On Nash Equilibria for Multi-Objective Normal-Form Games under Scalarised Expected Returns versus Expected Scalarised Returns

Willem Röpke
Vrije Universiteit Brussel
Brussels, Belgium
willem.ropke@vub.be

Diederik M. Roijers
Vrije Universiteit Brussel &
HU Utrecht Univ. of Applied Sciences
Brussels (BE) & Utrecht (NL)
diederik.roijers@vub.be

Ann Nowé
Vrije Universiteit Brussel
Brussels, Belgium
ann.nowe@vub.be

Roxana Rădulescu
Vrije Universiteit Brussel
Brussels, Belgium
roxana.radulescu@vub.be

ABSTRACT

Recent studies on multi-objective multi-agent systems, and Multi-Objective Normal-Form Games (MONFGs) in particular, have yielded new insights by making the different optimisation criteria – the scalarised expected returns (SER) and expected scalarised returns (ESR) – for these games explicit. In this paper, we study the existence and relation between Nash equilibria in MONFGs under ESR and SER. In particular we introduce new theorems regarding the portability of Nash equilibria (NE) between these criteria. Firstly, we prove by construction that the number of equilibria under SER and ESR can differ when both settings have at least one NE and that no equilibrium need necessarily exist in both criteria for the same game. Secondly, we analyse whether pure strategy NE persist from the ESR criterion to the SER criterion and vice versa. Specifically, we formally show that pure strategy NE under SER must necessarily also be NE under ESR, while the same does not hold the other way around. However, if we make the additional assumption that all utility functions that are used in the game are convex, pure strategy NE do persist from ESR to SER.

KEYWORDS

Game Theory, Multiple objectives, MONFGs

1 INTRODUCTION

Many real-world decision making settings involve multiple independent actors that have their own interests and influence each other's behaviour [5]. Further complicating these already complex settings is the observation that actors often also have multiple and frequently conflicting objectives [13]. With autonomous agents powered by artificial intelligence algorithms becoming ever more prevalent, as well as impactful, it becomes key for agents that operate in such settings to explicitly consider these multiple objectives to align with human needs [8, 21].

Multi-objective multi-agent decision problems are complex, and what constitutes a solution to such problems depends on multiple factors [18]. Not only is it important whether the agents receive the same reward vector, i.e., a *team reward* versus an *individual reward* setting, but also whether the agents value the received rewards differently, i.e., *team utility* versus *individual utility*. In this paper, we take a look at the team reward, individual utility settings, i.e,. the agents may or may not value the received rewards differently.

Another key aspect in multi-objective multi-agent decision making is the choice of optimality criterion. Recent studies on multi-objective multi-agent settings with individual utilities, and *multi-objective normal-form games (MONFGs)* in particular, have yielded new insights by making the different optimisation criteria –the scalarised expected returns (SER) and expected scalarised returns (ESR) optimisation criteria – for these games explicit. Rădulescu et al. [19] show that under SER, even with known utility functions, Nash equilibria need not exist, while under ESR they always do.

In this paper, we continue this line of work by explicitly looking at whether equilibria under one of the optimality criteria transfer over to the other or not, and under which conditions. We prove by construction that when the game has at least one NE under both SER and ESR, the amount of equilibria under SER and ESR can still differ, and that none of these equilibria need be an equilibrium under both criteria at the same time. Furthermore, we analyse whether pure strategy NE persist from the ESR criterion to the SER criterion and vice versa. Here, we formally show that pure strategy NE under SER must necessarily also be NE under ESR, while the same does not hold the other way around. If we make the additional assumption that all utility functions that are used in the game are convex, pure strategy NE do persist from ESR to SER.

It is our hope that these properties and their formal proofs will contribute to novel algorithms of MONFGs in the future, and make the MONFG more widely used as a model for real-world applications.

MODeM positioning

We position this paper with respect to the multi-objective multi-agent decision making field [18]. Specifically, we study settings with multiple agents with either team or individual rewards and individual utility, i.e., even if the agents receive the same reward vectors, they may value these vectors differently. Furthermore, we take the perspective of finding stable solutions, i.e., Nash equilibria. We allow for stochastic strategies unless otherwise indicated.

Proc. of the 1st Multi-Objective Decision Making Workshop (MODeM 2021), Hayes, Mannion, Vamplew (eds.), July 14-16, 2021, Online, http://modem2021.cs.nuigalway.ie. 2021.

2 BACKGROUND

In this section, we discuss the necessary background. We start by defining MONFGs and utility functions. We then move to explain multi-objective optimisation criteria and the impact on possible solutions.

2.1 Multi-Objective Normal-Form Games

A *Multi-Objective Normal-Form Game (MONFG)* [3] is similar to a (single-objective) normal-form game in all but one aspect: the payoff received by the agents is not a scalar value, but rather a vector of payoffs. In these payoff vectors, each element in the vector then corresponds to the value of a different objective. We can formally define an MONFG as follows:

Definition 2.1 (Multi-objective normal-form game). A (finite, n-person) multi-objective normal-form game is a tuple $(N, \mathcal{A}, \boldsymbol{p})$ with $n \geq 2$ and $d \geq 2$ objectives, where:

- N is a finite set of n players, indexed by i;
- $\mathcal{A} = A_1 \times \cdots \times A_n$, where A_i is a finite set of actions available to player i. Each vector $a = (a_1, \ldots, a_n) \in \mathcal{A}$ is called an action profile;
- $\boldsymbol{p} = (\boldsymbol{p_1}, \ldots, \boldsymbol{p_n})$ where $\boldsymbol{p_i} : A_i \to \mathbb{R}^d$ is the vectorial payoff of player i, given an action profile.

In this paper, we follow a utility-based approach [16]. In this approach, each agent i derives a utility from the payoff vector by applying their own scalarisation function, called the utility function $u_i : \mathbb{R}^d \to \mathbb{R}$, to this vector. A common simple example of a utility function is a linear utility-function which simply assigns a weight $w_o \in [0, 1]$ to each objective o, i.e., $u_i(\mathbf{p}) = \mathbf{w} \cdot \mathbf{p}$. In practice however, utility functions often are non-linear, for example because they have interaction components between the objectives, or have decreasing additional utility as the amount of payoff in an objective increases.

In general, we assume that each utility function is at least monotonically increasing in all objectives, which intuitively means that an agent will always prefer more of any objective over less given equal payoff values for the other objectives:

$$(\forall o, p_o^{\pi} \geq p_o^{\pi'}) \implies u(\boldsymbol{p}^{\pi}) \geq u(\boldsymbol{p}^{\pi'}) \tag{1}$$

with two policies π and π'. While this definition of utility seems straightforward, it does leave the question about what the agent should optimise for. Specifically, do we want to compute the utility over the expected payoff, or apply the utility function to the payoffs directly. When we want to optimise for the utility of a single play of a MONFG, we apply the utility function to the payoff vectors directly, resulting in the expected scalarised returns (ESR) optimisation criterion [7, 15, 19]:

$$p_{u,i} = \mathbb{E}\left[u\left(\boldsymbol{p}_i^{\boldsymbol{\pi}}\right)\right] \tag{2}$$

with $p_{u,i}$ the expected utility for agent i with utility function u and $\boldsymbol{p}_i^{\boldsymbol{\pi}}$ the payoff vector for agent i under the joint strategy $\boldsymbol{\pi}$.

Alternatively, it is possible that the agent cares about optimising for the utility it can derive from several plays of the game, in which case we first calculate the expectation over the payoff vectors before applying the utility function. This is called the scalarised expected returns (SER) criterion[16, 19]:

$$p_{u,i} = u\left(\mathbb{E}\left[\boldsymbol{p}_i^{\boldsymbol{\pi}}\right]\right) \tag{3}$$

where $p_{u,i}$ is now the utility of the expected payoff vector.

As a concrete illustration to show the differences between the ESR and SER optimisation criteria, consider treatments for a medical condition. There are multiple objectives in this setting, such as maximising the probability of a cure, minimising side-effects, and minimising costs. In this example, a patient might have very different utility for a 50 percent chance of a cure, but with side-effects and 50 percent no effects whatsoever, than for the same 50 percent chance of a cure without any side effects and 50 percent chance of side-effects without a cure (i.e., their preferred optimisation criterion is ESR). This may stand in contrast to the distributors of the means (such as medicines) for the treatment, who might mainly present the averages (i.e., according to the SER criterion). Therefore, the researchers that aim to optimise the treatment are faced with the difficult choice of whether to optimise for the average utility of each treatment outcome (i.e., ESR) or attempting to optimise the utility for the average outcome (i.e., SER); as this can result in very different treatment plans. The choice between these optimisation criteria is thus important to consider, as it has also been shown that ESR and SER are not equivalent under non-linear utility functions [19]. This work further showed that in stateless settings ESR can be reduced to a single-objective problem when the utility functions are known, implying that regular techniques can be used to solve such problems. The SER criterion on the other hand cannot easily be solved by traditional methods and has been understudied thus far.

2.2 Solution Concepts

Game-theoretic equilibria – such as the well-known solution concept of the Nash equilibrium [14] – determine a set of outcomes in strategic interactions from which players have no incentives to deviate. When we adapt the original single-objective formulation of Nash equilibria (NE) to multi-objective games under SER we obtain the following definition [18]:

Definition 2.2 (Nash equilibrium for scalarised expected returns). A joint policy $\boldsymbol{\pi}^{NE}$ leads to a Nash equilibrium under the scalarised expected returns criterion if for each agent $i \in 1, \cdots, n$ and for and all alternative policies $\pi_i \in \Pi_i$:

$$u_i\left(\mathbb{E}\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) \geq u_i\left(\mathbb{E}\boldsymbol{p}_i\left(\pi_i, \boldsymbol{\pi}_{-i}^{NE}\right)\right)$$

i.e. $\boldsymbol{\pi}^{NE}$ is a Nash equilibrium under SER if no agent can increase the *utility of its expected payoffs* by deviating unilaterally from $\boldsymbol{\pi}^{NE}$.

This definition is similar to the definition of a NE in a single-objective NFG. We must note that although every single-objective NFG has at least one NE, it has been proven that in MONFGs under SER, Nash equilibria need not exist [19]. This is because the expectation over payoff vectors is computed before the utility function is applied, leading to more freedom for the agents to obtain better results in expectation. Please see [18] for an example and formal proof that NE may not exist in MONFGs under SER.

The same problem, i.e., the possible non-existence of NE, does not exist under ESR.

Definition 2.3 (Nash equilibrium for expected scalarised returns). A joint policy $\boldsymbol{\pi}^{NE}$ is a Nash equilibrium in a MONFG under ESR if for all players $i \in \{1, ..., N\}$ and all alternative policies $\pi_i \in \Pi_i$:

$$\mathbb{E}\, u_i \left(\mathbf{p}_i \left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) \geq \mathbb{E}\, u_i \left(\mathbf{p}_i \left(\pi_i, \boldsymbol{\pi}_{-i}^{NE}\right)\right)$$

i.e. $\boldsymbol{\pi}^{NE}$ is a Nash equilibrium under ESR if no agent can increase the *expected utility of its payoffs* by deviating unilaterally from $\boldsymbol{\pi}^{NE}$.

We can see that under ESR, as in single-objective NFGs, an NE does always exist, because the utility functions are applied to the payoff vectors before the expectation. Therefore, if the utility functions are known, these can be used to scalarise the MONFG into an equivalent single-objective NFG, for which the existence of NE is known.

An open question that remains is in what cases NEs do exist under SER and if they do, whether a Nash equilibrium under SER is also a Nash equilibrium under ESR, and vice versa, and under which conditions.

3 THE EXISTENCE OF NASH EQUILIBRIA UNDER BOTH ESR AND SER

Single-objective normal-form games [14] have been a popular model for a long time. As a consequence, much is known about them, including the fact that each NFG must have at least one mixed strategy NE. The study of multi-objective NFGs on the other hand, and specifically using a utility-based approach, has received much less attention. As is the case with single-agent multi-objective models, work on MONFGs has been fragmented, as different assumptions about the setting in which this model is used can lead to vastly different outcomes. Recently however, with a survey on multi-objective multi-agent decision making [18], attention has been drawn to the fact that the two optimisation criteria – ESR and SER – are not equivalent, and a taxonomy has been offered on the basis of payoffs, utility and the type of desired outcomes.

In this paper we focus on individual utility, i.e., even if agents receive the same payoff vector they may value this payoff vector differently. Furthermore, we assume that no social welfare mechanism is employed, and that we are looking for stable outcomes in settings with self-interested agents. For this setting, in MONFGs, it has been shown recently that under SER no NE need necessarily exist [19]. We build upon this work by providing a further study of the existence of NE in MONFGs under both optimisation criteria. In this section, we aim to show that the total number of NE under SER and under ESR if both have at least one NE need not be equal (Theorem 3.1) and that no NE need be shared between the two different optimisation criteria (Theorem 3.2).

Theorem 3.1. *In a (finite, n-person) multi-objective normal-form game with at least one Nash equilibrium under both criteria, the size of the sets of Nash equilibria under the scalarised expected returns criterion and under the expected scalarised returns criterion need not be equal.*

Proof. We can prove this theorem by constructing a MONFG that has this exact property. The MONFG we use for this purpose can be seen in Table 1. We show next to this MONFG, the single-objective NFG resulting from directly applying the following utility function to the payoffs, assuming that both agents use this same utility function under ESR:

$$u(p_1, p_2) = 0.1 * p_1 + max(0, p_1) * max(0, p_2) \tag{4}$$

	A	B
A	(1, 0); (1, 0)	(0, 1); (0, 1)
B	(0, 1); (0, 1)	(-10, 0), (-10, 0)

(a) The multi-objective reward vectors.

	A	B
A	0.1; 0.1	0; 0
B	0; 0	-1; -1

(b) The ESR utility for both agents.

Table 1: A MONFG with team vector-valued payoffs (top), and (bottom) the scalarised single-objective NFG with individual rewards resulting from applying the utility function in Equation 4 directly to the payoff vectors (as per ESR) to the upper MONFG for both agents. This MONFG shows by construction the two properties in Theorem 3.1 and 3.2. The highlighted cell is a pure Nash equilibrium.

Let us first show the NE in the MONFG under ESR. We do this by first applying the utility functions for each agent – which in this case happens to be the same – directly to the payoff vectors in the MONFG, resulting in the single-objective NFG in Table 1b. We then observe that only the pure strategy profile (A, A) results in utilities above 0 for both agents. As such, there is no incentive for agents to play a mixed strategy when the other agent plays A at least part of the time, leading to the pure strategy NE of (A, A). Additionally, (B, B) is not a NE, as there is an incentive for either agent to play A, which increases their utility. This then again leads both agents to adapt their strategies to the NE of (A, A), making it the only NE of the MONFG under ESR.

Next, we discuss the NE for the MONFG under SER (1a). First note that the pure strategy NE of (A, A) under ESR is not a NE under SER. To see this, observe that when one agent plays A deterministically, the best response for the other agent is to play a mixed strategy with probability $\frac{11}{20}$ for action A and probability $\frac{9}{20}$ for action B. This results in an expected return of $(\frac{11}{20}, \frac{9}{20})$ and a utility of $0.1 \cdot \frac{11}{20} + \frac{11}{20} \cdot \frac{9}{20} = 0.3025$ for both agents. In fact, this constitutes a NE under SER for this game, as no agent has an incentive to deviate from this strategy. A second NE occurs when the agents switch strategies, resulting in the same payoffs. Please note that both agents receive the same expected payoff vectors, and apply the same utility function to these. We can also show that the pure strategy (B, B) is not a NE, as this can be improved upon by either agent deterministically playing A. As such, the MONFG in Table 1 has at least two mixed strategy NE under SER and no pure strategy NE.

In this MONFG, both the game under SER and ESR have NE. However, we can see that they have a different number of NE, proving Theorem 3.1. □

Our second finding pertaining to Nash equilibria in MONFGs states that when both SER and ESR have a Nash equilibrium, no NE must necessarily be shared. We formalise this in Theorem 3.2.

Theorem 3.2. *In a (finite, n-person) multi-objective normal-form game with at least one Nash equilibrium under both criteria, the set of Nash equilibria under the scalarised expected returns criterion and the set of Nash equilibria under the expected scalarised returns criterion may be disjoint.*

Proof. Theorem 3.2 can be shown by using the same example MONFG of Table 1, and the NE already mentioned in the proof for Theorem 3.1. It is clear from that in this example no joint strategy is a Nash equilibrium under both SER and ESR. □

4 PURE STRATEGY NASH EQUILIBRIA

As previously noted, SER and ESR are not equivalent in general and no Nash equilibrium need necessarily exist under SER [19]. Furthermore, we have proven in the previous section that even when both under SER and under ESR there are in fact Nash equilibria, it can still be that there is no joint strategy that is an NE under both SER and ESR. One important open question that remains however is whether and under which circumstances Nash equilibria can persist under both criteria.

In this section, we first show that a pure strategy Nash equilibrium under SER must always be a pure strategy Nash equilibrium under ESR as well. Furthermore, we show that the inverse does not hold by providing a counter example. However, we show that adding the assumption that all utility functions in the MONFG are convex does ensure that pure strategy NE under ESR are also NE under SER. Proving these equivalence relations is of importance as it means that approaches to calculating NE under one criterion could potentially be applied to the other criterion as well. Equivalence relations from ESR to SER in specific could be extremely useful as a MONFG under ESR can be reduced to a single-objective NFG for which there are several well-performing algorithms that are able to calculate one or all NE in the game [6, 9, 11].

In order to show that a pure strategy NE under SER must necessarily be a pure strategy NE under ESR, we first introduce a necessary concept in Lemma 4.1. This lemma states that the utility of a pure strategy profile under SER is the same as the utility of that pure strategy profile under ESR.

Lemma 4.1 (Utility of a pure strategy). *Given a pure strategy profile in a (finite, n-person) multi-objective normal-form game, the expectation of the payoff will always be the observed payoff, as the expectation of a constant is equal to that constant:*

$$\mathbb{E}[\boldsymbol{p}] = \boldsymbol{p}$$

and given a utility function u, the expected utility will also equal the observed utility by the same reasoning

$$\mathbb{E}[u(\boldsymbol{p})] = u(\boldsymbol{p})$$

We can thus say that for a pure strategy profile, the utility of a payoff under SER equals the utility under ESR:

$$u(\mathbb{E}[\boldsymbol{p}]) = u(\boldsymbol{p}) = \mathbb{E}[u(\boldsymbol{p})]$$

Given this lemma, we can now define the first theorem of this section which states that a pure strategy Nash equilibrium under SER, must always be a pure Nash equilibrium under ESR as well.

Theorem 4.2. *In a (finite, n-person) multi-objective normal-form game, a pure strategy Nash equilibrium under the scalarised expected returns criterion must necessarily also be a Nash equilibrium under the expected scalarised returns criterion.*

Proof. Given a pure strategy Nash equilibrium under SER $\boldsymbol{\pi}^{\boldsymbol{NE}}$, we can say that:

$$\begin{aligned}
u_i\left(\mathbb{E}\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) &\geq u_i\left(\mathbb{E}\boldsymbol{p}_i\left(\pi_i, \boldsymbol{\pi}_{-i}^{NE}\right)\right)\\
\iff u_i\left(\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) &\geq u_i\left(\mathbb{E}\boldsymbol{p}_i\left(\pi_i, \boldsymbol{\pi}_{-i}^{NE}\right)\right)\\
\implies \forall a \in A_i : u_i\left(\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) &\geq u_i\left(\boldsymbol{p}_i\left(a, \boldsymbol{\pi}_{-i}^{NE}\right)\right)\\
\iff u_i\left(\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) &\geq \max_{\boldsymbol{\alpha}} \sum_{a \in A_i} \alpha_a u_i\left(\boldsymbol{p}_i\left(a, \boldsymbol{\pi}_{-i}^{NE}\right)\right)\\
\iff \mathbb{E}u_i\left(\boldsymbol{p}_i\left(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}\right)\right) &\geq \mathbb{E}u_i\left(\boldsymbol{p}_i\left(\pi_i, \boldsymbol{\pi}_{-i}^{NE}\right)\right)\\
&\iff \text{A pure Nash equilibrium under ESR}
\end{aligned}$$

□

The proof starts with the general definition of a pure strategy Nash equilibrium under SER and removes the expected values where possible in line two. In line three, we remark that if the pure strategy profile is an NE, it must necessarily also be better than unilaterally playing another pure strategy. In line four, this leads us to state that the utility of the pure strategy NE is greater or equal to the optimal stochastic mixture of the utilities of the other pure strategies. In line five, we can freely introduce the expected value again in the left hand side of the inequality and rewrite the right hand side such that it now reflects the expected scalarised returns. This final inequality is also the definition of a Nash equilibrium under ESR. Given this positive result, it is alluring to believe that the inverse, so going from ESR to SER, would also hold. However, this is not actually the case as we can only guarantee that the utility of a pure strategy profile is greater or equal to the optimal stochastic mixture of scalar utilities. We can not guarantee that it is better than the utility of the optimal stochastic mixture of reward vectors.

Theorem 4.3. *In a (finite, n-person) multi-objective normal-form game, a pure strategy Nash equilibrium under the expected scalarised returns criterion need not also be a Nash equilibrium under the scalarised expected returns criterion.*

Proof. We show this theorem formally by using the same MONFG and utility functions as presented in the previous section in Table 1a. Recall that in this game, there was a pure NE under ESR but no pure NE under SER. □

We add that an additional assumption can be made to remedy this negative result. Concretely, by making the assumption that all utility functions used by the players in the game are convex, we are still able to show that a pure strategy NE under ESR, must also be a NE under SER.

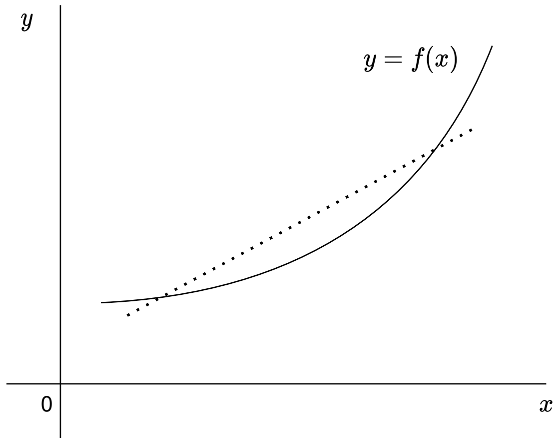

Figure 1: An example of a convex function. The dotted line denotes the fact that the line segment between any two points lies above the graph between them.

Theorem 4.4. *In a (finite, n-person) multi-objective normal-form game where all player utility functions are convex, a pure strategy Nash equilibrium under the expected scalarised returns criterion must necessarily also be a Nash equilibrium under the scalarised expected returns criterion.*

We provide a formal definition of a convex function below. In simple terms, a convex function can be defined as a function for which the line segment between any two points lies above the graph between these two points. We show a visual example of such a function in Figure 1.

Definition 4.5. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain is a convex set and for all $\boldsymbol{x}_1, \boldsymbol{x}_2$ in its domain, and all $t \in [0, 1]$, we have: $f(t\boldsymbol{x}_1 + (1-t)\boldsymbol{x}_2) \leq tf(\boldsymbol{x}_1) + (1-t)f(\boldsymbol{x}_2)$

We can apply this definition to highlight that the utility functions used to show Theorem 4.3 are not convex. If we take for example $\boldsymbol{x}_1 = (1, 0)$, $\boldsymbol{x}_2 = (0, 1)$ and $t = 0.5$, then we get $f((0.5, 0.5)) = 0.05 + 0.25 = 0.3$ for the left hand side and $0.5f((1, 0)) + 0.5f((0, 1)) = 0.05 + 0 = 0.05$ for the right hand side. It is clear then that this is not a convex utility function, as 0.3 is larger than 0.05.

Proof. Given Jensen's inequality, we know that if u_i is convex:

$$\mathbb{E}\left[u_i\left(\boldsymbol{p}_i\left(\boldsymbol{\pi}\right)\right)\right] \geq u_i\left(\mathbb{E}\left[\boldsymbol{p}_i\left(\boldsymbol{\pi}\right)\right]\right)$$

Then if we have a pure Nash equilibrium under ESR and if u_i is convex for every player i:

$$\begin{aligned}
&\mathbb{E}[u_i(\boldsymbol{p}(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE}))] \geq \mathbb{E}[u_i(\boldsymbol{p}(\pi_i, \boldsymbol{\pi}_{-i}^{NE}))] \\
\implies\ & u_i(\mathbb{E}[\boldsymbol{p}(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE})]) \geq \mathbb{E}[u_i(\boldsymbol{p}(\pi_i, \boldsymbol{\pi}_{-i}^{NE}))] \\
\implies\ & u_i(\mathbb{E}[\boldsymbol{p}(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE})]) \geq \mathbb{E}[u_i(\boldsymbol{p}(\pi_i, \boldsymbol{\pi}_{-i}^{NE}))] \\
&\qquad \geq u_i(\mathbb{E}[\boldsymbol{p}(\pi_i, \boldsymbol{\pi}_{-i}^{NE})]) \\
\implies\ & u_i(\mathbb{E}[\boldsymbol{p}(\pi_i^{NE}, \boldsymbol{\pi}_{-i}^{NE})]) \geq u_i(\mathbb{E}[\boldsymbol{p}(\pi_i, \boldsymbol{\pi}_{-i}^{NE})]) \\
&\implies \text{A pure Nash equilibrium under SER}
\end{aligned}$$

□

This proof first introduces Jensen's inequality [10] to show that when all utility functions are convex, the expected scalarised returns are always greater or equal to the scalarised expected returns. In the first line we write the definition of a NE under ESR. We then note that when the NE is a pure strategy profile, we can place the expectation inside the utility as it is equal. In the third and fourth line, we introduce a new element to the inequality by using Jensens formula. Lastly, we remove the inner part of the inequality. By doing this, we have arrived at the definition of a NE under SER, proving our statement.

5 RELATED WORK

Multi-objective games (also known as multicriteria games in the literature) were introduces by Blackwell et al. [4] and have been discussed extensively in the literature throughout the years. We highlight below a few of these works, together with the difference in perspective in comparison to our approach.

A lot of previous work in multi-objective games considers the case in which agents do not know their utility-function, and thus define utility-function agnostic equilibria. For example, Shapley and Rigby [20] extend and characterise the set of mixed-strategy agnostic Nash equilibria for multicriteria two-person zero-sum games for linear utility functions. Important here to note is their remark that if the utility functions differ, the scalarised game (implicitly assuming ESR) can possibly be no longer zero-sum.

Bergstresser and Yu [2] bring up the idea that utility functions could also be non-linear. However, in their practical analysis, they only consider linear utility functions and apply the ESR criterion to obtain the resulting trade-off game and corresponding solution points.

Finally, Lozovanu et al. [12] formulate an algorithm for finding Pareto-Nash equilibria in multi-objective non-cooperative games. More precisely, for every linear utility function for which the weights sum to one, they compute the trade-off game (i.e., implicitly assume the ESR criterion), then find its NE.

Consequently, given the use of linear utility functions, there was no distinction to be made between the ESR and SER optimisation criteria in the game theory literature. Rădulescu et al. [19] made the choice between an ESR and SER perspective explicit, and showed that this choice has profound consequences on the set of Nash and correlated equilibria [1] in MONFGs, when considering non-linear utility functions. In this paper, we have continued this line of work by explicitly looking at whether equilibria under one of the optimality criteria transfer over to the other or not, and under which conditions.

6 CONCLUSIONS

In this paper, we have analysed *Nash Equilibria (NE)* in *Multi-Objective Normal-Form Games (MONFGs)*, under two optimality criteria: *scalarised expected returns (SER)*, and *expected scalarised returns (ESR)*. Specifically, we have shown five properties of NE under ESR compared to NE under SER.

We have constructed an example to show that the sets of Nash equilibria under ESR and SER for the same MONFG may be disjoint.

From this same example, we further observe that the size of these sets of NE may differ as well.

We subsequently take a look at pure strategy NE. Firstly, we have shown that a pure strategy NE under SER must necessarily also be an NE under ESR. Secondly we have shown that the reverse does not hold, i.e., a pure strategy NE under ESR need not be an NE under SER. Thirdly, we have shown that if we restrict the utility functions of the agents to convex functions only, the pure strategy NE under ESR do transfer over to SER. Specifically, given that the utility functions of all agents in an MONFG are convex, a pure strategy NE under ESR must necessarily also be an NE under SER.

It is our hope that these properties may form the basis of algorithms to efficiently identify NE in MONFGs. In future work, we aim to construct such algorithms.

We would also like to mention a couple directions we believe to be promising for future theoretical work. Firstly, recently we have shown that by allowing agents to communicate preferred actions or strategies, cyclic Nash equilibria can be reached in MONFGs [17]. We hope to contribute theoretical properties of cyclic equilibria as well. Secondly, we have assumed in our definitions of Nash equilibria that both agents follow the same optimality criterion (ESR or SER). However, this need not be the case in practice. Thirdly, we have restricted the class of utility functions to prove our final theorem. It would be interesting to see whether different restrictions on the utility functions, or indeed restrictions on the payoff vectors, would lead to other useful properties.

ACKNOWLEDGMENTS

This research was supported by funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen".

REFERENCES

[1] Robert J Aumann. 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica: Journal of the Econometric Society* (1987), 1–18.

[2] K Bergstresser and PL Yu. 1977. Domination structures and multicriteria problems in N-person games. *Theory and Decision* 8, 1 (1977), 5–48.

[3] David Blackwell. 1954. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.* 6, 1 (1954), 1–8. https://doi.org/10.2140/pjm.1956.6.1

[4] David Blackwell et al. 1956. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.* 6, 1 (1956), 1–8.

[5] L Busoniu, R Babuska, and B De Schutter. 2008. A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 2 (2008), 156–172. https://doi.org/10.1109/TSMCC.2007.913919

[6] Federico Echenique. 2007. Finding all equilibria in games of strategic complements. *Journal of Economic Theory* 135, 1 (2007), 514–532. https://doi.org/10.1016/j.jet.2006.06.001

[7] Conor F. Hayes, Mathieu Reymond, Diederik M. Roijers, Enda Howley, and Patrick Mannion. 2021. Distributional Monte Carlo Tree Search for Risk-Aware and Multi-Objective Reinforcement Learning. In *The 20th International Conference on Autonomous Agents and Multiagent Systems.*

[8] Conor F. Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M. Zintgraf, Richard Dazeley, Fredrik Heintz, Enda Howley, Athirai A. Irissappane, Patrick Mannion, Ann Nowé, Gabriel Ramos, Marcello Restelli, Peter Vamplew, and Diederik M. Roijers. 2021. A Practical Guide to Multi-Objective Reinforcement Learning and Planning. *arXiv preprint arXiv:2103.09568* (2021).

[9] P. Jean Jacques Herings and Ronald Peeters. 2005. A globally convergent algorithm to compute all nash equilibria for n-person games. *Annals of Operations Research* 137, 1 (2005), 349–368. https://doi.org/10.1007/s10479-005-2265-4

[10] J. L.W.V. Jensen. 1906. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica* 30, 1 (1906), 175–193. https://doi.org/10.1007/BF02418571

[11] C E Lemke and Jr. Howson J. T. 1964. Equilibrium Points of Bimatrix Games. *J. Soc. Indust. Appl. Math.* 12, 2 (June 1964), 413–423. https://doi.org/10.1137/0112033

[12] Dmitrii Lozovanu, D Solomon, and A Zelikovsky. 2005. Multiobjective games and determining pareto-nash equilibria. *Buletinul Academiei de Ştiinţe a Republicii Moldova. Matematica* 3 (2005), 115–122.

[13] Patrick Mannion. 2017. *Knowledge-Based Multi-Objective Multi-Agent Reinforcement Learning.* Ph.D. Dissertation.

[14] John Nash. 1951. Non-Cooperative Games. *The Annals of Mathematics* 54, 2 (1951), 286. https://doi.org/10.2307/1969529

[15] Diederik M. Roijers, Denis Steckelmacher, and Ann Nowé. 2018. Multi-objective reinforcement learning for the expected utility of the return. In *Proceedings of the Adaptive and Learning Agents workshop at AAMAS (FAIM).*

[16] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. 2013. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research* 48 (2013), 67–113.

[17] Willem Röpke, Roxana Rădulescu, Diederik M. Roijers, and Ann Nowé. 2021. Communication Strategies in Multi-Objective Normal-Form Games. In *Adaptive and Learning Agents Workshop at AAMAS.*

[18] Roxana Rădulescu, Patrick Mannion, Diederik M. Roijers, and Ann Nowé. 2020. Multi-objective multi-agent decision making: a utility-based analysis and survey. *Autonomous Agents and Multi-Agent Systems* 34, 1 (April 2020), 10. https://doi.org/10.1007/s10458-019-09433-x

[19] Roxana Rădulescu, Patrick Mannion, Yijie Zhang, Diederik M Roijers, and Ann Nowé. 2020. A utility-based analysis of equilibria in multi-objective normal-form games. *The Knowledge Engineering Review* 35 (2020), e32. https://doi.org/10.1017/S0269888920000351

[20] Lloyd S. Shapley and Fred D. Rigby. 1959. Equilibrium points in games with vector payoffs. *Naval Research Logistics Quarterly* 6, 1 (1959), 57–61.

[21] Peter Vamplew, Richard Dazeley, Cameron Foale, Sally Firmin, and Jane Mummery. 2018. Human-aligned artificial intelligence is a multiobjective problem. *Ethics and Information Technology* 20, 1 (2018), 27–40.