Endogenous Energy Reactive Modules Games: Modelling Side Payments Among Resource-Bounded Agents

Julian Gutierrez[1] , **David Hyland**[2] , **Muhammad Najib**[3] ,
Giuseppe Perelli[4] and **Michael Wooldridge**[2]
[1]Monash University
[2]University of Oxford
[3]Heriot-Watt University
[4]Sapienza University of Rome

Abstract

We introduce *Energy Reactive Modules Games* (ERMGs), an extension of Reactive Modules Games (RMGs) in which actions incur an energy cost (which may be positive or negative), and the choices that players make are restricted by the energy available to them. In ERMGs, each action is associated with an energy level update, which determines how their energy level is affected by the performance of the action. In addition, agents are provided with an initial energy allowance. This allowance plays a crucial role in shaping an agent's behaviour, as it must be taken into consideration when one is determining their strategy: agents may only perform actions if they have the requisite energy. We begin by studying rational verification for ERMGs and then introduce *Endogenous* ERMGs, where agents can choose to transfer their energy to other agents. This exchange may enable equilibria that are impossible to achieve without such transfers. We study the decision problem of whether a stable outcome exists under both the Nash equilibrium and Core solution concepts.

1 Introduction

The paradigm of *rational verification* is an important approach to verifying the possible behaviours of multi-agent systems [Gutierrez *et al.*, 2017; Abate *et al.*, 2021; Gutierrez *et al.*, 2023b]. Rational verification draws inspiration from the well-known formal verification paradigm of model checking, which is concerned with automatically checking whether or not a given system satisfies certain properties, expressed as formulae of temporal logic [Baier and Katoen, 2008]. Rational verification differs from model checking in that it assumes that system components (agents) are rational actors, making choices in pursuit of their personal goals, and taking into account the strategic behaviours of other agents – goals are typically captured by associating with each agent a temporal logic formula that it desires to see satisfied. Since agents are assumed to be rational, game theory provides a natural framework through which to understand collective rational action: a classic decision problem in RV involves asking whether a given temporal logic property holds on some run of the system that arises by agents choosing strategies that constitute (for example) a Nash equilibrium [Abate *et al.*, 2021].

Many variations of rational verification have now been studied [Gutierrez *et al.*, 2018; Gutierrez *et al.*, 2023b; Bruyère *et al.*, 2022; Brice *et al.*, 2023]. In this work, we introduce a variation of the problem in which agents act under energy resource bounds. Specifically, we assume agents are given some initial energy endowment, and subsequently, all actions that the agent performs are assumed to affect this endowment. Actions may generate energy (leading to an increase in the endowment) or consume energy (reducing the endowment). Crucially, agents can only perform actions for which they have sufficient energy. Note that energy plays a secondary role in agents' preferences: agents are concerned with achieving their goal, and energy affects preferences only indirectly (by affecting the actions they can perform).

To capture this setting, we introduce a variation of *Reactive Modules Games* (RMGs) [Gutierrez *et al.*, 2017]. In our new variation, individual agent actions (specified via guarded commands) are associated with an energy value, which may either increase or decrease the agent's energy endowment. We begin our study by showing that this framework, while providing a very natural platform through which to model resource-bounded multi-agent systems, can in fact be reduced to "classic" RMGs, and as a consequence, the key non-cooperative and cooperative decision problems in rational verification for our new setting are no harder than in the "classic" setting.

We then study an extension of our model called *Endogenous ERMGs*, in which agents may transfer energy to other agents. Such offers change the possible actions that agents may perform, and hence may change the underlying strategic structure of the game: for example, it may make sense for me to "donate" energy to another agent so that it will choose actions that are of benefit to me. This leads to a two-stage game, with a pre-play "offer" phase. Such an exchange may facilitate equilibria that are impossible to achieve without such an exchange. We study the decision problem of whether a stable outcome and offer profile exists under both the Nash equilibrium and Core solution concepts, again showing that the complexities of the associated problems for both solution concepts in Endogenous ERMGs are also 2EXPTIME-c.

2 Preliminaries

We use classical propositional logic, defined over a finite and non-empty set Φ of Boolean variables. Each variable in Φ may take the values of truth $\top$ or falsity $\bot$, which we represent with the binary values 1 and 0, respectively. Our language includes the classical propositional logic connectives $\neg$ ("not"), $\vee$ ("or"), $\wedge$ ("and"), $\rightarrow$ ("implies"), and $\leftrightarrow$ ("iff"). A *valuation* $\vec{v} \in \{0,1\}^{|\Phi|}$ over Φ is a binary string representing an assignment of truth values to all variables in Φ. We say that a valuation $\vec{v}$ *satisfies* a propositional formula φ defined over Φ, written $\vec{v} \models \varphi$, if φ is true under $\vec{v}$. Where $\vec{x} = (x_1, \ldots, x_n)$ is an n-tuple and $A \subseteq \{1, \ldots, n\}$, we write $\vec{x}_A = (x_i)_{i \in A}$ to denote the $|A|$-tuple consisting of the elements in $\vec{x}$ that are indexed by those in A. Similarly, we write $\vec{x}_{-A} = (x_i)_{i \in \{1,\ldots,n\} \setminus A}$ to denote the tuple of elements in $\vec{x}$ which are not indexed by elements in A. Finally, where $\vec{x}$ and $\vec{y}$ are two disjoint tuples, we write $(\vec{x}, \vec{y})$ for the tuple formed by merging $\vec{x}$ and $\vec{y}$.

Linear Temporal Logic (LTL). We make extensive use of the standard framework of *Linear Temporal Logic* (LTL), which is an extension of propositional logic with tensed modal operators for expressing properties of infinite linear sequences of states [Pnueli, 1977]. Specifically, in addition to the usual stock of classical connectives as above, LTL includes the unary operators "**X**" (next), "**F**" (sometime), and "**G**" (always), and the binary "until" operator, "**U**". Given a set of variables Φ, let $LTL(\Phi)$ be the set of LTL formulae over Φ; where the variable set Φ is clear from the context, we simply write LTL. LTL formulae are interpreted with respect to infinite sequences of valuations, which we refer to as *runs*, typically denoted ρ, ρ', etc. Where $\rho = \vec{v}_0\vec{v}_1\vec{v}_2\ldots$ is a run, and $t \in \mathbb{N}$ is a temporal index into ρ, we write $(\rho, t) \models \varphi$ to mean that $\varphi \in LTL$ is true at time $t \in \mathbb{N}$ on run ρ. Additionally, we use the notation $\rho[t]$ to denote the valuation $\vec{v}_t$ in ρ at time point $t \in \mathbb{N}$. We write $\rho \models \varphi$ as a shorthand for $(\rho, 0) \models \varphi$, in which case we say that ρ *satisfies* φ. The *size* of an LTL formula φ, written $|\varphi|$, is given by the number of subformulae in φ. We refer to the reader to [Emerson, 1990; Baier and Katoen, 2008] for full details on the syntax and semantics of LTL.

Simple Reactive Modules. We use an extension of the *Simple Reactive Modules Language* (SRML) [van der Hoek *et al.*, 2006] to model agents, which we refer to as *modules*. An SRML module consists of:

1. An *interface*, which defines the module's name and the Boolean variables under the control of the module; and
2. Two sets of *guarded commands*, which define the choices available to the module at every state.

Interfaces are specified by the syntax **module** m_i **controls** Φ_i, where m_i is the name of the module and $\Phi_i \subseteq \Phi$ is the set of variables under its exclusive control. Guarded commands consist of two parts: a precondition for executing the command, known as the *guard*, and the actual *command*, which specifies how the value of (some of) the variables under the module's control are updated when the command is executed.

In the extension of SRML we work with in this paper, we augment guarded commands with an additional *energy value*, which specifies how the module i's *energy level* at time t, denoted E_i^t, is changed if the command is executed. A positive energy value will increase available energy at time $t+1$; a negative value will decrease available energy. Given this, the general form of a guarded command in our augmented SRML is of the form `[energy] guard` $\rightsquigarrow$ `command` where `energy` is the associated energy cost/gain. More formally, a *guarded command* g for a module m_i controlling Boolean variables $\Phi_i \subseteq \Phi$ is an expression:

$$[e]\, \varphi \rightsquigarrow x_1' :- \psi_1; \cdots ; x_k' :- \psi_k,$$

where $e \in \mathbb{Z}$, φ and each ψ_j is a propositional logic formula over Φ, and every x_j' represents the value of the variable $x_j \in \Phi_j$ after the command is executed. The value of any variable in Φ_i that does not appear in a guarded command g is unchanged by the execution of g. We also require that $x_a' \neq x_b'$ for all $a \neq b \in \{1, \ldots, k\}$, i.e., no variable's value is reassigned twice in the same guarded command. For a guarded command $g = [e]\, \varphi \rightsquigarrow x_1' :- \psi_1; \cdots ; x_k' :- \psi_k$, $\eta(g) = e$ represents the change in energy level associated with g, $\mathsf{guard}(g) = \varphi$ represents the guard of g, and $\mathsf{evl}(g) = x_1' :- \psi_1; \cdots ; x_k' :- \psi_k$ represents the command of g. For an agent i with energy level E_i and a valuation $\vec{v}$, we say that a guarded command g_i for i is *enabled* if both $-E_i \leq \eta(g_i)$ and $\vec{v} \models \mathsf{guard}(g_i)$ and we write $\mathsf{enable}_i(\vec{v}, E_i)$ for the set of all guarded commands g_i which are enabled for i under $\vec{v}$.

For example, the guarded command $[-2]\ (p \vee q) \rightsquigarrow p' :- \bot; q' :- \top$ for a module m_i can be read as "if p or q are true and i has at least two units of energy, then *one of the actions available to* m_i is to set p to $\bot$ and q to $\top$, which incurs an energy cost of 2 units for m_i."

There are two kinds of guarded commands for a module: those used to *initialise* the variables under the module's control, and those used to subsequently *update* the variables. These are represented as sets of guarded commands **`init`** and **`update`**, respectively.

To ensure that agents always have a well-defined action available, we assume that in each agent's **`init`** set, at least one initial guarded command has a non-negative energy value. We also equip every module with a **`skip`** guarded command as part of its **`update`** set, which is always available and does nothing to the variables under its control. This can be explicitly written as $[e_i^{skip}]\ \top \rightsquigarrow \emptyset$, where $e_i^{skip} \in \mathbb{N}$ (including 0) and we use $\emptyset$ to denote a command which does nothing to the variables under the module's control. We require that the energy cost for the **`skip`** command be non-negative to ensure that at least one command is always available to every agent, which entails that runs are always well-specified.

Formally, an *SRML module*, m_i is given by a triple

$$m_i = (\Phi_i, \mathbf{init}_i, \mathbf{update}_i), \text{ where:}$$

- $\Phi_i \subseteq \Phi$ is the set of variables controlled by m_i;
- $\mathbf{init}_i$ is a finite set of *initialisation guarded commands* for m_i; and
- $\mathbf{update}_i$ is a finite set of *update guarded commands* for m_i.

An SRML arena is then simply a collection of agents, their representative modules, and the specification of each agent's initial and maximum energy values:

$$A = (N, \Phi, (m_i)_{i\in N}, (e_i^{max})_{i\in N}, (E_i^0)_{i\in N}),$$

where:

- $N = \{1, \ldots, n\}$ is a finite, non-empty set of *agents*;
- $\Phi = \bigcup_{i\in N} \Phi_i$ is a finite, non-empty set of *propositional variables*, where the sets Φ_i are all pairwise disjoint;
- $m_i = (\Phi_i, \mathtt{init}_i, \mathtt{update}_i)$ is an SRML module over Φ that defines the choices available to agent $i \in N$;
- $e_i^{max} \in \mathbb{N}$ is the *maximum energy capacity* of agent i [1]; and
- $E_i^0 \in \{0, \ldots, e_i^{max}\}$ is the *initial energy level* of i.

module m_i **controls** Φ_i
init
$\left[e_i^1\right]\ \top \rightsquigarrow \Phi_i' :-\ v_i^1$
$\cdots$
$\left[e_i^{m_1}\right]\ \top \rightsquigarrow \Phi_i' :-\ v_i^{m_1}$
update
$\left[e_i^{m_1+1}\right]\ \varphi_{m_1+1} \rightsquigarrow \Phi_i' :-\ v_i^{m_1+1}$
$\cdots$
$\left[e_i^{k_1}\right]\ \varphi_{k_1} \rightsquigarrow \Phi_i' :-\ v_i^{k_1}$
skip

Figure 1: A Reactive Module

An agent module m_i for an agent i is defined by augmenting the set of original variables Φ_i under their control with a set of *command variables*, which are used to identify precisely which action the agent took at each point in time. Given this, the *agent module* for i takes the following general form as in Figure 1:

(i) for any set $\Psi \subseteq \Phi$ of Boolean variables and a Boolean literal $v \in \{\top, \bot\}$, $\Psi :- v$ denotes the assignment where all variables in Ψ are set to v; (ii) for any ordered set $\Psi = \{p_1, \ldots, p_m\} \subseteq \Phi$ of Boolean variables and a binary string $B = b_1 b_2 \ldots b_m$, $\Psi' :- B$ denotes the assignment $p_1' :- b_1$; $p_2' :- b_2$; $\ldots$; $p_m' :- b_m$. Finally, for a given energy level E_i of an agent i, we let $\mathtt{init}_i|_E$ and $\mathtt{update}_i|_E$ denote the set of initial and update guarded commands for i respectively whose corresponding energy values are at least $-E$:

$$\begin{aligned} \mathtt{init}_i|_E &:- \{g \in \mathtt{init}_i \mid \eta(g) \geq -E\} \\ \mathtt{update}_i|_E &:- \{g \in \mathtt{update}_i \mid \eta(g) \geq -E\}. \end{aligned}$$

These sets will be useful in identifying the set of guarded commands that are available to an agent, given their current energy level at any point in a run.

3 Energy Reactive Modules Games

With these definitions in place, we can define the model for concurrent games that we focus on in this study. Formally, an *Energy Reactive Modules Game* (ERMG) is a structure

$$\mathcal{G} = (A, \gamma_1, \ldots, \gamma_n),$$

where A is an SRML arena and γ_i is the LTL goal of agent $i \in N$. At the beginning of a game, each agent $i \in N$ selects an initial guarded command $g_i^0 \in \mathtt{init}_i|_{E_i^0}$. The energy level of each agent i at timestep 1 is then updated as $E_i^1 = \min(E_i^0 + \eta(g_i^0), e_i^{max})$ and the initial valuation $\vec{v}_0$ of the variables in Φ is set according to the commands chosen, i.e., $\mathsf{evl}(g_i^0)$.[2] Then, at every step $t \in \mathbb{Z}^+$ of the execution, each agent i selects an enabled update guarded command $g_i^t \in \mathtt{update}_i|_{E_i^t}$ to execute, which updates the values of the variables in Φ_i according to $\mathsf{evl}(g_i^t)$ and updates the agent's energy level from E_i^t to $E_i^{t+1} = \min(E_i^t + \eta(g_i^t), e_i^{max})$. This update gives rise to a new valuation $\vec{v}_t$, and then the next round proceeds in the same manner. This process repeats indefinitely, giving rise to a *run*, which in an ERMG is an infinite sequence of valuations $\rho = \vec{v}_0\vec{v}_1\ldots$ where for all $t \in \mathbb{N}$, we have that $\vec{v}_{t+1}$ is obtained from the execution of enabled guarded commands by all agents $i \in N$. Note that given the restrictions on allowable actions at each time point, each agent's energy level will always be a non-negative integer throughout any run.

A particular class of games which can be represented using ERMGs are *regular RMGs*, which are exactly the same as ERMGs, except that no energy values are involved. Regular RMGs can thus be specified by an arena $A = (N, \Phi, (m_i)_{i\in N})$ and LTL goals $(\gamma_i)_{i\in N}$. Key decision problems in rational verification under the Nash equilibrium solution concept for regular RMGs have been well-studied in prior work [Gutierrez *et al.*, 2017].

We model strategies as deterministic Mealy machines, which are known to be sufficient for optimality in our setting [Gutierrez *et al.*, 2017]. A strategy for a player $i \in N$ with associated module $m_i = (\Phi_i, \mathtt{init}_i, \mathtt{update}_i)$ is a Mealy machine (i.e., a finite state machine with output) $\sigma_i = (Q_i, q_i^0, \delta_i, \tau_i)$, where Q_i is a finite set of machine states, q_i^0 is the initial state, and for all $q \in Q_i, \vec{v} \in \{0,1\}^{|\Phi|}$, we have:

- $\delta_i : Q_i \times \{0,1\}^\Phi \to Q_i$ is the strategy's deterministic state update function such that $\delta_i(q, \vec{v}) \neq q_i^0$, i.e., the initial state is never revisited;
- $\tau_i(q, \vec{v}) \in \mathsf{enable}_i(\vec{v}, E_i)$ is the output function that specifies which enabled guarded command is selected by player i with energy level E_i, such that $\tau_i(q, \vec{v}) \in \mathtt{init}_i$ iff $q = q_i^0$, i.e., initial guarded commands are only selected at the started of the game;

For each player $i \in N$, we let Σ_i represent their set of possible strategies and write $\Sigma = \prod_{i\in N} \Sigma_i$ represent the set of all *strategy profiles*, i.e., tuples of strategies for each player. Since we consider deterministic strategies in this setting, a strategy profile $\vec{\sigma}$ deterministically generates a run $\rho(\vec{\sigma}, \mathcal{G})$ in an ERMG $\mathcal{G}$, which consists of the infinite sequence of valuations generated by the execution of enabled guarded commands by modules at each time step.

Given this, we are now in a position to define preferences and utility functions over runs in ERMGs. We assume that a player $i \in N$ has the sole objective of satisfying their LTL

[1] We assume that each agent is equipped with a device that has a finite capacity for storing energy.

[2] It is also possible to consider settings with unbounded energy capacities, but this is not physically realistic and is likely to lead to undecidable verification problems [Bulling and Farwer, 2010].

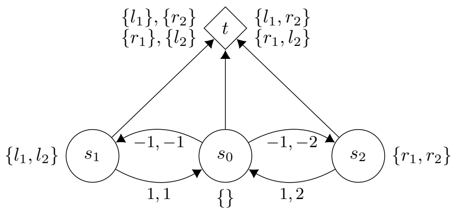

Figure 2: The diagram for Example 1. The losses/gains of energy for B_1 and B_2 are represented by edge labels. To maintain clarity, some labels are not shown. The nodes s_0, s_1, s_2 symbolise the positions of both robots in M, L, R respectively. The node t is a macrostate encompassing several states, as indicated by its labels. Each of these (macro)states also includes a self-loop edge (not shown) corresponding to the skip command.

goal γ_i. Given this, we say that player $i \in N$ *prefers* the run ρ over ρ', written $\rho \succeq_i \rho'$, if and only if $\rho' \models \gamma_i$ implies that $\rho \models \gamma_i$. This can be extended to the strict preference relation $\succ_i$ and the indifference relation $\sim_i$ in the usual manner (see [Maschler *et al.*, 2013]). Finally, since strategy profiles give rise to fixed runs in an ERMG, we can write $\vec{\sigma} \succeq_i \vec{\sigma}'$ as a shorthand for $\rho(\vec{\sigma}, \mathcal{G}) \succeq_i \rho(\vec{\sigma}', \mathcal{G})$. Notice that energy does not feature in the definition of preferences: agents are only secondarily concerned about energy consumption, in the sense that it places restrictions on the choices they can make.

Finally, we can define the first solution concept considered in this study. A strategy profile $\vec{\sigma}$ is a *Nash equilibrium* (or Nash-stable strategy profile) of an ERMG $\mathcal{G}$ if for all players $i \in N$ and all alternative strategies $\sigma_i' \in \Sigma_i$, it holds that $\vec{\sigma} \succeq_i (\vec{\sigma}_{-i}, \sigma_i')$. We let $\mathrm{NE}(\mathcal{G})$ denote the set of Nash Equilibria of an ERMG $\mathcal{G}$ and for an LTL formula φ, let $\mathrm{NE}_\varphi(\mathcal{G}) = \{\vec{\sigma} \in \mathrm{NE}(\mathcal{G}) : \rho(\vec{\sigma}, \mathcal{G}) \models \varphi\}$.

Example 1. *Consider a scenario involving two robots, B_1 and B_2, and three locations: middle (M), left (L), and right (R). From each location, each robot can either stay in the same position by choosing the skip command or move to another location. From M, each robot can move to either L or R. However, from L or R, they can only return to M. A move from one location to another can only be initiated when both robots are at the same location. Initially, both robots are positioned at M. For B_1, the cost of each move is -1. For B_2, moving left costs -1 but moving right costs -2. B_1 (resp. B_2) is assigned to service L (resp. R) at least once. However, for the service to be carried out, both robots must be at the same location. After completing the service and returning to M, B_1 and B_2 receive an energy recharge equal to the amount they have spent. For example, if B_1 and B_2 both visit R and then return to M, B_1 gains 1 unit of energy while B_2 gains 2 units. For $i \in \{1,2\}$, B_i is represented by the following module.*

module B_i ***controls*** $\{l_i, r_i\}$
init
$[0]\ \top \rightsquigarrow l_i' :=\ \bot; r_i' :=\ \bot$
update
$[-1]\ \neg(l_i \vee r_i \vee l_{3-i} \vee r_{3-i}) \rightsquigarrow l_i' :=\ \top$
$[-i]\ \neg(l_i \vee r_i \vee l_{3-i} \vee r_{3-i}) \rightsquigarrow r_i' :=\ \top$
$[1]\ l_i \wedge l_{3-i} \rightsquigarrow l_i' :=\ \bot$
$[i]\ r_i \wedge r_{3-i} \rightsquigarrow r_i' :=\ \bot$
skip

Each robot controls two variables l_i and r_i. If these variables are set to true, it indicates that the robot moves to the left or right. Let $E^0_{B_i} = e^{max}_{B_1} = 1, E^0_{B_2} = e^{max}_{B_2} = 2$. B_i's goal is given by $\gamma_{B_i} = \mathbf{F}\ s_i$. A graphical representation of the game is shown in Fig. 2. Consider a strategy profile which gives rise to the initial sequence $\{\}, \{l_1, l_2\}, \{\}, \{r_1, r_2\}$. This is a Nash equilibrium (NE), as the run satisfies $\gamma_{B_1} \wedge \gamma_{B_2}$, and the robots have enough energy to execute this strategy profile. Now, consider the sequence of assignments $\{\}, \{l_1, l_2\}, \{\}, \{r_1\}$. Although the robots also have enough energy to execute this run, it is not a NE. This is because B_2 can choose to set r_2 to true on the fourth round and achieve its goal.

4 Rational Verification

Before exploring mechanisms for agents to make energy transfers, we will briefly recapitulate the central questions of interest in rational verification, which are concerned with verifying whether a given temporal logic property holds in some or all rational outcomes of a multiplayer game. The crucial assumption here is that agents are rational and self-interested; that is, we can rule out executions of the game which are not stable with respect to deviations by individuals or coalitions of players.

4.1 Non-cooperative Games

The main non-cooperative rational verification problems we will deal with are as follows:

Given: Game $\mathcal{G}$, strategy profile $\vec{\sigma}$.
NASH-MEMBERSHIP: Is it true that $\vec{\sigma} \in \mathrm{NE}(\mathcal{G})$?

Given: Game $\mathcal{G}$, LTL goal φ.
E-NASH: Is it true that $\mathrm{NE}_\varphi(\mathcal{G}) \neq \emptyset$?
A-NASH: Is it true that $\mathrm{NE}_\varphi(\mathcal{G}) = \mathrm{NE}(\mathcal{G})$?

Using a polynomial time transformation from ERMGs into regular RMGs, we can establish that these rational verification problems are no harder than their counterparts in regular RMGs. Thus, we have the following results:

Proposition 1. NASH-MEMBERSHIP *for ERMGs is PSPACE-c, while* E-NASH *and* A-NASH *for ERMGs are 2EXPTIME-c.*

4.2 Cooperative Games

A central assumption in non-cooperative game theory is that agents act independently. In cooperative games, however, we assume that agents are able to form binding agreements with others – that is, agents can form *coalitions*, which may deviate from an outcome if it is mutually beneficial to do so. Here, we consider the well-known α-core, which was introduced in [Aumann, 1961], and remains a foundational solution concept in cooperative game theory. This solution concept assumes that any non-deviating players may respond to a coalitional deviation by trying to 'block' it, i.e., prevent the deviators from universally improving their utility.

Under this assumption, a strategy profile $\vec{\sigma}$ is said to be *core-stable* in $\mathcal{G}$ if for all coalitions $C \subseteq N$ such that $\rho(\vec{\sigma}, \mathcal{G}) \models \bigwedge_{i \in C} \neg\gamma_i$ and partial strategy profiles $\vec{\sigma}'_C = (\sigma_i')_{i \in C}$, there exists a response by the remaining players

$\vec{\sigma}'_{-C} = (\sigma'_i)_{i \in N \setminus C}$ such that for some player $i \in C$, it holds that $\rho((\vec{\sigma}'_C, \vec{\sigma}'_{-C}), \mathcal{G}) \not\models \gamma_i$. We let $\text{CORE}(\mathcal{G})$ denote the set of core-stable strategy profiles in $\mathcal{G}$ and define $\text{CORE}_\varphi(\mathcal{G}) = \{\vec{\sigma} \in \text{CORE}(\mathcal{G}) : \rho(\vec{\sigma}, \mathcal{G}) \models \varphi\}$. In the same manner, as with the Nash equilibrium solution concept, we can now define the cooperative rational verification problems as follows:

> *Given*: Game $\mathcal{G}$, LTL goal φ.
> E-CORE: Is it the case that $\text{CORE}_\varphi(\mathcal{G}) \neq \emptyset$?
> A-CORE: Does it hold $\text{CORE}_\varphi(\mathcal{G}) = \text{CORE}(\mathcal{G})$?

Example 2. *To illustrate the difference between NE and the core, let us revisit Example 1. Consider a strategy profile where both robots remain at location M forever. This strategy profile is a NE (albeit a suboptimal one), as there is no unilateral deviation by B_i that satisfies γ_{B_i}: any change in B_i's strategy would result in the game transitioning to state t. However, this strategy profile is not in the core, as the (grand) coalition $\{B_1, B_2\}$ has a strategy profile that ensures the satisfaction of $\gamma_{B_1} \wedge \gamma_{B_2}$. In fact, all strategy profiles in the core satisfy $\gamma_{B_1} \wedge \gamma_{B_2}$. As a result, an* A-NASH *query with $\varphi = \gamma_{B_1} \wedge \gamma_{B_2}$ would return a negative answer, while an* A-CORE *query would return a positive one.*

Because no results exist for cooperative rational verification in the context of regular RMGs, we solve these problems by reductions to and from concurrent game structures, for which these problems have been studied [Gutierrez *et al.*, 2023a]. Thus, we establish here the first results for cooperative rational verification in the RMGs model as well as in our extension to settings with bounded resources.

Proposition 2. E-CORE *and* A-CORE *for ERMGs and regular RMGs are 2EXPTIME-c.*

5 Endogenous ERMGs

We now turn to the primary focus of our study, which is to understand how energy offers can be strategically utilised by agents to achieve better outcomes in a stable manner. To this end, we first present a framework for modelling negotiations by introducing a pre-play offer phase, in which the agents can offer resource transfers to other agents before the game begins. Crucially, we maintain that such offers should be stable with respect to deviations by (coalitions of) players in the same way that strategies in the ensuing ERMGs are.

Formally, an *offer* by agent i is a function

$$\omega_i : N \setminus \{i\} \to \{0, \ldots, E_i^0\}, \text{ such that:}$$

1. $\mathsf{totoff}_i :- \sum_{j \neq i} \omega_i(j) \leq E_i^0$; and
2. for each agent $j \in N \setminus \{i\}$, we have $\omega_i(j) + E_j^0 \leq e_j^{max}$.

The first condition states that the total amount of energy an agent offers to other agents does not exceed their initial endowment, and the second condition states that offers must be made within the capacity constraints of their recipients. An offer thus specifies how much energy agent i proposes to transfer to each other agent in the game in the pre-play negotiation phase. Denote by Ω_i the set of all valid initial offers for an agent $i \in N$. Then, an *offer profile* $\vec{\omega} = (\omega_1, \ldots, \omega_n)$ is simply a tuple of offers for each agent $i \in N$, and we write $\Omega = \prod_{i \in N} \Omega_i$ for the set of all offer profiles. We write $\vec{\omega}^0$ for the *empty offer profile* in which no players offer any energy to any other player.

Given this, an *Endogenous Energy Reactive Modules Game* (EERMG) proceeds in the following manner:

Stage 1: Each agent $i \in N$ chooses a valid offer ω_i, giving rise to an offer profile $\vec{\omega}$.

Stage 2: Each agent chooses a strategy σ_i in the ERMG $G^{\vec{\omega}} = (A', \gamma_1, \ldots, \gamma_n)$, where A' is exactly the same as A, except that for each $i \in N$, we update i's initial energy as $E_i^{0'} = \min(E_i^0 - \mathsf{totoff}_i + \sum_{j \neq i} \omega_j(i), e_i^{max})$, for each agent i to reflect the resource transfer offers made in $\vec{\omega}$. The game $G^{\vec{\omega}}$ is then played according to the strategy profile $\vec{\sigma} = (\sigma_i)_{i \in N}$.

Given this two-stage game, the notion of a stable outcome can be ambiguous. This is because agents make two types of decisions, the former affecting the latter. Here, we will assume that when agents reason about offer profiles, they consider the stable outcomes induced by such offers. We use the classical Nash equilibrium and core stability notions for each stage separately. This means that we require both the energy offers to be stable against the appropriate kind of deviations *and* the strategy profile in the resulting stage 2 game to be stable to deviations. Given that we consider both unilateral and coalitional deviations, this approach allows us to combine the different solution concepts in different ways for each stage. Specifically, this gives rise to four possible combinations: Nash-Nash, Nash-Core, Core-Nash, and Core-Core. This decoupling between stability concepts for the first and second stages of the game allows one greater flexibility in finding a suitable model for the situation that they are trying to capture, because agents may have differing capabilities for communicating with one another and coordinating their behaviours in different stages of a game.

5.1 Stable Offers

To define a notion of stability over offer profiles, we require a way for the agents to rank such offers in terms of preferability. However, this is difficult to specify precisely, because different offer profiles can induce different sets of stable strategy profiles in stage 2 of the game. Since we do not make assumptions about how the agents resolve the equilibrium selection problem, we propose a minimal preference relation $\succ_i^S$ for each agent $i \in N$ over offer profiles that we should expect such agents to have. This relation is defined based on whether (i) A stable stage 2 outcome exists, (ii) A stable stage 2 outcome exists which satisfies i's goal, and (iii) All (and at least one) stable stage 2 outcomes satisfy i's goal. More concretely, for an agent $i \in N$ and a stability concept $S \in \{\text{NASH}, \text{CORE}\}$, let $\emptyset_i^S :- \{\vec{\omega} \in \Omega \mid S_{\gamma_i}(G^{\vec{\omega}}) = \emptyset\}$, $\exists_i^S :- \{\vec{\omega} \in \Omega \mid S(G^{\vec{\omega}}) \neq S_{\gamma_i}(G^{\vec{\omega}}) \neq \emptyset\}$, and $\forall_i^S :- \{\vec{\omega} \in \Omega \mid S(G^{\vec{\omega}}) = S_{\gamma_i}(G^{\vec{\omega}}) \neq \emptyset\}$.

Given any offer profiles $\vec{\omega}_1 \in \emptyset_i^S, \vec{\omega}_2 \in \exists_i^S$, and $\vec{\omega}_3 \in \forall_i^S$, we define $\succ_i^S$ such that $\vec{\omega}_3 \succ_i^S \vec{\omega}_2 \succ_i^S \vec{\omega}_1$. Moreover, for any two offer profiles $\vec{\omega}, \vec{\omega}'$ in the same set ($\emptyset_i^S, \exists_i^S$, or $\forall_i^S$), we assume that i is indifferent between the two offer profiles, i.e., $\vec{\omega} \sim_i^S \vec{\omega}'$. This preference relation captures the intuition that an agent should (i) prefer an offer profile in which it is *possible* for their goal to be achieved in some stable outcome over one in which it is not possible to do so, and (ii) prefer an offer profile in which their goal is *guaranteed* to be achieved in any stable outcome over one in which it is merely possible.

With this, we can define Nash equilibrium in the usual way. An offer profile $\vec{\omega}$ is *Nash-S-stable* for the stage 1 negotiation phase if for all agents $i \in N$ and all alternative offers $\omega'_i \in \Omega_i$, it holds that $\vec{\omega} \succeq_i^S (\vec{\omega}_{-i}, \omega'_i)$. For the case of core stability, the definition is not as immediate. The reason for this is that once a coalition $C \subseteq N$ deviates by making alternative offers, we assume that the remaining players $N \setminus C$ have the opportunity to *respond* and block the deviation from being profitable for all of the deviating players. In the context of offer profiles, however, a deviation can change the set of offers the remaining players can respond with. In this study, we will assume that a response takes into account the new offers made by the deviating players. However, we will also assume that for both the deviating and responding players, the option of withdrawing all previously made offers and re-allocating the withdrawn offers is always available. Given this, we will say that an offer profile $\vec{\omega}$ is *Core-S-stable* for stage 1 if for all coalitions $C \subseteq N$ and alternative partial offer profiles $\vec{\omega}'_A = (\omega'_i)_{i \in A}$, there is some response $\vec{\omega}'_{-A} = (\omega'_i)_{i \in N \setminus A}$ such that for some $i \in N$, it holds that $\vec{\omega} \succeq_i^S (\vec{\omega}'_A, \vec{\omega}'_{-A})$.

Example 3. *Consider the following modification of Example 1. Let $E^0_{B_1} = E^0_{B_2} = 0$ and for $i \in \{1, 2\}, \gamma_{B_i} = \mathbf{F}s_i \wedge \bigwedge_{j=0}^{2} \mathbf{G}(s_j \rightarrow \mathbf{X}\neg s_j)$, i.e., both B_1 and B_2 also do not want to stay in the same location twice in a row. We then introduce two additional robots B_3 and B_4 with $E^0_{B_3} = e^{max}_{B_3} = 2, E^0_{B_4} = e^{max}_{B_4} = 1$. Unlike the other robots, these additional robots are immobile and can only transfer their energy to others. Suppose that their goals are defined as $\gamma_{B_3} = \mathbf{X}s_1, \gamma_{B_4} = \mathbf{X}s_2 \vee \mathbf{X}t$. Consider the offer profile in which B_3 offers B_1 and B_2 1 unit each, and B_4 offers B_2 1 unit of energy. This profile is* CORE-CORE-*stable, as any possible offer deviation by B_3 does not result in γ_{B_3} being satisfied in all strategy profiles in the core of the second stage game (e.g., if B_3 withdraws the offer to B_2, B_4 can counter by also withdrawing its offer). On the other hand, in this scenario, there is no* NASH-CORE-*stable offer profile, since any valid offer profile will eventually enter a cycle of deviations.*

Example 3 illustrates that stable offer profiles may not exist in EERMGs. This example suggests that some games are inherently unstable during the first stage. Therefore, the stability of the negotiation phase becomes a critical issue to address. In the following sections, we examine some decision problems related to this and provide several algorithms for solving them. We demonstrate that solving such problems is no harder than standard rational verification.

5.2 Decision Problems

Now, we turn to the central question in this study, which is to determine whether a stable offer profile exists in a given EERMG. To this end, we introduce the following decision problems:

Given: Game $\mathcal{G}$, agent $i \in N$, offer profiles $\vec{\omega}_1, \vec{\omega}_2$, solution concept S for stage 2.
S-OFFER-PREFERENCE: Is it true that $\vec{\omega}_1 \succeq_i^S \vec{\omega}_2$?

Given: Game $\mathcal{G}$, solution concept S^1 for stage 1, solution concept S^2 for stage 2.
S^1-S^2-OFFER-EXISTENCE: Does there exist an S^1-S^2-stable offer profile $\vec{\omega}$ in $\mathcal{G}$?

It is worth noting that, in general, an EERMG is not guaranteed to have a NASH-S- or CORE-S-stable offer profile, since the preference relations $\succ_i^S$ implicitly group offer profiles into more than two categories, thus allowing "deviation cycles" to exist.

Turning to S-OFFER-PREFERENCE, for the upper-bound, we employ an algorithm that simply runs a stable profile check over the give offer profiles. For the lower bound, we reduce the problem of deciding whether an S-stable strategy profile exists in a regular RMG, hereafter called the S-NON-EMPTINESS problem.

Theorem 3. *For $S \in \{\text{NASH}, \text{CORE}\}$, S-OFFER-PREFERENCE is 2EXPTIME-c.*

Using this result, we can settle the complexity of the remaining problems. We show that they are also 2EXPTIME-c, and hence, reasoning about the existence of stable offers with desirable properties is no harder than the underlying rational verification problems. The upper bound for the NASH-S-OFFER EXISTENCE problem is established by checking, for every agent, whether there is no beneficial deviation from the given strategy profile, which can be done by employing a suitable variant of LTL synthesis. The lower bound again uses a reduction from the S-NON-EMPTINESS problem.

Theorem 4. *For $S \in \{\text{NASH}, \text{CORE}\}$, NASH-$S$-OFFER-EXISTENCE is 2EXPTIME-c.*

Finally, we turn to the cooperative setting and study the existence of core-stable offer profiles. Here, we establish the upper bound with a reasoning similar to the one for Nash-S-Offer-Existence, and reduce from S-NON-EMPTINESS to obtain the lower bound.[3]

Theorem 5. *For $S \in \{\text{NASH}, \text{CORE}\}$, CORE-$S$-OFFER-EXISTENCE is 2EXPTIME-c.*

6 Related Work

Endogenous games and side-payments: Jackson and Wilkie discuss the stability of transfers in their seminal work on endogenous games [Jackson and Wilkie, 2005]. In their work, transfers are dependent on the outcome of the resulting game and thus allow agents to make conditional offers to each other. In the context of infinite games, however, energy values may constrain the set of strategies that are available to players in an ERMG, so transfers should be made *before* the game begins to be of use to the agents. This motivates the

[3] Note that although we reduce from the same decision problem in these three cases, the constructions required in each case are very distinct, due to the differences in the considered solution concepts as well as the fact that the S^1-S^2-OFFER-EXISTENCE problems require one to consider the set of all possible offer profiles.

consideration of *unconditional* energy transfers in the setting we study. One could also extend our model to settings involving dynamic energy transfers, where each agent can transfer energy to other agents at every round of the game. We leave this as an open problem for future investigation.

Much work has been done on exploring pre-play negotiations and side payments in the context of strategic form games [Clercq *et al.*, 2016; Goranko and Turrini, 2016; Goranko, 2022; Renou, 2009; Turrini, 2016]. Of these, our work aligns most closely with the study conducted by Turrini (2016) [Turrini, 2016]. These studies consider one-shot *strategic-form* or *Boolean games* as their focus. In contrast, RMGs are played for an infinite number of rounds, which allows agents' goals to be modelled using expressive logics like LTL. Furthermore, unlike the mentioned approaches, energy transfers in our model do not directly influence the utility functions of the recipient players. Instead, these transfers serve only to improve an agent's capability to achieve their goal. Thus, our study offers a different perspective and a complementary framework for examining resource transfers that primarily aid in goal achievement without being intrinsically valued or optimised by the agents.

Resource-bounded games and logics: Resource-bounded games have attracted considerable attention and have been explored in various contexts. *Energy games* [Chakrabarti *et al.*, 2003; Bouyer *et al.*, 2008] and their subsequent extensions are typically played in a two-player, turn-based, zero-sum setting. In *energy parity games* [Chatterjee and Doyen, 2012], the objective of Player 1 is to satisfy a qualitative parity condition while maintaining a positive energy level. This aligns closely with our model here, as LTL formulae can always be translated into parity conditions [Piterman, 2007]. However, the game graphs in [Chatterjee and Doyen, 2012] are singly-weighted, preventing a direct reduction of our games to theirs. Energy games with reachability [Hélouët *et al.*, 2022] and ω-regular objectives [Amram *et al.*, 2021] are also studied in the literature, but again focus only on two-player games. Energy games played on multi-weighted graphs are considered in [Velner *et al.*, 2015; Kupferman and Halevy, 2022], but these settings only consider a quantitative condition, i.e., the players' energy levels. The work [Maubert *et al.*, 2019] is also relevant, studying the existence of *winning* strategies for a team of agents to achieve some LTL formula in one form of concurrent game structures. This is very similar to the notion of a beneficial deviation in the cooperative setting we consider, but we focus on the existence of strategy profiles which are *stable against deviations*.

Many logics have been developed for reasoning about resource-bounded games (see [Alechina and Logan, 2020] for an extensive overview of such logics). For instance, pe-ATL [Della Monica and Murano, 2018] is an extension of the logic ATL, which can be used to reason about energy parity games involving multiple agents. Alechina et al. introduced RB-ATL [Alechina *et al.*, 2010; Nguyen *et al.*, 2018], another extension which assumed that resources could only be consumed and not replenished. This was then expanded by RB $\pm$ ATL(*) [Alechina *et al.*, 2017; Alechina *et al.*, 2018], allowing for both resource consumption and production. Pertaining to our work, note that while (RB $\pm$)ATL* can encode statements about the core, it cannot express statements regarding the existence of Nash equilibria.

Rational Verification: Our work builds upon the existing literature on rational verification and synthesis [Fisman *et al.*, 2010; Wooldridge *et al.*, 2016; Kupferman *et al.*, 2016; Almagor *et al.*, 2018]. While many studies have used RMGs as the model for rational verification (e.g., [Gutierrez *et al.*, 2017; Gutierrez *et al.*, 2018; Najib, 2019; Steeples *et al.*, 2021]), few have addressed resource-boundedness. Electric Boolean games [Harrenstein *et al.*, 2015; Oualhadj and Troquard, 2016] consider resource-bounded games like our approach. While these models study resource redistributions, they operate under the assumption that a central authority exists that is capable of redistributing energy endowments. In contrast, the key novel aspect of our model is that we assume agents are independent and can choose their redistributions, giving rise to strategic considerations regarding transfers.

7 Conclusion

We have introduced ERMGs, a model of concurrent games with resource-constrained players using the reactive modules framework, and settled the complexity of both the key cooperative and non-cooperative rational verification questions for this model, showing that they are no harder than for regular RMGs. This therefore expands the range of practical applications which can be succinctly modelled using the reactive modules framework to analyse situations related to bounded resources, with almost no additional complexity cost. We then introduced Endogenous ERMGs, which include a preplay phase that allows players to make energy offers to one another, along with a notion of preferences over offer profiles. This preference relation enables us to study the stability of offers in both cooperative and non-cooperative settings. We then settled the problem of deciding whether a stable offer profile exists in a given game under all combinations of stability concepts, which remains 2EXPTIME-c. This initial study of the Endogenous ERMG model thus sheds light on the strategic considerations an agent faces when their actions in one stage of a game may affect the possible rational outcomes in the second stage of the game.

One natural extension of our work is to study how a top-down incentive designer [Gutierrez *et al.*, 2019; Hyland *et al.*, 2023] could modify the energy levels of agents to shape the set of resulting stable outcomes. Relatedly, one could take an approach akin to *parameterised* resource-bounded ATL [Alechina *et al.*, 2018; Alechina *et al.*, 2020], which can express formulae about the minimal amount of energy a coalition needs to achieve a particular goal, by asking whether energy subsidies or taxes can be introduced to enable specific outcomes. Finally, a direct next step would be to study whether some or all stable offer profiles induce games with (un)desirable properties. This becomes especially pertinent when considering the presence of malicious agents who might collude by exchanging energies to bring about an undesirable outcome. These results on the existence of stable offer profiles lay the groundwork for these future developments.

Acknowledgments

This research has been supported by the PNRR MUR project PE0000013-FAIR. Perelli was partially supported also by MUR under the PRIN 2020 projects PINPOINT and Sapienza University of Rome under the "*Progetti Grandi di Ateneo*" programme, grant RG123188B3F7414A (ASGARD - Autonomous and Self-Governing Agent-Based Rule Design). Wooldridge was supported by the UKRI under a Turing AI World Leading Research Fellowship (EP/W002949/1). Gutierrez was supported by the DARPA Assured Neuro Symbolic Learning and Reasoning (ANSR) program under award number FA8750-23-2-1016. The authors would like to extend special thanks to Paul Harrenstein for valuable feedback on an early draft of this paper.

References

[Abate *et al.*, 2021] Alessandro Abate, Julian Gutierrez, Lewis Hammond, Paul Harrenstein, Marta Kwiatkowska, Muhammad Najib, Giuseppe Perelli, Thomas Steeples, and Michael Wooldridge. Rational verification: game-theoretic verification of multi-agent systems. *Applied Intelligence*, 51(9):6569–6584, 2021.

[Alechina and Logan, 2020] Natasha Alechina and Brian Logan. State of the art in logics for verification of resource-bounded multi-agent systems. In *Fields of Logic and Computation III - Essays Dedicated to Yuri Gurevich on the Occasion of His 80th Birthday*, volume 12180 of *Lecture Notes in Computer Science*, pages 9–29. Springer, 2020.

[Alechina *et al.*, 2010] Natasha Alechina, Brian Logan, Nguyen Hoang Nga, and Abdur Rakib. Resource-bounded alternating-time temporal logic. In *AAMAS'10*, pages 481–488. Citeseer, 2010.

[Alechina *et al.*, 2017] Natasha Alechina, Brian Logan, Hoang Nga Nguyen, and Franco Raimondi. Model-checking for resource-bounded ATL with production and consumption of resources. *J. Comput. Syst. Sci.*, 88:126–144, 2017.

[Alechina *et al.*, 2018] Natasha Alechina, Nils Bulling, Stéphane Demri, and Brian Logan. On the complexity of resource-bounded logics. *Theoretical Computer Science*, 750:69–100, 2018.

[Alechina *et al.*, 2020] Natasha Alechina, Stéphane Demri, and Brian Logan. Parameterised resource-bounded atl. In *AAAI'20*, volume 34, pages 7040–7046, 2020.

[Almagor *et al.*, 2018] Shaull Almagor, Orna Kupferman, and Giuseppe Perelli. Synthesis of controllable nash equilibria in quantitative objective game. In Jérôme Lang, editor, *IJCAI'18*, pages 35–41. ijcai.org, 2018.

[Amram *et al.*, 2021] Gal Amram, Shahar Maoz, Or Pistiner, and Jan Oliver Ringert. Efficient algorithms for omega-regular energy games. In Marieke Huisman, Corina S. Pasareanu, and Naijun Zhan, editors, *FM'21*, volume 13047 of *Lecture Notes in Computer Science*, pages 163–181. Springer, 2021.

[Aumann, 1961] Robert J Aumann. The core of a cooperative game without side payments. *Transactions of the American Mathematical Society*, 98(3):539–552, 1961.

[Baier and Katoen, 2008] Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.

[Bouyer *et al.*, 2008] Patricia Bouyer, Uli Fahrenberg, Kim G Larsen, Nicolas Markey, and Jiří Srba. Infinite runs in weighted timed automata with energy constraints. In *FORMATS'08*, pages 33–47. Springer, 2008.

[Brice *et al.*, 2023] Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Rational verification for nash and subgame-perfect equilibria in graph games. In *MFCS'23*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2023.

[Bruyère *et al.*, 2022] Véronique Bruyère, Jean-François Raskin, and Clément Tamines. Pareto-rational verification. In *CONCUR'22*, 2022.

[Bulling and Farwer, 2010] Nils Bulling and Berndt Farwer. On the (un-)decidability of model checking resource-bounded agents. In *ECAI'10*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, pages 567–572. IOS Press, 2010.

[Chakrabarti *et al.*, 2003] Arindam Chakrabarti, Luca De Alfaro, Thomas A Henzinger, and Mariëlle Stoelinga. Resource interfaces. In *International Workshop on Embedded Software*, pages 117–133. Springer, 2003.

[Chatterjee and Doyen, 2012] Krishnendu Chatterjee and Laurent Doyen. Energy parity games. *Theoretical Computer Science*, 458:49–60, 2012.

[Clercq *et al.*, 2016] Sofie De Clercq, Steven Schockaert, Ann Nowé, and Martine De Cock. Formalizing Commitment-Based Deals in Boolean Games. In *ECAI'16*, volume 285 of *Frontiers in Artificial Intelligence and Applications*, pages 329–337. IOS Press, 2016.

[Della Monica and Murano, 2018] Dario Della Monica and Aniello Murano. Parity-energy atl for qualitative and quantitative reasoning in mas. In *AAMAS'18*, page 1441–1449, Richland, SC, 2018. International Foundation for Autonomous Agents and Multiagent Systems.

[Emerson, 1990] E. A. Emerson. Temporal and modal logic. In *Handbook of Theoretical Computer Science Volume B: Formal Models and Semantics*, pages 996–1072. Elsevier, 1990.

[Fisman *et al.*, 2010] Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In *TACAS*, volume 6015 of *LNCS*, pages 190–204. Springer, 2010.

[Goranko and Turrini, 2016] Valentin Goranko and Paolo Turrini. Two-player preplay negotiation games with conditional offers. *IGTR*, 18(1):1550017:1–1550017:31, 2016.

[Goranko, 2022] Valentin Goranko. Preplay negotiations with unconditional offers of side payments in two-player strategic-form games: Towards non-cooperative cooperation. *Mathematics*, 10(14):2518, 2022.

[Gutierrez *et al.*, 2017] Julian Gutierrez, Paul Harrenstein, and Michael Wooldridge. From model checking to equilibrium checking: Reactive modules for rational verification. *Artificial Intelligence*, 248:123–157, 2017.

[Gutierrez *et al.*, 2018] Julian Gutierrez, Giuseppe Perelli, and Michael Wooldridge. Imperfect information in reactive modules games. *Information and Computation*, 261:650–675, 2018.

[Gutierrez *et al.*, 2019] Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael J. Wooldridge. Equilibrium Design for Concurrent Games. In *CONCUR'19*, 2019.

[Gutierrez *et al.*, 2023a] Julian Gutierrez, Szymon Kowara, Sarit Kraus, Thomas Steeples, and Michael Wooldridge. Cooperative concurrent games. *Artificial Intelligence*, 314:103806, 2023.

[Gutierrez *et al.*, 2023b] Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael Wooldridge. On the complexity of rational verification. *Annals of Mathematics and Artificial Intelligence*, 91(4):409–430, 2023.

[Harrenstein *et al.*, 2015] Paul Harrenstein, Paolo Turrini, and Michael Wooldridge. Electric boolean games: redistribution schemes for resource-bounded agents. In *AAMAS'15*, pages 655–663. ACM, 2015.

[Hélouët *et al.*, 2022] Loïc Hélouët, Nicolas Markey, and Ritam Raha. Reachability games with relaxed energy constraints. *Inf. Comput.*, 285(Part):104806, 2022.

[Hyland *et al.*, 2023] David Hyland, Julian Gutierrez, and Michael J. Wooldridge. Incentive engineering for concurrent games. In Rineke Verbrugge, editor, *TARK'23*, volume 379 of *EPTCS*, pages 344–358, 2023.

[Jackson and Wilkie, 2005] Matthew O Jackson and Simon Wilkie. Endogenous games and mechanisms: Side payments among players. *The Review of Economic Studies*, 72(2):543–566, 2005.

[Kupferman and Halevy, 2022] Orna Kupferman and Naama Shamash Halevy. Energy games with resource-bounded environments. In *CONCUR'22*, volume 243 of *LIPIcs*, pages 19:1–19:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[Kupferman *et al.*, 2016] Orna Kupferman, Giuseppe Perelli, and Moshe Y. Vardi. Synthesis with rational environments. *Ann. Math. Artif. Intell.*, 78(1):3–20, 2016.

[Maschler *et al.*, 2013] Michael Maschler, Eilon Solan, and Shmuel Zamir. *Game Theory*, chapter Utility theory, page 9–38. Cambridge University Press, 2013.

[Maubert *et al.*, 2019] Bastien Maubert, Aniello Murano, Pasquale Perillo, Sasha Rubin, and Alessio Spasiano. Towards a tool for LTL synthesis with bounded-energy constraints. In *ICTCS'19*, volume 2504 of *CEUR Workshop Proceedings*, pages 229–234. CEUR-WS.org, 2019.

[Najib, 2019] M Najib. *Rational verification in multi-agent systems*. PhD thesis, University of Oxford, 2019.

[Nguyen *et al.*, 2018] Hoang Nga Nguyen, Natasha Alechina, Brian Logan, and Abdur Rakib. Alternating-time temporal logic with resource bounds. *J. Log. Comput.*, 28(4):631–663, 2018.

[Oualhadj and Troquard, 2016] Y Oualhadj and Nicolas Troquard. Rational verification in iterated electric boolean games. In *SR'16*, volume 218, page 11. Open Publishing Association, 2016.

[Piterman, 2007] Nir Piterman. From nondeterministic b́uchi and streett automata to deterministic parity automata. *Logical Methods in Computer Science*, 3, 2007.

[Pnueli, 1977] A. Pnueli. The temporal logic of programs. In *Proceedings of the Eighteenth IEEE Symposium on the Foundations of Computer Science (FOCS'77)*, pages 46–57. The Society, 1977.

[Renou, 2009] Ludovic Renou. Commitment games. *Games and Economic Behavior*, 66(1):488–505, 2009.

[Steeples *et al.*, 2021] Thomas Steeples, Julian Gutierrez, and Michael Wooldridge. Mean-payoff games with ω-regular specifications. In *AAMAS'21*, pages 1272–1280, 2021.

[Turrini, 2016] Paolo Turrini. Endogenous games with goals: side-payments among goal-directed agents. *AAMAS'16*, 30:765–792, 2016.

[van der Hoek *et al.*, 2006] W. van der Hoek, A. Lomuscio, and M. Wooldridge. On the complexity of practical ATL model checking. In H. Nakashima, M. P. Wellman, G. Weiss, and P. Stone, editors, *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems (AAMAS'06)*, pages 201–208. ACM, 2006.

[Velner *et al.*, 2015] Yaron Velner, Krishnendu Chatterjee, Laurent Doyen, Thomas A Henzinger, Alexander Rabinovich, and Jean-François Raskin. The complexity of multi-mean-payoff and multi-energy games. *Information and Computation*, 241:177–196, 2015.

[Wooldridge *et al.*, 2016] Michael J. Wooldridge, Julian Gutierrez, Paul Harrenstein, Enrico Marchioni, Giuseppe Perelli, and Alexis Toumi. Rational Verification: From Model Checking to Equilibrium Checking. In Dale Schuurmans and Michael P. Wellman, editors, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*, pages 4184–4191. AAAI Press, 2016.