile [182, 183] to thwart drug trafficking, contraband, and illegal entry. To this end, both day and night shift patrols along the border are arranged by the Carabineros according to distinct requirements.

The night shift patrols are specially considered. To make it practically implementable, the region is partitioned into some police precincts, some of which are paired up when scheduling the patrol because of the vast expanse and harsh landscape at the border and the manpower limitation. In addition, a set of vantage locations, which are suited for conducting surveillance with high-tech equipment, such as heat sensors and night goggles, have been identified by the Carabineros along the border of the region. A night shift action means the deployment of a joint detail with personnel from two paired precincts to conduct overnight vigil at the vantage locations within the realm of the paired precincts. Meanwhile, given the logistical constraints, a joint detail is deployed for every precinct pair to a surveillance location once a week. Figure 5 illustrates the case with 3 pairings, 7 precincts, and 10 locations.

In general, the border patrol problem is formulated as an MILP (possibly different in different literature), e.g., the following form in [291]:

$$\max_{x, s^k, q^k, r^k} \quad \sum_{k=1}^{p} \varpi^k r^k \tag{32}$$

$$\text{s.t.} \quad x^\top \mathbf{1} = 1,\ x \geqslant \mathbf{0}, \tag{33}$$

$$(q^k)^\top \mathbf{1} = 1,\ q^k \in \{0,1\}^{|A_f|}, \quad \forall k \in F, \tag{34}$$

$$r^k \leqslant \sum_{i=1}^{|A_l|} R_{ij}^k x_i + M(1 - q_j^k), \quad \forall k \in F,\ \forall j \in [|A_f|], \tag{35}$$

$$0 \leqslant s^k - \sum_{i=1}^{|A_l|} C_{ij}^k x_i \leqslant M(1 - q_j^k), \quad \forall k \in F,\ \forall j \in [|A_f|], \tag{36}$$

where r^k and s^k are the expected utilities for the leader and follower $k \in F$ when facing each other, respectively, $M > 0$ is a large constant relevant to the highest utility value which makes the constraints redundant if $q_j^k = 0$, and the other notations are the same as those in Subsection 2.2. As an application of SSGs, the concept solution is SSEs, which can be computed by prevalent algorithms, as introduced in Subsection 4.2. Meanwhile, other methods may also exist, such as the sampling method in [182], which is a two-stage approximate method based on random sampling and optimization. Interested readers can refer to [182] for more details.

Example 3 (Pursuit-evasion problems)**.** Pursuit-evasion problems are one of the prevalent applications of zero-sum DGs, which have been widely applied to many practical problems, such as surveillance and navigation, robotics, and aerospace. In pursuit-evasion problems, there usually exists a collection of pursuers and evaders (one pursuer and one evader in the simplest case), possibly with a moving target or stationary target set/area, and the pursuers aim to capture or intercept the evaders, who have opposed objectives [204]. As a concrete example, in [292], the authors considered a case in the plane where there exists one pursuer (or defender) that protects a maritime coastline or border (say, the x-axis) from the attacks of two slower evaders (or attackers), which is played in the open half-plane $y > 0$. The pursuer needs to sequentially pursue the evaders and intercept them as far as possible from the coastline.

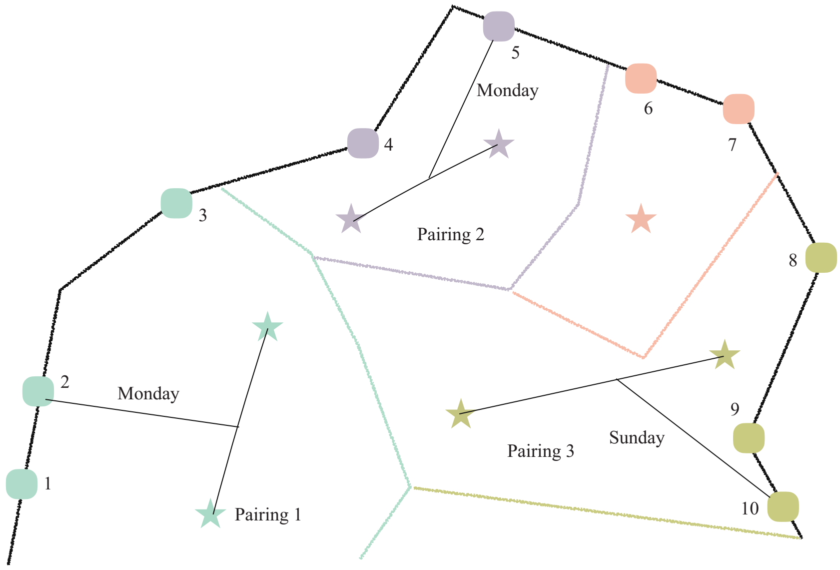

Figure 5 (Color online) Feasible schedule for a week, where the stars and squares denote the precinct headquarters and border outposts, respectively, cited from [182].

Meanwhile, the two evaders can collaborate and minimize their combined distance to the coastline before they are intercepted. The states of the pursuer and the two evaders are specified by position coordinates $x_P = (x_P, y_P)$, $x_1 = (x_1, y_1)$, and $x_2 = (x_2, y_2)$, respectively, and the complete state of the game is given as $z = (x_p, y_P, x_1, y_1, x_2, y_2) \in \mathbb{R}^6$. The two evaders are assumed to have the same speed v, which is slower than the pursuer, i.e., $v < v_P$. The controls of the pursuer and the two evaders are their instantaneous heading angles ψ, ϕ_1, ϕ_2. The dynamics $\dot{z} = f(z, u, v)$ as in DG (10) are given by

$$\begin{aligned}
&\dot{x}_P = \cos(\psi), \quad \dot{y}_P = \sin\psi, \\
&\dot{x}_1 = v_r \cos(\phi_1), \quad \dot{y}_1 = v_r \sin\phi_1, \\
&\dot{x}_2 = v_r \cos(\phi_2), \quad \dot{y}_2 = v_r \sin\phi_2,
\end{aligned} \tag{37}$$

where $u = \{\psi\}$, $v = \{\phi_1, \phi_2\}$, $v_r := v/v_P$, and $\psi, \phi_1, \phi_2 \in [-\pi, \pi)$. Let t' and t_f denote the time instants when the first and second evaders are captured by the pursuer, respectively. Then, the terminal cost function is given as

$$J(u(t), v(t), z_0) = \Phi(z(t'), z(t_f)), \tag{38}$$

$$\Phi(z(t'), z(t_f)) := y_P(t') + y_P(t_f), \tag{39}$$

where z_0 is the initial state of the system. It is a zero-sum DG whose solution concept is the NE, meaning optimal instantaneous headings for all three players at each time. In [292], the solution is resolved by resorting to the Apollonius circle and the HJI equation.

Example 4 (Generative adversarial networks (GANs))**.** As mentioned in Section 3, game theory plays a vital role in machine learning [156]. As one of the prevalent applications in machine learning, the success of GANs is substantially underpinned by zero-sum games. As a sort of deep learning architecture, GAN was invented by Goodfellow et al. in 2014 [293], which involves two neural networks (called generator and discriminator) competing with each other in a game. By training images, the goal of GANs is to generate new images that are indistinguishable to humans. GANs have been widely used in unsupervised learning, semi-supervised learning, fully supervised learning, and reinforcement learning. For example, in unsupervised learning, GANs have become one of the most prevailing methods for implicitly learning the underlying distribution of a given dataset [293]. The objective of the generator is to generate phony images from random noise, while the discriminator aims to correctly classify phony and real images, which, in general, is mathematically modeled as

$$\max_{\mathcal{D}} \min_{\mathcal{G}} V(\mathcal{G}, \mathcal{D}), \tag{40}$$

where $\mathcal{G}$ and $\mathcal{D}$ represent the generator and discriminator networks, respectively, and V is the cost function of training, defined by

$$V(\mathcal{G},\mathcal{D}) := \mathbb{E}_{p_{\rm data}(x)} \log \mathcal{D}(x) + \mathbb{E}_{p_g(x)} \log(1-\mathcal{D}(x)). \tag{41}$$

In (41), $\mathbb{E}(\cdot)$ means the mathematical expectation, $p_{\rm data}(x)$ denotes the probability density function over random vector x, and $p_g(x)$ stands for the distribution of the vectors produced by the generator network. From (41), one can easily see that the problem is modeled as a two-player zero-sum game between a discriminator and a generator whose solution is generally the NE which can be computed by any efficient method for solving two-player zero-sum games. For more details and applications to deep learning, the readers can refer to recent surveys [156, 294].

6 Possible future directions

In view of some challenges in adversarial games, this section attempts to present and discuss potential future research directions.

• Efficient algorithm design. Even though a wide range of algorithms have been proposed in the literature, as discussed previously, efficient, fast, and optimal algorithms with limited computing, storage, and memory capabilities are still the overarching research directions in (adversarial) games and AI, which are far from fully explored, including a plethora of scenarios, e.g., equilibrium computation [250], real-time strategy (RTS) making [295], exploiting suboptimal opponents [296], and attack resiliency [297]. In addition, for adversarial games with large action spaces and/or infosets, practical limitations, such as limited computing resources, impose the need for efficient algorithm designs amenable to implementation with limited computation, storage, and even communication [298].

• Last-iterate convergence. In general, no-regret learning can guarantee the convergence of the empirical distribution of play (i.e., time-average convergence) for each player to the set of NEs. However, the last-iterate convergence fails in general [266, 267], although restricted classes of games indeed have the last-iterate convergence by no-regret learning algorithms, such as two-player zero-sum games [96]. Note that the last-iterate convergence is important in many practical applications, e.g., GANs [299] and economics [255], which have been receiving considerable interest in recent years [300].

• Imperfect information. Imperfect information, as a possible main feature of many practical adversarial games (e.g., the card game of poker), inflicts a major challenge in adversarial games due to the existence of hidden information [126, 301]. Currently, many popular methods (e.g., CFR-based methods), including AI approaches, have been developed to cope with imperfect-information games, especially II-EFGs [116]. However, it is far from being fully explored under various scenarios, such as computationally light algorithms, domain-free algorithms, and the case with a large number of players.

• Incomplete information. Incomplete information is one of the main hallmarks of many adversarial games and one of the challenge sources. In general, game uncertainties, such as parameter uncertainty, action outcome uncertainty, and underlying world state uncertainty, can be subsumed in the category of incomplete information. The main investigated models are Bayesian and interval models [25, 302, 303].

• Bounded rationality. Completely rational players are often assumed in the study of games. Nonetheless, irrational players naturally appear in practice, which has triggered an increasing interest in games with bounded rationality, e.g., behavior models (e.g., lens-QR models), prospect theory inspired models, and quantal response models [304–306].

• Dynamic environments. Most of the games have been investigated as static ones, i.e., with time-invariant game rules. However, because of the possible dynamic characteristics of the environment within which players compete, online games (or time-varying games), where each player's utility function is time-varying or even adversarial without any distribution assumptions, need to be investigated further in the future [21, 251–254].

• Hybrid games. Many realistic adversarial games involve both continuous and discrete physical dynamics that govern players' motion or changing rules, which can be framed in the framework of hybrid games [307, 308]. In this respect, how to combine the game theory with control dynamics is an important yet challenging research area.

• AI in games. Recent years have witnessed considerable progress in the success of AI methods applied in games, which can integrate some advanced approaches of reinforcement learning, neural networks, and meta-learning [146, 309–311]. With the advent of modern high-tech and big-data complex missions, AI methods provide an effective manner for the implementation of RTSs by solely exploiting offline or real-time streaming data [150].

7 Conclusion

Adversarial games play a significant role in practical applications, for which this survey provided a systematic overview of three main categories, i.e., zero-sum normal-form and extensive-form games, Stackelberg (security) games, and zero-sum DGs. To this end, several distinct angles have been employed to anatomize adversarial games, ranging from game models, solution concepts, problem classification, research frontiers, prevailing algorithms, and real-world applications to potential future directions. In general, this survey has attempted to review the existing research in an intact manner, although the references are too vast to cover in its entirety. To the best of our knowledge, this survey is the first to present a systematic overview of adversarial games. Finally, future potential directions have also been discussed.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2022ZD0119702), National Natural Science Foundation of China (Grant Nos. 62003243, 62103305, 62088101), Shanghai Pujiang Program (Grant No. 21PJ1413100), and Shanghai Municipal Science and Technology Major Project (Grant No. 2021SHZDZX0100). The authors are grateful to the associate editor and the anonymous reviewers for their insightful suggestions.

References

1 von Neumann J, Morgenstern O. Theory of Games and Economic Behavior. 2nd ed. Princeton: Princeton University Press, 1947
2 Nash J F. Equilibrium points in n-person games. Proc Natl Acad Sci USA, 1950, 36: 48–49
3 Nash J. Non-cooperative games. Ann Math, 1951, 54: 286–295
4 Fudenberg D, Tirole J. Game Theory. Cambridge: MIT Press, 1991
5 Osborne M J, Rubinstein A. A Course in Game Theory. Cambridge: MIT Press, 1994
6 Başar T, Zaccour G. Handbook of Dynamic Game Theory. Berlin: Springer International Publishing, 2018
7 Meng M, Li X X. On the linear convergence of distributed Nash equilibrium seeking for multi-cluster games under partial-decision information. Automatica, 2023, 151: 110919
8 Aumann R J, Maschler M, Stearns R E. Repeated Games with Incomplete Information. Cambridge: MIT Press, 1995
9 Bard N, Hawkin J, Rubin J, et al. The annual computer poker competition. AI Mag, 2013, 34: 112–114
10 Leyton-Brown K, Shoham Y. Essentials of Game Theory: A Concise Multidisciplinary Introduction. Cham: Springer, 2008. 1–88
11 Sinha A, Fang F, An B, et al. Stackelberg security games: looking beyond a decade of success. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Stockholm, 2018. 5494–5501
12 Nguyen T H, Kar D, Brown M, et al. Towards a science of security games. In: Proceedings of the Mathematical Sciences with Multidisciplinary Applications, 2016. 347–381
13 Etesami S R, Başar T. Dynamic games in cyber-physical security: an overview. Dyn Games Appl, 2019, 9: 884–913
14 Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search. Nature, 2016, 529: 484–489
15 Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of Go without human knowledge. Nature, 2017, 550: 354–359
16 Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science, 2018, 362: 1140–1144
17 Aradillas-López A. The econometrics of static games. Annu Rev Econ, 2020, 12: 135–165
18 Li H L, Wang X, Jia F W, et al. A survey of Nash equilibrium strategy solving based on CFR. Arch Computat Methods Eng, 2021, 28: 2749–2760
19 Sohrabi M K, Azgomi H. A survey on the combined use of optimization methods and game theory. Arch Computat Methods Eng, 2020, 27: 59–80
20 Wang J R, Hong Y T, Wang J L, et al. Cooperative and competitive multi-agent systems: from optimization to games. IEEE CAA J Autom Sin, 2022, 9: 763–783
21 Li X X, Xie L H, Li N. A survey on distributed online optimization and game. 2022. ArXiv:2205.00473
22 Ho E, Rajagopalan A, Skvortsov A, et al. Game theory in defence applications: a review. Sensors, 2022, 22: 1032
23 Shishika D, Kumar V. A review of multi-agent perimeter defense games. In: Proceedings of the International Conference on Decision and Game Theory for Security, College Park, 2020. 472–485
24 Zhu M, Anwar A H, Wan Z L, et al. A survey of defensive deception: approaches using game theory and machine learning. IEEE Commun Surv Tutorials, 2021, 23: 2460–2493
25 Insua D R, Ruggeri F, Soyer R, et al. Advances in Bayesian decision making in reliability. Eur J Operational Res, 2020, 282: 1–18
26 Lanctot M, Zambaldi V, Gruslys A, et al. A unified game-theoretic approach to multiagent reinforcement learning. In: Proceedings of the Advances in Neural Information Processing Systems, Long Beach, 2017. 30
27 Littman M L. Markov games as a framework for multi-agent reinforcement learning. In: Proceedings of the Machine Learning Proceedings, 1994. 157–163
28 Zamir S. Bayesian Games: Games With Incomplete Information. Technical Report, 2008
29 Chen X, Deng X T, Teng S-H. Settling the complexity of computing two-player Nash equilibria. J ACM, 2009, 56: 1–57
30 Daskalakis C, Goldberg P W, Papadimitriou C H. The complexity of computing a Nash equilibrium. SIAM J Comput, 2009, 39: 195–259
31 Rubinstein A. Hardness of Approximation Between P and NP. Williston: Morgan & Claypool, 2019
32 Aumann R J. Subjectivity and correlation in randomized strategies. J Math Economics, 1974, 1: 67–96

33 Hannan J. Approximation to Bayes risk in repeated play. In: Contributions to the Theory of Games. Princeton: Princeton University Press, 1957. 3: 97–139
34 Neumann J V. Zur theorie der gesellschaftsspiele. Math Ann, 1928, 100: 295–320
35 Farina G, Bianchi T, Sandholm T. Coarse correlation in extensive-form games. In: Proceedings of AAAI Conference on Artificial Intelligence, 2020. 34: 1934–1941
36 Celli A, Coniglio S, Gatti N. Computing optimal coarse correlated equilibria in sequential games. 2019. ArXiv:1901.06221
37 Celli A, Gatti N. Computational results for extensive-form adversarial team games. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2018. 32
38 von Stengel B, Koller D. Team-maxmin equilibria. Games Economic Behav, 1997, 21: 309–321
39 Omidshafiei S, Papadimitriou C, Piliouras G, et al. α-rank: multi-agent evaluation by evolution. Sci Rep, 2019, 9: 9937
40 von Stackelberg H. Marktform und Gleichgewicht. Berlin: Springer-Verlag, 1934
41 An B, Ordóñez F, Tambe M, et al. A deployed quantal response-based patrol planning system for the U.S. coast guard. Interfaces, 2013, 43: 400–420
42 Casorrán C, Fortz B, Labbé M, et al. A study of general and security Stackelberg game formulations. Eur J Operational Res, 2019, 278: 855–868
43 Conitzer V, Sandholm T. Computing the optimal strategy to commit to. In: Proceedings of the 7th ACM Conference on Electronic Commerce, Michigan, 2006. 82–90
44 Leitmann G. On generalized Stackelberg strategies. J Optim Theor Appl, 1978, 26: 637–643
45 von Stackelberg H. Market Structure and Equilibrium. Berlin: Springer Science & Business Media, 2011
46 Başar T, Olsder G J. Dynamic noncooperative game theory. In: Proceedings of Society for Industrial and Applied Mathematics, 1998
47 Isaacs R. Differential Games. New York: Wiley, 1965
48 Lewis F L, Vrabie D, Syrmos V L. Optimal Control. Hoboken: John Wiley & Sons, 2012
49 Buckdahn R, Cardaliaguet P, Quincampoix M. Some recent aspects of differential game theory. Dyn Games Appl, 2011, 1: 74–114
50 Friedman A. Differential Games. Mineola: Courier Corporation, 2013
51 Jiang H, Zhang H G. Iterative ADP learning algorithms for discrete-time multi-player games. Artif Intell Rev, 2018, 50: 75–91
52 Qiao J F, Li M H, Wang D. Asymmetric constrained optimal tracking control with critic learning of nonlinear multiplayer zero-sum games. IEEE Trans Neural Networks Learning Syst, 2022. doi: 10.1109/TNNLS.2022.3208611
53 Liu M S, Wan Y, Lewis F L, et al. Stochastic two-player zero-sum learning differential games. In: Proceedings of the 15th International Conference on Control and Automation (ICCA), Edinburgh, 2019. 1038–1043
54 Li D X, Cruz J J B, Schumacher C J. Stochastic multi-player pursuit-evasion differential games. Intl J Robust Nonlinear, 2008, 18: 218–247
55 Barman S, Ligett K. Finding any nontrivial coarse correlated equilibrium is hard. Sigecom Exch, 2015, 14: 76–79
56 Garg J, Jiang A X, Mehta R. Bilinear games: polynomial time algorithms for rank based subclasses. In: Proceedings of the International Workshop on Internet and Network Economics, Singapore, 2011. 399–407
57 Lemke C E, Howson J J T. Equilibrium points of bimatrix games. J Soc Industrial Appl Math, 1964, 12: 413–423
58 Anagnostides I, Penna P. Solving zero-sum games through alternating projections. 2020. ArXiv:2010.00109
59 Dinh L C, Yang Y D, Tian Z, et al. Online double oracle. 2021. ArXiv:2103.07780
60 Murhekar A. Approximate Nash equilibria of imitation games: algorithms and complexity. In: Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, 2020. 887–894
61 Borel E. La théorie du jeu et les équations intégralesa noyau symétrique. Comptes rendus de l'Académie des Sciences, 1921, 173: 1304–1308
62 Howson J J T. Equilibria of polymatrix games. Manage Sci, 1972, 18: 312–318
63 Sengodan G, Arumugasamy C. Linear complementarity problems and bi-linear games. Appl Math, 2020, 65: 665–675
64 Deligkas A, Fasoulakis M, Markakis E. A polynomial-time algorithm for 1/3-approximate Nash equilibria in bimatrix games. 2022. ArXiv:2204.11525
65 Deligkas A, Fearnley J, Savani R. Tree polymatrix games are PPAD-hard. 2020. ArXiv:2002.12119
66 Seddighin S. Campaigning via LPs: solving blotto and beyond. Dissertation for Ph.D. Degree. College Park: University of Maryland, 2019
67 Mehta R. Constant rank two-player games are PPAD-hard. SIAM J Comput, 2018, 47: 1858–1887
68 Boodaghians S, Brakensiek J, Hopkins S B, et al. Smoothed complexity of 2-player Nash equilibria. In: Proceedings of the Annual Symposium on Foundations of Computer Science, 2020. 271–282
69 Behnezhad S, Blum A, Derakhshan M, et al. Optimal strategies of Blotto games: beyond convexity. In: Proceedings of the ACM Conference on Economics and Computation, Phoenix, 2019. 597–616
70 Behnezhad S, Dehghani S, Derakhshan M, et al. Fast and simple solutions of Blotto games. Oper Res, 2022. doi: 10.1287/opre.2022.2261
71 Beaglehole D. An efficient approximation algorithm for the Colonel Blotto game. 2022. ArXiv:2201.10758
72 Leon V, Etesami S R. Bandit learning for dynamic Colonel Blotto game with a budget constraint. 2021. ArXiv:2103.12833
73 Vu D Q, Loiseau P, Silva A. Approximate equilibria in generalized Colonel Blotto and generalized Lottery Blotto games. 2019. ArXiv:1910.06559
74 Boix-Adserá E, Edelman B L, Jayanti S. The multiplayer Colonel Blotto game. Games Economic Behav, 2021, 129: 15–31
75 Vlatakis-Gkaragkounis E V, Flokas L, Piliouras G. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games. In: Proceedings of the Advances in Neural Information Processing Systems, Vancouver, 2019. 1–12
76 Zhang G D, Wang Y H, Lessard L, et al. Near-optimal local convergence of alternating gradient descent-ascent for minimax optimization. In: Proceedings of the International Conference on Artificial Intelligence and Statistics, 2022, 7659–7679
77 Hamedani E Y, Aybat N S. A primal-dual algorithm with line search for general convex-concave saddle point problems. SIAM J Optim, 2021, 31: 1299–1329
78 Tominin V, Tominin Y, Borodich E, et al. On accelerated methods for saddle-point problems with composite structure. 2021. ArXiv:2103.09344
79 Xie G Z, Han Y Z, Zhang Z H. DIPPA: an improved method for bilinear saddle point problems. 2021. ArXiv:2103.08270
80 Kovalev D, Gasnikov A, Richtárik P. Accelerated primal-dual gradient method for smooth and convex-concave saddle-point

problems with bilinear coupling. 2021. ArXiv:2112.15199
81 Thekumparampil K K, He N, Oh S. Lifted primal-dual method for bilinearly coupled smooth minimax optimization. 2022. ArXiv:2201.07427
82 Gidel G, Jebara T, Lacoste-Julien S. Frank-wolfe algorithms for saddle point problems. In: Proceedings of the International Conference on Artificial Intelligence and Statistics, 2017. 362–371
83 Chen C, Luo L, Zhang W N, et al. Efficient projection-free algorithms for saddle point problems. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 10799–10808
84 Li H C, Tian Y, Zhang J Z, et al. Complexity lower bounds for nonconvex-strongly-concave min-max optimization. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 1–13
85 Hsieh Y P, Mertikopoulos P, Cevher V. The limits of min-max optimization algorithms: convergence to spurious non-critical sets. In: Proceedings of the International Conference on Machine Learning, 2021. 4337–4348
86 Wei C Y, Lee C W, Zhang M X, et al. Linear last-iterate convergence in constrained saddle-point optimization. In: Proceedings of the International Conference on Learning Representations, 2021. 1–12
87 Bistritz I, Zhou Z Y, Chen X, et al. No weighted-regret learning in adversarial bandits with delays. J Machine Learning Res, 2022, 23: 1–43
88 Fiez T, Sim R, Skoulakis S, et al. Online learning in periodic zero-sum games. 2021. ArXiv:2111.03377
89 Gao H C, Wang X Q, Luo L, et al. On the convergence of stochastic compositional gradient descent ascent method. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2021. 1–7
90 Beznosikov A, Scutari G, Rogozin A, et al. Distributed saddle-point problems under data similarity. 2021. ArXiv:2107.10706
91 Balduzzi D, Garnelo M, Bachrach Y, et al. Open-ended learning in symmetric zero-sum games. In: Proceedings of the International Conference on Machine Learning, 2019. 434–443
92 Vlatakis-Gkaragkounis E V, Flokas L, Piliouras G. Solving min-max optimization with hidden structure via gradient descent ascent. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34: 1–14
93 Goktas D, Greenwald A. Convex-concave min-max Stackelberg games. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34
94 Xefteris D. Symmetric zero-sum games with only asymmetric equilibria. Games Economic Behav, 2015, 89: 122–125
95 Cai Y, Daskalakis C. On minmax theorems for multiplayer games. In: Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, 2011. 217–234
96 Anagnostides I, Panageas I, Farina G, et al. On last-iterate convergence beyond zero-sum games. 2022. ArXiv:2203.12056
97 Bailey J P. $o(1/t)$ time-average convergence in a generalization of multiagent zero-sum games. 2021. ArXiv:2110.02482
98 Fiez T, Sim R, Skoulakis S, et al. Online learning in periodic zero-sum games: von Neumann vs Poincaré, 2021. ArXiv:2111.03377
99 Skoulakis S, Fiez T, Sim R, et al. Evolutionary game theory squared: evolving agents in endogenously evolving zero-sum games. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2021. 1–9
100 Hughes E, Anthony T W, Eccles T, et al. Learning to resolve alliance dilemmas in many-player zero-sum games. 2020. ArXiv:2003.00799
101 Ganzfried S. Fast complete algorithm for multiplayer Nash equilibrium. 2020. ArXiv:2002.04734
102 Anagnostides I, Daskalakis C, Farina G, et al. Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games. 2021. ArXiv:2111.06008
103 Anagnostides I, Farina G, Kroer C, et al. Faster no-regret learning dynamics for extensive-form correlated and coarse correlated equilibria. 2022. ArXiv:2202.05446
104 Gidel G. Multi-player games in the era of machine learning. Dissertation for Ph.D. Degree. Montreal: Université de Montréal, 2020
105 Zhang Y Z, An B. Converging to team-maxmin equilibria in zero-sum multiplayer games. In: Proceedings of the International Conference on Machine Learning, 2020. 11033–11043
106 Kalogiannis F, Vlatakis-Gkaragkounis E V, Panageas I. Teamwork makes von Neumann work: min-max optimization in two-team zero-sum games. 2021. ArXiv:2111.04178
107 HansenK A, Hansen T D, Miltersen P B, et al. Approximability and parameterized complexity of minmax values. In: Proceedings of the International Workshop on Internet and Network Economics, 2008. 684–695
108 Borgs C, Chayes J, Immorlica N, et al. The myth of the Folk Theorem. Games Economic Behav, 2010, 70: 34–43
109 Gharesifard B, Cortés J. Distributed convergence to Nash equilibria in two-network zero-sum games. Automatica, 2013, 49: 1683–1692
110 Lou Y C, Hong Y G, Xie L H, et al. Nash equilibrium computation in subnetwork zero-sum games with switching communications. IEEE Trans Automat Contr, 2016, 61: 2920–2935
111 Huang S J, Lei J L, Hong Y G, et al. No-regret distributed learning in two-network zero-sum games. In: Proceedings of the IEEE Conference on Decision and Control, Austin, 2021. 924–929
112 Zhang Y Z, An B. Computing team-maxmin equilibria in zero-sum multiplayer extensive-form games. In: Proceedings of AAAI Conference on Artificial Intelligence, 2020. 34: 2318–2325
113 Carminati L, Cacciamani F, Ciccone M, et al. Public information representation for adversarial team games. 2022. ArXiv:2201.10377
114 Farina G, Celli A, Gatti N, et al. Faster algorithms for optimal ex-ante coordinated collusive strategies in extensive-form zero-sum games. 2020. ArXiv:2009.10061
115 Zhang B H, Sandholm T. Team correlated equilibria in zero-sum extensive-form games via tree decompositions. 2021. ArXiv:2109.05284
116 Brown N. Equilibrium finding for large adversarial imperfect-information games. Dissertation for Ph.D. Degree. Pittsburgh: Carnegie Mellon University, 2020
117 Tammelin O, Burch N, Johanson M, et al. Solving heads-up limit Texas Hold'em. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2015. 645–652
118 Moravčík M, Schmid M, Burch N, et al. DeepStack: expert-level artificial intelligence in heads-up no-limit poker. Science, 2017, 356: 508–513
119 Brown N, Sandholm T. Superhuman AI for heads-up no-limit poker: libratus beats top professionals. Science, 2018, 359: 418–424
120 Munos R, Perolat J, Lespiau J B, et al. Fast computation of Nash equilibria in imperfect information games. In: Proceedings of the International Conference on Machine Learning, 2020. 7119–7129

121 Farina G, Kroer C, Sandholm T. Better regularization for sequential decision spaces: fast convergence rates for Nash, correlated, and team equilibria. 2021. ArXiv:2105.12954
122 Brown N, Sandholm T. Safe and nested subgame solving for imperfect-information games. In: Proceedings of the Advances in Neural Information Processing Systems, 2017. 30: 1–11
123 Brown N, Sandholm T, Amos B. Depth-limited solving for imperfect-information games. In: Proceedings of the Advances in Neural Information Processing Systems, 2018. 31: 1–12
124 Marsland T A. A review of game-tree pruning. J Int Comput Games Assoc, 1986, 9: 3–19
125 Sandholm T. Solving imperfect-information games. Science, 2015, 347: 122–123
126 Schmid M. Search in imperfect information games. 2021. ArXiv:2111.05884
127 Kovařík V, Milec D, Šustr M, et al. Fast algorithms for poker require modelling it as a sequential Bayesian game. 2021. ArXiv:2112.10890
128 Farina G, Lee C W, Luo H P, et al. Kernelized multiplicative weights for 0/1-polyhedral games: bridging the gap between learning in extensive-form and normal-form games. 2022. ArXiv:2202.00237
129 Meng L J, Gao Y. Generalized bandit regret minimizer framework in imperfect information extensive-form game. 2022. ArXiv:2203.05920
130 Bai Y, Jin C, Mei S, et al. Near-optimal learning of extensive-form games with imperfect information. 2022. ArXiv:2202.01752
131 Kozuno T, Ménard P, Munos R, et al. Model-free learning for two-player zero-sum partially observable Markov games with perfect recall. 2021. ArXiv:2106.06279
132 Brown N, Sandholm T. Superhuman AI for multiplayer poker. Science, 2019, 365: 885–890
133 Blair A, Saffidine A. AI surpasses humans at six-player poker. Science, 2019, 365: 864–865
134 Wu B. Hierarchical macro strategy model for MOBA game AI. In: Proceedings of AAAI Conference on Artificial Intelligence, 2019. 33: 1206–1213
135 Tian Y D, Gong Q C, Jiang Y. Joint policy search for multi-agent collaboration with imperfect information. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 33: 19931–19942
136 Ganzfried S, Laughlin C, Morefield C. Parallel algorithm for Nash equilibrium in multiplayer stochastic games with application to naval strategic planning. In: Proceedings of the International Conference on Distributed Artificial Intelligence, 2020. 1–13
137 Yuan W L, Hu Z Z, Luo J R, et al. Imperfect information game in multiplayer no-limit Texas Hold'em based on mean approximation and deep CFVnet. In: Proceedings of the China Automation Congress, 2021. 2459–2466
138 Heinrich J, Silver D. Deep reinforcement learning from self-play in imperfect-information games. 2016. ArXiv:1603.01121
139 Li H, Hu K L, Zhang S H, et al. Double neural counterfactual regret minimization. In: Proceedings of the International Conference on Learning Representations, 2019. 1–13
140 Farnia F, Ozdaglar A. Do GANs always have Nash equilibria? In: Proceedings of the International Conference on Machine Learning, 2020. 3029–3039
141 Gruslys A, Lanctot M, Munos R, et al. The advantage regret-matching actor-critic. 2020. ArXiv:2008.12234
142 Ye D H, Chen G B, Zhang W, et al. Towards playing full MOBA games with deep reinforcement learning. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 33: 621–632
143 Ye D, Liu Z, Sun M, et al. Mastering complex control in MOBA games with deep reinforcement learning. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 34: 6672–6679
144 Schmid M, Moravcik M, Burch M, et al. Player of games. 2021. ArXiv:2112.03178
145 Phillips P. Reinforcement learning in two-player zero-sum simultaneous action games. 2021. ArXiv:2110.04835
146 Fu H B, Liu W M, Wu S, et al. Actor-critic policy optimization in a large-scale imperfect-information game. In: Proceedings of the International Conference on Learning Representations, 2021. 1–12
147 Wang X R, Cerny J, Li S X, et al. A unified perspective on deep equilibrium finding. 2022. ArXiv:2204.04930
148 Feng X D, Slumbers O, Wan Z Y, et al. Neural auto-curricula in two-player zero-sum games. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34
149 Feng X D, Slumbers O, Yang Y D, et al. Discovering multi-agent auto-curricula in two-player zero-sum games. 2021. ArXiv:2106.02745
150 Yin Q Y, Yang J, Ni W C, et al. AI in games: techniques, challenges and opportunities. 2021. ArXiv:2111.07631
151 Celli A, Marchesi A, Bianchi T, et al. Learning to correlate in multi-player general-sum sequential games. In: Proceedings of the Advances in Neural Information Processing Systems, 2019. 32
152 Celli A, Marchesi A, Farina G, et al. No-regret learning dynamics for extensive-form correlated equilibrium. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 33: 7722–7732
153 Song Z, Mei S, Bai Y. Sample-efficient learning of correlated equilibria in extensive-form games. 2022. ArXiv:2205.07223
154 Wei C Y, Lee C W, Zhang M X, et al. Last-iterate convergence of decentralized optimistic gradient descent/ascent in infinite-horizon competitive Markov games. In: Proceedings of the Annual Conference on Learning Theory, 2021. 4259–4299
155 Mao W C, Başar T. Provably efficient reinforcement learning in decentralized general-sum Markov games. Dyn Games Appl, 2023, 13: 165–186
156 Hazra T, Anjaria K. Applications of game theory in deep learning: a survey. Multimed Tools Appl, 2022, 81: 8963–8994
157 Erdem A, Pelillo M. Graph transduction as a non-cooperative game. In: Proceedings of the International Workshop on Graph-Based Representations in Pattern Recognition, 2011. 195–204
158 Casorrán-Amilburu C. Formulations and algorithms for general and security Stackelberg games. Dissertation for Ph.D. Degree. Bruxelles: Université libre de Bruxelles; Santiago: Universidad de Chile, 2017
159 Fritz I A A. Benders decomposition based algorithms for general and security Stackelberg games. Dissertation for Master's Degree. Santiago: Universidad de Chile, 2021
160 Dempe S. Bilevel Optimization: Theory, Algorithms and Applications, Volume 3. Freiberg: TU Bergakademie Freiberg, 2018
161 Li T, Sethi S P. A review of dynamic Stackelberg game models. Discret Contin Dynam Syst-B, 2017, 22: 125–159
162 Maharjan S, Zhu Q, Zhang Y, et al. Dependable demand response management in the smart grid: a Stackelberg game approach. IEEE Trans Smart Grid, 2013, 4: 120–132
163 Yu M, Hong S H. A real-time demand-response algorithm for smart grids: a Stackelberg game approach. IEEE Trans Smart Grid, 2015, 7: 879–888
164 Yang D J, Xue G L, Zhang J, et al. Coping with a smart jammer in wireless networks: a Stackelberg game approach. IEEE Trans Wireless Commun, 2013, 12: 4038–4047

165 Guzmán C, Riffo J, Telha C, et al. A sequential Stackelberg game for dynamic inspection problems. Eur J Operational Res, 2022, 302: 727–739
166 Jiang Y N, Zhong Y, Ge X H. IIoT data sharing based on blockchain: a multileader multifollower Stackelberg game approach. IEEE Internet Things J, 2022, 9: 4396–4410
167 Leyffer S, Munson T. Solving multi-leader-common-follower games. Optimization Methods Software, 2010, 25: 601–623
168 Zhang H Q, Xiao Y, Cai L X, et al. A multi-leader multi-follower Stackelberg game for resource management in LTE unlicensed. IEEE Trans Wireless Commun, 2017, 16: 348–361
169 Mallozzi L, Messalli R. Multi-leader multi-follower model with aggregative uncertainty. Games, 2017, 8: 25
170 Tran T D, Le L B. Resource allocation for multi-tenant network slicing: a multi-leader multi-follower Stackelberg game approach. IEEE Trans Veh Technol, 2020, 69: 8886–8899
171 Castiglioni M, Marchesi A, Gatti N. Committing to correlated strategies with multiple leaders. Artif Intelligence, 2021, 300: 103549
172 Pita J, Jain M, Tambe M, et al. Robust solutions to Stackelberg games: addressing bounded rationality and limited observations in human cognition. Artif Intelligence, 2010, 174: 1142–1171
173 Bai Y, Jin C, Wang H, et al. Sample-efficient learning of Stackelberg equilibria in general-sum games. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34
174 Korzhyk D, Conitzer V, Parr R. Complexity of computing optimal Stackelberg strategies in security resource allocation games. In: Proceedings of the AAAI Conference on Artificial Intelligence, Georgia, 2010. 805–810
175 Kiekintveld C, Jain M, Tsai J, et al. Computing optimal randomized resource allocations for massive security games. In: Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, Budapest, 2009. 689–696
176 Jain M, Tsai J, Pita J, et al. Software assistants for randomized patrol planning for the LAX airport police and the federal air marshal service. Interfaces, 2010, 40: 267–290
177 Fang F, Nguyen T H. Green security games: apply game theory to addressing green security challenges. SIGecom Exch, 2016, 15: 78–83
178 Brown M, Sinha A, Schlenker A, et al. One size does not fit all: a game-theoretic approach for dynamically and effectively screening for threats. In: Proceedings of the AAAI Conference on Artificial Intelligence, Arizona, 2016. 30
179 Zhang C, Gholami S, Kar D, et al. Keeping pace with criminals: an extended study of designing patrol allocation against adaptive opportunistic criminals. Games, 2016, 7: 15
180 Dasgupta P, Collins J B, Mittu R. Adversary-Aware Learning Techniques and Trends in Cybersecurity. Berlin: Springer, 2021
181 Galinkin E. Information security games: a survey. 2021. ArXiv:2103.12520
182 Bucarey V, Casorrán C, Figueroa O, et al. Building real Stackelberg security games for border patrols. In: Proceedings of the International Conference on Decision and Game Theory for Security, Vienna, 2017. 193–212
183 Bucarey V, Casorrán C, Labbé M, et al. Coordinating resources in Stackelberg security games. Eur J Operational Res, 2021, 291: 846–861
184 Lou J, Vorobeychik Y. Equilibrium analysis of multi-defender security games. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, 2015. 596–602
185 Mutzari D, Aumann Y, Kraus S. Robust solutions for multi-defender Stackelberg security games. 2022. ArXiv:2204.14000
186 Li Y Q, Conitzer V, Korzhyk D. Catcher-evader games. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), New York, 2016. 329–337
187 Wang B R, Zhang Y, Zhou Z-H, et al. On repeated Stackelberg security game with the cooperative human behavior model for wildlife protection. Appl Intell, 2019, 49: 1002–1015
188 Ma W J, Liu W, McAreavey K, et al. A decision support framework for security resource allocation under ambiguity. Int J Intell Syst, 2021, 36: 5–52
189 Fiez T, Chasnov B, Ratliff L J. Convergence of learning dynamics in Stackelberg games. 2019. ArXiv:1906.01217
190 Kulkarni A A, Shanbhag U V. An existence result for hierarchical Stackelberg v/s Stackelberg games. IEEE Trans Automat Contr, 2015, 60: 3379–3384
191 Goktas D, Zhao J Y, Greenwald A. Robust no-regret learning in min-max Stackelberg games. 2022. ArXiv:2203.14126
192 Maffioli M. Dealing with partial information in follower's behavior identification. Dissertation for Master's Degree. Milano: Politecnico di Milano, 2019
193 Cheng Z Y, Chen G P, Hong Y G. Single-leader-multiple-followers Stackelberg security game with hypergame framework. IEEE Trans Inform Forensic Secur, 2022, 17: 954–969
194 Birmpas G, Gan J R, Hollender A, et al. Optimally deceiving a learning leader in Stackelberg games. J Artif Intell Res, 2021, 72: 507–531
195 Sinha A. AI and security: a game perspective. In: Proceedings of the International Conference on Communication Systems & Networks (COMSNETS), Bangalore, 2022. 393–396
196 Ding N N, Fang Z X, Duan L J, et al. Incentive mechanism design for distributed coded machine learning. In: Proceedings of the IEEE Conference on Computer Communications, Vancouver, 2021. 1–10
197 Cheng C, Zhu Z Q, Xin B, et al. A multi-agent reinforcement learning algorithm based on Stackelberg game. In: Proceedings of the Data Driven Control and Learning Systems, 2017. 727–732
198 Rajeswaran A, Mordatch I, Kumar V. A game theoretic framework for model based reinforcement learning. In: Proceedings of the International Conference on Machine Learning, 2020. 7953–7963
199 Zheng L Y, Fiez T, Alumbaugh Z, et al. Stackelberg actor-critic: a game-theoretic perspective. In: Proceedings of the AAAI Workshop on Reinforcement Learning and Games, 2021. 1–9
200 Huang P D, Xu M D, Fang F, et al. Robust reinforcement learning as a Stackelberg game via adaptively-regularized adversarial training. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2022. 3099–3106
201 Lukes D L, Russell D L. A global theory for linear-quadratic differential games. J Math Anal Appl, 1971, 33: 96–123
202 Engwerda J. Linear quadratic differential games: an overview. In: Proceedings of the Advances in Dynamic Games and Their Applications, 2009. 1–34
203 Shinar J, Turetsky V, Glizer V Y, et al. Solvability of linear-quadratic differential games associated with pursuit-evasion problems. Int Game Theor Rev, 2008, 10: 481–515
204 Weintraub I E, Pachter M, Garcia E. An introduction to pursuit-evasion differential games. In: Proceedings of the American Control Conference (ACC), Denver, 2020. 1049–1066

205 Gibali A, Kelis O. An analytic and numerical investigation of a differential game. Axioms, 2021, 10: 66
206 Huang Y H, Chen J T, Zhu Q Y. Defending an asset with partial information and selected observations: a differential game framework. In: Proceedings of IEEE Conference on Decision and Control (CDC), Austin, 2021. 2366–2373
207 Huang Y H, Zhu Q Y. A pursuit-evasion differential game with strategic information acquisition. 2021. ArXiv:2102.05469
208 Li X, Shi J T, Yong J M. Mean-field linear-quadratic stochastic differential games in an infinite horizon. 2020. ArXiv:2007.06130
209 Garcia E, Casbeer D W, Pachter M, et al. A two-team linear quadratic differential game of defending a target. In: Proceedings of American Control Conference (ACC), Denver, 2020. 1665–1670
210 Song R Z, Wei Q L, Song B. Neural-network-based synchronous iteration learning method for multi-player zero-sum games. Neurocomputing, 2017, 242: 73–82
211 Zhang H G, Wei Q L, Liu D R. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. Automatica, 2011, 47: 207–214
212 Song R Z, Li J S, Lewis F L. Robust optimal control for disturbed nonlinear zero-sum differential games based on single NN and least squares. IEEE Trans Syst Man Cybern Syst, 2020, 50: 4009–4019
213 Lukoyanov N Y. Functional Hamilton-Jacobi type equations with ci-derivatives in control problems with hereditary information. Nonlinear Funct Anal Appl, 2003, 8: 535–555
214 Plaksin A. On Hamilton-Jacobi-Bellman-Isaacs equation for time-delay systems. IFAC-PapersOnLine, 2019, 52: 138–143
215 Meng W J, Shi J T. A linear quadratic stochastic Stackelberg differential game with time delay. 2021. ArXiv:2012.14145
216 Gomoyunov M I. Dynamic programming principle and Hamilton-Jacobi-Bellman equations for fractional-order systems. SIAM J Control Optim, 2020, 58: 3185–3211
217 Moon J, Basar T. Zero-sum differential games on the Wasserstein space. 2019. ArXiv:1912.06084
218 Liu D R, Wei Q L. Multiperson zero-sum differential games for a class of uncertain nonlinear systems. Int J Adaptive Control Signal Process, 2014, 28: 205–231
219 Fridovich-Keil D, Ratner E, Peters L, et al. Efficient iterative linear-quadratic approximations for nonlinear multi-player general-sum differential games. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Paris, 2020. 1475–1481
220 Garcia E, Casbeer D W, von Moll A, et al. Multiple pursuer multiple evader differential games. IEEE Trans Automat Contr, 2021, 66: 2345–2350
221 Xu Y H, Yang H, Jiang B, et al. Multiplayer pursuit-evasion differential games with malicious pursuers. IEEE Trans Automat Contr, 2022, 67: 4939–4946
222 Garcia E, Casbeer D W, Pachter M. Optimal strategies for a class of multi-player reach-avoid differential games in 3D space. IEEE Robot Autom Lett, 2020, 5: 4257–4264
223 Yan R, Duan X M, Shi Z Y, et al. Matching-based capture strategies for 3D heterogeneous multiplayer reach-avoid differential games. Automatica, 2022, 140: 110207
224 Shi J T, Wang G C. A linear-quadratic Stackelberg differential game with mixed deterministic and stochastic controls. 2020. ArXiv:2004.00653
225 Megahed A E M A. The Stackelberg differential game for counter-terrorism. Qual Quant, 2019, 53: 207–220
226 Lee D G, Tomlin C J. Hamilton-Jacobi equations for two classes of state-constrained zero-sum games. 2021. ArXiv:2106.15006
227 Elliott R J, Davis M H A. Optimal play in a stochastic differential game. SIAM J Control Optim, 1981, 19: 543–554
228 Moon J, Duncan T E, Basar T. Risk-sensitive zero-sum differential games. IEEE Trans Automat Contr, 2019, 64: 1503–1518
229 Sun J. Two-person zero-sum stochastic linear-quadratic differential games. SIAM J Control Optim, 2021, 59: 1804–1829
230 Li J, Li W Q, Zhao H Z. On the value of a general stochastic differential game with ergodic payoff. 2021. ArXiv:2106.15894
231 Shi J T, Wang G C, Xiong J. Linear-quadratic stochastic Stackelberg differential game with asymmetric information. Sci China Inf Sci, 2017, 60: 092202
232 Moon J. Linear-quadratic stochastic Stackelberg differential games for jump-diffusion systems. SIAM J Control Optim, 2021, 59: 954–976
233 Sun J R, Wang H X, Wen J Q. Zero-sum Stackelberg stochastic linear-quadratic differential games. 2021. ArXiv:2109.14893
234 Huang J H, Wang S J, Wu Z. Robust Stackelberg differential game with model uncertainty. IEEE Trans Automat Contr, 2022, 67: 3363–3380
235 Zheng Y Y, Shi J T. Stackelberg stochastic differential game with asymmetric noisy observations. Int J Control, 2022, 95: 2510–2530
236 Evans L C, Souganidis P E. Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. Indiana Univ Mathem J, 1984, 33: 773–797
237 Altarovici A, Bokanowski O, Zidani H. A general Hamilton-Jacobi framework for non-linear state-constrained control problems. ESAIM Contr Optim Ca, 2013, 19: 337–357
238 Mitchell I M, Bayen A M, Tomlin C J. A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. IEEE Trans Automat Contr, 2005, 50: 947–957
239 Margellos K, Lygeros J. Hamilton-Jacobi formulation for reach-avoid differential games. IEEE Trans Automat Contr, 2011, 56: 1849–1861
240 Fisac J F, Chen M, Tomlin C J, et al. Reach-avoid problems with time-varying dynamics, targets and constraints. In: Proceedings of the International Conference on Hybrid Systems: Computation and Control, Seattle, 2015. 11–20
241 Asri B E, Lalioui H. Deterministic differential games in infinite horizon involving continuous and impulse controls. 2021. ArXiv:2107.03524
242 Moon J. Linear-quadratic mean field stochastic zero-sum differential games. Automatica, 2020, 120: 109067
243 Sun J R, Wang H X, Wu Z. Mean-field linear-quadratic stochastic differential games. J Differ Equ, 2021, 296: 299–334
244 Hart S, Mas-Colell A. A simple adaptive procedure leading to correlated equilibrium. Econometrica, 2000, 68: 1127–1150
245 Tammelin O. Solving large imperfect information games using CFR+. 2014. ArXiv:1407.5042
246 Brown G W. Iterative solution of games by fictitious play. In: Activity Analysis of Production and Allocation. New York: John Wiley & Sons, Inc., 1951. 13: 374–376
247 Ganzfried S. Fictitious play outperforms counterfactual regret minimization. 2020. ArXiv:2001.11165
248 McMahan H B, Gordon G J, Blum A. Planning in the presence of cost functions controlled by an adversary. In: Proceedings of the International Conference on Machine Learning, Washington, 2003. 536–543
249 Xu X, Zhao Q. Distributed no-regret learning in multiagent systems: challenges and recent developments. IEEE Signal Process Mag, 2020, 37: 84–91

250 Zhang H, Lerer A, Brown N. Equilibrium finding in normal-form games via greedy regret minimization. 2022. ArXiv:2204.04826
251 Lu K H, Li G Q, Wang L. Online distributed algorithms for seeking generalized Nash equilibria in dynamic environments. IEEE Trans Automat Contr, 2021, 66: 2289–2296
252 Meng M, Li X X, Hong Y G, et al. Decentralized online learning for noncooperative games in dynamic environments. 2021. ArXiv:2105.06200
253 Meng M, Li X X, Jie Chen J. Decentralized Nash equilibria learning for online game with bandit feedback. 2022. ArXiv:2204.09467
254 Zhang M X, Zhao P, Luo H P, et al. No-regret learning in time-varying zero-sum games. 2022. ArXiv:2201.12736
255 Daskalakis C, Fishelson M, Golowich N. Near-optimal no-regret learning in general games. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34: 1–13
256 Hsieh Y G, Antonakopoulos K, Mertikopoulos P. Adaptive learning in continuous games: optimal regret bounds and convergence to Nash equilibrium. In: Proceedings of the Annual Conference on Learning Theory, 2021. 2388–2422
257 Zinkevich M, Johanson M, Bowling M, et al. Regret minimization in games with incomplete information. In: Proceedings of the Advances in Neural Information Processing Systems, 2007. 20: 1–8
258 Bowling M, Burch N, Johanson M, et al. Heads-up limit hold'em poker is solved. Science, 2015, 347: 145–149
259 Brown N, Sandholm T. Solving imperfect-information games via discounted regret minimization. In: Proceedings of AAAI Conference on Artificial Intelligence, 2019. 33: 1829–1836
260 Brown N, Lerer A, Gross S, et al. Deep counterfactual regret minimization. In: Proceedings of the International Conference on Machine Learning, 2019. 793–802
261 Li H L, Wang X, Qi S H, et al. Solving imperfect-information games via exponential counterfactual regret minimization. 2020. ArXiv:2008.02679v2
262 Xu H, Li K, Fu H B, et al. AutoCFR: learning to design counterfactual regret minimization algorithms. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2022. 1–8
263 Neller T W, Lanctot M. An introduction to counterfactual regret minimization. In: Proceedings of the Model AI Assignments, the 4th Symposium on Educational Advances in Artificial Intelligence, 2013. 11
264 Muller P, Omidshafiei S, Rowland M, et al. A generalized training approach for multiagent learning. In: Proceedings of the International Conference on Learning Representations, 2020. 1–13
265 Steinberger E. Single deep counterfactual regret minimization. 2019. ArXiv:1901.07621
266 Mertikopoulos P, Papadimitriou C, Piliouras G. Cycles in adversarial regularized learning. In: Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, 2018. 2703–2717
267 Vlatakis-Gkaragkounis E V, Flokas L, Lianeas T, et al. No-regret learning and mixed Nash equilibria: they do not mix. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 1380–1391
268 Daskalakis C, Panageas I. Last-iterate convergence: zero-sum games and constrained min-max optimization. 2018. ArXiv:1807.04252
269 Abernethy J, Lai K A, Wibisono A. Last-iterate convergence rates for min-max optimization. 2019. ArXiv:1906.02027
270 Golowich N, Pattathil S, Daskalakis C, et al. Last iterate is slower than averaged iterate in smooth convex-concave saddle point problems. In: Proceedings of the Annual Conference on Learning Theory, 2020. 1758–1784
271 Conitzer V, Korzhyk D. Commitment to correlated strategies. In: Proceedings of the AAAI Conference on Artificial Intelligence, California, 2011. 632–637
272 Benders J F. Partitioning procedures for solving mixed-variables programming problems. Numer Math, 1962, 4: 238–252
273 Farkas J. Theorie der einfachen Ungleichungen. J für die reine und angewandte Mathematik, 1902, 1902: 1–27
274 Fischetti M, Salvagnin D, Zanette A. Minimal infeasible subsystems and benders cuts. https://www.dei.unipd.it/~fisch/papers/Benders_mis_extended_draft.pdf
275 Gomory R E. Outline of an algorithm for integer solutions to linear programs. Bull Amer Math Soc, 1958, 64: 275–278
276 Land A H, Doig A G. An automatic method of solving discrete programming problems. Econometrica, 1960, 28: 497–520
277 Ruder S. An overview of gradient descent optimization algorithms. 2016. ArXiv:1609.04747
278 Gottipati S, Paruchuri P. A genetic algorithm approach to compute mixed strategy solutions for general Stackelberg games. In: Proceedings of the IEEE Congress on Evolutionary Computation, Krakow, 2021. 1648–1655
279 de Nittis G, Trovo F. Machine learning techniques for Stackelberg security games: a survey. 2016. ArXiv:1609.09341
280 Tran H V. Hamilton-Jacobi Equations: Theory and Applications. Volume 213. Washington: American Mathematical Society, 2021
281 Li M, Qin J H, Freris N M, et al. Multiplayer Stackelberg-Nash game for nonlinear system via value iteration-based integral reinforcement learning. IEEE Trans Neural Netw Learn Syst, 2022, 33: 1429–1440
282 Ontanón S, Synnaeve G, Uriarte A, et al. A survey of real-time strategy game AI research and competition in starCraft. IEEE Trans Comput Intell AI Games, 2013, 5: 293–311
283 Davidai S, Ongis M. The politics of zero-sum thinking: the relationship between political ideology and the belief that life is a zero-sum game. Sci Adv, 2019, 5: 12
284 von Moll A, Garcia E, Casbeer D, et al. Multiple-pursuer, single-evader border defense differential game. J Aerospace Inf Syst, 2020, 17: 407–416
285 Gao X B, Akyol E, Basar T. Communication scheduling and remote estimation with adversarial intervention. IEEE CAA J Autom Sin, 2019, 6: 32–44
286 Na X X, Cole D. Theoretical and experimental investigation of driver noncooperative-game steering control behavior. IEEE CAA J Autom Sin, 2021, 8: 189–205
287 Albert L A, Nikolaev A, Jacobson S H. Homeland security research opportunities. IISE Trans, 2023, 55: 22–31
288 Song X F, Willett P, Zhou S L, et al. The MIMO radar and jammer games. IEEE Trans Signal Process, 2012, 60: 687–699
289 Li H Y, Han Z W, Pu W Q, et al. Counterfactual regret minimization for anti-jamming game of frequency agile radar. 2022. ArXiv:2202.10049
290 Bachmann D J, Evans R J, Moran B. Game theoretic analysis of adaptive radar jamming. IEEE Trans Aerosp Electron Syst, 2011, 47: 1081–1100
291 Paruchuri P, Pearce J P, Marecki J, et al. Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems, 2008. 895–902
292 Garcia E, Von Moll A, Casbeer D W, et al. Strategies for defending a coastline against multiple attackers. In: Proceedings

of the IEEE Conference on Decision and Control (CDC), Nice, 2019. 7319–7324
293 Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Proceedings of the Advances in Neural Information Processing Systems, 2014. 2672–2680
294 Creswell A, White T, Dumoulin V, et al. Generative adversarial networks: an overview. IEEE Signal Process Mag, 2018, 35: 53–65
295 Lelis L H S. Planning algorithms for zero-sum games with exponential action spaces: a unifying perspective. In: Proceedings of the International Conference on International Joint Conferences on Artificial Intelligence, 2021. 4892–4898
296 Liu Q H, Wang Y H, Jin C. Learning Markov games with adversarial opponents: efficient algorithms and fundamental limits. 2022. ArXiv:2203.06803
297 Banik S, Bopardikar S D. Attack-resilient path planning using dynamic games with stopping states. IEEE Trans Robot, 2022, 38: 25–41
298 Henderson H. Cybered competition, cooperation, and conflict in a game of imperfect information. The Cyber Defense Rev, 2021, 6: 43–60
299 Wang K F, Gou C, Duan Y J, et al. Generative adversarial networks: introduction and outlook. IEEE CAA J Autom Sin, 2017, 4: 588–598
300 Lee C W, Kroer C, Luo H P. Last-iterate convergence in extensive-form games. In: Proceedings of the Advances in Neural Information Processing Systems, 2021. 34: 1–13
301 Perolat J, Munos R, Lespiau J B, et al. From Poincaré recurrence to convergence in imperfect information games: finding equilibrium via regularization. In: Proceedings of the International Conference on Machine Learning, 2021. 8525–8535
302 Costikyan G. Uncertainty in Games. Cambridge: MIT Press, 2013
303 Xu L. Learning and planning under uncertainty for green security. In: Proceedings of the International Joint Conference on Artificial Intelligence, 2021. 1–3
304 Kar D, Fang F, Fave F D, et al. A game of thrones: when human behavior models compete in repeated Stackelberg security games. In: Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, 2015. 1381–1390
305 Caballero W N, Lunday B J, Uber R P. Identifying behaviorally robust strategies for normal form games under varying forms of uncertainty. Eur J Operational Res, 2021, 288: 971–982
306 Tsiotras P. Bounded rationality in learning, perception, decision-making, and stochastic games. In: Proceedings of the Handbook of Reinforcement Learning and Control, 2021. 491–523
307 Platzer A. Differential game logic. ACM Trans Comput Logic, 2015, 17: 1–51
308 Iyer M, Gilby B. Modeling an adversarial poacher-ranger hybrid game. https://lfcps.org/course/lfcps21/projects/bgilby_maiai.pdf
309 Brown N, Bakhtin A, Lerer A, et al. Combining deep reinforcement learning and search for imperfect-information games. In: Proceedings of the Advances in Neural Information Processing Systems, 2020. 33: 17057–17069
310 Li K, Xu H, Zhang M, et al. OpenHoldem: an open toolkit for large-scale imperfect-information game research. 2020. ArXiv:2012.06168
311 Oh I, Rho S, Moon S, et al. Creating pro-level AI for a real-time fighting game using deep reinforcement learning. IEEE Trans Games, 2022, 14: 212–220