

## Article

# Games over Probability Distributions Revisited: New Equilibrium Models and Refinements

Stefan Rass <sup>1,2,\*</sup> , Sandra König <sup>3</sup>  and Stefan Schauer <sup>3</sup> <sup>1</sup> LIT Secure and Correct Systems Lab, Johannes Kepler University, 4040 Linz, Austria<sup>2</sup> Institute for Artificial Intelligence and Cybersecurity, University of Klagenfurt, 9020 Klagenfurt, Austria<sup>3</sup> Austrian Institute of Technology, Center for Digital Safety & Security, Giefinggasse 4, 1210 Vienna, Austria

\* Correspondence: stefan.rass@jku.at

**Abstract:** This article is an overview of recent progress on a theory of games, whose payoffs are probability distributions rather than real numbers, and which have their equilibria defined and computed over a (suitably restricted yet dense) set of distributions. While the classical method of defining game models with real-valued utility functions has proven strikingly successful in many domains, some use cases from the security area revealed shortcomings of the classical real-valued game models. These issues motivated the use of probability distributions as a more complex object to express revenues. The resulting class of games displays a variety of phenomena not encountered in classical games, such as games that have continuous payoff functions but still no equilibrium, or games that are zero-sum but for which fictitious play does not converge. We discuss suitable restrictions of how such games should be defined to allow the definition of equilibria, and show the notion of a lexicographic Nash equilibrium, as a proposed solution concept in this generalized class of games.

**Keywords:** game theory; security strategy; generalized game; loss distributions; decision making; Nash equilibrium; lexicographic optimization



**Citation:** Rass, S.; König, S.; Schauer, S. Games over Probability Distributions Revisited: New Equilibrium Models and Refinements. *Games* **2022**, *13*, 80. <https://doi.org/10.3390/g13060080>

Academic Editors: Kjell Hausken and Ulrich Berger

Received: 25 October 2022

Accepted: 22 November 2022

Published: 1 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the year 2015, a theory of games was proposed that does not use real-valued payoffs, but rather achieves optimization using stochastic orders on rewards that are probability distributions. Ever since, a couple of intricacies—even pathologies—have been found in such generalized games, and this article is a compilation of the recent findings, advantages, but also pitfalls to avoid when working with game theory over the abstract space of probability distributions. The theory was originally motivated by applications of games in security risk management, where payoffs are hardly crisp or accurately quantifiable; thus, there is a desire to include uncertainty in the game model already before computing solutions. This uncertainty is not in the actions, as is the case for refined equilibrium concepts such as trembling hands or perfect equilibria, but rather in the rewards themselves. Reducing the random outcomes to their averages (expectations) or other representative statistics necessarily sacrifices information that is, in security risk management, generally scarce already. This creates an additional desire to use information about possible effects of an action to the maximum extent available.

The story behind all our upcoming considerations evolves around a defending player 1 who seeks to guard a system against a rational or irrational attacker from outside. In a typical generic setting, the defender may be the chief security officer (CISO) of some large enterprise, and is committed to business continuity management and the most important line of defense against external competitors and adversaries (hackers, and others).

In this situation, the defender often has to consider multiple goals simultaneously, including damage or loss of sensitive data, the possibility of physical damage to equipment

and people (for example, if a production line gets hacked and a robot damages goods or hurts a person), up to issues of reputation, damage compensations, and many others. A security game in such contexts is thus generally a multi-criteria optimization problem, with hardly quantifiable losses to be minimized. For example, if a production line stops for some time, we may have an average estimate of how much that costs per hour, but this value is not quantifiable in an exact way, since the duration of the outage remains random. Insurance and actuarial science [1,2] offer lots of probabilistic models to describe extreme events (the whole theory of extreme value distributions, up to catastrophe theory, can be useful here), but game theory can become difficult to apply if we need to define real-valued, and hence exact, utility functions for optimization.

Summarizing the challenges, we have the following situation: A player is minimizing losses by taking strategic action. The losses are inherently random, and admit a quantification only up to the point of modeling a probability distribution that is conditional on the other players' actions. Specifically, we are unable to give an exact valuation of payoffs to a player due to intrinsic uncertainty. We assume, however, that we can reasonably model the players' random payoff as a (conditional) probability distribution, with loss distribution models that are assumed to exist and be known for the given application context. As an example (taken from [3]), such a conditional distribution can be the probability to "catch an intruder", conditional on the (random) location of the security guard to be at the same location (by coincidence). Similar games have been proposed for border protection, airport protection and coast guarding [4].

Besides the problem with uncertain revenues, some additional quantities that are difficult to bring into classical game models may be interesting. For example, since games typically optimize an average (expected) payoff, what are the chances of receiving more or less than this optimal average? This is important to know when buying insurance or building up backup resources, and is non-trivial to handle with classical game models. We discuss this situation as a concrete motivation for the generalized class of games in Section 1.3.

### 1.1. Our Contribution

This study compiles recent work, findings and progress towards a theory of games for strategic decision making in non-cooperative situations, when consequences are intrinsically random and real valued (and hence crisp) payoff values are not available or are unreliable. This is essentially different from situations in which the imprecision is in the gameplay, such as trembling hands or perfect equilibria capture [5,6], and is also different from situations in which the uncertainty is about the adversary, such as Bayesian games [7] cover by making assumptions about different types of players, e.g., based on levels of rationality [8]. Defining games with "uncertain opponents" technically leads to uncertain utilities, as we consider, but the optimization itself is generally again over real values being expectations of (conditional) distributions. We intentionally want to avoid derandomization by averaging, since the expectation  $E(U) \in \mathbb{R}$  contains less information than the random variable  $U$ . In situations in which information is notoriously scarce, such as in security, unnecessary loss of information should be avoided. This is a common circumstance in risk management, where qualitative and overly precise valuations of an actions consequence are not only hard, but even explicitly discouraged [9]. As an illustration, let us consider quantitative risk management in engineering. In this context, a risk manager makes a list of possible threats to its assets (enterprise business values, human beings or similar), where each threat, if it manifests in reality, may have an impact (e.g., monetary losses, reputational damages, etc.), and occurs with a likelihood to be estimated. The term "risk" is defined by a simple formula to be

$$\text{risk} = \text{impact} \times \text{likelihood}, \quad (1)$$

and, provided that the risk manager can assign reliable values to both variables, nicely lends itself to optimization and game theory [10] (indeed (1) is interpretable as “expected damage” and can be taken as an adversary’s expected utility). If the threats are caused by a rational adversary trying to hack or otherwise hit a system, the threat list is nothing else than an action set of a hostile opponent. For the defender, engaging as the other player in the game has its own action set to protect against the possible threats. Risk management standards such as ISO27000 or ISO31000 provide long lists of actions (termed “controls” in this context), which are the defender’s actions to choose from strategically. The above formula can then be operationalized by playing a *non-cooperative game* between two players, featuring the defender versus the attacker. The player’s expected utility can be taken directly as the risk = impact  $\times$  likelihood. From the defender’s perspective, the “likelihood” would be the (optimized) probabilities of actions taken by the opponent, which is nothing other than a mixed equilibrium strategy. Likewise, the defender’s equilibrium then becomes an optimal randomized choice of actions to protect against all threats simultaneously. This mixed equilibrium is then convertible into an optimal resource allocation strategy for the defender, thus providing a provably optimal use of (typically scarce) defense budgets, against a rationally and strategically acting adversary. Recent work on game theoretic cyber security has independently pointed out the lack of consensual interpretations of mixed strategies in security, and our work addresses the resource allocation- and mixed strategy interpretation problem, both noted by [11], on a common ground. Reference [3] provided an interesting application for security surveillance monitoring, where the optimization must consider multiple criteria, even including the coverage level of the surveillance, but also the level of inconvenience that surveillance may cause to people at work.

Game theory offers powerful possibilities in letting us quantify likelihoods for risk management directly as optimal choice rules, for example, or interpreting mixed strategies as optimal resource allocation rules. However, the difficulty of applying games over real values is the need to “accurately quantify” the impact, which some authors even discourage in light of much and reliable data to be available. Reference [9] explains this recommendation using an example of the decision about whether or not to buy protection against lightning strikes. Based on long-term statistical evidence, the likelihood of a lightning strike is accurately quantifiable (to be  $\approx 1.24 \times 10^{-6}$  in the region around Munich/Germany), thus providing a seemingly reliable value for the risk Formula (1). Similarly, the impact of a lightning strike is also not difficult to quantify, since we most likely know how expensive a repair to the facility would be. Thus, despite both values being accurately estimable (e.g., 10,000\$), their product numerically evaluates to an expected loss of  $\approx 0.124$ \$. Basing a decision on this value to not invest in lightning protection is clearly implausible, and this example (among many others that are similar), leads to official recommendations to not estimate probabilities or impacts with seeming precision numerically, where there is no absolute accuracy.

The problem appears more widespread than only in the risk-management context, since any strategic decision under uncertain consequences of actions will need a well-founded account for uncertainty. For example, the effects of fake news spreading over social media certainly call for strategic intervention, but the effect of campaigns against fake news and disinformation are easily quantified in terms of costs, but are almost impossible to quantify in terms of effects. This makes the definition of crisp, i.e., real-valued utilities, generally hard in many practical instances of security management, and creates a need for a more general concept of games that can work with vague, fuzzy or generally uncertain outcomes of strategic actions.

This work, and its contributions, are thus motivated by the following five aspects collected from practical project work, which touch upon all of the above points:

First, the quantification of impacts for strategic optimization of actions should be avoided by official recommendations [9]. Our work addresses this by letting games be played over uncertain utility values.

Second, strategic interventions in non-cooperative situations where the response dynamics is unknown, intrinsically uncertain or simply not describable in precise mathematical terms, such as social risk response—for example, [12,13]—may not lend themselves to the definition of crisp values. Agent-based simulations, as one possibility to anticipate social or community replies to actions of authorities, deliver a multitude of possible scenarios, and compiling a single value as a utility to optimize actions and interventions would be accompanied by a considerable loss of valuable information. Games using probability distributions are “conservative” here in using all information embodied in random simulations (e.g., Monte Carlo) for decision making. This also relates our work to empirical game theory, making statistical models such as those in actuarial science [1,14] useful for strategic decision making using games.

Third, optimizing averages offers no information, nor guarantee, about the fluctuations around the optimal mean. If the game is played to minimize the “expected damage”, one may consider buying insurance to cover for such expected losses. However, what if the game overshoots the losses in one or several rounds? How much more than the expected loss should we anticipate to buy insurance for? Real-valued game models are only concerned with optimizing the average, but do not give information about how “stable” this average outcome is. Playing the game over the whole distribution object, we can also consider the variance, skewness and other properties of the distribution in the decision-making process. One such property is the *disappointment rate*, which would lead to a discontinuous utility function in a real-valued instance of a game, and hence Nash equilibria may no longer exist. A contribution of this work is showing how probability distributions for payoffs in a game model can avoid such existence issues, and how to play games not only for the best outcome, but also for the least disappointment. This is a new form of equilibrium refinement, which much other work (not only in security) approaches by perfect equilibria or other (more technical) means.

Fourth, it has long been recognized that the utility maximization paradigm is not necessarily consistent with human decision making [15]. Procedural theories, prospect theory and many other paradigms of decision making have been proposed ever since. As an example, reference [3] has proposed a method to define utilities with regard to subjective risk appetite (aversion or affinity to risk). This construction leads to vector-valued utility functions, and recent work [16] has demonstrated that such games do not necessarily have Nash equilibria. One contribution of this work is showing a proper solution concept in this case.

Finally, along the lines of vector-optimization, multi-criteria games are long known, and concepts such as Pareto–Nash equilibria are definable and useful [17]. However, Pareto optimization needs a weighting among the multitude of criteria to scalarize the utility and thereby reduce the problem to a game with an aggregate, but scalar, utility. Since the lexicographic order is known to not be representable by a continuous real-valued function, this work contributes a method of defining an equilibrium over games with multiple goals in strict priority. A domain of practical application is in robotics, where the first priority is to avoid human coworkers being hurt by a malfunctioning robot, and the cost and functional correctness of the robot come as secondary and ternary goals.

The contributions made here, relative to related work, are as follows:

- The proposal to model utilities not only as numbers, but whole distributions. This was first proposed in [18], and applied in [3,10,19]. The theoretical pathologies arising with this, however, were not discussed in these past studies, until reference [16] first identified theoretical difficulties. This prior reference did not provide solutions for some of the issues raised, which our work does.
- The use disappointment rates in game theory. Prior work did either not consider disappointment in the context of games [20,21], or focused on the applicability of conventional equilibria by means of approximation or using endogenous sharing [22]. Our study differs by describing and proving how to make disappointment rates

continuous, and for the first time, proposes disappointment as a (novel) equilibrium selection criterion.

- We show the *lexicographic Nash equilibrium* as a solution concept that is well defined for vector-valued games that past work has shown to not necessarily have a classical Nash equilibrium [16]. Our work picks up this past example of a game without an equilibrium and shows how to define and compute a meaningful solution.
- We show and explain a case of non-convergence of Fictitious Play (FP) in zero-sum games that has been reported in [23] but was left unexplained ever since in past research. This phenomenon appears, as so-far known, only in games with distributions as utilities, since FP is known to converge for zero-sum games by classical results [24].

### 1.2. Preliminaries and Notation

We let normal font lower-case letters be scalars, and bold-face letters denote vectors. Upper-case letters in normal font denote sets or random variables, and bold print means matrices. The symbol  $X \sim F$  means that the random variable has the distribution  $F$ . For the expectation of  $X$ , we write  $E_F(X)$  or briefly  $E(X)$  if the distribution is unambiguous from the context. For a bounded set  $X$ , we write  $\Delta(X)$  to mean the set (simplex) of all probability distributions supported on  $X$ . If  $X$  is finite with cardinality  $|X| = n$ , we have  $\Delta(X) \subset [0, 1]^n$  as the set  $\Delta(X) = \{(p_1, \dots, p_n) \in [0, 1]^n \mid p_1, \dots, p_n \geq 0; \sum_{i=1}^n p_i = 1\}$ . The symbol  $AS$  will hereafter mean the “action set” of a player. It contains all *pure strategies*, and the corresponding set of *mixed strategies* is denoted as  $S = \Delta(AS)$ , with annotations to make the players explicit. That is,  $AS_i, S_i$  are the  $i$ -th player’s pure and mixed strategies, whereas  $AS_{-i}, S_{-i}$  is the Cartesian product of the strategy spaces of player  $i$ ’s opponents. An asterisk annotation denotes an “optimal” value, e.g., an equilibrium or general best response.

### 1.3. Disappointment Rates

The so-called *disappointment rate* [20–22], is, for a random payoff  $X$ , the probability  $\Pr(X > E(X))$ , where the expectation is with regard to the equilibrium strategy in the game. That is, a minimizing player may not only be interested in the least possible loss  $X$ , but may also seek to minimize chances to suffer more than the expected loss  $E(X)$ . In a security application, where the game is used to minimize expected damage, we may thus face the following challenge:

- The game is designed to advise the defender to best protect its system against an attacker. The optimization will thus be a minimization of the expected loss  $E_{\mathbf{x}, \mathbf{y}}(L)$  for the defender, accomplished at the equilibrium strategy  $\mathbf{x} = \mathbf{x}^*$  and best reply to it  $\mathbf{y}^*$  in a security game model. See [25] for a collection of examples.
- Knowing that it has to prepare for an expected loss of  $v = E(L)$ , the defender may build up backup resources to cover for cases of a loss  $> v$ , one way of which is buying insurance for it. However,  $E(L)$  is an average, and naturally, there may be infinitely many incidents where the loss overshoots the optimized expectation  $v$  that an insurance may cover. Hence, naturally, the defender will strive to minimize the chances for the current loss being  $> v$ . This is where disappointment rates come into the game.

Modeling payoffs as entire probability distributions put more information into the game model and can help in cases in which additional quantities besides the expected payoff are of interest, such as disappointment events. Adding this optimization goal as a second dimension to the game, however, introduces a discontinuity in the payoff function. For a finite (matrix) two-player game, the defender would optimize the function

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n x_i \cdot \mathbb{I}(a_{ij} > \mathbf{x}^\top \cdot \mathbf{A} \cdot \mathbf{y}) \cdot y_j,$$

if the loss, i.e., the highest priority goal, is described by the payoff structure  $\mathbf{A} \in \mathbb{R}^{n \times m}$ . The inner indicator function  $\mathbb{I}$  is a discontinuous part, and as such can render the classical

equilibrium existence results inapplicable. Nash’s theorem and all theory that builds upon it relies on continuity of payoffs, and [26] gave an explicit example of a game with discontinuous payoffs that does not have a Nash equilibrium.

While games with discontinuous payoffs may not admit equilibria in general, computing the disappointment rate for a fixed equilibrium can be as simple as in Example 1.

**Example 1.** Consider the zero-sum two-player game with payoff structure

$$A = \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}$$

whose equilibrium is  $\mathbf{x}^* = (0.8, 0.2)$  and  $\mathbf{y}^* = (0.4, 0.6)$  for player 1 being a minimizer. The saddle point value is  $v = (\mathbf{x}^*)^\top \cdot A \cdot \mathbf{y}^* = 3.6$ . For computing player 1’s disappointment rate  $d$ , only those entries that are larger than  $v$  are relevant; all others count as zero, i.e., we get the matrix of disappointment indicators as

$$D = (\mathbb{I}(a_{ij} > v))_{ij=1\dots 2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The equilibrium average of indicators in  $D$  is the equilibrium disappointment, given by  $d(\mathbf{x}^*, \mathbf{y}^*) = (\mathbf{x}^*)^\top D \mathbf{y}^* = 0.56$ . Thus, when playing the equilibrium strategy, we have a 56% chance that we will “suffer” more than the expected loss  $v$ , and hence are “disappointed”.

In the simple case of matrix games, Example 1 may serve as an *equilibrium selection criterion*; namely, among several possibilities, a player may choose the equilibrium of smallest disappointment. If both players do this, we are back at a minimax optimization problem, yet only on the disappointment rates instead, and over the equilibria that exist in the original game. If there are finitely many equilibria, and each has its uniquely associated disappointment rate, we arrive at nothing else than another finite game about equilibrium selection in a prior game. Two facts about this observation are important to emphasize: first, the disappointment rate cannot serve as a goal in its own right, since if a player is just seeking to avoid disappointment, the disappointment rate can play towards maximizing the losses to avoid being disappointed (effectively making  $\Pr(X > E(X))$  zero and hence minimal). So, disappointments are only meaningful as a secondary goal that depends on a more important utility. This is the second important observation: the idea of first computing an equilibrium in a given game, and then moving on to a secondary game to select from the perhaps many equilibria that we found before, lets us handle multi-criteria decisions with strict priority orders on the goals. In some cases, such as for disappointment rates, the priority ordering is natural, since disappointment can occur about some utility of primary interest. In other cases, however, goals may not sub-quantify other goals, but can be considered on their own and more or less important than other goals (according to a priority ranking). The idea of optimizing lexicographically will hereafter turn out to be fruitful to solve a generalized class of games whose payoffs are more than numbers, and the rewards—not just the mixed strategies—come as whole probability distributions.

One motivation for expressing payoffs as entire probability distributions is, in fact, avoiding the discontinuity problem of disappointment rates, as we show in the following section.

#### 1.4. Making the Disappointment Rate a Continuous Function

We let  $X$  be the random reward from a game in which the players act according to a joint mixed, and hence random, equilibrium strategy  $\mathbf{s}^*$ . In this section, we let  $\mathbf{s}^*$  be a multivariate probability distribution over the Cartesian product of all strategy spaces of all players (thus, including the case of  $n$ -person games with  $n > 2$ ). How can we optimize the quantity  $\Pr(X > E(X)) = E_{\mathbf{s}^*}(X > E_{\mathbf{s}^*}(X))$ , if it is, by expressing it as an indicator function, essentially discontinuous?

Several answers are possible, among them the use of more complex methods such as endogenous sharing [22], or smooth approximations of the disappointment rate, e.g., for two-player games, by changing the indicator to a continuous function such as  $d(\mathbf{x}, \mathbf{y}) = \max\{0, \mathbf{x}^\top \cdot \mathbf{A} \cdot \mathbf{y} - v\}$ , which is zero if we get less than  $v$ , and only has a penalizing effect if more than the expected minimum is paid.

An even more straightforward solution is setting up the game with not only a payoff value for each strategy, but instead defining a whole payoff distribution that is conditional on the players' strategies. That is, for each joint (pure) strategy profile  $(\mathbf{x}_i, \mathbf{x}_{-i})$  of the  $i$ -th player and its group of opponents (denoted by  $-i$  in the subscript), we would classically define the payoff value  $u_i(\mathbf{x}_i, \mathbf{x}_{-i})$ , or its expectation in case of the mixed extension. In both cases,  $u_i$  is typically a real number. However, if we let  $u_i$  be a probability distribution over the set of (real-valued) utilities, we may define  $u_i(\mathbf{x}_i, \mathbf{x}_{-i}) = F_i(X|\mathbf{x}_i, \mathbf{x}_{-i})$  as the distribution function of the random reward to be obtained if player  $i$  picks (pure) strategy  $i$ , and the opponents jointly play the strategy profile  $\mathbf{x}_{-i}$ .

Let us illustrate this modeling for the special case of a finite two-player game: we let  $AS_1, AS_2$  be finite sets of pure strategies hereafter. If the game matrix is populated with conditional payoff distributions  $F_{ij} = F(\cdot|i, j)$ , such that player 1 receives the random payoff  $X \sim F_{ij}$  whenever it takes action  $i \in AS_1$  and Player 2 plays action  $j \in AS_2$ , then the overall payoff, by the law of total probability, for player 1 has the distribution

$$F_s(t) = \sum_{i,j} \Pr(L \leq t|i, j) \cdot \Pr(i, j), \tag{2}$$

where  $\Pr(i, j)$  is the probability of the action profile  $(i, j) \in AS_1 \times AS_2$  chosen at random by the players. Assuming independence of actions, we have  $\Pr(i, j) = \Pr(i) \cdot \Pr(j) = x_i \cdot y_j$ , when  $\mathbf{x}, \mathbf{y}$  denote the mixed strategies. Furthermore,  $\Pr(X \leq x|i, j)$  is exactly a payoff distribution  $F_{ij}$  put into the game's payoff matrix. Thus, we can continue (2) by writing

$$F_s(t) = \sum_{i,j} x_i \cdot x_j \cdot F_{ij} = \mathbf{x}^\top \cdot \mathbf{A} \cdot \mathbf{y}, \tag{3}$$

and we are back at the familiar payoff functional for matrix games, only now having the matrix  $\mathbf{A} = (F_{ij})_{i,j=1}^{n,m}$  defined with all distribution functions for the payoffs, rather than single (crisp) values.

The case of modeling with numbers is (only) included as a special case, since we can take the expected value  $E_F(X)$  of the random variable  $X \sim F(\cdot|\mathbf{x}_i, \mathbf{x}_{-i})$  and receive a classical model that uses only numbers in the payoff structure.

If we now add disappointment rates as an equilibrium selection criterion, i.e., a subordinate goal, we can compute its value from the distributions. Letting  $\mathbf{s} \in \Delta(\prod_{i=1}^n AS_i)$  denote the joint mixed strategy space of all players, the disappointment rate for the  $i$ -th player is

$$d(\mathbf{s}) = \Pr(X > E_s(X)) = 1 - F(u(\mathbf{s})|\mathbf{s}), \tag{4}$$

where  $F$  is the cumulative distribution function of the random payoff  $X$  to player  $i$ , and  $u(\mathbf{s})$  is the utility function for this player. Since the formula would be the same for all players, we omit the indication of (4) for a (fixed)  $i$ -th player in the following.

Now, we are interested in how  $d$  is depending on the joint mixed strategies embodied in  $\mathbf{s}$ . For a distribution-valued modeling, the disappointment rate is continuous again:

**Proposition 1.** *Let the Cartesian product of all strategies in an  $n$ -player game be the set  $\Omega$ , and let  $\Omega \subset \mathbb{R}^n$  be compact. Moreover, let the utility function  $u : \Omega \rightarrow \mathbb{R}$  be continuous. Let each player have its own (possibly distinct) function  $F(\cdot|\mathbf{s})$  that is a conditional distribution of the random reward  $X \sim F(\cdot|\mathbf{s})$  that this player receives when  $\mathbf{s}$  is the joint action profile of all players.*

*Then, the disappointment function  $d$  for each player, as defined in (4), is continuous in  $\mathbf{s} \in \Omega$ .*

The proof of Proposition 1 is given in Appendix A.

Modeling games with payoff distributions rather than payoff values naturally raises the question of whether we can carry over the theory of classical games to this generalized setting. In fact, this transfer is possible via a field extension from  $\mathbb{R}$  to a strict superset of real numbers, in which payoff distributions can be mapped into and totally ordered stochastically. However, the price for this generality is significant, as we discover various unusual, and partly unpleasant, phenomena through this transition. We dedicate the rest of this article to a description thereof.

## 2. Games with Rewards Expressed as Whole Distributions

Following the classical axiomatic construction that von Neuman and Morgenstern have put forth (a very concise and readable account is given in [27], (Section 2)), it is enough to have a certain ordering  $\leq$  on the reward space  $R$ , into which we let all utility functions map. Under certain assumptions on the ordering  $\leq$ , we can define a real-valued utility function  $u$  such that a best decision for a player is characterized as carrying the maximum utility value over possible choices in  $R$ . This is the von Neumann–Morgenstern construction, which, in the way we state it here, requires  $R$  to be a vector space over  $\mathbb{R}$ , and the ordering  $\leq$  should be (i) *complete*, meaning that for all  $\forall r_1, r_2 \in R$ , we either have  $r_1 \leq r_2$ ,  $r_1 > r_2$  or  $r_1 = r_2$ , (ii) *transitive*, meaning that  $r_1 \leq r_2$  and  $r_2 \leq r_3$  implies  $r_1 \leq r_3$ , (iii) *continuous*, meaning that if  $r_1 \leq r_2 \leq r_3$ , then there is some  $p \in [0, 1]$ , such that  $p \cdot r_1 + (1 - p) \cdot r_3 = r_2$ , and (iv) *independent*, meaning that whenever  $r_1 \leq r_2$ , then for any  $s \in R$  and  $p \in [0, 1]$  we have  $p \cdot r_1 + (1 - p) \cdot s \leq p \cdot r_2 + (1 - p) \cdot s$ .

Under these hypotheses, the existence of a (continuous) utility function can be established, such that the utility-maximizing principle to find best decisions becomes applicable. A related similar result is the Debreu representation theorem [28], which asserts the existence of utility functions under a set of slightly different (yet all topological) assumptions on the space  $R$ .

It is not difficult to verify that the axioms of von Neumann and Morgenstern all apply to the real numbers  $\mathbb{R}$  as the reward space, and hence most game models of today are formulated in terms of real numbers.

Now, the idea that first appeared in [23] was to simply replace the space  $\mathbb{R}$  with the space of hyperreal number  ${}^*\mathbb{R}$ , which is a strict extension field of  $\mathbb{R}$  to include infinitesimal and infinitely large numbers. It does so by modeling numbers as sequences  $\hat{a} = (a_n)_{n \in \mathbb{N}} \in {}^*\mathbb{R}$ , with the intuition that an infinitesimally small number would be such that  $\lim_{n \rightarrow \infty} a_n = 0$ , while an infinitely large number  $\hat{a} = (a_n)_{n \in \mathbb{N}}$  have  $\lim_{n \rightarrow \infty} a_n = \infty$ .

**Remark 1** (Notation for the hyperreal space and hyperreal numbers). *It is common in the literature to denote the hyperreal space by a superscript  $*$  preceding  $\mathbb{R}$ , i.e., to use the symbol  ${}^*\mathbb{R}$ . The superscripted star is, however, also commonly used to denote optimal strategies in game theory, such as  $\mathbf{x}^*$  as the best strategy for some players. To avoid confusion about the rather similar notation here, we therefore write  $\hat{x}$  to mean a hyperreal element, and reserve the  $*$ -annotation for variables to mark them as “optima” for some minimization or maximization problem, with the symbol  ${}^*\mathbb{R}$  being the only exception, but without inducing ambiguities.*

The space  ${}^*\mathbb{R}$  is actually a quotient structure, with the equivalence  $\hat{a} = \hat{b}$  to hold if and only if certain sets of indices match. The exact information of which indices matter is given by an *ultrafilter*  $\mathcal{U}$  over  $\mathbb{N}$ . This is a *filter*<sup>1</sup> over  $\mathbb{N}$  that is maximal with regard to  $\supseteq$ . To define  ${}^*\mathbb{R}$ , we additionally require that the intersection of all elements in  $\mathcal{U}$  is empty (so that  $\mathcal{U}$  is called *free*); then  ${}^*\mathbb{R} = \mathbb{R}^\infty / \mathcal{U}$ , where the equivalence relation to define this quotient set is  $(a_1, a_2, \dots) = (b_1, b_2, \dots)$  if and only if  $\{i \in \mathbb{N} \mid a_i = b_i\} \in \mathcal{U}$ . In other words,  $\mathcal{U}$  is a family of subsets that specifies which indices matter for a comparison of two hyperreals  $\hat{a}, \hat{b}$ . Likewise, we can define  $\leq, <, \geq, \dots$  relations, over  ${}^*\mathbb{R}$  in this form. Moreover,  ${}^*\mathbb{R}$  inherits the field structure from  $\mathbb{R}$ , so it has a well-defined arithmetic (addition, subtraction, multiplication and division) just as  $\mathbb{R}$ .



The crucial point about  ${}^*\mathbb{R}$  is made by Łos’ theorem (see, e.g., [29], elsewhere also called the transfer principle [30]), which intuitively states that all first-order logical formulas  $\phi$  that are true in  $\mathbb{R}$  remain analogously true in  ${}^*\mathbb{R}$ , if all variables are replaced by a hyperreal counterpart. In other words,  ${}^*\mathbb{R}$  behaves (logically) just as  $\mathbb{R}$ , and therefore many statements proven in  $\mathbb{R}$  carry over to their hyperreal counterparts in  ${}^*\mathbb{R}$ , by the transfer principle (actually Łos’ theorem).

The idea put forth in [31] was to use  ${}^*\mathbb{R}$  to “encode” a whole probability distribution  $F$  supported on  $\Omega$  by associating it with the hyperreal number that is just the sequence of moments  $\hat{f} = (m_1(F), m_2(F), \dots, m_k(F), \dots)$ , where  $m_k(F) = \int_{\Omega} x^k dF$ , provided that the moments of all orders exist. Then, one would impose the ordering on the hyperreals as an ordering on distribution functions, i.e., put two distributions  $F_1, F_2$  into an order  $F_1 \leq_{hr} F_2$  if and only if  $\hat{f}_1 \leq \hat{f}_2$  holds for their hyperreal representatives, i.e., moment sequences.

**Definition 1** (Hyperreal  $\leq_{hr}$  stochastic order). *Let  ${}^*\mathbb{R}$  be a space of hyperreals (induced by an arbitrary but fixed ultrafilter). Let  $X \sim F, Y \sim G$  be random variables with distributions  $F, G$  that have moments of all orders. To both,  $X$  and  $Y$ , associate the hyperreal number  $\hat{f} = (E(X^k))_{k \in \mathbb{N}}$  and  $\hat{g} = (E(Y^k))_{k \in \mathbb{N}}$ . The hyperreal stochastic order  $\leq_{hr}$  is defined between  $X \leq_{hr} Y$  as  $X \leq_{hr} Y \stackrel{def}{\iff} \hat{f} \leq \hat{g}$  in the given instance of  ${}^*\mathbb{R}$ .*

Applying this idea to game theory is then straightforward:

- We let the payoffs be distributions that have moments of all orders.
- These distributions are replaced by their hyperreal representatives.
- The reward space is  ${}^*\mathbb{R}$ , with its total ordering  $\leq$  on it.
- By the transfer principle,  ${}^*\mathbb{R}$  “behaves like  $\mathbb{R}$ ”, and hence the concepts and results from game theory “carry over”.

Let us apply the transfer principle to the simple equilibrium condition in a finite two-player zero-sum game: let  $\mathbf{A} \in \mathbb{R}^{n \times m}$  be a payoff matrix, and let  $\mathbf{x} \in [0, 1]^n, \mathbf{y} \in [0, 1]^m$  be probability distributions corresponding to mixed strategies. The pair  $(\mathbf{x}^*, \mathbf{y}^*)$  is an equilibrium (for a minimizing row- and maximizing column-player) if and only if

$$\mathbf{x}^\top \mathbf{A} \mathbf{y}^* \geq (\mathbf{x}^*)^\top \mathbf{A} \mathbf{y}^* \geq (\mathbf{x}^*)^\top \mathbf{A} \mathbf{y} \tag{5}$$

The classical result of von Neumann and Morgenstern, later generalized by Nash, establishes the existence of an equilibrium in every finite game (also for more than two players and not necessarily zero-sum). Applying the transfer principle, we can rely on the same assurance that an equilibrium exists, but this time for the hyperreal “version” of (5) stated as follows:

$$\hat{\mathbf{x}}^\top \hat{\mathbf{A}} \hat{\mathbf{y}}^* \geq (\hat{\mathbf{x}}^*)^\top \hat{\mathbf{A}} \hat{\mathbf{y}}^* \geq (\hat{\mathbf{x}}^*)^\top \hat{\mathbf{A}} \hat{\mathbf{y}}, \tag{6}$$

inside which

- $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n), \hat{\mathbf{y}} = (\hat{y}_1, \dots, \hat{y}_n)$  are vectors of hyperreals that satisfy the same conditions as all (categorical) probability distributions do, i.e.,  $\hat{x}_i \geq \hat{0}$  and  $\hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_n = \hat{1}$  (likewise for the  $\hat{y}_i$ s); only to be taken inside the hyperreal space.
- $\hat{\mathbf{A}}$  is the matrix of payoffs, again composed from hyperreal numbers.

For our case of games with probability distributions to replace the utility functions, the hyperreal numbers in  $\hat{\mathbf{A}}$  would simply be the moment sequences of the payoff distributions, which, under the usual regularity conditions on characteristic functions, uniquely represent the distribution  $F_{ij}$  as a hyperreal number. This number, by the transfer principle, behaves “like a real value”, but represents the more complex distribution object. Thus, in playing the game over the hyperreals, we would actually (physically) play the game over whole distributions. This has the appeal of giving a hyperreal version of the whole theory seemingly “for free”, within the richer structure that contains infinite sequences rather than just real numbers. Such sequences are, obviously, able to carry much more information than a single value, and, despite them being “vectors”, can still be totally ordered (a feature

that the finite dimensional space  $\mathbb{R}^n$  would lack for every  $n > 1$ , making vector-valued games much more involved to handle in practice).

### 2.1. Obtaining the Payoff Distributions

Aside from the theoretical possibilities of using a probability distribution as a payoff, the practical question of where to get the distribution from is a separate issue with multiple possible answers. If the game is about strategic protection against natural disasters or threats, actuarial science [1] provides various distribution models that can be instantiated using empirical data. Strategic actions in contexts of very complex dynamics, such as large traffic systems, social networks, networks of critical interdependent infrastructures, or similar, may be open to simulations (event-based, agent-based, Monte Carlo, etc.). Such methods can bring up large amounts of data and different possible outcome scenarios of a single action, which may be compiled into a distribution over a set of possible effects that an action can have. Repeating such a simulation for all choices of a strategic action can deliver a set of distributions to play a game with. Finally, some contexts may just admit subjective probabilities to be specified, up to specifying perhaps whole “ranges” of possible outcomes (e.g., stating that the damage of an action will be “between  $x$ \$ and  $y$ \$, but being no more precise than this; coming to a uniform distribution as the payoff). In most cases of practical risk management guidelines for different countries [32–35], impacts (and hence utilities) are specified on categorical scales, which then comes to categorical distributions being subjectively defined.

The list of possibilities sketched here is not exhaustive or comprehensive, and the entire area of statistical modeling is to be considered as a toolbox to construct probability distributions appropriately, including methods of model diagnostics and goodness-of-fit testing. We leave this as a separate problem of clear relevance, but outside the main scope of this article.

### 2.2. Computing Equilibria Is (Not Automatically) Possible

The hyperreals appear as an easy way of obtaining a total order on the space of probability distributions, while, at the same time, inheriting the rich structure of  $\mathbb{R}$  to define equilibria and other aspects of game theory in  ${}^*\mathbb{R}$  without further ado, just by relying on the transfer principle. The practical caveat, however, comes in at the point of algorithmically computing equilibria, since the transfer from  ${}^*\mathbb{R}$  back into  $\mathbb{R}$  is much less trivial.

To see the challenge, let us return to (6) and take a second look at how the variables are to be understood: While the interpretation of the entries of  $\hat{A}$  is straightforward, as its entries are based on payoff distributions, the interpretation of  $\hat{x}, \hat{y}$  is much less obvious, since these are no longer values but, likewise, infinite sequences of values that represent a “hyperreal probability distribution”. Given the practical difficulties of interpreting the meaning and use of probabilities, using the extended objects here will not likely ease matters of practical decision making. An additional practical difficulty comes in by the structure of  ${}^*\mathbb{R}$  itself, since although arithmetic is well-defined in  ${}^*\mathbb{R}$ , we cannot practically carry it out without knowledge of  $\mathcal{U}$ . Alas, until today, the existence of  ${}^*\mathbb{R}$  has been established only non-constructively, thus making practical arithmetic in  ${}^*\mathbb{R}$  difficult, if not ultimately impossible.

For the computation of equilibria, however, iterative schemes such as FP can compute equilibria (approximately) without requiring arithmetic, and only need a method to algorithmically decide the preference relation, possibly even without alluding to the ultrafilter  $\mathcal{U}$  at all. To this end, we first review how to decide the order  $\leq_{hr}$  as induced by the hyperreals.

### 2.3. The Hyperreal $\leq_{hr}$ as a Tail Order

Without much loss of generality, we can assume the random reward  $X \in \mathbb{R}$  to be  $\geq 1$ , and, for all moments to exist, additionally be bounded, so that the support  $\Omega \subset \mathbb{R}$  is a compact interval. Additionally, to make the Lebesgue–Stieltjes integrals well defined, we

assume a continuous density function if the distribution is itself continuous. This provides the following conditions (see [36]) to define a loss as a random variable  $X \sim F$  with (loss) distribution  $F$  satisfying the conditions (7):

$$\left. \begin{aligned} &0 \leq a \leq X \leq b < \infty \text{ with constants } a, b \in \mathbb{R}, \text{ and} \\ &\text{if } X \text{ is continuous, then it has a continuous density function } f. \end{aligned} \right\} \quad (7)$$

A random variable  $X$  has, under (7), all finite moments, and a respective game could reward a player with a random payoff  $X$ , which the player seeks to strategically minimize. For a risk management game, this would correspond to taking an optimal control against a threat from the adversarial second player by taking strategic action.

To decide the  $\leq_{hr}$ -relation, the following consideration was proposed [23,36]: let two losses  $X \sim F_1, Y \sim F_2$  given with probability densities  $f_1, f_2$ , both categorical or both continuous<sup>2</sup> on the (common) interval  $\Omega = [a, b]$ . Now, suppose that within a left neighborhood of  $b$ , i.e., within an interval  $[b - \varepsilon, b]$  for some  $\varepsilon > 0$ , we have the strict inequality  $f_1(x) < f_2(x)$  for all  $x \in [b - \varepsilon, b]$  (the case when the distributions are categorical here corresponds to the interval  $[b - \varepsilon, b]$  containing at least one up to perhaps several categories). Then, it is a matter of simple algebra to verify the difference of moments  $\int_{\Omega} x^k d(f_2 - f_1) \rightarrow \infty$  as  $k \rightarrow \infty$ ; intuitively explained by the fact that  $f_2$  puts more mass on the tail region  $[b - \varepsilon, b]$  which ultimately will dominate and hence determine the growth rate. More importantly, if the moment sequence  $\hat{a} = (m_k(X))_{k \in \mathbb{N}}$  of  $X$ , grows faster than the moment sequence  $\hat{b} = (m_k(Y))_{k \in \mathbb{N}}$ , then, as hyperreal numbers, they satisfy  $\hat{a} \leq \hat{b}$  and thereby induce the preference  $X \leq_{hr} Y$ . Recalling that the condition that led to this behavior was  $f_2$  putting more mass on the interval  $[b - \varepsilon, b]$  than  $X$  does, we end up with the observation that  $\leq_{hr}$ , as induced by the hyperreal structure, is actually a stochastic tail ordering, which *prefers decisions under which larger losses are less likely*. This makes  $\leq_{hr}$  a good candidate for risk management, which is about controlling large losses (similar control of losses on the small end of the scale is left to be covered by business continuity management). Algorithmically, the decision of whether  $X \leq_{hr} Y$  or  $Y \leq_{hr} X$  would then come to the mere check, which of the two distribution density functions belonging to  $X$  or  $Y$  is larger on an entire right neighborhood  $[b - \varepsilon, b]$ .

The authors of [16] constructed examples of distributions for which this simple decision criterion fails to apply. Specifically, they gave distribution functions that are compactly supported, but for which the difference  $f_1(x) - f_2(x)$  of the density functions oscillates infinitely often as  $x \rightarrow b$  within the interval  $[a, b]$ . Their construction works with piecewise analytic, even infinitely often differentiable distributions and possibly strictly monotone density functions. However, an infinitely often (and hence unboundedly fast) alternation within a finite range is beyond all representability under finite accuracy, such as within a computer, and hence has only limited practical relevance. Moreover, such pathological densities are easy to exclude even analytically during the modeling, if we extend the requirements on a loss distribution according to demand that a random loss  $X \sim F$  with distribution  $F$  is such that [37]:

$$\left. \begin{aligned} &X \text{ satisfies condition (7) and, if it is continuous,} \\ &\text{has a density that is piecewise polynomial on some} \\ &\text{finite partition of the support } \Omega = \overline{\{x \mid f(x) > 0\}} \end{aligned} \right\}, \quad (8)$$

where the overbar means the topological closure.

The requirement of the density function to be composed of only finitely many polynomials precludes the possibility of infinite oscillations, and, at the same time, makes matters of constructing the loss distribution from data much simpler, since we can just approximate histograms by polynomials over different regions of the loss scale. The additional requirement that (8) imposes over (7) is analytically mild, since using Weierstraß' approximation theorem, it is not difficult to show that in fact, every distribution with a compact support can be approximated, up to arbitrary precision, by a distribution that satisfies (8). We give

the precise result here as Lemma 1, letting the proof follow in Appendix B for convenience of the reader.

**Lemma 1** ([37]). *Let  $f : [a, b] \rightarrow \mathbb{R}$  be a continuous probability density function supported on the compact interval  $[a, b] \subset \mathbb{R}$ . Then, for every  $\varepsilon > 0$ , there is a piecewise polynomial probability density  $g$  that uniformly approximates  $f$  as  $\|f - g\|_\infty < \varepsilon$ .*

More importantly, condition (8) restores the interpretation  $\leq_{hr}$  as being a tail order comparison of loss distributions, and at the same time, makes the preference independent of the ultrafilter  $\mathcal{U}$  underlying  ${}^*\mathbb{R}$ . This is especially important, since it renders the ordering “natural” in a way, independent of possibly non-isomorphic instances of  ${}^*\mathbb{R}$ , whose existence has, until today, neither been proven nor refuted. Specifically, (8) will imply that among any two distributions  $f, g$ , one will necessarily dominate the other in some tail region, or the two must be identical. This means that the hyperreal ordering is determined by the moment sequences exhibiting asymptotically different growth rates, meaning that ultimately, one sequence remains strictly larger than the other, putting the respective index sequence into any ultrafilter (since the complement of the index set is finite). Imposing a tail-ordering on the distributions by demanding a piecewise polynomial representation thus makes the ordering (i) easy to decide algorithmically, and (ii) well defined, in the sense of not “depending” on the hyperreals as such, but only depending on the properties of the distributions. For the comparison of two non-degenerate continuous or categorical distributions, we have the following criteria:

C1: Criterion for categorical distributions: let  $f = (p_1, \dots, p_n), g = (q_1, \dots, q_n)$  be categorical distributions on a common and strictly ordered support  $\Omega = \{1, 2, \dots, n\}$ . Then, we have

$$f \leq_{hr} g \iff (p_n, p_{n-1}, \dots, p_1) \leq_{lex} (q_n, q_{n-1}, \dots, q_1), \tag{9}$$

where  $\leq_{lex}$  is the (standard) lexicographic order on vectors, evaluated from left to right<sup>3</sup>.

C2: Criterion for continuous distributions ([37], Proposition 4.3): let  $f$  satisfy condition (8). For every such  $f$ , we can calculate a finite-dimensional vector  $\mathbf{v}_f \in \mathbb{R}^n$ , where  $n$  depends on  $f$ , with the following property: given two density functions  $f, g$  with computed vectors  $\mathbf{v}_f = (v_{f,1}, \dots, v_{f,n})$  and  $\mathbf{v}_g = (v_{g,1}, \dots, v_{g,m})$ , we have  $f \leq_{hr} g$  if and only if  $\mathbf{v}_f \leq_{lex} \mathbf{v}_g$ , taking absent coordinates to be zero when  $n \neq m$ . If the two vectors are lexicographically equal, then  $f = g$ .

These conditions first appeared in [23], with a flaw in the proof later identified by [16], and with a corresponding correction being published as [37].

It is tempting to use these algorithmic criteria in a canonic generalization of Nash- and other equilibria, but as [16] pointed out, even the existence of Nash equilibria for such games on categorical distributions is not trivially assured by Nash’s theorem any more.

Specifically, [16] gave an example of a finite game that has (i) continuous yet vector-valued payoffs, but (ii) has no Nash equilibrium in terms of the ordering condition C1. We will look into this next.

### 3. A Game with Continuous Vector-Valued Payoffs but without a Nash Equilibrium

Consider a  $2 \times 2$  zero-sum game composed of categorical distributions on the common support  $\Omega = \{1, 2, 3\}$ , given by the following payoff structure, with the lexicographically maximizing row player having strategies  $r_1, r_2$  and the lex-minimizing column player having strategies  $c_1, c_2$ . The lexicographic order is herein taken from right to left.

<b><math>\Gamma</math></b>	$c_1$	$c_2$	(10)
$r_1$	(0.3, 0.2, 0.5)	(0.6, 0.3, 0.1)	
$r_2$	(0.8, 0.1, 0.1)	(0.3, 0.2, 0.5)	

The payoff functions can be defined as vector-valued mappings,  $u_1 : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  sending the mixed strategies

$$(\mathbf{x}, \mathbf{y}) \mapsto \begin{pmatrix} (\mathbf{x})^\top \mathbf{A}_1 \mathbf{y} \\ (\mathbf{x})^\top \mathbf{A}_2 \mathbf{y} \\ (\mathbf{x})^\top \mathbf{A}_3 \mathbf{y} \end{pmatrix}^\top,$$

in which

$$\mathbf{A}_1 = \begin{pmatrix} 0.3 & 0.6 \\ 0.8 & 0.3 \end{pmatrix}, \mathbf{A}_2 = \begin{pmatrix} 0.2 & 0.3 \\ 0.1 & 0.2 \end{pmatrix}, \text{ and } \mathbf{A}_3 = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix}$$

are the matrices containing only the first, second and third payoff coordinate (like projections). Observe that the mapping  $u_1$  is hereby *continuous*, since all its coordinate functions are continuous.

The point observed by V. Bürgin et al. [16] was that a Nash equilibrium in this game is necessarily also optimal on the matrix game composed of only the third coordinates, with the payoff matrix  $\mathbf{A}_3 = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix}$ . This game admits only one Nash equilibrium for a maximizing row-player as  $\mathbf{x}^* = \mathbf{y}^* = (0.5, 0.5)$  with an average payoff of  $(\mathbf{x}^*)^\top \cdot \mathbf{\Gamma} \cdot \mathbf{y}^* = (0.5, 0.2, 0.3)$ ; here, we let the multiplication with the vector  $\mathbf{\Gamma}$  of matrices be understood element wise. However, the mixed strategy to this payoff is not optimal, since a component-wise revenue that is lexicographically “larger” is obtained by playing the pure strategy  $\mathbf{x}' = (1, 0)$  to receive  $(1, 0)^\top \cdot \mathbf{\Gamma} \cdot \mathbf{y}^* = (0.35, 0.25, 0.3)$ , contrary to what a Nash equilibrium would suggest. From this, Ref. [16] concluded that there is no Nash equilibrium at all in this game. We can intuitively substantiate this claim by recalling that the (geometric) cone that the lexicographic order induces is not topologically closed (hence, no maximum needs to exist), or alternatively, by remembering that the lexicographic order relation is not continuous<sup>4</sup>, so that it does not admit a continuous utility function representation by von Neumann’s or Debreu’s representation theorems. Nonetheless, by Łos’ theorem, we can safely say that a Nash equilibrium exists as  $(\hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*) \in {}^*\mathbb{R}^n \times {}^*\mathbb{R}^m$ , since Nash’s result carries over into  ${}^*\mathbb{R}$ , with all variables (including the equilibrium) syntactically changed to hyperreals. Apparently, however, the equilibrium  $(\hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*)$  has no correspondence to an equilibrium in  $\mathbb{R}^n \times \mathbb{R}^m$ , and [38] gave possible theoretical explanations as to why this happens.

#### 4. Lexicographic Nash Equilibria as a Solution Concept

To define a proper solution concept for games such as (10), let us take a look as to why the unilateral deviation is apparently rational in this case, although the definition of an equilibrium says exactly the opposite. The point is the assumption that, if player 1 would unilaterally deviate towards  $\mathbf{x}' \neq \mathbf{x}^*$ , player 2 would *keep on playing* the same strategy  $\mathbf{y}^*$ . However, in reality, player 1’s deviation from  $\mathbf{x}^* \rightsquigarrow \mathbf{x}'$  may incentivize player 2 to also change their strategy from  $\mathbf{y}^*$  to a best response to  $\mathbf{x}'$ . In that sense, the two thus get into an (iterated) learning process, which may eventually carry to convergence, but this needs to be proven. We will go into detail about iterative online learning later in Section 4.2, but for now, we continue pondering about how player 2 would react if player 1 tried to get more by playing  $\mathbf{x}' = (1, 0)$ . Indeed, once player 2 notices this, it can instantly change behavior towards playing  $\mathbf{y}' = (0, 1)$ , to make player 1’s payoff now be 0.1, and hence worse than under strategy  $\mathbf{x}^*$ . This is exactly what happened in the game of Section 3, and the observed effect is consistent with an equilibrium being a “stable” situation in which no player has an incentive to deviate. However, this incentive (more precisely its absence) is based on the payoff in a single game, but the game (10) puts the two players into a parallel play of three games.

A traditional method of vector optimization is to scalarize the vector-valued payoffs, and then find an optimal point on the resulting Pareto front. Formally, this means that any deviation for the sake of increasing one goal  $u_i$  will necessarily decrease the payoffs in another dimension  $u_j$ . We adapt this idea to model the above effect of any deviation of a player towards increasing the payoff in the  $i$ -th goal (in the above game, player 1 would have tried to increase the payoff in the coordinate with index  $i = 2$ ), which would enable the second player to increase their own gains in a goal that is “more important” than goal  $i$ .

This is different to a conventional Nash equilibrium, where there is no incentive to play other than optimal, because this would only decrease one’s payoffs. A lexicographic Nash equilibrium  $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbb{R}^n \times \mathbb{R}^m$  for a two-player zero-sum game will:

- Consider several goals  $u_1 \succ u_2 \succ \dots \succ u_d$  in a strict order of priority  $\succ$ , with  $u_1$  being the most important goal, and all these goals mapping mixed strategies into  $\mathbb{R}$ -valued payoffs.
- Provide a strategy such that any attempt to increase the revenue in  $u_i$  will enable the other player to find a strategy  $\mathbf{y}'_i$ , so that the expected payoffs for player 1 become reduced to  $u_j(\mathbf{x}', \mathbf{y}'_i) < u_j(\mathbf{x}^*, \mathbf{y}^*)$  for a more important goal  $u_j \succ u_i$ .

Can we re-frame this intuition as a player’s “best reply” to a known strategy profile of its opponents? An algorithm described in [39] provides us with the first intuition for finite two-player games: call  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_d \in \mathbb{R}^{n \times m}$  the payoff matrices ordered according to their priority for a total of  $d$  goals that two players can control by strategic choices from their strategy sets  $AS_1$  of size  $n$  and  $AS_2$  of size  $m$ . The two players start by computing a (conventional) Nash equilibrium for the 1st, i.e., most important, game using the payoff matrix  $\mathbf{A}_1$  only.

If the game only has a single Nash equilibrium, the procedure terminates, since any unilateral deviation from the equilibrium for  $\mathbf{A}_1$  would only make the respective players’ situations worse, if the second player reacts on this and also deviates. This is exactly what happens in the example of Section 3, and is to be clearly expected in a security setting where players are naturally adaptive to all changes in the behavior of their opponents.

If there is more than one equilibrium, their entirety, as a set, is convex (see [39] for a proof) and compact. So, the set of Nash equilibria lends itself as a (new) strategy space to play the second game with payoff matrix  $\mathbf{A}_2$  over it, and will by construction retain the optimal payoff in the last payoff dimension  $\mathbf{A}_1$ . In other words, we simply use  $\mathbf{A}_2$  as a tie-breaker to refine the so far known set of equilibria. In a different view, contributed by [40], we may think of  $\mathbf{A}_1$  to define the overall game, and think of  $\mathbf{A}_2$  as defining a subgame over the equilibria in  $\mathbf{A}_1$ . In this view, the lexicographic optimization only delivers a subgame perfect equilibrium, and we may repeat this with  $\mathbf{A}_3, \mathbf{A}_4$  until  $\mathbf{A}_d$ .

More formally, let us convert the payoff matrix  $\mathbf{A}$  into a utility function  $u : \Delta(AS_1) \times \Delta(AS_2) \rightarrow \mathbb{R}, (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x}^\top \mathbf{A} \mathbf{y}$ , and hereafter consider the utility functions as priority-ordered goals  $u_1 \succ u_2 \succ \dots \succ u_d$  as before.

To signify the above lexicographic optimization as some fixed point to a best response correspondence, let us iteratively define the (lexicographically) best reply for less important goals to be the best response, conditional on all higher (more important) coordinates having been optimized. For our notation, we will write  $u_{i:j}$  to mean the payoff vector for goals  $1 \leq i, i + 1, \dots, j \leq d$ ; with the convention that  $u_{1:1} := u_1$ . Then, we define the sets of *best replies* for a *minimizing* player  $k$  as

$$BR_{u_{1:i}}(\mathbf{x}_k, \mathbf{x}_{-k}) := \begin{cases} \operatorname{argmin}_{\mathbf{x} \in \Delta(AS_1)} u_1(\mathbf{x}, \mathbf{x}_{-k}) & \text{if } i = 1, \\ \operatorname{argmin}_{\mathbf{x} \in \Delta(BR_{u_{1:i-1}}(\mathbf{x}_k, \mathbf{x}_{-k}))} u_i(\mathbf{x}, \mathbf{x}_{-k}) & \text{for } 2 \leq i \leq d, \end{cases} \quad (11)$$

and look at player  $k$ ’s best reply set  $BR_{u_{1:d}}(\mathbf{x}_k, \mathbf{x}_{-k})$ , in which  $u_{1:d}$  then includes all goal dimensions of this player (we omit the annotation of  $k$  at the utility function of this player here, for ease of notation, but without ambiguity, since no other utilities than those of this player will appear).

**Proposition 2.** *Let a finite two-player zero-sum game be given with pure strategy sets  $AS_1, AS_2$  and vector-valued payoff functions  $\mathbf{u}$  mapping into  $\mathbb{R}^d$  for player 1 and  $-\mathbf{u}$  for player 2. Furthermore, let the coordinate functions of  $\mathbf{u}$  be ordered by decreasing priority (i.e.,  $u_1$  is the most important goals for the players).*

*The correspondence that maps  $(\mathbf{x}, \mathbf{y}) \in \Delta(AS_1) \times \Delta(AS_2)$  to its best reply set  $BR_{u_{1:d}}(\mathbf{x}, \mathbf{y}) \times BR_{-u_{1:d}}(\mathbf{y}, \mathbf{x})$  has a fixed point.*

**Proof.** We need to prove that some  $(\mathbf{x}^*, \mathbf{y}^*)$  is a best-reply to itself, i.e.,

$$(\mathbf{x}^*, \mathbf{y}^*) \in BR_{u_{1:d}}(\mathbf{x}^*, \mathbf{y}^*) \times BR_{-u_{1:d}}(\mathbf{y}^*, \mathbf{x}^*).$$

The proof is by induction over  $d$ : for  $d = 1$ , the game is a conventional matrix game in which a fixed point exists by Nash’s classical existence result about equilibria, since the game is finite. For  $i > 1$ , (11) prescribes find an optimum over the so-far known set of equilibria, any convex combination of which is again an equilibrium<sup>5</sup>. This set is compact, so the game with (modified) strategy spaces  $\Delta(BR_{u_{1:i-1}})$  for player 1, and  $\Delta(BR_{-u_{1:i-1}})$  for player 2, and payoff functions  $u_i, -u_i$ , has again an equilibrium as a fixed point by Glicksberg’s theorem [41]. Since the equilibrium is a best response to itself, this puts  $(\mathbf{x}_i^*, \mathbf{y}_i^*) \in BR_{u_{1:i}}(\mathbf{x}^*, \mathbf{y}^*) \times BR_{-u_{1:i}}(\mathbf{y}^*, \mathbf{x}^*)$ , and completes the induction step.  $\square$

Formally, we have the following definition appearing in [37] earlier, but now reframed into a fixed-point formulation:

**Definition 2** (Lexicographic Nash equilibrium). *Let  $\{\mathbf{A}_i \in \mathbb{R}^{n \times m} \mid i = 1, 2, \dots, d\}$  be a finite collection of matrices that define payoff functions  $\mathbf{u}$  in a two-player zero-sum game, all over the same strategy spaces  $AS_1, AS_2$  for the players, and listed in descending order of priority (i.e.,  $\mathbf{A}_1$  is the most important, and  $\mathbf{A}_d$  is the least important goal dimension). We call a strategy profile  $(\mathbf{x}^*, \mathbf{y}^*) \in \Delta(AS_1) \times \Delta(AS_2)$  a lexicographic Nash equilibrium in mixed strategies, if it is best reply to itself, i.e., a fixed point of the best response correspondence  $(\mathbf{x}, \mathbf{y}) \mapsto BR_{u_{1:d}}(\mathbf{x}, \mathbf{y}^*) \times BR_{-u_{1:d}}(\mathbf{y}^*, \mathbf{x}^*)$ , with  $BR$  defined by (11).*

Proposition 2 then establishes the non-emptiness of Definition 2, and can be interpreted as the assertion that a lexicographic Nash equilibrium is indeed a best reply (in the sense of (11)) to itself.

By (11), the one-dimensional case is the requirement that

$$(\mathbf{x}^*, \mathbf{y}^*) \in \left[ \underset{\mathbf{x} \in \Delta(AS_1)}{\operatorname{argmin}} u(\mathbf{x}, \mathbf{y}^*) \right] \times \left[ \underset{\mathbf{y} \in \Delta(AS_2)}{\operatorname{argmin}} -u(\mathbf{x}^*, \mathbf{y}) \right],$$

which is equivalent to  $u(\mathbf{x}^*, \mathbf{y}^*) \leq u(\mathbf{x}, \mathbf{y}^*)$  for all  $\mathbf{x}$  and  $u(\mathbf{x}^*, \mathbf{y}) \geq u(\mathbf{x}^*, \mathbf{y}^*)$  for all  $\mathbf{y}$  or more compactly,

$$u(\mathbf{x}, \mathbf{y}^*) \geq u(\mathbf{x}^*, \mathbf{y}^*) \geq u(\mathbf{x}^*, \mathbf{y}), \tag{12}$$

i.e., the usual definition of an equilibrium; cf. (5). However, as [16] demonstrated for the game from Equation (10), we cannot simply rewrite (12) to use  $\leq_{\text{lex}}$  instead of  $\leq$ , since we may lose the existence of equilibria. The concept of a lexicographic Nash equilibrium from Definition 2, together with Proposition 2 provides us with an equilibrium concept that does exist, at least in the zero-sum case.

**Remark 2.** *Alternatively to an equilibrium, a player can also look for a security strategy [42,43], which is the best that they can do, presuming any behavior of the opponent. For a general  $n$ -person game with strategy space  $AS_i$  and utility function  $u_i$  for the  $i$ -th player, a (mixed) security strategy is found as the solution to the following optimization problem:*

$$\mathbf{a}_i^* \in \underset{\mathbf{a} \in \Delta(AS_i)}{\operatorname{argmin}} \min_{\mathbf{a}_{-i} \in \Delta(AS_{-i})} u_i(\mathbf{a}, \mathbf{a}_{-i}), \tag{13}$$

In two-player games, a security strategy is computable by a player replacing the opponent's utility function  $u_2$  simply by the negative version of its own utility to assume a worst-case behavior. That is, the game may be non-zero sum in allowing the players to have distinct utilities  $u_1 \neq -u_2$ , and a security strategy for player 1 is simply a Nash equilibrium in the substitute zero-sum game assuming  $u_2 := -u_1$  for player 1, and conversely for the other player.

The point made here is that a security strategy is, in some cases, a fixed point of some properly modified best-response correspondence (namely that of the assumed zero-sum game). Proposition 2 makes the same fixed point assertion about lexicographic Nash equilibria, thus establishing a likewise similarity between classical security strategies, and lexicographic security strategies from past literature [39].

As for the computation of equilibria, we may consider direct procedures using linear optimization, but also online learning schemes such as fictitious play. It turns out that the latter, unlike in the classical case, does not necessarily converge, at least not to a Nash equilibrium in the hyperreal space.

#### 4.1. Convergence (Failure) of FP in Zero-Sum Games

Returning to the question of whether an online learning process will carry the players to convergence, we discover another unusual behavior of games played over probability distributions: FP does not necessarily converge even for zero-sum games, in contrast to what we would expect according to classical results [24].

This failure of convergence is easily demonstrated by another  $2 \times 2$ -game reported in [23], and given later as Example 2. Before demonstrating the problem, let us briefly recall the idea behind FP: it is essentially an endless repetition of the game in which each player keeps record of the opponent's moves, and correspondingly plays a best reply to the empirical mixed strategy that was observed so far. We delegate a detailed description to Appendix C. J. Robinson has shown [24] that this process will always carry to convergence if the game is zero sum. Later, the convergence was also established for various other classes of games, including non-degenerate  $2 \times 2$ -games, potential games, games that can be solved by iterated elimination of dominated strategies,  $2 \times n$ -games, and several others.

Example 2 shows that the convergence no longer holds if the game is played with distributions as rewards instead of numbers, although it is still zero sum.

**Example 2** (FP not converging to an equilibrium, although the game is zero sum [23]). Let the expected payoffs in a hypothetical  $2 \times 2$ -game be given as

$$\mathbf{A} = \begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} \quad (14)$$

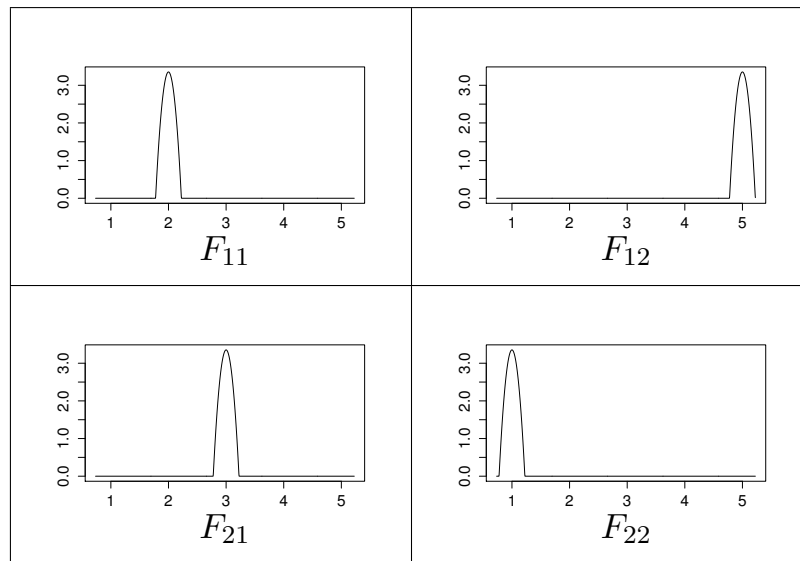
and let us assume that uncertainty is to be included here by allowing a bounded random deviation from these means. This variation is describable in various ways, but let us assume that the modeler has collected empirical data to construct a non-parametric estimator using Epanechnikov kernels

$$k(x) := \begin{cases} \frac{3}{4}(1 - x^2), & |x| \leq 1 \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

and replaces the entry  $a_{ij}$  by the distribution  $F_{ij}$  with mean  $a_{ij}$  and some small variance.

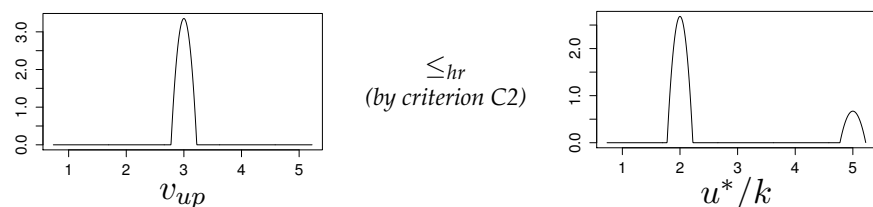
The respective game with distributions as payoffs thus is again a  $2 \times 2$ -matrix, but with probability density functions in the cells; see Figure 1.





**Figure 1.** Distribution-valued game as a version of (14) with uncertainty around the expected payoffs (on the abscissa of each probability density).

It is easy to compute the Nash equilibrium for the exact matrix  $\mathbf{A}$  as  $v(\mathbf{A}) = 2.6$ , obtained by the mixed equilibrium strategies  $\mathbf{x}^* = (0.4, 0.6)$  and  $\mathbf{y}^* = (0.8, 0.2)$  for both players. Since our setting shall merely capture our uncertainty about the payoffs, we would thus naturally expect a somewhat similar result when working on the payoff distributions in Figure 1. Unfortunately, however, running FP according to the algorithm from Appendix C demonstrably fails. After a few iterations, the algorithm gets stuck in always choosing the first row for player 1, since  $v_{up}$  is always  $\leq_{hr}$ -preferable over  $u^*/k$ , which is immediately obvious from plotting the two distributions:



Observe that choosing the upper row in the payoff structure adds probability mass to lower damages, but leaves the tail of the distribution unchanged. Thus, although the overall damage accumulates, this effect is not noticeable by the stochastic ordering based on the hyperreal  $\leq_{hr}$  relation. Consequently, the algorithm will counter-intuitively come to the conclusion that  $\mathbf{x} = (0, 1)$  is a pure equilibrium, which is not plausible or meaningful.

Where does the process fail? The answer is found by a deeper inspection of Robinson’s proof [24], bearing in mind that all variables are to be taken as hyperreals. The sequence enters an  $\hat{\epsilon}$ -neighborhood for  $\hat{\epsilon} > 0$  once the iteration count exceeds the lower bound  $\hat{n}_0 = 8\hat{a}\hat{t}^*/\hat{\epsilon}$ , where  $\hat{a}$  is the maximum (hyperreal) absolute value of the entries in the payoff matrix. Since this matrix contains all distributions, the respective  $k$ -th order moment sequences diverge towards infinity. Hence,  $\hat{a}$  is an infinitely large hyperreal, making  $\hat{n}_0$  itself infinitely large. Hence, no iteration counter from within  $\mathbb{N}$  can ever exceed  $\hat{n}_0$ , which rules out convergence towards a Nash equilibrium for any algorithm running on a conventional computer, although the sequence would converge if it runs towards hyperreal infinity.

However, if we assume that the distributions all have a common support, i.e., there are no regions with zero probability mass over a whole interval  $\Omega = [a, b] \subset \mathbb{R}$ , then the individual preferences will depend on the values that the density functions in  $\mathbf{A}$  take on at  $x = b$ . That is, if the matrix  $\mathbf{A} = (f_{ij}(x) : \Omega \rightarrow \mathbb{R}) \in \mathcal{F}^{n \times m}$  is composed from all densities, then running FP boils down to running the iteration on the real-valued matrix

$\mathbf{A}(b) = (f_{ij}(b)) \in \mathbb{R}^{n \times m}$ . The game remains zero sum, and hence the process *does converge* according to [24].

What we have discovered is thus a slightly unexpected case of a sequence (of hyperreal numbers) that is convergent, but nonetheless admits subsequences that do not converge to the limit. Based on the example from Section 3, we may not reliably claim that the limit to which FP converges for a distribution-valued game is a Nash equilibrium, but we can identify it as a lexicographic Nash equilibrium by looking at what the process of online learning does in more detail.

#### 4.2. Online Learning Lexicographic Nash Equilibria in Zero-Sum Games

Without loss of generality, let us consider an arbitrary payoff distribution to be represented as a vector of values  $(p_1, \dots, p_d)$ , which are either direct probability masses for a categorical distribution (see criterion C1 above) or otherwise representative values computable for a continuous distribution (see criterion C2 above). So, the payoff matrix for a game with payoffs represented as full probability distributions is a matrix of vectors  $\mathbf{A} = (\mathbb{R}^d)^{n \times m}$ , and in which the payoffs are strictly ordered lexicographically.

The matrix  $\mathbf{A}$  of vectors translates into a total of  $d$  real-valued matrices, just as in the setting just described, only that the projected matrices  $\mathbf{A}_1, \dots, \mathbf{A}_d$  corresponding to the respective  $(d + 1 - i)$ -th coordinates of the probability mass vectors appear in reverse order of priority (so that  $\mathbf{A}_1$  is the matrix of all tail masses  $p_d$  for each strategy combination  $i, j$ ,  $\mathbf{A}_2$  is the matrix of masses  $p_{d-1}$  for each strategy combination, and so on). Letting the players be minimizers, they will, among two choices, prefer the distribution with less tail mass. This means nothing other than letting the FP process initially run on the matrix  $\mathbf{A}_1$  of last coordinates only, and from there, it will converge to a best reply. Once FP has converged, it will continue by changing strategies again to optimize the value on the second-largest coordinate. However, exactly this update forces the algorithm to then return to the last coordinate again and restore the optimum there. Here, we have another explanation as to why FP cannot converge over a countable sequence of iterations, since it would take a theoretical infinitude of steps to optimize the last coordinate, followed by a single optimization step on the second-to-last coordinate, and then taking the next infinitude of steps to restore optimality on the last coordinate again. Fortunately, we can prevent the computation from having to re-optimize over and over again, if we compute the equilibrium using plain linear programming first, and then add the constraint of not losing the optimal gains on the previous coordinate when moving to the optimization of the next coordinate.

#### 4.3. Exact Computation of Lexicographic Nash Equilibria in Two-Player Zero-Sum Games

We can directly translate (11) into a series of optimization problems, with more and more constraints being added: initially, on the first coordinate, we just set up a standard linear program,

$$\left. \begin{array}{l} \text{minimize } v \\ \text{subject to } v \geq \sum_{i=1}^n (\mathbf{A}_1)_{ij} \cdot x_i \quad \forall j = 1, \dots, m; \\ x_1 + x_2 + \dots + x_n = 1; \\ x_i \geq 0 \quad \forall i = 1, 2, \dots, n; \end{array} \right\} \quad (\text{LP})$$

in which the symbol  $(\mathbf{A}_1)_{ij}$  denotes the  $ij$ -th entry of the payoff matrix  $\mathbf{A}_1 \in \mathbb{R}^{n \times m}$ . To showcase how to proceed to the next coordinate, let us simplify and thereby better distinguish the payoff matrices by putting  $\mathbf{A} := \mathbf{A}_1 \in \mathbb{R}^{n \times m}$  and writing  $\mathbf{B} := \mathbf{A}_2 \in \mathbb{R}^{n \times m}$ , to mean the most important (matrix  $\mathbf{A}$  of tail masses  $p_d$ ) and second-most important payoff ( $\mathbf{B}$  of probability masses  $p_{d-1}$ ).

A more compact version of (LP) in matrix form, using the symbols  $\mathbf{1}_{p \times q}$  to denote  $(p \times q)$ -matrix of all 1 s, is

$$\begin{aligned} & \text{minimize} && \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\top \cdot (v_1, \mathbf{x}_1) && (16) \\ & \text{subject to} && \begin{pmatrix} \mathbf{1}_{m \times 1} & -\mathbf{A}^\top \\ 0 & \mathbf{1}_{1 \times n} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \mathbf{x}_1 \end{pmatrix} \begin{matrix} \geq 0 \\ = 1 \end{matrix} \end{aligned}$$

where we added the subscript “1” as a reminder that  $v_1$  is the saddle-point payoff in the (most important) coordinate, and attained by playing the strategy  $\mathbf{x}_1^*$ . This corresponds to solving the “ $i = 1$ ”-case of (11).

Next, we go for the second (butlast) coordinate with payoff structure **B**. We end up with the same linear programming problem (LP) as above, but in addition, the strategy  $\mathbf{x}_2^*$  computed in this new LP should, on the payoff structure **A**, perform at least as well as  $\mathbf{x}_1^*$  so far, meaning that we want all coordinates of  $\mathbf{x}_2^\top \cdot \mathbf{A} \geq v_1$ .

Intuitively, any behavior different from  $\mathbf{x}' \neq \mathbf{x}_1^*$  would end up with a possibility for the opponent to cause less revenue than  $v_1$  for the first player by the equilibrium property (5), meaning that  $(\mathbf{x}')^\top \cdot \mathbf{A} \cdot \mathbf{y}^* \geq (\mathbf{x}_1^*)^\top \cdot \mathbf{A} \cdot \mathbf{y}^*$ , since player 1 is a minimizer.

So, among the remaining equilibria, we would pick one that optimizes the second-to-last coordinate, while it retains the performance in the coordinate(s) that we had so far. This solves the “ $i = 2$ ”-case of (11) if player 1 adds the constraint  $(\mathbf{x}_2)^\top \cdot \mathbf{A} \cdot \mathbf{e}_i \leq v_1$  for  $i = 1, 2, \dots, m$ , in which  $\mathbf{e}_i$  is the  $i$ -th unit vector, to the program. The next linear program that player 1 solves is thus

$$\begin{aligned} & \text{minimize} && \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\top \cdot (v_2, \mathbf{x}_2) && (17) \\ & \text{subject to} && \begin{pmatrix} \mathbf{1}_{m \times 1} & -\mathbf{B}^\top \\ 0 & \mathbf{1}_{1 \times n} \\ 0 & \mathbf{A}^\top \end{pmatrix} \cdot \begin{pmatrix} v_2 \\ \mathbf{x}_2 \end{pmatrix} \begin{matrix} \geq 0 \\ = 1 \\ \leq v_1 \cdot \mathbf{1}_{m \times 1} \end{matrix} \end{aligned} \tag{18}$$

where, likewise, we have the saddle point  $v_2$  on the second coordinate, by playing strategy  $\mathbf{x}_2$ , which remains optimal for the first game due to  $\mathbf{A}^\top \cdot \mathbf{x}_2$  giving at most the value  $v_1$  whatever player 2 does on its  $m$  column strategies. Observe that the goals (16) and (17) are only different in terms of the index, which we can safely drop in an implementation; the only change relates to the constraint matrix that grows by yet another block  $(0 \ \mathbf{A}_i^\top)$  upon moving to the  $i$ -th coordinate for  $i = 2, 3, \dots, d$ . The additional lines in (18) thus implement the condition to optimize over  $\mathbf{x} \in \Delta(BR_{u_1, i-1})$  in (11).

Each of these linear programs remains feasible, since the first solution  $\mathbf{x}_1$  is trivially feasible for (18), and so on, so that we will finally obtain an optimum  $\mathbf{x}_d = \mathbf{x}^*$  as one part of the fixed point that Proposition 2 guarantees. The complementary other part  $\mathbf{y}^*$  is obtained by solving the sequence of dual programs to (16) and (17). Along these lines, we obtain a sequence of saddle point values  $v_1, v_2, \dots, v_d$  that are best accomplished under the assumption that player 2 will adapt to a change in behavior observed for player 1. This adaptation is reflected by letting player 1 anticipate it and hence add the constraint to retain its so-far-accomplished saddle point payoffs, appearing as the added constraint in (17). Without this constraint, player 1 would simply run another equilibrium computation on the second coordinate, but just as Section 3 has shown, player 2 can then adapt to it, and reduce the payoffs for player 1 in the first game on  $\mathbf{A}_1$ . This possibility is removed by the added constraint in (17).

This process has been practically implemented in a package for R [44]. Concerning the complexity, the repeated linear programming has, using interior point methods, an overall polynomial complexity in the size of the game and number of dimensions to optimize. On the contrary, fictitious play has worse complexity than in the classical case (where it is worst-case exponential already), but may take even longer if the process is played over vectors or distributions. For this reason, the practical recommendation and implementation is using linear programming in the finite case.

**Remark 3.** Reference [37] has defined a lexicographic Nash equilibrium as a situation in which an unilateral deviation of player 1 will enable player 2 to adapt to this, and reduce player 1's revenue accordingly. Prior work has not explained why this can be called an "equilibrium", since the typical definition of a Nash equilibrium assumes unilateral deviations to be not responded to, i.e., the optimum holds given that all other players keep their strategies (even if mixed) unaltered. In Proposition 2, we have provided an alternative fixed-point characterization of a lexicographic Nash equilibrium, which we believe better justifies the naming as "equilibrium" in the sense of a best reply for all players, than [37] does.

## 5. Summary

Defining games with payoffs that are entire distributions rather than just real numbers has a variety of pitfalls to avoid, but is, to the best of our current knowledge, possible according to the following steps:

- Define the payoff distributions over a joint and compact support  $\Omega = [a, b]$  that is a closed interval over the reals, or a common discrete interval  $[a, a + 1, \dots, b - 1, b] \subset \mathbb{N}$ , with  $a \geq 1$  in both cases.
- If the payoff distributions are continuous, they should either have piecewise polynomial density functions, or be approximated by such functions, which is possible up to arbitrary precision by Lemma 1.
- The games themselves may not admit Nash equilibria as mixed strategies with real-valued probabilities, but do have lexicographic Nash equilibria that are efficiently computable by a series of linear optimizations (see Section 4.3) using existing software [44], at least for finite zero-sum games. Care has to be taken to not misinterpret the resulting strategy profiles as conventional Nash equilibria. In particular, the solution does not need to be lexicographically optimal (as shown in Section 3), but is a best reply to itself under an accordingly modified reply-correspondence (11).

## 6. Outlook

Some authors [16] have argued to leave the route of game theory over hyperreal spaces aside, as it appears to be a dead end due to practical difficulties. We instead propose taking these difficulties as challenges for further research, substantiated by the diverse phenomena encountered along these lines, as there are several interesting things to discover.

### 6.1. Phenomena That May Merit Future Studies

- Convergence failure of FP in the sense that the iteration may converge to an accumulation point that is not an equilibrium, despite the fact that the iteration, when carried out towards hyperreal infinity, would converge to an equilibrium; only it is not reachable from within iterating in  $\mathbb{N}$ , resp.  $\mathbb{R}$ . This phenomenon is due to the fact that  $\mathbb{N}$  is a strict subset of its hyperreal counterpart  ${}^*\mathbb{N}$ , in which FP does converge (by the transfer principle).
- Games that lack a classical Nash equilibrium, although they do have continuous payoffs, with the only difference that they are vector valued. In the absence of conventional Nash equilibria, lexicographic optimization and equilibria, as proposed in Definition 2, may be a solution concept to consider. Intuitively, we hereby treat multiple goals as equilibrium selection criteria, signifying one goal as the most important, and using the subordinate goals to merely refine the equilibrium strategies if they are ambiguous.

Overall, the study of games with distributions as rewards has led to new thoughts on equilibrium selection, and solution concepts for multi-criteria games up to algorithms to compute them. In addition to the above, we get extended possibilities due to the modeling of payoffs as whole distributions. Concepts such as disappointment rates may be useful, but remain to be studied, as explanatory mechanisms of bounded rationality. Speculating on why people may, in reality, choose to deviate from the utility maximization paradigm [15], the anticipation of disappointment may be one possible cause. Moreover, also formally,

disappointment rates are an example of functions that are discontinuous under the classical modeling paradigm, but can be made continuous if the game is defined with distributions (see Section 1.4).

The concept of a lexicographic Nash equilibrium is in itself a simple method of equilibrium refinement by tie-breaking using more properties of the payoff distributions that would not be encoded if the game is defined with crisp values. For example, we cannot only optimize the expected revenue, but also the fluctuation (variance) around it to select the “most stable” equilibrium if there are several.

## 6.2. Conclusions

Generalizations of the theory laid out in this work are straightforward and manifold, yet can be expected to be non-trivial. For example, Proposition 2 is given for two-player zero-sum games, but open to extend to  $n$ -player games and the non-zero-sum case (Note 5 shows the argument that could fail in the more general case). Likewise, the algorithms to compute Nash equilibria (and hence also lexicographic Nash equilibria) are open for extensions to the generalized setting.

Given the intricacy of phenomena that occur in hyperreal games, we argue, on the contrary to previous authors (see citations above), that the so-far discovered oddities justify indeed a deeper study of games inside the hyperreal space, even if just for the sake of scientific curiosity.

**Author Contributions:** Validation, S.S.; Formal analysis, S.R. and S.K.; Investigation, S.K. and S.S.; Writing—original draft, S.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** Open Access Funding by the University of Linz.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We are grateful for many helpful discussions and contributions by Jeremias Epperlein, Vincent Bürgin, Fabian Wirth, Tatjana Cobit, Peter Occil and Jasmin Wachter. A special thanks goes to Paul Schweinzer, for having drawn our attention to subgame perfectness as a concept related to the lexicographic equilibrium notions in this work, and we are also indebted to Ali Alshawish, whose insightful testing and applications of the theory led to invaluable insights and new ideas. The input of the aforementioned people, discussions and deep thinking about the theory of games with distributions as payoffs has been invaluable to improve the theory over earlier publications. Finally, we are indebted to the anonymous reviewers, who gave very good comments and suggestions to improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

FP! Fictitious Play.

## Appendix A. Proof of Proposition 1

The proof given here has been taken from [45]. Intuition suggests that  $d$  is a composition of continuous functions, and as such should be itself continuous. We make this rigorous: take distinct  $s, s' \in \mathbb{R}^{n+m}$  and consider  $d(s') - d(s)$ , which equals

$$\begin{aligned} F_s(u(s)) - F_{s'}(u(s)) &= F_s(u(s)) - F_{s'}(u(s')) + F_{s'}(u(s)) - F_{s'}(u(s)) \\ &= \underbrace{[F_s(u(s)) - F_{s'}(u(s))]}_{\text{term (I)}} - \underbrace{[F_{s'}(u(s')) - F_{s'}(u(s))]}_{\text{term (II)}} \end{aligned}$$

We look at the left and right terms in square brackets separately:

**Term (I):** put  $t = u(s) = E_{F_s}(L)$  and assume  $\|s' - s\|_\infty < \delta_1$ . We have  $\sum_{ij} x_i y_j F_{ij}(t) - \sum_{ij} x'_i y'_j F_{ij}(t) = \sum_{ij} (x_i y_j - x'_i y'_j) F_{ij}(t)$  and furthermore  $|x_i y_j - x'_i y'_j| = |x_i y_j - x'_i y'_j - x'_i y_j + x'_i y_j| \leq |y_j(x_i - x'_i)| + |x'_i(y_j - y'_j)| \leq \delta_1 |y_j| + \delta_1 |x'_i| \leq 2\delta_1$ , because  $|y_j|, |x'_i| \leq 1$  for every categorical distribution (to which  $y_j$  and  $x'_i$  are the respective probability masses).

Applying this bound on the sum gives

$$\begin{aligned} \left| \sum_{ij} x_i y_j F_{ij}(t) - \sum_{ij} x'_i y'_j F_{ij}(t) \right| &= \left| \sum_{ij} (x_i y_j - x'_i y'_j) F_{ij}(t) \right| \\ &\leq \sum_{ij} |x_i y_j - x'_i y'_j| F_{ij}(t) \\ &\leq 2\delta_1 \cdot \sum_{ij} F_{ij}(t) \leq 2\delta_1 \cdot n \cdot m =: \varepsilon, \end{aligned}$$

since  $F_{ij}(t) \leq 1$  as  $F_{ij}$  is a (cumulative) distribution function. Since  $n, m$  are constants, we can require  $\|s' - s\|_\infty < \delta_1 := \frac{\varepsilon}{2mn}$  to imply  $|F_s(u(s)) - F_{s'}(u(s))| < \varepsilon$  for every  $\varepsilon > 0$ .

**Term (II):** the mapping  $s' = (x', y') \mapsto F(t) = \sum_{ij} x'_i y'_j F_{ij}(t)$  delivers a function, which is uniformly continuous as a function of  $t$  on its entire support, since:

- All  $F_{ij}$  were assumed to be continuous and compactly supported,
- So they are all uniformly continuous, and  $\varepsilon > 0$  admits individual values  $\delta_{ij} > 0$ , so that  $|t - t'| < \delta_{ij}$  implies  $|F(t) - F(t')| < \varepsilon$  for all  $t, t'$  within the support.
- Since the game is finite, we can put  $\delta_2 := \min\{\delta_{ij}\}$  to conclude that  $\|t - t'\|_\infty < \delta_2$  further implies  $|F_{s'}(t) - F_{s'}(t')| < \varepsilon$ , i.e., the function  $F$  is, independently of  $s'$  continuous with the same  $\varepsilon, \delta_2$ , since:  $F_s(t) = \sum_{ij} x_i y_j F_{ij}(t)$  and so  $|F_s(t) - F_{s'}(t)| \leq \sum_{ij} x'_i y'_j |F_{ij}(t) - F_{ij}(t')| < \sum_{ij} x'_i y'_j \cdot \varepsilon = \varepsilon \cdot \sum_{ij} x'_i y'_j = \varepsilon$ , as long as  $|t - t'| < \delta_2$ .

Hence, the family  $\{F_s(t) : s \in \Delta(AS_1) \times \Delta(AS_2)\}$  is in fact equicontinuous in  $t$  (over its support).

Now, we can put things together: take any  $\varepsilon > 0$  and let  $\|s - s'\|_\infty < \min\{\delta_1, \delta_2\}$ . Then,  $|F_s(u(s)) - F_{s'}(u(s'))| < 2\varepsilon \cdot n \cdot m + \varepsilon$ , which can be made as small as desired. Hence, the function  $d$  is continuous.

### Appendix B. Proof of Lemma 1

The proof as appears below is taken from [37].

Pick any  $\delta > 0$ . It is straightforward to use Weierstraß' approximation theorem to get a polynomial  $p$  that uniformly  $\delta$ -approximates  $f$  on the given interval. The issue is that (i)  $p$  may take on negative values, and (ii),  $p$  is not necessarily normalized to be a probability distribution. To fix both, we define the sought function  $g(x) := \alpha \cdot p^+(x)$  with  $p^+(x) = \max(0, f(x))$  for  $x \in [a, b]$ , with a normalization constant  $\alpha > 0$  chosen to make  $\int_a^b g(t) dt = 1$ . Let us postpone the role of  $\alpha$  until a little later, and look at how well the function  $p^+$  approximates  $f$ .

By definition,  $p^+$  is different from  $p$  only at positions  $x \in [a, b]$  when  $p(x) < 0$ , and otherwise identical. Since  $f$  is bounded from below by zero,  $p^+$  can only get "closer" to the graph of  $f$ , and hence also satisfies  $\|f - p^+\|_\infty < \delta$ . We will make use of this later. Now, let us normalize  $p^+$  into a probability density, i.e., choose  $\alpha > 0$  such that  $\int_a^b \alpha \cdot p^+(t) dt = 1$ ,

and look at the maximal error  $\|f - \alpha \cdot p^+\|_\infty$  on the interval  $[a, b]$ . We apply the triangle inequality twice after expanding the inner difference,

$$\begin{aligned} \|f - \alpha \cdot p^+\|_\infty &= \|f - \alpha \cdot p^+ - p + p + \alpha \cdot p - \alpha \cdot p\|_\infty \\ &= \|(f - p) + (\alpha p - \alpha p^+) + (p - \alpha p)\|_\infty \\ &\leq \|f - p\|_\infty + |\alpha| \cdot \|p - p^+\|_\infty + \|p - \alpha p\|_\infty. \end{aligned}$$

Therein, the first term is  $< \delta$  by construction (Weierstraß’ theorem). The second term is the maximum difference between  $p$  and  $p^+$ , which is also bounded by  $\delta$  since  $p$  cannot fall below  $-\delta$  as  $f$  is bounded to be  $\geq 0$ , and  $p$  is “ $\delta$ -bound” to  $f$ .

The third term  $p - \alpha p$  attains a maximum where its first order derivative  $p' - \alpha \cdot p' = (1 - \alpha)p'$  vanishes. Now, if  $\alpha = 1$ , then  $p^+$  is normalized already and we are finished. Otherwise,  $\alpha \neq 1$ , and we can just look for a maximum of  $p$ . Again, since  $p$  is bound to a deviation from  $f$  that is smaller than  $\delta$ , its maximum must be in an  $\delta$ -neighborhood of the maximum of  $f$ , and we can approximately locate the extreme value  $p_{\max} \in [\|f\|_\infty - \delta, \|f\|_\infty + \delta]$ . Wherever this maximum is attained, the same position  $x = x_{\max}$  will also maximize the deviation  $\alpha \cdot p(x)$ . Thus, the maximum possible deviation between  $p_{\max}$  and  $\alpha \cdot p_{\max}$  is  $\leq |\alpha(\|f\|_\infty + \delta) - (\|f\|_\infty - \delta)| \leq |\alpha \cdot \|f\|_\infty - \|f\|_\infty| + |\alpha\delta - \delta| = |\alpha - 1| \cdot (\|f\|_\infty + \delta)$ .

Combining all three bounds, we find

$$\|f - \alpha \cdot p^+\|_\infty \leq \delta + |\alpha| \cdot \delta + |\alpha - 1| \cdot (\|f\|_\infty + \delta) \tag{A1}$$

To exhibit this bound to ultimately become arbitrarily small, let us finally estimate the value  $\alpha$ . To this end, we return to our previous observation that  $f - \delta \leq p^+ \leq f + \delta$  due to the uniform approximation property. Integrating the inequalities from  $a$  to  $b$ , we find

$$1 - \delta(b - a) \leq \int_a^b p^+(t)dt = \frac{1}{\alpha} \leq 1 + \delta(b - a)$$

Taking the reciprocal gives us

$$\frac{1}{1 + \delta(b - a)} \leq \alpha \leq \frac{1}{1 - \delta(b - a)}$$

Letting  $\delta \rightarrow 0$ , will make  $\alpha \rightarrow 1$  (sandwich theorem), and thereby also let the bound (A1) become  $< \varepsilon$  ultimately (for any  $\varepsilon > 0$  that we can choose in advance), thus proving the claim.

### Appendix C. Fictitious Play Algorithm

The following algorithm assumes a minimizing first player, and lets all minima, maxima and  $\leq$ -relation between probability distributions be decided according to the hyperreal representation of two random variables (i.e., by their moment sequences), according to criteria C1 and C2 from Section 2.3. It is an adapted version from [23].

**Algorithm A1** Fictitious Play

---

**Require:** an  $(n \times m)$ -matrix  $\mathbf{A}$  of payoff distributions  $\mathbf{A} = (F_{ij})$

**Ensure:** an approximation  $(\mathbf{x}, \mathbf{y})$  of an equilibrium pair  $(\mathbf{p}^*, \mathbf{q}^*)$  and two distributions  $v_{low}, v_{up}$  so that  $v_{low} \leq_{hr} F(\mathbf{p}^*, \mathbf{q}^*) \leq_{hr} v_{up}$ . Here,  $F(\mathbf{p}^*, \mathbf{q}^*)(r) = \Pr(R \leq r) = \sum_{i,j} F_{ij}(r) \cdot p_i^* q_j^*$ .

- 1: initialize  $\mathbf{x} \leftarrow \mathbf{0} \in \mathbb{R}^n$ , and  $\mathbf{y} \leftarrow \mathbf{0} \in \mathbb{R}^m$
- 2:  $v_{low} \leftarrow$  the minimum over all column-maxima
- 3:  $r \leftarrow$  the row index giving  $v_{low}$
- 4:  $v_{up} \leftarrow$  the maximum over all row-minima
- 5:  $c \leftarrow$  the column index giving  $v_{up}$
- 6:  $\mathbf{u} \leftarrow (F_{1,c}, \dots, F_{n,c})$
- 7:  $y_c \leftarrow y_c + 1$   $\triangleright \mathbf{y} = (y_1, \dots, y_m)$
- 8:  $\mathbf{v} \leftarrow \mathbf{0}$   $\triangleright$  initialize  $\mathbf{v}$  with  $m$  functions that are zero everywhere
- 9: **for**  $k = 1, 2, \dots$  **do**
- 10:    $u^* \leftarrow$  the minimum of  $\mathbf{u}$
- 11:    $r \leftarrow$  the index of  $u^*$  in  $\mathbf{u}$
- 12:    $v_{up} \leftarrow$  the maximum of  $\{u^*/k, v_{up}\}$   $\triangleright$  pointwise scaling of the distribution  $\mathbf{u}^*$
- 13:    $\mathbf{v} \leftarrow \mathbf{v} + (F_{r,1}, \dots, F_{r,m})$   $\triangleright$  pointwise addition of functions
- 14:    $x_r \leftarrow x_r + 1$   $\triangleright \mathbf{x} = (x_1, \dots, x_n)$
- 15:    $v_* \leftarrow$  the maximum of  $\mathbf{v}$
- 16:    $c \leftarrow$  the index of  $v_*$  in  $\mathbf{v}$
- 17:    $v_{low} \leftarrow$  the minimum of  $\{v_*/k, v_{low}\}$   $\triangleright$  pointwise scaling of the distribution  $v_*$
- 18:    $\mathbf{u} \leftarrow \mathbf{u} + (F_{1,c}, \dots, F_{n,c})$   $\triangleright$  pointwise addition of functions
- 19:    $y_c \leftarrow y_c + 1$   $\triangleright \mathbf{y} = (y_1, \dots, y_m)$
- 20:   exit the loop upon convergence  $\triangleright$  concrete condition given below
- 21: **end for**
- 22: Normalize  $\mathbf{x}, \mathbf{y}$  to unit total sum  $\triangleright$  turn  $\mathbf{x}, \mathbf{y}$  into probability distributions.
- 23: **return**  $\mathbf{p}^* \leftarrow \mathbf{x}, \mathbf{q}^* \leftarrow \mathbf{y}$ , and  $F(\mathbf{p}^*, \mathbf{q}^*) \leftarrow \sum_{i,j} F_{ij}(r) \cdot x_i \cdot y_j$   $\triangleright \approx (\mathbf{p}^*)^\top \mathbf{A} \mathbf{q}^*$

---

**Notes**

- 1 i.e., a set such that  $\emptyset \notin \mathcal{U}, \mathbb{N} \in \mathcal{U}$ , closed under  $\supseteq$ -relation, and closed under finite intersections.
- 2 The mixed case is discussed in [23], but was found to be practically not meaningful, since it would be difficult to interpret the semantics of a comparison of categories to real numbers that are not ranks.
- 3 Inductively, we can define  $a \leq_{\text{lex}} b : \iff a \leq b$  for  $a, b \in \mathbb{R}$ , and for  $(a_n, \dots, a_1), (b_n, \dots, b_1)$  we put  $\mathbf{a} \leq_{\text{lex}} \mathbf{b} : \iff [(a_n \leq b_n) \wedge (a_{n-1}, \dots, a_1) \leq_{\text{lex}} (b_{n-1}, \dots, b_1)]$
- 4 Instantly visible by considering any null-sequence  $0 < x_n \rightarrow 0$  as  $n \rightarrow \infty$ , satisfying  $(x_n, 0) >_{\text{lex}} (0, 1)$ , while  $\lim_{n \rightarrow \infty} (x_n, 0) = (0, 0) \leq_{\text{lex}} (0, 0)$ .
- 5 For zero-sum games, two arbitrary equilibria  $(\mathbf{x}^*, \mathbf{y}^*)$  and  $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{y}}^*)$  give rise to two more equilibria  $(\mathbf{x}^*, \tilde{\mathbf{y}}^*)$  and  $(\tilde{\mathbf{x}}^*, \mathbf{y}^*)$ , which may not be the case for non-zero-sum games. A proof of convexity of the set for zero-sum games is found in [39].

**References**

1. Herath, H.S.B.; Herath, T.C. Copula-Based Actuarial Model for Pricing Cyber-Insurance Policies. *Insur. Mark. Co.* **2011**, *2*, 14.
2. Hogg, R.V.; Klugman, S.A. *Loss Distributions*; Wiley Series in Probability and Mathematical Statistics Applied Probability and Statistics; Wiley: New York, NY, USA, 1984. [[CrossRef](#)]
3. AlShawish, A. Risk-Based Security Management in Critical Infrastructure Organizations. Ph.D. Thesis, University of Passau, Passau, Germany, 2021.
4. Tambe, M. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*; Cambridge University Press: Cambridge, NY, USA, 2012.
5. Pawlick, J.; Zhu, Q. *Game Theory for Cyber Deception: From Theory to Applications*; Static & Dynamic Game Theory: Foundations & Applications; Springer International Publishing: Cham, Switzerland, 2021. [[CrossRef](#)]
6. Nguyen, T.; Xu, H. When Can the Defender Effectively Deceive Attackers in Security Games? In Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence, Online, 22 February–1 March 2022; pp. 9405–9412.
7. Zhang, M.; Li, N.; Adepu, S.; Kang, E.; Jin, Z. A Game-Theoretical Self-Adaptation Framework for Securing Software-Intensive Systems. *arXiv* **2021**, arXiv:2112.07588. [[CrossRef](#)]
8. Nguyen, T.H.; Sinha, A.; He, H. Partial Adversarial Behavior Deception in Security Games. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI), Yokohama, Japan, 11–17 July 2020; pp. 283–289.
9. Münch, I. Wege Zur Risikobewertung. In Proceedings of the DACH Security 2012, Graz, Austria, 25 September 2012; Schartner, P., Taeger, J., Eds.; Syssec: Bochum, Germany, 2012; pp. 326–337.



10. Rass, S.; Schauer, S.; König, S.; Zhu, Q. *Cyber-Security in Critical Infrastructures: A Game-Theoretic Approach*; Springer Nature: Berlin/Heidelberg, Germany, 2020.
11. Iqbal, A.; Gunn, L.J.; Guo, M.; Ali Babar, M.; Abbott, D. Game Theoretical Modelling of Network/Cybersecurity. *IEEE Access* **2019**, *7*, 154167–154179. [[CrossRef](#)]
12. Busby, J.S.; Onggo, B.; Liu, Y. Agent-Based Computational Modelling of Social Risk Responses. *Eur. J. Oper. Res.* **2016**, *251*, 1029–1042. [[CrossRef](#)]
13. Busby, J.; Duckett, D. Social Risk Amplification as an Attribution. *J. Risk Res.* **2012**, *15*, 1049–1074. [[CrossRef](#)]
14. Klugman, S.A.; Panjer, H.H.; Willmot, G.E. *Loss Models*; A Wiley-Interscience Publication; Wiley: New York, NY, USA, 1998.
15. Starmer, C. Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *J. Econ. Lit.* **2000**, *38*, 332–382. [[CrossRef](#)]
16. Bürgin, V.; Epperlein, J.; Wirth, F. Remarks on the Tail Order on Moment Sequences. *J. Math. Anal. Appl.* **2022**, *512*, 126135. [[CrossRef](#)]
17. Lozovanu, D.; Solomon, D.; Zelikovsky, A. Multiobjective Games and Determining Pareto-Nash Equilibria. *Bul. Acad. Stiinte Republicii Mold. Mat.* **2005**, *3*, 115–122.
18. Rass, S. On Game-Theoretic Risk Management (Part One)-towards a Theory of Games with Payoffs That Are Probability-Distributions. *arXiv* **2015**, arXiv:1506.07368.
19. Rass, S.; Schauer, S. (Eds.) *Game Theory for Security and Risk Management; Static & Dynamic Game Theory; Foundations & Applications*, Birkhäuser: Cham, Switzerland, 2018.
20. Chauveau, T. Subjective Risk and Disappointment. Documents de travail du Centre d’Economie de la Sorbonne, Université Panthéon-Sorbonne (Paris 1), Centre d’Economie de la Sorbonne, 2012 Technical Report 12063. Available online: <https://ideas.repec.org/p/mse/cesdoc/12063.html> (accessed on 21 November 2022).
21. Gul, F. A Theory of Disappointment Aversion. *Econometrica* **1991**, *59*, 667. [[CrossRef](#)]
22. Wachter, J.; Rass, S.; König, S.; Schauer, S. Disappointment-Aversion in Security Games. In Proceedings of the International Conference on Decision and Game Theory for Security, Seattle, WA, USA, 29–31 October 2018; pp. 314–325. [[CrossRef](#)]
23. Rass, S. On Game-Theoretic Risk Management (Part Two)—Algorithms to Compute Nash-Equilibria in Games with Distributions as Payoffs. *arXiv* **2020**, arXiv:1511.08591. [[CrossRef](#)]
24. Robinson, J. An Iterative Method for Solving a Game. *Ann. Math.* **1951**, *54*, 296–301. [[CrossRef](#)]
25. Alpcan, T.; Başar, T. *Network Security: A Decision and Game Theoretic Approach*; Cambridge University Press: Cambridge, UK, 2010.
26. Sion, M.; Wolfe, P. On a Game without a Value. In *Contributions to the Theory of Games (AM-39), Volume III*; Princeton University Press: Princeton, NJ, USA, 1957; pp. 299–306.
27. Robert, C.P. *The Bayesian Choice*; Springer: New York, NY, USA, 1994.
28. Debreu, G. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, 19th ed.; Number 17 in Monograph/Cowles Foundation for Research in Economics at Yale University; Yale University Press: New Haven, CT, USA, 1987.
29. Schoutens, H. Ultraproducts and Łoś’ Theorem. In *The Use of Ultraproducts in Commutative Algebra*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 1999, pp. 7–27. [[CrossRef](#)]
30. Robinson, A. *Non-Standard Analysis*; Princeton Landmarks in Mathematics and Physics; Princeton University Press: Princeton, NJ, USA, 1996.
31. Rass, S.; König, S.; Schauer, S. Uncertainty in Games: Using Probability-Distributions as Payoffs. In *Decision and Game Theory for Security, Proceedings of the 6th International Conference, GameSec 2015, London, UK, 4–5 November 2015*; Khouzani, M., Panaousis, E., Theodorakopoulos, G., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 346–357. [[CrossRef](#)]
32. Schauer, S.; Palensky, B.; Latzenhofer, M.; Stierle, M. GeRBA. Gesamtstaatliche Risiko- Und Bedrohungsanalyse. In *Studie Im Rahmen Des KIRAS-Forschungsprogrammes*; Technical Report; Wien: Vienna, Austria, 2015.
33. Bundesamt für Bevölkerungsschutz (BABS) Methode zur nationalen Risikoanalyse. Katastrophen und Notlagen Schweiz 2020. Version 2.0. BABS. Available online: <https://www.babs.admin.ch/de/aufgabenbabs/gefaehdrisiken/natgefaehrdanalyse.html> (accessed on 21 November 2022).
34. Cottin, C.; Döhler, S. *Risikoanalyse. Modellierung, Beurteilung Und Management von Risiken Mit Praxisbeispielen*; Vieweg+Teubner: Wiesbaden, Germany, 2009.
35. Bundesministerium Für Inneres (BMI), Austria. *Risikomanagement Im Katastrophenmanagement Leitfadens*; Technical Report; Bundesministerium Für Inneres (BMI): Vienna, Austria, 2018.
36. Rass, S.; König, S.; Schauer, S. Decisions with Uncertain Consequences-A Total Ordering on Loss-Distributions. *PLoS ONE* **2016**, *11*, e0168583. [[CrossRef](#)] [[PubMed](#)]
37. Rass, S.; König, S.; Schauer, S.; Bürgin, V.; Epperlein, J.; Wirth, F. On Game Theory Using Stochastic Tail Orders. *arXiv* **2021**, arXiv:2108.00680.
38. Wachter, J. (Universitaet Klagenfurt, Klagenfurt am Wörthersee, Austria). Personal communication, August 2022.
39. Rass, S.; Wiegele, A.; König, S. Security Games over Lexicographic Orders. In *Decision and Game Theory for Security*; Zhu, Q., Baras, J.S., Poovendran, R., Chen, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 12513, pp. 422–441. [[CrossRef](#)]
40. Schweinzer, P. (Universitaet Klagenfurt, Klagenfurt am Wörthersee, Austria). Personal communication, 7 September 2022.

41. Glicksberg, I.L. A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points. *Proc. Am. Math. Soc.* **1952**, *3*, 170–174. [[CrossRef](#)]
42. Acosta Ortega, F.; Rafels, C. *Security Strategies and Equilibria in Multiobjective Matrix Games*; Facultat de Ciències Econòmiques i Empresariales, Universitat de Barcelona: Barcelona, Spain, 2005.
43. Rass, S.; Rainer, B. Numerical Computation of Multi-Goal Security Strategies. In Proceedings of the Decision and Game Theory for Security, Los Angeles, CA, USA, 6–7 November 2014; Poovendran, R., Saad, W., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 118–133. [[CrossRef](#)]
44. Rass, S.; König, S.; Alshawish, A. *R Package 'HyRiM': Multicriteria Risk Management Using Zero-Sum Games with Vector-Valued Payoffs That Are Probability Distributions*, Version 2.0.0; Austrian Institute of Technology (AIT): Seibersdorf, Austria, 2020.
45. Busby, J.; Gouglidis, A.; König, S.; Rass, S. Analyzing Infrastructure Cybersecurity Risk as Signaling Ecology. 2018, *unpublished manuscript*.