Games of Incomplete Information: a Framework Based on Belief Functions

Pierre Pomeret-Coquot, Hélène Fargier, Érik Martin-Dorel

Games of Incomplete Information: A Framework Based on Belief Functions *

Pierre Pomeret-Coquot[a,b,*], Helene Fargier[a,b], Érik Martin-Dorel[b]

[a] *ANITI, B612, 3, rue Tarfaya, 31400 Toulouse*
[b] *IRIT, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 9*

Abstract

This paper proposes a model for incomplete-information games where the knowledge of the players is represented by a Dempster-Shafer belief function. Beyond an extension of the classical definitions, it shows such a game can be transformed into an equivalent hypergraphical complete-information game (without uncertainty), thus generalizing Howson and Rosenthal's theorem to the framework of belief functions and to any number of players. The complexity of this transformation is finally studied and shown to be polynomial in the degree of k-additivity of the mass function.

Keywords: Game theory, Incomplete-information games, Belief functions, Choquet integrals

---

1. Introduction

Game theory [1, 2] proposes a powerful framework to capture decision problems involving several agents. In non-cooperative games of complete information, the players do not coordinate their actions but each of them knows everything about the game: the players, their available actions and all their utilities. This assumption of complete knowledge cannot always be satisfied. In the real world indeed, players are not so well informed and have only limited knowledge

---

*Corresponding author – `pierre.pomeret@irit.fr`
*Pierre Pomeret-Coquot and Helene Fargier have benefited from the AI Interdisciplinary Institute ANITI. ANITI is funded by the French "Investing for the Future – PIA3" program under the Grant agreement n°ANR-19-PI3A-0004.

about the game. This is why Bayesian games of incomplete information have been proposed [3]. Nevertheless, the Bayesian hypothesis is strong, and requires a good knowledge of the environment. For instance, in case of ignorance, the Bayesian way can lead to a model that does not fit with the agents' behavior (e.g., see Ellsberg's paradox [4] or Von Mises' wine/water paradox [5]).

In the present paper, we propose a new kind of game of incomplete information, which we call a *Bel game*. Agents have a partial knowledge, represented by a Dempster-Shafer belief function [6, 7], and cardinal utilities, but do not necessarily make the equiprobability assumption. The underlying decision rule is generally the Choquet integral based on the *Bel* measure [8], which amounts to the maximization of the worst expected utility [9, 10]. Bel games as defined here are also compatible with the transferable belief model [11], which amounts to extracting the pignisitic probability when the decision is to be made, and with Jaffray's linear utility [12].

We then follow the line defined by Howson and Rosenthal [13] who have shown that any 2-player Bayesian game can be transformed into a complete-information *polymatrix game* [14]. In this paper, we show that such a transformation is possible for Bel games, and for any number of agents, producing a *hypergraphical game* [15]. A notable consequence of this result is that the algorithmics developed for hypergraphical games [16, 17] can be reused for the search of Nash equilibria in Bel games.

2. Background and motivations

To illustrate and motivate our work, we will use the following example inspired by the *murder of Mr. Jones* [11], where the suspects are *Peter, Paul and Mary.*

Example 1 (Peter, Quentin and Rose). *Two agents, named Agent* 1 *and Agent* 2*, are independently looking for a business association, with either Peter (P), Quentin (Q), or Rose (R). The point is that a crime has been committed, for which these three people are suspected.*

On the one hand, a classification algorithm was run on the surveillance video on which the murderer appears, but the poor quality did not allow to obtain a better result than the murderer has 50% chance to be a man, and 50% chance to be a woman. On the other hand, the police investigation deduced, by elimination, that the only three suspects are P, Q and R.
As to the interest of the associations, making the deal with an innocent leads to a payoff of \$6k (to be shared between the people making the deal), while associating with a guilty person produces no payoff (\$0k).
Moreover, Agent 1 is investigating about P and will eventually know whether he is guilty or not. Similarly, Agent 2 will know if R is guilty before making the decision.

The Bayesian approach is not appropriate here. Indeed, if Agent 1 learns that P is innocent, the probability of guilt should become 1/2 for Q and 1/2 for R. However, in a purely Bayesian view, equiprobability would be applied and the prior probability of guilt would be 1/4 to P and 1/4 to Q. Then, after conditioning, Agent 1 would assign a probability of 1/3 to Q and 2/3 to R.

2.1. Dempster-Shafer's theory of evidence

Let us first look at the epistemic aspect of the problem. The prior knowledge is simply that $P(\{P,Q\}) = P(\{M\}) = \frac{1}{2}$, and nothing more. The kind of knowledge at work here is well captured by Dempster-Shafer's theory of evidence, which does not restrict probability assignments to elements of the frame of discernment:

Definition 1 (Mass function). *A mass function for a frame of discernment Ω (or "bpa" for basic probability assignment) is a function $m : 2^\Omega \to [0,1]$ such that $m(\emptyset) = 0$ and $\sum_{A \subseteq \Omega} m(A) = 1$.*

A set with a nonzero mass is called a *focal element* and the set of focal elements is denoted $\mathcal{S}_m$. Two dual measures on 2^Ω derive from m:

$$\mathrm{Bel}(A) = \sum_{B \in \mathcal{S}_m, B \subseteq A} m(B) \qquad \text{and} \qquad \mathrm{Pl}(A) = \sum_{B \in \mathcal{S}_m, B \cap A \neq \emptyset} m(B).$$

Bel(A) (resp. Pl(A)) estimates to what extent A is implied by (resp. is compatible with) the knowledge captured by m. Since a belief function is directly encoded by its mass function, its spatial complexity is the number of focal elements: $\mathrm{Size}(m) = |\mathcal{S}_m|$.

Probabilities correspond to the special case of belief functions where mass functions are 1-additive: the focal elements are singletons. k-additivity is more generally defined as follows:

Definition 2 (k-additivity). *A mass function whose largest focal element is of size k is said to be k-additive, i.e., $k = \max_{B \in \mathcal{S}_m} |B|$.*

Probability theory is recovered as a special case when the available information is uncertain, but precise. On the contrary, when there is only one focal element B, i.e., when the belief function describes a piece of evidence that tells us that ω is in B for sure, and nothing more, the information is certain but imprecise. Following this interpretation, the mass function is seen as a generalized set [18].

Belief function can alternatively be understood as a particular case of imprecise probability theory. A belief function Bel and its dual Pl indeed delimit the lower and upper bounds of a probability family $\mathcal{F} = \{\Pr \mid \forall A, \mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)\}$ – $\mathcal{F}$ is called the credal set of m. Since this set is defined by linear constraints, it is convex, i.e., any element of $\mathcal{F}$ can be obtained by distributing each mass $m(B)$ among the elements of B. But on the other side, not any probability family is bound by a belief function and its dual [19].

Example 2 (Belief function modelling). *In Example 1, there are three possible "states of the world": one for each potential murderer. So, $\Omega = \{P, Q, R\}$. The knowledge at work here says that there is 50% chance that the murderer is a man (thus, P or Q), and that there is 50% chance that the murderer is a woman (thus, R).*

- *In the evidential interpretation: this knowledge is directly captured by the mass function m: $m(\{R\}) = 0.5$; $m(\{P, Q\}) = 0.5$ (cf. Figure 1).*

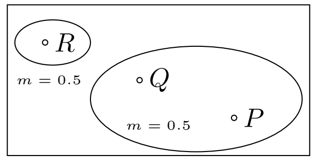

Figure 1: Example 1's prior knowledge expressed by a mass function; circles denote focal elements

- *In the credal interpretation: m, as well as* Bel *and* Pl*, are just tools to delimit the probability family* $\mathcal{F} = \{\Pr \mid \Pr(\{P, Q\}) = 0.5 \quad \wedge \quad \Pr(\{R\}) = 0.5\}$.

In this toy example, the belief function arises by mixing variability and ignorance (the probabilistic 50-50 result and the logical result of having those three suspects). Belief functions also arise in data-driven approaches: yielded by statistical inference Belief functions also arise in data-driven approaches: yielded by statistical inference [6, 20, 21] or by belief-function-based algorithms [22].

2.2. Decision making with belief functions

Let us now consider belief functions in a (mono-agent) decision making context. A common setting in decision theory is to consider that a decision (or "action") is a function $a : \Omega \to X$ where Ω is the set of possible states, as previously, and X is the set of possible outcomes. The preferences of an agent are represented by an utility function $u : X \to \mathbb{R}$. As soon as the global preference of the decision maker is supposed to be complete and transitive, it can be captured by an aggregated utility – the higher the global utility of an action, the more the decision maker prefers this action (the reader shall refer to [23] for a survey of decision making with belief functions). In the following, we consider three aggregation functions that generalize expected utility to belief functions, yielding a complete order over the decision maker's preferences.

The multi prior model and the discrete Choquet integral

A belief function Bel and its dual Pl delimit the lower and upper bounds of a probability family $\mathcal{F} = \{Pr \mid \forall A, \mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)\}$, and thus, for each decision, a full range of values for its expected utility. Gilboa and Schmeidler [10, 9] consider a pessimistic agent and propose to evaluate each decision by the minimum of its possible expected utilities.

Definition 3 (MPEU). $\mathrm{MPEU}(a) = \min_{\Pr \in \mathcal{F}} \left(\sum_\omega \Pr(\omega).u(a(\omega))\right)$

Another line of thought is to consider that belief functions are particular capacity measures, and to compute the global merit of an act on the basis of its Choquet value (for a theoretical justification of the use of Choquet integral in decision making under (non-probabilitistic) uncertainty, see [24, 25]). When the capacity is a belief function, the Choquet expectation can be written as follows.

Definition 4 (Choquet expected utility (CEU)). *Let $\Lambda(a) = \{\lambda_1 \leq \cdots \leq \lambda_{|\Lambda(a)|}\}$ be the set of utility values reached by an action a, labelled by increasing order, and $E_{\lambda_i}(a) = \{\omega \mid u(a(\omega)) \geq \lambda_i\}$ denote the set of worlds for which the utility of action a is at least λ_i. The Choquet Expected Utility of a is:*

$$\mathrm{CEU}(a) = \lambda_1 + \sum_{i=2}^{|\Lambda(a)|} (\lambda_i - \lambda_{i-1}) \times \mathrm{Bel}(E_{\lambda_i}(a)).$$

The CEU has a simple expression in terms of the mass function:

$$\mathrm{CEU}(a) = \sum_{B \in \mathcal{S}_m} m(B) \times \min_{\omega \in B} u(a(\omega)).$$

Example 3 (Choquet expected utility). *Consider a decision situation much simpler than that of Example 1: there is only one agent who chooses a partner without any investigation. In this case, the agent chooses to partner with either Peter, Quentin or Rose, depending only on the knowledge that men, like women, have a 50% chance of being the murderer, i.e., $m\big(\{\omega_P, \omega_Q\}\big) = m\big(\{\omega_R\}\big) = \frac{1}{2}$.*

The CEU *function of the different possible actions is:*

$$\begin{aligned}
\mathrm{CEU}(P) &= \frac{1}{2} \times \min\big(u(P,\omega_P), u(P,\omega_Q)\big) + \frac{1}{2} \times \min\big(u(P,\omega_R)\big) &&= 1.5\\
\mathrm{CEU}(Q) &= \frac{1}{2} \times \min\big(u(Q,\omega_P), u(Q,\omega_Q)\big) + \frac{1}{2} \times \min\big(u(Q,\omega_R)\big) &&= 1.5\\
\mathrm{CEU}(R) &= \frac{1}{2} \times \min\big(u(R,\omega_P), u(R,\omega_Q)\big) + \frac{1}{2} \times \min\big(u(R,\omega_R)\big) &&= 1.5
\end{aligned}$$

Since $\mathrm{CEU}(P) = \mathrm{CEU}(Q) = \mathrm{CEU}(R)$*, the agent is indifferent w.r.t. the possible partners: they consider indeed that every potential partner has 50% chance of being the murderer, which amounts to the worst case scenario for each choice.*

Gilboa and Schmeidler [26] have shown that the MPEU value of a decision is equal to its CEU value when the imprecise probability yields a belief function. We recover here the double interpretation of belief functions, in terms of generalized set or in terms of imprecise probability. It should nevertheless be recalled that they do differ, especially in a dynamic context, where a conditioning must be applied.

For the sake of completeness, let us remark that the Choquet expectation could be defined w.r.t the plausibility measure in the former equation of Definition 4, which then amounts to considering the max operator in the latter equation. This would capture the behaviour of an optimistic agent, and coincides with the maximum of possible expected utility in the multi prior model.

*Jaffray's expected utility (*JEU*)*

Jaffray's [12] expected utility (JEU) is defined directly in terms of the mass function, and generalizes the CEU by allowing one to modulate the agent's pessimism locally, using a series of Hurwicz coefficients. Denoeux and Shenoy also provide an axiomatic justification for Jaffray's linear utility (and for an even more general notion of it) in the Dempster-Shafer theory, i.e., for the evidential interpretation [27].

$$\mathrm{JEU}(a) = \sum_{B \in \mathcal{S}_m} m(B) \times \big(\alpha_{B_*(a),B^*(a)} \times B_*(a) + (1 - \alpha_{B_*(a),B^*(a)}) \times B^*(a)\big)$$

where $B_*(a) = \min_{\omega \in B} u(a(\omega))$ and $B^*(a) = \max_{\omega \in B} u(a(\omega))$.

The α_{x_i,x_j} coefficients represent the agent's pessimism and have to be elicited for each pair (x_i, x_j) with $i < j$; there is a quadratic number of such pairs (w.r.t. the possible utility values). Note that JEU = CEU if all coefficients α_{x_i,x_j} are equal to 1. If all coefficients are equal to 0, it coincides with the CEU w.r.t. the plausibility measure.

*The transferable belief model (*TBEU*)*

This model, due to Smets and Kennes [11], proceeds in two steps: at the credal level, knowledge is represented by a belief function and revised by Dempster's rule of conditioning; at the pignistic level (when a decision takes place), a probability law named BetP is constructed: $\mathrm{BetP}(\omega) = \sum_{\omega \in B \in \mathcal{S}_m} m(B)/|B|$ (which amounts to the Shapley power index computed for the game represented by Bel [28]). The agent then maximizes his expected utility with regards to BetP.

2.3. Knowledge revision with belief functions

The classical probabilistic conditioning makes it possible to revise knowledge expressed by a probability measure. In the context of this paper, we simply need to recall that if $\Pr : 2^\Omega \to [0,1]$ is a probability measure whose distribution is $p : \Omega \to [0,1]$, then $p(\omega \mid C) = p(\omega)/\Pr(C)$ if $\omega \in C$ and 0 otherwise. It means that when learning that event C holds, one only considers the chances that were assigned to elements $\omega \in C$ – up to a normalization by $\Pr(C)$.

The evidential and the credal views of belief functions lead to different generalizations of the probabilistic conditioning [29, 30]. In a pure DS theory, in the conditioning of m by C, the mass assigned to a focal set B is transferred to their non-empty intersection. The conditioning at work here is Dempster's rule [6] (see [31] for more details):

Definition 5 (Dempster conditioning). *For any nonempty $A, C \subseteq \Omega$, with* $\mathrm{Pl}(C) > 0$ *(at least one focal element intersects C),*

$$m_{|C}^{Dem}(A) := K_C \cdot \sum_{\substack{B \in \mathcal{S}_m \\ C \cap B = A}} m(B),$$

where $K_C = 1/\mathrm{Pl}(C)$ is a normalization factor, constant for a given subset $C \subseteq \Omega$.

Masses of $m_{|C}^{\mathrm{Dem}}$ can be computed using a simple algorithm. The mass $m(B)$ of any focal element $B \in \mathcal{S}_m$ is transferred to the subset $B \cap C$ if it is nonempty, and discarded otherwise. Thus all masses of $m^{\text{Dem}}_{|C}$ can be computed through a single loop over $\mathcal{S}_m$, followed by a normalization, in linear time w.r.t. the size of m.

It is worthwhile noticing that Demspter conditioning preserves the size and the k-additivity of m: $|\mathcal{S}_m| \geq |\mathcal{S}_{m^{\text{Dem}}_{|C}}|$ and if m is k-additive, then $m^{\text{Dem}}_{|C}$ is at most k-additive

Example 4 (Dempster conditioning). *The conditioning at work in our running example is Dempster conditioning. Let us say that ω^* denote the actual state of the world, which agents don't know. Now, consider the case of Agent 1 learning that Peter is not the murderer: Agent 1 learns that $\omega^* \notin \{P\}$, i.e. $\omega^* \in \{Q, R\}$. In this case, the evidence concerning men now only concerns Quentin, so from the viewpoint of Agent 1, the knowledge becomes $m^{Dem}_{|\{Q,R\}}(\{Q\}) = m^{Dem}_{|\{Q,R\}}(\{R\}) = 0.5$ (Figure 2, center, the mass assigned to $\{P, Q\}$ was transferred to $\{P, Q\} \cap \{Q, R\} = \{Q\}$). On the contrary, now consider that Agent 2 learns that Rose is not the murderer: Agent 2 learns $\omega^* \notin \{R\}$, i.e. $\omega^* \in \{P, Q\}$. In this case, the mass concerning women has to be discarded, so the knowledge becomes $m^{Dem}_{|P,Q}(\{P, Q\}) = 1$ (Figure 2, right, the mass assigned to $\{R\}$ was discarded since $\{R\} \cap \{P, Q\} = \emptyset$).*

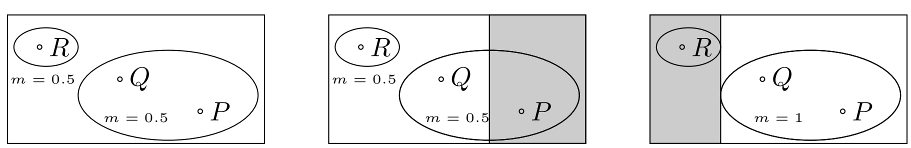

Figure 2: Prior knowledge (left) and revised knowledge given $\{Q, R\}$ (center) and given $\{P, Q\}$ (right). White and gray areas denote possible and impossible events – circles denote focal elements.

On the contrary, The Fagin-Halpern conditioning [32] derives from the interpretation of belief functions in the theory of imprecise probabilities.

Definition 6 (Fagin-Halpern conditioning). *For any non-empty C with* $\text{Bel}(C) > 0$ *(i.e. at least one focal element is included in C),*

$$\mathrm{Bel}(A|C) = \inf_{\Pr\in\mathcal{F}_m} \Pr(A|C) = \frac{\mathrm{Bel}(A\cap C)}{\mathrm{Bel}(A\cap C)+\mathrm{Pl}(\bar{A}\cap C)}$$

Fagin and Halpern [32] have shown that the above-defined measure is a Bel measure. The number of focal elements of $m^{FH}_{|C}$ may be much greater than the number of focal elements of the original Bel and the FH conditioning does not preserve nor the size, neither the k-additivity of the mass function[2]. FH conditioning does not benefit low degrees of k-additivity, however computing $m^{FH}_{|C}$'s values had been shown possible in $O(2^{|\Omega|})$ by Polpitiya et al. [33].

Gong and Meng [34] show that there is a range of conditionings from which Fagin-Halpern's and Dempster's conditioning are the two extremes. To complete the picture, let us cite two alternative rules of conditioning proposed for the credal interpretation, the Strong conditioning [7, 18, 35, 34], also called "geometrical conditioning", which amounts to applying Jeffrey's rule when learning the categorical mass function m' such that $m'(C) = 1$ and the weak conditioning [36], seldom used because leading to strange results (for example $\mathrm{Bel}^{\mathrm{Weak}}(C \mid C) = \mathrm{Bel}(C)/\mathrm{Pl}(C) \leq 1$).

$$m^{\mathrm{Strong}}_{|C}(B) = \begin{cases} m(B)/\mathrm{Bel}(C) & \text{if } B \subseteq C \\ 0 & \text{otherwise} \end{cases}$$

$$m^{\mathrm{Weak}}_{|C}(B) = \begin{cases} m(B)/\mathrm{Pl}(C) & \text{if } B \cap C \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Like Dempster conditioning, both strong and weak conditioning are linear in the size of the original bpa and preserve the k-additivity.

[2]Consider for instance a frame of discernment $\Omega = \{\omega_1, \ldots, \omega_m\}$ and a 2-additive mass function m such that $m(\{\omega_i\}) > 0$ for all i and $m(\{\omega_i, \omega_j\}) > 0$ for all $i \neq j$. Then, for any nonempty $C \subset \Omega$, each subset of $B \subseteq C$ is a focal element of $\mathrm{Bel}(\cdot \mid C)$ – thus $|\mathcal{S}_{m^{FH}_{|C}}| = 2^{|C|}$ and $\mathrm{Bel}(\cdot \mid C)$ is $|C|$-additive.

2.4. Game theory

A simultaneous (or strategic) game of complete information models a situation where each agent makes a decision (the term "action" is rather used in game theory) without coordination with the other agents – and the final utility of each agent depends on the actions chosen by all agents.

Definition 7 (Complete-information game).
A simultaneous game of complete information (also called complete-information game*) is a tuple* $G = \big(N, (A_i)_{i \in N}, (u_i)_{i \in N}\big)$ *where:*

- $N = \{1, \ldots, n\}$ *is a finite set of agents (or "players"),*
- A_i *is the finite set of actions of Agent* i*; the set* $A := \prod_{i \in N} A_i$ *contains all the possible combinations of actions a.k.a. "profiles",*
- $u_i : A \to \mathbb{R}$ *is the utility function of Agent* i*.*

A mixed strategy *for player* i *is a probability distribution on* A_i*. The strategy is said to be* pure *when only one action receives a non-zero probability.*
A pure (resp. mixed) strategy profile *is a vector* $p = (p_1, \ldots, p_n)$ *which assigns a pure (resp. mixed) strategy for each player.*

The game is said to be in standard normal form (SNF) *iff the utility functions are given by tables.*

In the following, we will use the following notations: for any vector $v = (v_1, \ldots, v_n)$ in some product domain $V = \prod_{i \in N} V_i$ and for any $e \subseteq N$, v_e is the restriction of v to e and $V_e = \prod_{i \in e} V_i$. By abuse of notation, we write v_i for $v_{\{i\}}$. For any i, $-i$ denotes the set $N \setminus \{i\}$, i.e. $v_{-i} = (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \in V_{-i} = \prod_{j \neq i} V_j$. Thus, v_{-i} is the restriction of v to all players but i. Finally, "." denotes the concatenation, e.g., $v'_i.v_{-i} = (v_1, \ldots, v_{i-1}, v'_i, v_{i+1}, \ldots, v_n)$. Hence $a = a_i.a_{-i}$ belongs to A and given two profiles $a, a' \in A$, $a'_i.a_{-i}$ denotes the profile a where a_i is replaced with a'_i.

Example 5 (Complete-information game). *Let us consider the following variant of the decision situation of Example 1: suppose that it is known by all agents that Peter is the murderer. Thus, agents will not earn anything if they choose Peter as a partner, and may earn \$2k or \$3k by choosing Quentin or Rose as a partner, depending on the other agent's choice. This situation can be fully described by the complete-information game* $G = \big(N, (A_i, u_i)_{i\in N}\big)$ *where:*

- $N = \{1, 2\}$ *is the set of agents*
- $A_i = \{P_i, Q_i, R_i\}$ *is the set of Agent* i*'s actions*
- $u_i : A_1 \times A_2 \to \mathbb{R}$ *is the utility function of Agent* i*, whose values are given in Table 1.*

	P_2	Q_2	R_2
P_1	$(0, 0)$	$(0, 3)$	$(0, 3)$
Q_1	$(3, 0)$	$(2, 2)$	$(3, 3)$
R_1	$(3, 0)$	$(3, 3)$	$(2, 2)$

Table 1: Utility matrix of Example 5's complete-information game. The tuple at the intersection of row a_1 and column a_2 denotes $\big(u_1(a_1, a_2), u_2(a_1, a_2)\big)$

Because the strategies can be randomized, the global utility for a player of a mixed strategy profile p is defined as the expected utility (EU) of u_i according to the probability distribution it induces over A (obviously, when the strategy is pure, EU_i is equal to the utility value given by u_i):

Definition 8 (Utility of a strategy). *Given a strategy profile* p *in a complete-information game* $\big(N, (A_i)_{i\in N}, (u_i)_{i\in N}\big)$*, the expected utility of player* i *is defined by:*

$$\mathrm{EU}_i(p) = \sum_{a\in A} \left(\prod_{j\in N} p_j(a_j) \right) \times u_i(a).$$

Nash equilibria are the strategy profiles where no agent can have an increase in utility by switching to another strategy:

Definition 9 (Nash equilibrium [2]). *A strategy profile p is a* Nash equilibrium *iff for any $i \in N$, there exists no p'_i such that* $\mathrm{EU}_i(p'_i.p_{-i}) > \mathrm{EU}_i(p)$.

Example 6 (Strategy profile and Nash equilibrium). *Consider Example 5's complete-information game.*

- *The vector $a = (R_1, Q_2) \in A$ is a pure strategy profile, assigning a pure strategy to each agent. It is a Nash equilibrium since no agent has any incentive to deviate unilaterally. Indeed, a change of strategy for an agent would lead to a loss for this agent (a payoff of \$0 or \$2k instead of \$3k). On the contrary, the pure strategy profile $a' = (P_1, Q_2)$ is not a Nash equilibrium since Agent 1 prefers to move from P_1 to R_1 when Agent 2 plays Q_2.*

- *The vector $p = \big(\ (Q_1/0.5, R_1/0.5) \ , \ (Q_2/0.5, R_2/0.5) \ \big)$ is a mixed strategy profile, where $(Q_1/0.5, R_1/0.5)$ denote the mixed strategy of Agent 1, assigning a probability of 0.5 to actions Q_1 and R_1 – and likewise for $(Q_2/0.5, R_2/0.5)$ for Agent 2. Agent 1's expected utility for p is:*
  $\mathrm{EU}_1(p) = 0.25 \times u_1(Q_1, Q_2) + 0.25 \times u_1(Q_1, R_2) + 0.25 \times u_1(R_1, Q_2) + 0.25 \times u_1(R_1, R_2) = 2.5.$
  Similarly, $\mathrm{EU}_2(p) = 2.5$. *p is a mixed Nash equilibrium, since switching one's strategy cannot yield more than an expected utility of* 2.5.

- *It can also be noted that $a = (R_1, Q_2)$ is also a mixed strategy profile, where Agent 1 assigns probability 1 to R_1 and Agent 2 probability 1 to Q_2.*

When the utility functions are described in an explicit way, G is said to be in standard normal form (SNF). SNF representations become spatially costly when the number of players increases ($O(n\alpha^n)$ for a game with n players and α actions per player). More succinct forms have been proposed, that suit cases where utility functions can be decomposed as a sum of smaller utility functions, namely hypergraphical games [15] and their particular cases, polymatrix games [14] and graphical games [37].

Definition 10 (Hypergraphical game, polymatrix game).

A hypergraphical game *is a tuple* $G = \big(N, E, (A_i)_{i\in N}, (u_i^e)_{e\in E, i\in e}\big)$ *where:*

- N *is a set of players,*
- $E = \{e_1, \dots e_m\}$ *is a multiset of subsets of* N *(*(N, E) *is a hypergraph)*
- *For each* $e \in E$, $\big(e, (A_i)_{i\in e}, (u_i^e)_{i\in e}\big)$ *is a complete-information game.*

The global utility of each Agent i *sums the local utilities:* $u_i(a) = \sum_{e\in E} u_i^e(a_e)$.

A polymatrix game *is a hypergraphical game with 2-player local games:* $\forall e \in E, |e| = 2$.

Any hypergraphical game $\big(N, E, (A_i)_{i\in N}, (u_i^e)_{e\in E, i\in e}\big)$ is a complete-information game – it is a succinct representation of the SNF game $\big(N, (A_i, u_i)_{i\in N}\big)$ where u_i denotes the global utility of *Agent* i; conversely any complete-information game in SNF can be described at least by the trivial hypergraphical game $\big(N, \{N\}, (A_i)_{i\in N}, (u_i^N)_{i\in N}\big)$, in which there is only one local game involving all players.

These frameworks assume that each player knows everything about the game: the players, the actions available to each player, all their utilities for each combination of actions, etc. The assumption of complete knowledge cannot always be satisfied. In the real world indeed, players have only a limited knowledge about the outcomes of their strategies – the final outcomes may depend on an ill-known event (in Example 1, the payoff for making the deal with one of P, Q, or R depends on whether they are guilty or innocent).

Harsanyi [3] proposed games of incomplete information as a way to capture problems pervaded with a probabilistic uncertainty (see also [38], for more details). A game of incomplete information can be first understood as a set of possible classical games (of complete information) – one for each possible world $\omega \in \Omega$. Players don't know exactly which world is the real one, but may have some knowledge about it. Harsanyi extends the framework to the dynamical case, considering that just before playing, each player i will receive some information $\tau_i(\omega^*)$ about the real world ω^*. τ_i maps any world to an element θ_i of a set Θ_i called the set of "types" of Agent i. After having observed $\tau_i(\omega^*)$, Agent i knows more about the real game, but several games may still be plausible. The player then conditions the prior knowledge on $\tau_i(\omega^*)$ and decides which action to play. Notice that the different agents may receive different pieces of information and thus have a different posterior knowledge. The question is then, for each player, to determine a strategy (either an action, or a probabilistic strategy) for each possible type.

Harsanyi has shown that such games can be described on the space of types $\Theta = \Theta_1 \times \cdots \times \Theta_n$ (the underlying worlds are omitted). The idea of Harsanyi when defining types is that this concept can encapsulate every piece of information agents may have access to. It includes the agent-observable world status, but also their beliefs on other agents and their introspective mental states.

However, incomplete-information games, as defined by Harsanyi, hold only for the probabilistic approach, hence the name Bayesian game. That is, Bayesian games can only model decision situations under risk – i.e. where the probability law is perfectly known, which is a strong hypothesis that cannot always be satisfied. As for our running example, we aim at modeling such decision situations with a partial knowledge, represented by belief functions.

3. Bel games

Let us first consider the static decision problem when agents don't acquire any piece of information – in our example, agents do not investigate on Peter nor on Rose: they will choose a partner considering the prior knowledge only. The agents play a complete-information game, but this complete-information game is ill-known: there are several possible such games, one for each possible state of the world – and a common belief function describes the knowledge of the agents. All agents aim to maximize their CEU (resp. JEU or TBEU).

This kind of situation can easily be reduced to a single complete-information game, where agents' utility functions assign the CEU value (resp. JEU or TBEU value) to each action profile.

Example 7 (Static incomplete-information game). *Consider the situation described in Example 1, but suppose that agents are not investigating either Peter or Rose: they will choose a partner based only on prior knowledge. There are three possible complete-information games, one for each $\omega \in \Omega$ – these games, G^P, G^Q and G^R are listed in Table 2.*

G^P				G^Q				G^R			
	P_2	Q_2	R_2		P_2	Q_2	R_2		P_2	Q_2	R_2
P_1	$(0,0)$	$(0,3)$	$(0,3)$	P_1	$(2,2)$	$(3,0)$	$(3,3)$	P_1	$(2,2)$	$(3,3)$	$(3,0)$
Q_1	$(3,0)$	$(2,2)$	$(3,3)$	Q_1	$(0,3)$	$(0,0)$	$(0,3)$	Q_1	$(3,3)$	$(2,2)$	$(3,0)$
R_1	$(3,0)$	$(3,3)$	$(2,2)$	R_1	$(3,3)$	$(3,0)$	$(2,2)$	R_1	$(0,3)$	$(0,3)$	$(0,0)$

Table 2: Three possible complete-information games G^ω, depending on the actual state of the world $\omega \in \{\omega_P, \omega_Q, \omega_R\}$.

Consider the strategy where Agent 1 *chooses to partner with Peter, while Agent* 2 *chooses to partner with Rose. In this case, the* CEU *value for Agent* 1 *is:*

$$\mathrm{CEU}_1(P_1, R_2) = \frac{1}{2} \times \min\left(u_1^P(P_1, R_2), u_1^Q(P_1, R_2)\right) + \frac{1}{2} \times u_1^R(P_1, R_2) = 1.5$$

The CEU value computed for each pure strategy is given by Table 3. This table directly defines a complete-information game – in other terms, in static problems where the agents maximize their CEU *without getting any new information are perfectly captured by a complete-information game. The same reasoning can be made considering the* JEU *value and the* TBEU *value.*

	P_2	Q_2	R_2
P_1	$(1,1)$	$(\frac{3}{2}, \frac{3}{2})$	$(\frac{3}{2}, \frac{3}{2})$
Q_1	$(\frac{3}{2}, \frac{3}{2})$	$(1,1)$	$(\frac{3}{2}, \frac{3}{2})$
R_1	$(\frac{3}{2}, \frac{3}{2})$	$(\frac{3}{2}, \frac{3}{2})$	$(1,1)$

Table 3: The complete information game capturing the incomplete-information *static problem of Example 7 (*CEU *maximization)*

3.1. Bel games

In the general case, agents can learn information. Such problems cannot be reduced to a complete game where the utility values are the CEU values. Each agent may indeed make a different decision, depending on the private information received (depending on the "type" received).

Bayesian games are games of incomplete information where prior knowledge is captured by a probability measure. To capture problems where the Bayesian assumption is not appropriate (as in our motivating example), we propose the more general framework of Bel games:

Definition 11 (Bel game). *A simultaneous (or strategic) Bel game G is defined as a tuple $\big(N, (A_i)_{i \in N}, (\Theta_i)_{i \in N}, (u_i)_{i \in N}, m\big)$ where:*

- $N = \{1, \ldots, n\}$ *is a finite set of players,*
- A_i *is the finite set of actions of player i; $A = \prod_{i \in N} A_i$ denotes the set of all action profiles,*
- Θ_i *is the finite set of types of player i; $\Theta = \prod_{i \in N} \Theta_i$ denotes the set of all type configurations,*
- $m : 2^{\Theta} \to [0, 1]$ *is the mass function describing the common prior knowledge,*
- $u_i : A \times \Theta \to \mathbb{R}$ *is the utility function of Agent i.*

G is said to be in standard normal form (SNF) iff the values of utility functions u_i and those of the mass function m are given in tables.

Bel games generalize Bayesian games, which are recovered when m is a probability distribution.

Proposition 1 (Size of a Bel game). *A Bel game in SNF has a spatial complexity in $O\big(n(\alpha\beta)^n + kn\,\mathrm{Size}(m)\big)$, where $\alpha = \max_{i \in N} |A_i|$ and $\beta = \max_{i \in N} |\Theta_i|$.*

Proof. *See Proof 13 in Appendix.*

Example 8. *We are now equipped to fully capture the problem described by Example 1, namely by a Bel game* $G = \big(N, (A_i)_{i\in N}(\Theta_i)_{i\in N}, (u_i)_{i\in N}, m\big)$ *where:*

- $N = \{1, 2\}$*;*
- $A_1 = \{P_1, Q_1, R_1\}$*,* $A_2 = \{P_2, Q_2, R_2\}$ *(each agent chooses an associate).*
- $\Theta_1 = \{\mathtt{P}, \bar{\mathtt{P}}\}$*,* $\Theta_2 = \{\mathtt{R}, \bar{\mathtt{R}}\}$ *(Agent 1 investigates on Peter, Agent 2 investigates on Rose).*
- $m : 2^{\Theta} \to [0, 1]$ *has two focal elements:* $m\big(\{(\bar{\mathtt{P}}, \mathtt{R})\}\big) = 1/2$ *(the murderer is a woman, thus necessarily Rose – in this case Agent* 1 *will learn* $\bar{\mathtt{P}}$ *and Agent* 2 *will learn* $\mathtt{R}$*) and* $m\big(\{(\mathtt{P}, \bar{\mathtt{R}}), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\}\big) = 1/2$ *(the murderer is a man: Agent* 2 *necessarily learns* $\bar{\mathtt{R}}$ *but Agent* 1 *can learn either* $\bar{\mathtt{P}}$ *– which happens when Quentin is the murderer – or* $\mathtt{P}$ *– Peter is the murderer).*
- *Making a deal with a murderer has a utility value of* 0*, making a deal with an innocent leads to a utility of* $\frac{6}{2} = 3$*, unless the other agent approaches the same associate, in which case each agent receives* $\frac{6}{3} = 2$*. The utility functions depend on the configuration of types – they are summarized in Table 4. There are only three possible type configurations,* $(\bar{\mathtt{P}}, \bar{\mathtt{R}})$*,* $(\mathtt{P}, \bar{\mathtt{R}})$ *and* $(\bar{\mathtt{P}}, \mathtt{R})$*. The case where* $\theta = (\mathtt{P}, \mathtt{R})$ *(both* R *and* P *are guilty) is not a possible world – null values (in gray) are given this configuration.*

Following Harsyani's approach of incomplete-information games, we consider the "*ex interim*" setting where each player plans a strategy for each of the types he/she can receive. We thus adopt the definition of strategy proposed by Harsyani's in the general context of incomplete-information games:

Definition 12 (Pure and mixed strategies [3]). *A pure (resp. mixed) strategy for player* i *in a Bel game is a function* ρ_i *which maps each "type"* $\theta_i \in \Theta_i$ *to an action of (resp. a probability over)* A_i*.*

		$\theta_2 = \bar{\mathtt{R}}$				$\theta_2 = \mathtt{R}$		
		P_2	Q_2	R_2		P_2	Q_2	R_2
	P_1	$(0,0)$	$(0,3)$	$(0,3)$	P_1	$(0,0)$	$(0,0)$	$(0,0)$
$\theta_1 = \mathtt{P}$	Q_1	$(3,0)$	$(2,2)$	$(3,3)$	Q_1	$(0,0)$	$(0,0)$	$(0,0)$
	R_1	$(3,0)$	$(3,3)$	$(2,2)$	R_1	$(0,0)$	$(0,0)$	$(0,0)$
		P_2	Q_2	R_2		P_2	Q_2	R_2
	P_1	$(2,2)$	$(3,0)$	$(3,3)$	P_1	$(2,2)$	$(3,3)$	$(3,0)$
$\theta_1 = \bar{\mathtt{P}}$	Q_1	$(0,3)$	$(0,0)$	$(0,3)$	Q_1	$(3,3)$	$(2,2)$	$(3,0)$
	R_1	$(3,3)$	$(3,0)$	$(2,2)$	R_1	$(0,3)$	$(0,3)$	$(0,0)$

Table 4: Example 8: Utility matrices for each configuration of the types.

A pure (resp. mixed) strategy profile is a vector $p = (p_1, \ldots, p_n)$ which assigns a pure (resp. mixed) strategy for each player.

$\rho(\theta) = \big(\rho_1(\theta_1), \ldots, \rho_n(\theta_n)\big)$ denotes the profile which will be played if the configuration of types is θ.

The set of all pure strategy profiles is denoted $\Sigma = \prod_{i \in N}(\Theta_i \to A_i)$.

Example 9 (Pure strategy). *In our running example,*

- *$\rho_1 : \Theta_1 \to A_1$, defined by $\rho_1(\mathtt{P}) = R_1$ and $\rho_1(\bar{\mathtt{P}}) = P_1$ is a pure strategy of Agent 1; in this strategy Agent 1 will choose Rose when learning that Peter is the murderer, and choose Quentin otherwise.*

- *$\rho_2 : \Theta_2 \to A_2$, defined by $\rho_2(\mathtt{R}) = Q_2$ and $\rho_2(\bar{\mathtt{R}}) = R_2$ is a pure strategy of Agent 2; in this strategy Agent 2 will choose Quentin when learning that Rose is the murderer, and choose Rose otherwise.*

- *$\rho = (\rho_1, \rho_2)$ is a pure strategy profile, and $\rho(\mathtt{P}, \bar{\mathtt{R}}) = (R_1, R_2)$ is the action profile that will actually be played if Peter is the murderer (i.e., when Agent 1 learns $\mathtt{P}$ and Agent 2 learns $\bar{\mathtt{R}}$).*

In the *ex interim* approach of incomplete-information games, when receiving the type θ_i, Agent i revises the prior knowledge – in a Bel game, Agent i's posterior knowledge over the joint types' configuration is $m_{|\theta_i}$.

Let us first consider the case where the agents maximize their Choquet utility (this approach being compatible with both the evidential and the credal interpretation of belief functions). In this case the utility of a pure strategy profile for Agent i of type θ_i, shall thus be defined as:

Definition 13 (Choquet Expected Utility of a pure strategy profile). *The utility of a pure strategy profile $\rho = (\rho_1, \ldots, \rho_n)$, for Agent i of type θ_i, is:* $\mathrm{CEU}_{(i,\theta_i)}(\rho) = \sum_{B \in S_{m_{|\theta_i}}} m_{|\theta_i}(B) \times \min_{\theta' \in B} u_i(\rho(\theta'), \theta')$.

where $m_{|\theta_i}$ denotes the conditioning that is compatible with the interpretation of the belief function (e.g., Demspter's conditioning or Fagin–Halpern's one).

Let us now consider mixed strategies. A mixed strategy profile ρ defines a probability distribution $\mathrm{Pr}^\rho(\sigma) = \prod_{i \in N} \prod_{\theta_i \in \Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i))$ over the set Σ of pure strategy profiles. If we now merge Pr^ρ with m, we get a bpa m^ρ over $A \times \Theta$: to any element $X = \{(a, \theta), (a', \theta'), \ldots\} \subseteq A \times \Theta$ correspond both a set of type configurations $B_X := \{\theta \mid (a, \theta) \in X\} \subseteq \Theta$ and a set of compatible pure strategy profiles $S_X := \{\sigma \mid \forall (a, \theta) \in X,\ \sigma(\theta) = a\} \subseteq \Sigma$. The mass of X is $m^\rho(X) = m(B_X) \times \sum_{\sigma \in S_X} \mathrm{Pr}^\rho(\sigma)$, that is, X is focal if B_X is focal and some compatible pure strategy profiles are possible. Finally, Agent i receiving type θ_i conditions the prior knowledge which becomes $m^\rho_{|\theta_i}$. Hence the following definition of the utility of a mixed strategy profile:

Definition 14 (Choquet Expected Utility of a mixed strategy profile). *The utility of a mixed strategy profile $\rho = (\rho_1, \ldots, \rho_n)$, for player i of type θ_i, is:*

$$\mathrm{CEU}_{(i,\theta_i)}(\rho) = \sum_{\sigma \in \Sigma} \mathrm{Pr}^\rho(\sigma) \times \sum_{B \in \mathcal{S}_{m_{|\theta_i}}} m_{|\theta_i}(B) \times \min_{\theta' \in B} u_i\big(\sigma(\theta'), \theta'\big)$$

where $\mathrm{Pr}^\rho(\sigma) = \prod_{i \in N} \prod_{\theta_i \in \Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i))$

Proof (Proof of correctness). *See Proof 1 in Appendix.*

It can be checked that Definition 14 amounts to Definition 13 when ρ is a pure strategy profile.

Now, recall that a strategy is a Nash equilibrium iff no agent can have an increase in utility by switching to another strategy. This concept straightforwardly extends to Bel games:

Definition 15 (Nash equilibrium). *A mixed (resp. pure) strategy profile ρ is a Nash equilibrium for* CEU *iff, whatever (i, θ_i), there exists no mixed (resp. pure) strategy ρ'_i such that $\mathrm{CEU}_{(i,\theta_i)}(\rho'_i.\rho_{-i}) > \mathrm{CEU}_{(i,\theta_i)}(\rho)$.*

Example 10. *Let ρ be the pure strategy defined in Example 9: Agent 1 makes the deal with R when learning that P is guilty and with P otherwise, and Agent 2 joins Q when learning that R is guilty and R otherwise:*
$\rho_1(\mathtt{P}) = R_1,\ \rho_1(\bar{\mathtt{P}}) = P_1,\ \rho_2(\mathtt{R}) = Q_2$ *and* $\rho_2(\bar{\mathtt{R}}) = R_2$.
As usual with this "Peter, Paul and Mary" example, the Demspter rule of conditioning is used.

- *Consider Agent 1 receiving type $\mathtt{P}$: the conditioned bpa, $m^{Dem}_{|\mathtt{P}}$, has only one focal element $\{(\mathtt{P}, \bar{\mathtt{R}})\}$, $K_{\mathtt{P}} = 1/\frac{1}{2}$ and $m^{Dem}_{|\mathtt{P}}(\{(\mathtt{P}, \bar{\mathtt{R}})\}) = 1$. In short, Agent 1 knows that P is guilty and R is not. In the only possible configuration $(\mathtt{P}, \bar{\mathtt{R}})$, ρ prescribes $\rho_1(\mathtt{P}) = R_1$ for Agent 1 and $\rho_2(\bar{\mathtt{R}}) = R_2$ for Agent 2. Then,*
  $\mathrm{CEU}_{(1,\mathtt{P})}(\rho) = m^{Dem}_{|\mathtt{P}}\big(\{(\mathtt{P}, \bar{\mathtt{R}})\}\big) \times u_1\big((R_1, R_2), (\mathtt{P}, \bar{\mathtt{R}})\big) = 1 \times 2 = 2.$
- *Consider now Agent 1 receiving $\bar{\mathtt{P}}$: revised knowledge, $m^{Dem}_{|\bar{\mathtt{P}}}$, has two focal elements, $\{(\bar{\mathtt{P}}, \mathtt{R})\}$ and $\{(\bar{\mathtt{P}}, \bar{\mathtt{R}})\}$ (each with probability $\frac{1}{2}$, thus $K_{\bar{\mathtt{P}}} = 1$). The strategy prescribes $\rho(\bar{\mathtt{P}}) = P_1$ for Agent 1, who doesn't know whether Agent 2 learns $\mathtt{R}$ (and plays $\rho(\mathtt{R}) = Q_2$) or $\bar{\mathtt{R}}$ (and plays $\rho(\bar{\mathtt{R}}) = R_2$). Hence*
  $\mathrm{CEU}_{(1,\bar{\mathtt{P}})}(\rho) = \frac{1}{2} \times u_1\big((P_1, R_2), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\big) + \frac{1}{2} \times u_1\big((P_1, Q_2), (\bar{\mathtt{P}}, \mathtt{R})\big) = 3$
- *Similarly, the bpa of Agent 2 receiving $\mathtt{R}$, $m^{Dem}_{|\mathtt{R}}$, has only one focal element, $\{(\bar{\mathtt{P}}, \mathtt{R})\}$ (thus $K_{\mathtt{R}} = 1/\frac{1}{2}$) in which ρ prescribes P_1 for Agent 1 and Q_2 for Agent 2. Then*
  $\mathrm{CEU}_{(2,\mathtt{R})}(\rho) = 1 \times u_2\big((P_1, R_2), (\bar{\mathtt{P}}, \mathtt{R})\big) = 1 \times 3 = 3.$

- *Finally, the bpa of Agent* 2 *receiving* R*,* $m^{Dem}_{|\bar{\mathtt{R}}}$*, has one focal element,* $\{(\bar{\mathtt{P}}.\bar{\mathtt{R}}), (\mathtt{P}.\bar{\mathtt{R}})\}$ *and* $K_{\bar{\mathtt{R}}} = 1/\frac{1}{2}$*. Agent* 2 *does not know whether Agent* 1 *receives* $\bar{\mathtt{P}}$ *or* P*. Since* ρ *prescribes Agent* 1 *to play* P_1 *in the first case,* R_1 *in the second one and prescribes Agent* 2 *to play* R_2 *in both cases,*
$\mathrm{CEU}_{(2,\bar{\mathtt{R}})}(\rho) = 1 \times \min\big[u_2\big((P_1, R_2), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\big), u_2\big((R_1, R_2), (\mathtt{P}, \bar{\mathtt{R}})\big)\big]$
$= 1 \times \min(3, 2) = 2.$

In this strategy, Agent 1 *does not give the best possible response to Agent* 2*'s strategy: when learning that* P *is guilty, Agent* 1 *plays* R_1 *while knowing that in this case Agent* 2 *learns* $\bar{\mathtt{R}}$ *and thus plays* R_2*. Let Agent* 1 *make a change of strategy and play* Q_1 *when learning* P *– hence the strategy* ρ'*:*
$\rho'_1(\mathtt{P}) = Q_1,\ \rho'_1(\bar{\mathtt{P}}) = P_1,\ \rho'_2(\mathtt{R}) = Q_2,\ \rho'_2(\bar{\mathtt{R}}) = R_2.$

- $\mathrm{CEU}_{(1,\mathtt{P})}(\rho') = K_{\mathtt{P}} \times u_1\big((Q_1, R_2), (\mathtt{P}, \bar{\mathtt{R}})\big) = 1 \times 3 = 3,$
- $\mathrm{CEU}_{(1,\bar{\mathtt{P}})}(\rho') = K_{\bar{\mathtt{P}}} \times u_1\big((P_1, R_2), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\big) + K_{\bar{\mathtt{P}}} \times u_1\big((P_1, Q_2), (\bar{\mathtt{P}}, \mathtt{R})\big) = 3,$
- $\mathrm{CEU}_{(2,\mathtt{R})}(\rho') = K_{\mathtt{R}} \times u_2\big((P_1, Q_2), (\bar{\mathtt{P}}, \mathtt{R})\big) = 1 \times 3 = 3,$
- $\mathrm{CEU}_{(2,\bar{\mathtt{R}})}(\rho') = K_{\bar{\mathtt{R}}} \times \min\Big(u_2\big((P_1, R_2), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\big), u_2\big((Q_1, R_2), (\mathtt{P}, \bar{\mathtt{R}})\big)\Big) = 3.$

It can be checked that with ρ'*, each player gets the maximal possible utility ($3k) – no player has incentive to deviate:* ρ' *is a pure Nash equilibrium.*

3.2. Credal Games and Evidential games

Bel games as defined above can be understood under the DS theory or under the theory of imprecise probabilities. Because the Choquet expected utility is compatible with both theories, we have first considered the pessimistic Choquet integral as a way to evaluate the utility of the agents. Let us briefly investigate the model in each of the two interpretations, with respect to the different ways of conditioning and to the different decision rules.

3.2.1. CEU, JEU and Pignistic Games in the DS theory of evidence

Let us first consider problems having an interpretation in the DS theory and are thus based on the Dempster's rule of conditioning – we call these games

"Evidential games". Several decision rules can be used in this context, namely the Choquet integral (CEU) used in the previous section, Jaffray's linear utility (JEU) and the transferable belief model (TBEU). Let us capture all of them as particular cases of a generalized expected utility:

$$\mathrm{XEU}(a) = \sum_{B \in \mathcal{S}_m} m(B) \times f_{u \circ a}^{\mathrm{XEU}}(B)$$

If m is a probability distribution, then $EU(a) = \mathrm{XEU}(a)$ in all three cases. We find back the CEU, JEU and TBEU with:

- $f_{u \circ a}^{\mathrm{CEU}}(B) = \min_{\omega \in B} u(a(\omega))$
- $f_{u \circ a}^{\mathrm{JEU}}(B) = \alpha_B \min_{\omega \in B} u(a(\omega)) + (1 - \alpha_B) \max_{\omega \in B} u(a(\omega))$
- $f_{u \circ a}^{\mathrm{TBEU}}(B) = \sum_{\omega \in B} \frac{u(a(\omega))}{|B|}$

As to CEU (resp. JEU), the proof is trivial: one has just to rewrite $f_{u \circ a}^{\mathrm{CEU}}$ (resp. $f_{u \circ a}^{\mathrm{JEU}}$) in the equation to get back the on-focal-set expression.

As to TBEU, we need to go back to the expected utility model using the distribution

$$\begin{aligned}
\sum_{\omega \in \Omega} \mathrm{BetP}_m(\omega) \times u(a(\omega)) &= \sum_{\omega \in \Omega} \left(\sum_{\substack{B \subseteq \mathcal{S}_m \\ \omega \in B}} \frac{m(B)}{|B|} \right) \times u(a(\omega)) \\
&= \sum_{B \in \mathcal{S}_m} \sum_{\omega \in B} \frac{m(B)}{|B|} \times u(a(\omega)) \\
&= \sum_{B \in \mathcal{S}_m} m(B) \times \sum_{\omega \in B} \frac{u(a(\omega))}{|B|} \\
&= \sum_{B \in \mathcal{S}_m} m(B) \times f_{u \circ a}^{\mathrm{TBEU}}(B)
\end{aligned}$$

Definition 16 (Utility in an evidential game). *The utility of a mixed strategy profile $\rho = (\rho_1, \ldots, \rho_n)$, for player i of type θ_i, is:*

$$\mathrm{XEU}_{(i,\theta_i)}(\rho) = \sum_{\sigma \in \Sigma} \left(\prod_{i \in N} \prod_{\theta_i \in \Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{B \in \mathcal{S}_{m^{Dem}_{|\theta_i}}} m^{Dem}_{|\theta_i}(B) \times f_{v_i^\sigma}^{\mathrm{XEU}}(B)$$

where $\mathrm{XEU} \in \{\mathrm{CEU}, \mathrm{JEU}, \mathrm{TBEU}\}$ *and* $v_i^\sigma(\theta) = u_i(\sigma(\theta), \theta)$

Proof (Proof of correctness). *See Proof 2 in Appendix.*

It can be noted that we find back Definition 14 by setting XEU := CEU.

The definitions of pure and mixed Nash equilibria remain unchanged, i.e. ρ is a pure (resp. mixed) Nash equilibrium for XEU iff, whatever (i, θ_i), there exists no pure (resp. mixed) strategy ρ'_i such that $\mathrm{XEU}_{(i,\theta_i)}(\rho'_i.\rho_{-i}) > \mathrm{XEU}_{(i,\theta_i)}(\rho)$.

3.2.2. Bel games in the credal interpretation

In the credal interpretation, the mass distribution actually defines a family of probability over the combinations of types, hence the use of the Fagin-Halpern conditioning. This interpretation is compatible with the Choquet-based decision rule (when the capacity used is a Bel measure, the Choquet value of a decision is equal to the minimum value of the expected utilities provided by the different probabilities of the family). In [39], it is also shown to be echo compatible with Jaffray's linear utility (JEU).

Notice that Dempster's rule also receives an interpretation in the credal context: it leads to a family consisting of the conditionals of those probabilities in the family which are the most likely (assessing a maximal probability to the event C we now know for sure) – hence the name "Max likelihood conditioning".

So, in a credal game:

Definition 17 (Utility in a Credal game). *The utility of a mixed strategy profile $\rho = (\rho_1, \ldots, \rho_n)$, for player i of type θ_i, is:*

$$\mathrm{XEU}_{(i,\theta_i)}(\rho) = \sum_{\sigma \in \Sigma} \left(\prod_{i \in N} \prod_{\theta_i \in \Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{B \in \mathcal{S}_{m_{|\theta_i}}} m_{|\theta_i}(B) \times f^{\mathrm{XEU}}_{v_i^\sigma}(B)$$

where $\mathrm{XEU} \in \{\mathrm{JEU}, \mathrm{CEU}\}$ *and* $m_{|C} \in \{m^{FH}_{|C}, m^{Strong}_{|C}, m^{Weak}_{|C}, m^{Dem}_{|C}\}$

Proof (Proof of correctness). *See Proof 3 in Appendix.*

This modelling leaves the definitions of pure and mixed Nash equilibria unchanged, i.e. ρ is a pure (resp. mixed) Nash equilibrium for XEU iff, whatever (i, θ_i), there exists no pure (resp. mixed) strategy ρ'_i with a greater XEU.

4. From Bel games to complete-information games

One of the most prominent results about Bayesian games is Howson–Rosenthal's theorem [13]: any 2-player Bayesian game can be transformed into a (complete information) polymatrix game equivalent to the original one. This result is important from the computational point of view since it provides 2-player Bayesian games with practical resolution tools: to solve a 2-player Bayesian games, it is enough to use this theorem and to solve the resulting polymatrix game by using one of the algorithms proposed for such games [16, 17]. In the sequel, we generalize this theorem to Bel games and extend it to any number of players.

4.1. The direct transform

A first idea is to define from a Bel game G, a hypergraphical game $\tilde{G}$, the vertices (players) of which are pairs (i, θ_i) with action set A_i – to each pure strategy σ of G corresponds a unique pure strategy $\tilde{\sigma}$ of $\tilde{G}$ and conversely – we call $\tilde{\sigma}$ the Selten[3] transform of σ:[4]

Definition 18 (Selten transform of a pure strategy). *For any pure strategy σ of G, the Selten transform of σ is the vector $\tilde{\sigma}$ defined by $\tilde{\sigma}_{(i,\theta_i)} = \sigma(\theta_i)$.*

The local games of the hypergraphical game correspond to the focal elements of m. Roughly, (i, θ_i) plays in the local game corresponding to the focal element B if the type θ_i is plausible for B – i.e. if there exists $\theta' \in B$ such that $\theta'_i = \theta_i$. In this local game, (i, θ_i) obtains a local utility $K_{|\theta_i} \cdot m(B) \cdot f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B \cap \{\theta' \mid \theta'_i = \theta_i\})$.

Given a profile of actions $\tilde{\sigma}$, and a player (i, θ_i), the hypergraphical game sums (i, θ_i)'s local utilities over all the focal elements for which θ_i is plausible: the global utility for (i, θ_i) is equal to the XEU of the joint σ.

One may note that two pairs (i, θ_i) and (i, θ'_i) may play in the same local game – this happens when θ and θ' belong to the same focal element. In this

[3] Named after Selten, who proposed this definition for Bayesian games [3].

[4] We could use the notation ρ for both, but the pure strategies of the Bel game are vectors of functions $\sigma_i : \Theta_i \to A_i$ while the pure strategies of $\tilde{G}$ are vectors in $\prod_{i \in N} \prod_{\theta_i \in \Theta_i} A_i$. So, we keep the two notations σ and $\tilde{\sigma}$.

case, the utility of (i, θ'_i) does not depend on the action played by (i, θ_i) and conversely.

For any focal element B of m, let $Players(B) = \{(i, \theta_i) \mid \theta \in B, i \in N\}$ – $Players(B)$ denotes the future players involved in the local game corresponding to B. Let $\tilde{E}$ be the multiset $\tilde{E} := [Players(B) \mid B \in S_m]$. The elements e of $\tilde{E}$ and the focal elements in S_m are in one-to-one correspondence and we denote B_e the focal element of m which leads to e. These notations allow us to propose a first, direct generalization of Howson–Rosenthal's transform to Bel games:

Definition 19 (Direct transform of a Bel game). *The direct transform of a Bel game* $G = \big(N, (A_i, \Theta_i, u_i)_{i \in N}, m\big)$ *is the hypergraphical game* $\tilde{G} = \big(\tilde{N}, \tilde{E}, (\tilde{A}_{(i,\theta_i)})_{(i,\theta_i)\in\tilde{N}}, (\tilde{u}^e_{(i,\theta_i)})_{e\in\tilde{E},(i,\theta_i)\in e}\big)$ *where:*

- $\tilde{N} = \{(i, \theta_i) \mid i \in N, \theta_i \in \Theta_i\}$,
- $\tilde{A}_{(i,\theta_i)} = A_i$,
- $\tilde{E} = [Players(B) \mid B \in \mathcal{S}_m]$,
- *For each* $e \in \tilde{E}$, $(i, \theta_i) \in e$ *and* $\tilde{\sigma} \in \tilde{A}$, $\tilde{u}^e_{(i,\theta_i)}(\tilde{\sigma}_e) = K_{|\theta_i} \cdot m(B_e) \cdot f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B \cap \{\theta' \mid \theta'_i = \theta_i\})$, *using* $\tilde{v}_i^{\tilde{\sigma}}(\theta) = u_i(\tilde{\sigma}_\theta, \theta)$ *where we recall that* $\tilde{\sigma}_\theta = (\tilde{\sigma}_{(1,\theta_1)}, \ldots, \tilde{\sigma}_{(n,\theta_n)})$

It is straightforward to show that the XEU value of a pure strategy ρ in G and the global utility of $\tilde{\rho}$ in $\tilde{G}$ are equal, whatever is the couple (i, θ_i) considered.

Proposition 2. *Let* G *be a Bel game based on the Dempster's rule of conditioning and let* $\tilde{G}$ *be its direct transform. For any pure strategy* σ *of* G*, it holds that* $\mathrm{XEU}_{(i,\theta_i)}(\sigma) = \tilde{u}_{(i,\theta)}(\tilde{\sigma})$.

Proof. *See Proof 4 in Appendix.*

Let us extend the Selten transform to mixed strategies ρ of G: each $\tilde{\rho}_{(i,\theta_i)} = \rho_i(\theta_i)$ is then a probability distribution over A_i, and $\tilde{\rho}$ is then a vector of such distributions.

Proposition 3. *Let* G *be a Bel game and* $\tilde{G}$ *its direct transform. For any mixed strategy* ρ *of* G*, it holds that* $\mathrm{XEU}_{(i,\theta_i)}(\rho) = \tilde{u}_{(i,\theta)}(\tilde{\rho})$.

Proof. *See Proof 5 in Appendix.*

It can be checked that when m is a probability distribution, and G is a 2-player game, we get at most $|\Theta|$ local games, each involving two players (i, θ_i) and (j, θ_j): $\tilde{G}$ is a polymatrix game, and Howson–Rosenthal's Theorem is recovered. More generally, we prove:

Theorem 1 (Generalized Howson–Rosenthal's Theorem). *For any Bel game G based on a* XEU *utility and the Dempster's rule of conditioning, there exists a hypergraphical game $\tilde{G}$ such that ρ is a pure (resp. mixed) Nash equilibrium of G iff $\tilde{\rho}$ is a pure (resp. mixed) Nash equilibrium of $\tilde{G}$.*

Proof. *See Proof 6 in Appendix.*

Example 11. *Let us define the direct transform of the Bel game G corresponding to our running example (again, with Dempster conditioning and the Choquet expected utility). The set of players is: $\tilde{N} = \{(1, \mathtt{P}), (1, \bar{\mathtt{P}}), (2, \mathtt{R}), (2, \bar{\mathtt{R}})\}$. The set of actions are $\tilde{A}_{(i,\theta_i)} = \{P_i, Q_i, R_i\}$.*

Because m has two focal elements $B_1 = \{(\bar{\mathtt{P}}, \mathtt{R})\}$ and $B_2 = \{(\bar{\mathtt{P}}, \bar{\mathtt{R}}), (\mathtt{P}, \bar{\mathtt{R}})\}$ each with a mass of $\frac{1}{2}$, $\tilde{G}$ involves two local games. The set of players involved are respectively $e_1 = \{(1, \bar{\mathtt{P}}), (2, \mathtt{R})\}$ and $e_2 = \{(1, \bar{\mathtt{P}}), (1, \mathtt{P}), (2, \bar{\mathtt{R}})\}$. $\tilde{G}$'s hypergraph is drawn on Figure 3.

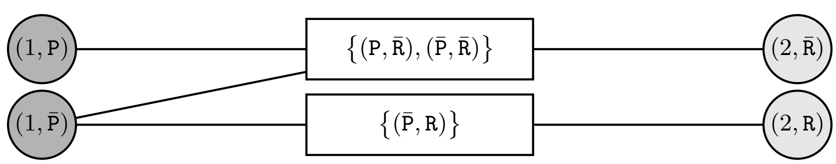

Figure 3: G's direct transform. Gray circles denote vertices (players; one shade per agent), white boxes denote hyperedges (local games; linked to the players involved).

Player $(2, \bar{\mathtt{R}})$ plays only in e_2, we have for instance:

$$\begin{aligned}\tilde{u}^{e_2}_{(2,\bar{\mathtt{R}})}(R_1, P_1, R_2) &= K_{\bar{\mathtt{R}}} \cdot m(B_2) \times \min\Big[u_2\big((R_1, R_2), (\mathtt{P}, \bar{\mathtt{R}})\big),\ u_2\big((P_1, R_2), (\bar{\mathtt{P}}, \bar{\mathtt{R}})\big)\Big] \\ &= 1 \times \min(2, 3) = 2.\end{aligned}$$

For player $(1,\bar{\mathtt{P}})$*, which plays in both local games, we have for instance:*

$$\tilde{u}^{e_1}_{(1,\bar{\mathtt{P}})}(P_1,P_2) = K_{\bar{\mathtt{P}}}\cdot m(B_1)\times u_1\big((P_1,P_2),(\bar{\mathtt{P}},\mathtt{R})\big) = 0.5\times 2 = 1$$
$$\tilde{u}^{e_2}_{(1,\bar{\mathtt{P}})}(a_{(1,\mathtt{P})},P_1,Q_2) = K_{\bar{\mathtt{P}}}\cdot m(B_2)\times u_1\big((P_1,Q_2),(\bar{\mathtt{P}},\bar{\mathtt{R}})\big) = 0.5\times 3 = 1.5$$

The Selten transform of the Nash equilibrium ρ' *described in Example 10 is:*

$$\tilde{\rho}'((1,\bar{\mathtt{P}})) = P_1,\ \tilde{\rho}'((1,\mathtt{P})) = Q_1,\ \tilde{\rho}'((2,\bar{\mathtt{R}})) = R_1,\ \tilde{\rho}'((2,\mathtt{R})) = Q_2.$$

One can check that:

$$\tilde{u}_{(1,\bar{\mathtt{P}})}(\tilde{\rho}') = \tilde{u}^{e_1}_{(1,\bar{\mathtt{P}})}\big((P_1,Q_2)\big) + \tilde{u}^{e_2}_{(1,\bar{\mathtt{P}})}\big((Q_1,P_1,R_2)\big) = \mathrm{CEU}_{(1,\bar{\mathtt{P}})}(\rho').$$

Notice that in the sum, one part of the utility of $(1,\bar{\mathtt{P}})$ *comes from subgame* e_1 *(i.e., from* B_1*) and the other part from local game* e_2 *(i.e., from* B_2*).*

As to the complexity of the transform, let α (resp. β) be the maximum number of actions (resp. types) per player in G and k the degree of additivity of m. It holds that G contains n utility tables of size $(\alpha\cdot\beta)^n$ and the size of the description of m is bounded by $k\cdot n\cdot|\mathcal{S}_m|$. So, Size$(G)$ is in $O\big(n\cdot(\alpha\cdot\beta)^n + k\cdot n\cdot|\mathcal{S}_m|\big)$.

$\tilde{G}$ contains $|\mathcal{S}_m|$ local games. Each of them involves at most $k\cdot n$ players (i,θ_i) – the size of their SNF representation is thus at most $k\cdot n\cdot\alpha^{kn}$ – hence a spatial cost for the representation of $\tilde{G}$ in $O(|\mathcal{S}_m|\cdot k\cdot n\cdot\alpha^{kn})$. Notice now that since m is k-additive, $|\mathcal{S}_m| < \beta^{k\cdot n}$. So, Size$(\tilde{G})$ is bounded by $k\cdot n\cdot(\alpha\cdot\beta)^{k\cdot n} \le n^k\cdot(\alpha\cdot\beta)^{k\cdot n}$. In short, we get:

Proposition 4 (Complexity of the direct transform). *The direct transform of a Bel game* G *has a temporal complexity in* $O\big(|\mathcal{S}_m|\cdot k^2\cdot n\cdot\alpha^{k\cdot n}\cdot\beta\big) \subseteq O\big(k\cdot\beta\cdot\mathrm{Size}(G)^k\big)$ *and a spatial complexity in* $O\big(|\mathcal{S}_m|\cdot k\cdot n\alpha^{k\cdot n}\big) \subseteq O\big(\mathrm{Size}(G)^k\big)$.

Proof. *See Proof 14 in Appendix.*

So, the degree of additivity of the bpa is the main factor of complexity. Hopefully, low degrees of additivity can be assumed – it has indeed been shown [40, 41] that such low values (typically, $k\le 3$) allow the description of many cases of interest. In such situations, the transform is quadratic or, at worst, cubic.

The direct transform, as defined above, holds for Demspter rule of conditioning. A variant can be used for each conditioning rule in which the focal elements of conditioned bpa are obtained by directly conditioning the original focal elements (each focal element B containing θ_i leads to a focal element $B_{|\theta_i}$). During strong or weak conditioning, masses aren't transferred but stay on the prior focal elements – i.e., $B_{|\theta_i} = B$. Modifying the local utility definition, switching the term $f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B \cap \{\theta' \mid \theta'_i = \theta_i\})$ to $f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B)$, captures both strong conditioning (with $K_{|\theta_i} = 1/\operatorname{Pl}(\{\theta' \mid \theta'_i = \theta_i\}))$ and weak conditioning (with $K_{|\theta_i} = 1/\operatorname{Bel}(\{\theta' \mid \theta'_i = \theta_i\}))$.

Except for very peculiar cases, this kind of transform cannot be used with Fagin-Halpern's rule of conditioning, in which the conditioned focal elements cannot be assumed to be subsets of the prior ones.

The following transform enables both kinds of conditioning.

4.2. The conditioned transform

In the previously defined transform, we compute the CEU over the prior focal set, which is not possible in general. On the contrary, for the following transform, we first compute the set of conditioned focal elements, which will all lead to a local game, even if they are not (subsets of) prior focal elements.

Let $\mathcal{S}_{\cup} = \bigcup_{i \in N, \theta_i \in \Theta_i} \mathcal{S}_{m_{|\theta_i}}$ be the set of all $m_{|\theta_i}$'s focal elements, that is, the union of focal sets obtained after all possible conditioning "given θ_i". The local games of the hypergraphical game $\tilde{G}$ correspond to the elements $B \in \mathcal{S}_{\cup}$. Again, (i, θ_i) plays in the local game corresponding to B if the type θ_i is plausible for B and obtains a local utility $m_{|\theta_i}(B) \times f^{\mathrm{XEU}}(B)$, equal to the amount of XEU which is computed over B.

Definition 20 (Conditioned transform).
The conditioned transform of a Bel game $G = \big(N, (A_i, \Theta_i, u_i)_{i \in N}, m\big)$ *is the hypergraphical game* $\tilde{G} = \big(\tilde{N}, \tilde{E}, (\tilde{A}_{(i,\theta_i)})_{(i,\theta_i) \in \tilde{N}}, (\tilde{u}^e_{(i,\theta_i)})_{e \in \tilde{E}, (i,\theta_i) \in e}\big)$ *where:*

- $\tilde{N} = \{(i, \theta_i) \mid i \in N, \theta_i \in \Theta_i\}$,
- $\tilde{A}_{(i,\theta_i)} = A_i$,
- $\tilde{E} = \Big[Players(B) \mid B \in \bigcup_{(i,\theta_i)\in\tilde{N}} \mathcal{S}_{m_{|\theta_i}}\Big]$,
- *For each* $e \in \tilde{E}$, $(i, \theta_i) \in e$ *and* $\tilde{\rho} \in \tilde{A}$, $\tilde{u}^e_{(i,\theta_i)}(\tilde{\rho}_e) = m_{|\theta_i}(B_e) \times f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B_e)$, *where* $\tilde{v}_i^{\tilde{\sigma}}(\theta) = u_i(\tilde{\sigma}_\theta, \theta)$.

It is straightforward to show that the XEU value of a pure strategy σ in G and the global utility of $\tilde{\sigma}$ in $\tilde{G}$ are equal, whatever is the couple (i, θ_i) considered. We also prove that:

Proposition 5. *Let G be a Bel game and $\tilde{G}$ its conditioned transform. For any pure or mixed strategy ρ of G, it holds that:*

(i) $\mathrm{XEU}_{(i,\theta_i)}(\rho) = \tilde{u}_{(i,\theta)}(\tilde{\rho})$

(ii) *ρ is a Nash equilibrium of G iff $\tilde{\rho}$ is a Nash equilibrium of $\tilde{G}$.*

Proof. *See Proof 6 in Appendix.*

Example 12. *The hypergraph of the conditioned transform of our running example is drawn on Figure 4.*

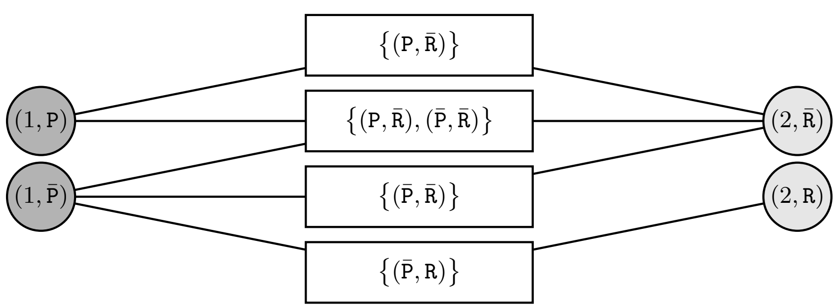

Figure 4: G's conditioned transform. Gray circles are vertices (players; one color per agent), white boxes are hyperedges (local games; linked to the involved players).

This transform can be applied with any rule of conditioning. Notice that the hypergraphical game it produces can be different from the one obtained with the direct transform (assuming the same conditioning in both, e.g., Dempster's rule of conditioning). In particular, it may produce more local games (with fewer players) than the direct transform.

Proposition 6 (Complexity of the conditioned transform).
The conditioned transform of a Bel game G has a temporal complexity in $O\big(n\cdot\beta\cdot T_{cond} + |\mathcal{S}_\cup|\cdot k'^2\cdot n\cdot\alpha^{k'\cdot n}\big)$ *and a spatial complexity in* $O\big(|\mathcal{S}_\cup|\cdot k'\cdot n\cdot\alpha^{k'\cdot n}\big)$*, where* $\mathcal{S}_\cup = \bigcup_{(i,\theta_i)\in\tilde{N}} \mathcal{S}_{m_{|\theta_i}}$, $k' = \max_{B\in\mathcal{S}_\cup}|B|$ *and* T_{cond} *is the temporal complexity of a conditioning of* m *"given* θ_i*".*
Using Dempster, Strong or Weak conditioning, it leads to a temporal complexity in $O\big(k\cdot \mathrm{Size}(G)^k\big)$ *and a spatial complexity of* $O\big(\mathrm{Size}(G)^k\big)$.

Proof. *See Proof 15 in Appendix.*

So, the conditioned transform leads to a different hypergraphical game than the direct one – but both represent the same problem, and both transforms have the same worst-case spatial complexity. In practice, the size of the transformed game depends on the structure of the belief function. Typically, if a focal element B involves only one type θ_i for a given agent i, both transforms will lead to the same local game $Players(B)$ (as $B = B_{|\theta_i}$), but the conditioned transform may produce (many) more local games and be less concise. If on the contrary, many types are compatible with a focal element B for any agent, the local game produced by the direct transform may have a bigger size. Consider for example a 2-player Bel game where m verify $m(\{(\theta_1,\theta_2),(\theta'_1,\theta'_2)\}) = m(\{(\theta_1,\theta'_2),(\theta'_1,\theta_2)\}) = 1/2$. With the direct transform, we get two 4-player local games, while the conditioned transform leads to four 2-player local games.

4.3. The TBM transform

In the framework of the Transferable Belief Model each agent first revises the prior knowledge using Dempster's rule of conditioning. Then, at the pignistic level (at the very moment the decision is made), the agent deduces a probabilistic distribution, by making the equiprobability assumption, and ranks the actions according to their expected utility. Of course, any of the two previous transformations can be used, letting XEU be the TBEU value in Definitions 19 and 20. However it is possible to exploit the 1-additivity of these pignistic probabilities in order to get a low-complexity hypergraph.

This last transform, dedicated to the TBM model, is simply defined by using a particular conditioning $m_{|C}^{\text{Pign}}$ in Definition 20: it summarizes the composition of the Dempster conditioning and the pignistic probability computation in one single step and directly derives a 1-additive mass function. Since any focal element is a singleton, the number of players in each local game is equal to the number of players in the original game. In particular, the resulting hypergraph is a polymatrix game when only two agents are involved in the original game.

For any C such that $\text{Pl}(C) > 0$ and any $\omega \in C$:

$$m_{|C}^{\text{Pign}}(\{\omega\}) = \frac{1}{\text{Pl}(C)} \times \sum_{\substack{B \in \mathcal{S}_m \\ \omega \in B}} \frac{m(B)}{|B \cap C|}$$

This leads to the following transform:

Definition 21 (TBM transform).

The TBM transform of a Bel game $G = \big(N, (A_i, \Theta_i, u_i)_{i \in N}, m\big)$ *is the hypergraphical game* $\tilde{G} = \big(\tilde{N}, \tilde{E}, (\tilde{A}_{(i,\theta_i)})_{(i,\theta_i) \in \tilde{N}}, (\tilde{u}^e_{(i,\theta_i)})_{e \in \tilde{E}, (i,\theta_i) \in e}\big)$ *where:*

- $\tilde{N} = \{(i, \theta_i) \mid i \in N, \theta_i \in \Theta_i\}$,
- $\tilde{A}_{(i,\theta_i)} = A_i$,
- $\tilde{E} = \big[Players\big(\{\theta\}\big) \mid \theta \in \Theta\big]$,
- *For each* $e \in \tilde{E}$*, let* θ *be the type configuration which led to* e*. For each* $(i, \theta_i) \in e$ *and* $\tilde{\rho} \in \tilde{A}$*,* $\tilde{u}^e_{(i,\theta_i)}(\tilde{\rho}_e) = m_{|\theta_i}^{Pign}(\{\theta\}) \times u_i(\tilde{\rho}_\theta, \theta)$.

Proposition 7. *Let* G *be a Bel game and* $\tilde{G}$ *its TBM transform. For any pure or mixed strategy* ρ *of* G*, it holds that:*

(i) $\text{TBEU}_{(i,\theta_i)}(\rho) = \tilde{u}_{(i,\theta)}(\tilde{\rho})$

(ii) ρ *is a Nash equilibrium of* G *iff* $\tilde{\rho}$ *is a Nash equilibrium of* $\tilde{G}$.

Proof. *See Proof 7 in Appendix.*

Proposition 8 (Complexity of the TBM transform).
The TBM transform of a Bel game G has a temporal complexity in $O\big(k \cdot \mathrm{Size}(G)\big)$ and a spatial complexity in $O\big(\,\mathrm{Size}(G)\big)$.

Proof. *See Proof 16 in Appendix.*

Example 13. *The hypergraphical game obtained by the TBM transform of our example is drawn on Figure 5.*

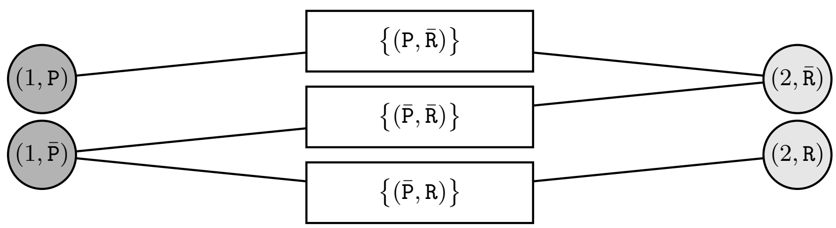

Figure 5: G's TBM transform. Gray circles are vertices (players; one color per agent), white boxes represent the local games.

4.4. Summary

Let us now summarize which transform can be used, depending on the nature of the prior knowledge (evidential or credal), on the conditioning rule and on the global utility criterion (CEU, JEU or TBEU):

- In a evidential view, the revision is made using Dempster conditioning, and the XEU may be the TBEU [11], JEU [27] or its special case CEU.
- In a credal view, the Dempster and Strong rules of conditioning hold when the knowledge is revised (i.e., when the agents learn facts) while FH conditioning holds when a focusing is to be performed (i.e., agents observe incidental events) – see [42, 31, 34] for further discussion about knowledge revision and focusing. Both the CEU and the JEU criteria are compatible with this view (the latter being a generalization of the former).

Table 5 indicates the transforms which suit each of these settings. In short, the conditioned transform is the only one which is suitable for all the settings. The direct transform holds only for Dempster's rule of conditioning while the TBM transform holds only for the TBEU.

Conditioning	XEU	Transform
Dempster's cond.	Any XEU	Direct transform
Any conditioning	Any XEU	Conditioned transform
Dempster's cond.	TBEU	TBM transform

Table 5: Valid conditionings and XEU depending on the transform.

Last, when several transforms are suitable for the setting, the choice of the transform to use can be guided by its complexity. In short: the TBM transform has a lower spatial complexity than the direct and the conditioned ones (Table 6). Except for the FH conditioning, the last two have the same worst-case complexity; they differ in that the direct one may have bigger local games (involving numerous agents) while the conditioned one may have more local games.

Transform	Temporal complexity	Spatial complexity
Direct transform	$O\big(k \cdot \mathrm{Size}(G)^k\big)$ *	$O\big(\,\mathrm{Size}(G)^k\big)$
Conditioned transform	$O\big(k \cdot \mathrm{Size}(G)^k\big)$ **	$O\big(\,\mathrm{Size}(G)^k\big)$ **
TBM transform	$O\big(k \cdot \mathrm{Size}(G)\big)$ *	$O\big(\,\mathrm{Size}(G)\big)$

Table 6: Complexity of the transforms – k is the degree of k-additivity – * Normalisation time excluded – ** Using Dempster's, Strong or Weak conditioning.

5. Conclusion

This article provides two main contributions. On the one hand, we define a model for simultaneous games of incomplete information based on belief functions. On the other hand, we introduce three transforms which make it possible to build an hypergraphical game (of complete information) equivalent to the initial Bel game, thus generalizing Howson–Rosenthal's theorem. The transformations preserve utilities, so the study of a Bel game can be reduced to that of a complete-information game. In particular, Nash equilibria are in correspondence: any equilibrium in one game is an equilibrium in the other. Furthermore, under some conditions (a low degree of additivity of the mass function or the use of the Transferable Belief Model framework), the transformation is polynomial in time and space; as a consequence, the algorithmic tools developed for hypergraphical games [17, 16] can be used to solve Bel games.

This work opens several research directions. First, we aim at extending this model to Choquet integrals based on any kind of capacity measure. It will allow for the definition of games based on other decision rules, for instance the rank-dependent utility rule [43]. Then, a finer complexity analysis can be conducted, based on the characterization of the conditioned mass functions. Finally, we would like to formalize those results with the Coq proof assistant [44] in order to build, with other on-going results, a modular, formal library on incomplete-information games and decision theory. This work will facilitate the extension of the model presented in this paper to other decision rules, including rules leading to partial orders (e.g., the interval-valued utility for belief functions [27]). In particular, we expect that the definitions could be extended straightforwardly and that the transforms should hold, up to a change in the range of utility functions from $\mathbb{R}$ to a non-totally-ordered set.

Appendix A. Proofs

Appendix A.1. Proofs of correctness

Proof 1 (Correctness of Definition 14 – Utility). *Particular case of Definition 17, using* CEU *and the Dempster conditioning (see Proof 3).*

Proof 2 (Correctness of Definition 16 – Utility). *Particular case of Definition 17, using the Dempster conditioning (see Proof 3).*

Proof 3 (Correctness of Definition 17 – Utility). *On the one hand, any mixed strategy profile* $\rho \in \prod_{i \in N}(\Theta_i \to \pi(A_i))$ *defines a probability* Pr^ρ *over the possible pure strategy profiles* $\sigma \in \Sigma = \prod_{i \in N}(\Theta_i \to A_i)$ *by:*[5]

$$\mathrm{Pr}^\rho(\sigma) = \prod_{i \in N} \prod_{\theta_i \in \Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i))$$

On the other hand, merging ρ *and* $m_{|\theta_i}$ *leads to a bpa* m^ρ *over* $A \times \Theta$*, from which any element* $X = \{(a, \theta), (a', \theta'), \dots\}$ *is focal iff* $B = \{\theta, \theta', \dots\}$ *is focal for* m *and (at least) one pure strategy profile* σ *is compatible with* X *and possible according to* Pr^ρ*; i.e.,* $\sigma(\theta) = a$*,* $\sigma(\theta') = a', \dots$ *and* $\mathrm{Pr}^\rho(\sigma) > 0$*.*

Let $g.B := \{(g(\theta), \theta) \mid \theta \in B\}$ *denote such focal element. By definition:*

$$m^\rho_{|\theta_i}(g.B) = m(B) \times \sum_{\substack{\sigma \in \Sigma \\ \forall \theta \in B, \sigma(\theta) = g(\theta)}} \mathrm{Pr}^\rho(\sigma)$$

Thus, the XEU *of a mixed strategy profile rewrites:*

$$\begin{aligned}
\mathrm{XEU}_{(i,\theta_i)}(\rho) &= \sum_{g.B \in \mathcal{S}_{m^\rho_{|\theta_i}}} m^\rho_{|\theta_i}(g.B) \times f^{\mathrm{XEU}}_{v_i^\sigma}(B) \\
&= \sum_{B \in \mathcal{S}_{m_{|\theta_i}}} \sum_{g : B \to A} m_{|\theta_i}(B) \times \left(\sum_{\substack{\sigma \in \Sigma \\ \forall \theta \in B, \sigma(\theta) = g(\theta)}} \mathrm{Pr}^\rho(\sigma) \right) \times f^{\mathrm{XEU}}_{v_i^\sigma}(B) \\
&= \sum_{B \in \mathcal{S}_{m_{|\theta_i}}} \sum_{g : B \to A} \sum_{\substack{\sigma \in \Sigma \\ \forall \theta \in B, \sigma(\theta) = g(\theta)}} m_{|\theta_i}(B) \times \mathrm{Pr}^\rho(\sigma) \times f^{\mathrm{XEU}}_{v_i^\sigma}(B)
\end{aligned}$$

Given any $B \subseteq \Theta$*, the set of functions* $g : B \to A$ *defines a partition of* Σ

[5] $\pi(X)$ denotes the set of probabilities over X

(which groups σ mappings whose images of B are identical), so:

$$
\begin{aligned}
\mathrm{XEU}_{(i,\theta_i)}(\rho) &= \sum_{B\in\mathcal{S}_{m_{|\theta_i}}} \sum_{\sigma\in\Sigma} m_{|\theta_i}(B) \times \mathrm{Pr}^{\rho}(\sigma) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B) \\
&= \sum_{\sigma\in\Sigma} \mathrm{Pr}^{\rho}(\sigma) \times \sum_{B\in\mathcal{S}_{m_{|\theta_i}}} \times m_{|\theta_i}(B) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B) \\
&= \sum_{\sigma\in\Sigma} \left(\prod_{i\in N} \prod_{\theta_i\in\Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{B\in\mathcal{S}_{m_{|\theta_i}}} m_{|\theta_i}(B) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B)
\end{aligned}
$$

Proof 4 (Proposition 2 – Direct transform). *Corollary of Proposition 3 in the case of pure strategies (see Proof 5).*

Proof 5 (Proposition 3 – Direct transform). *Let G be a Bel Game, $\tilde{G}$ be its conditioned transform and $E_{\theta_i} = \{\theta' \mid \theta'_i = \theta_i\}$ be the conditioning event "given θ_i".*

Recall that for any mixed strategy profile ρ of G, $\tilde{\rho}$ is its Selten transform and $\tilde{\rho}_{(i,\theta_i)} = \rho_i(\theta_i)$ is a probability distribution over $A_i = \tilde{A}_{(i,\theta_i)}$. Similarly, for any pure strategy profile σ, $\tilde{\sigma}$ is its Selten transform and $\tilde{\sigma}_{(i,\theta_i)} = \sigma_i(\theta_i) \in A_i = \tilde{A}_{(i,\theta_i)}$. Finally, since the Selten transform is bijective, we have:

$$
\begin{aligned}
\mathrm{XEU}_{(i,\theta_i)}(\rho) &= \sum_{\sigma\in\Sigma} \left(\prod_{i\in N} \prod_{\theta_i\in\Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{B\in\mathcal{S}_{m^{Dem}_{|\theta_i}}} m^{Dem}_{|\theta_i}(B) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B) \\
&= \sum_{\sigma\in\Sigma} \left(\prod_{i\in N} \prod_{\theta_i\in\Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{\substack{B\in\mathcal{S}_m \\ B\cap E_{\theta_i}\neq\emptyset}} K_{|\theta_i} \times m(B) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B\cap E_{\theta_i}) \\
&= \sum_{\sigma\in\Sigma} \left(\prod_{i\in N} \prod_{\theta_i\in\Theta_i} \rho_i(\theta_i)(\sigma_i(\theta_i)) \right) \times \sum_{\substack{e\in\tilde{E} \\ (i,\theta_i)\in e}} K_{|\theta_i} \times m(B_e) \times f^{\mathrm{XEU}}_{v_i^{\sigma}}(B_e\cap E_{\theta_i}) \\
&= \sum_{\tilde{\sigma}\in\tilde{A}} \left(\prod_{(i,\theta_i)\in\tilde{N}} \tilde{\rho}_{(i,\theta_i)}(\tilde{\sigma}_{(i,\theta_i)}) \right) \times \tilde{u}_{(i,\theta_i)}(\tilde{\sigma}) \\
&= EU_{(i,\theta_i)}(\tilde{\rho})
\end{aligned}
$$

Proof 6 (Proposition 5 – Conditioned transform). *Same remarks as for Proof 5. It leads to:*

$$\begin{aligned}
\mathrm{XEU}_{(i,\theta_i)}(\rho) &= \sum_{\sigma\in\Sigma}\left(\prod_{i\in N}\prod_{\theta_i\in\Theta_i}\rho_i(\theta_i)(\sigma_i(\theta_i))\right)\times\sum_{B\in\mathcal{S}_{m_{|\theta_i}}} m_{|\theta_i}(B)\times f^{\mathrm{XEU}}_{v_i^\sigma}(B)\\
&= \sum_{\tilde\sigma\in\tilde A}\left(\prod_{(i,\theta_i)\in\tilde N}\tilde\rho_{(i,\theta_i)}(\tilde\sigma_{(i,\theta_i)})\right)\times\sum_{B\in\mathcal{S}_{m_{|\theta_i}}} m_{|\theta_i}(B)\times f^{\mathrm{XEU}}_{\tilde v_i^{\tilde\sigma}}(B)\\
&= \sum_{\tilde\sigma\in\tilde A}\left(\prod_{(i,\theta_i)\in\tilde N}\tilde\rho_{(i,\theta_i)}(\tilde\sigma_{(i,\theta_i)})\right)\times\sum_{\substack{e\in\tilde E\\ \theta_i\in e}} m_{|\theta_i}(B_e)\times f^{\mathrm{XEU}}_{\tilde v_i^{\tilde\sigma}}(B_e)\\
&= \sum_{\tilde\sigma\in\tilde A}\left(\prod_{(i,\theta_i)\in\tilde N}\tilde\rho_{(i,\theta_i)}(\tilde\sigma_{(i,\theta_i)})\right)\times\tilde u_{(i,\theta_i)}(\tilde\sigma)\\
&= EU_{(i,\theta_i)}(\tilde\rho)
\end{aligned}$$

Proof 7 (Proposition 7 – TBM transform). *Same remarks as for Proof 5. It leads to:*

$$\begin{aligned}
\mathrm{TBEU}_{(i,\theta_i)}(\rho) &= \sum_{\sigma\in\Sigma}\left(\prod_{i\in N}\prod_{\theta_i\in\Theta_i}\rho_i(\theta_i)(\sigma_i(\theta_i))\right)\times\sum_{\theta\in\Theta}\mathrm{BetP}_{m^{Dem}_{|\theta_i}}(\theta)\times u_i(\sigma(\theta),\theta)\\
&= \sum_{\sigma\in\Sigma}\left(\prod_{i\in N}\prod_{\theta_i\in\Theta_i}\rho_i(\theta_i)(\sigma_i(\theta_i))\right)\times\left(\sum_{\substack{B\in\mathcal{S}_{m^{Dem}_{|\theta_i}}\\ \theta\in B}}\frac{m(B)}{|B|}\right)\times u_i(\sigma(\theta),\theta)\\
&= \sum_{\sigma\in\Sigma}\left(\prod_{i\in N}\prod_{\theta_i\in\Theta_i}\rho_i(\theta_i)(\sigma_i(\theta_i))\right)\times\left(\frac{1}{\mathrm{Pl}(E_{\theta_i})}\times\sum_{\substack{B\in\mathcal{S}_m\\ \theta\in B}}\frac{m(B)}{|B\cap E_{\theta_i}|}\right)\times u_i(\sigma(\theta),\theta)\\
&= \sum_{\sigma\in\Sigma}\left(\prod_{i\in N}\prod_{\theta_i\in\Theta_i}\rho_i(\theta_i)(\sigma_i(\theta_i))\right)\times m^{Pign}_{|\theta_i}(\{\theta\})\times u_i(\sigma(\theta),\theta)\\
&= \sum_{\tilde\sigma\in\tilde A}\left(\prod_{(i,\theta_i)\in\tilde N}\tilde\rho_{(i,\theta_i)}(\tilde\sigma_{(i,\theta_i)})\right)\times\tilde u_{(i,\theta_i)}(\tilde\sigma)\\
&= EU_{(i,\theta_i)}(\tilde\rho)
\end{aligned}$$

Proof 8 (Theorem 1 – Extended Howson–Rosenthal's theorem). *Direct* *corollary of Proof 6.*

Appendix A.2. Complexity lemmas – mass function operations

In this section we consider a k-additive mass function m. We denote $s = |\mathcal{S}_m|$. Set operations on focal elements (such as union, intersection and membership) involve $O(k)$ operations. Representing a mass function just involves a dictionary mapping at most s subsets (encoded by k bits) to numeric values. Hence a space complexity of $O(ks)$. In the sequel, we will denote such complexity statements by $\mathrm{Time}(\cap) \in O(k)$ and $\mathrm{Size}(m) \in O(ks)$ for example.

Lemma 1 (Dempster conditioning's complexity).

- *m is k-additive $\implies m_{|C}^{Dem}$ is at most k-additive*
- $\mathrm{Size}(m_{|C}^{Dem}) \leq \mathrm{Size}(m) \in O(ks)$,
- $\mathrm{Time}\left(m_{|C}^{Dem}\right) \in O(ks)$.

Proof 9 (Lemma 1 – Dempster conditioning's complexity). *First, note that every focal element of $m_{|C}^{Dem}$ is a subset of a focal element of m, so $m_{|C}^{Dem}$ is at most k-additive and $|\mathcal{S}_{m_{|C}^{Dem}}| \leq s$. Thus we have* $\mathrm{Size}(m_{|C}^{Dem}) \leq \mathrm{Size}(m) \in O(ks)$.

To compute all the values of $m_{|C}^{Dem}$, two loops over $\mathcal{S}_m$ suffice. Initialize $m_{|C}^{Dem}$ as a function which defaults to 0, and also a single variable $\mathrm{Pl}(C) := 0$.

- *First, compute both $m_{|C}^{Dem}$'s unnormalized values and the normalization factor $\mathrm{Pl}(C)$: for each $B \in \mathcal{S}_m$, if $B \cap C \neq \emptyset$, add $m(B)$ to $m_{|C}^{Dem}(B \cap C)$ and to $\mathrm{Pl}(C)$.*
- *Second, normalize those values: for each $B \in \mathcal{S}_{m_{|C}^{Dem}}$, $m_{|C}^{Dem}(B)$ becomes $m_{|C}^{Dem}(B)/\,\mathrm{Pl}(C)$.*

The first loop involves s tests thus is in $O(ks)$. The second one doesn't involve any test and is thus in $O(s)$. Finally, $\mathrm{Time}(m_{|C}^{Dem}) \in O(ks)$.

Lemma 2 (Weak conditioning's complexity).

- m *is* k*-additive* $\implies m_{|C}^{Weak}$ *is* k*-additive,*
- $\mathrm{Size}(m_{|C}^{Weak}) \leq \mathrm{Size}(m) \in O(ks)$*,*
- $\mathrm{Time}\left(m_{|C}^{Weak}\right) \in O(ks)$

Proof 10 (Lemma 2 – Weak conditioning's complexity). *The proof is similar to Proof 9 since the algorithm is almost identical: the only difference is that masses stays on* B *(they are not transferred to* $B \cap C$*). The complexity result holds:* $\mathrm{Time}(m_{|C}^{Weak})$ *and* $\mathrm{Size}(m_{|C}^{Weak})$ *are both in* $O(ks)$*.*

Lemma 3 (Strong conditioning's complexity).

- m *is* k*-additive* $\implies m_{|C}^{Strong}$ *is* k*-additive,*
- $\mathrm{Size}(m_{|C}^{Strong}) \leq \mathrm{Size}(m) \in O(ks)$*,*
- $\mathrm{Time}\left(m_{|C}^{Strong}\right) \in O(ks)$

Proof 11 (Lemma 3 – Strong conditioning's complexity). *The proof is similar to Proof 10 since the algorithm is almost identical: the only difference is that the test condition changes from* $B \cap C \neq \emptyset$ *to* $B \subseteq C$ *(thus the normalization factor is* $\mathrm{Bel}(C)$*). Since this test is also in* $O(k)$*, the complexity result holds:* $\mathrm{Time}(m_{|C}^{Strong})$ *and* $\mathrm{Size}(m_{|C}^{Strong})$ *are both in* $O(ks)$*.*

Lemma 4 (Fagin-Halpern conditioning's complexity).

- *In the worst case,* $m_{|C}^{FH}$ *can be* $|C|$*-additive, even if* m *is* k*-additive with* $k < |C|$
- $\mathrm{Size}\left(m_{|C}^{FH}\right)$ *may have* $2^{|C|}$ *focal elements even if* $\mathcal{S}_m < 2^{|C|}$

Proof 12 (Lemma 4 – Fagin-Halpern conditioning's complexity). *Fagin–Halpern conditioning does not preserve nor the size, neither the* k*-additivity of* m*. Consider for instance a frame of discernment* $\Omega = \{\omega_1, \dots, \omega_n\}$ *and a 2-additive mass function* m *such as* $m(\{\omega_i\}) > 0$ *for all* i *and* $m(\{\omega_i, \omega_j\}) > 0$ *for all* $i \neq j$*. Then, for any nonempty* $C \subset \Omega$*, each subset of* $B \subseteq C$ *is a focal element of* $\mathrm{Bel}(\cdot \mid C)$ *– thus* $|\mathcal{S}_{m_{|C}^{FH}}| = 2^{|C|}$ *and* $\mathrm{Bel}(\cdot \mid C)$ *is* $|C|$*-additive.*

Appendix A.3. Proofs of complexity – Games

In this section we consider a Bel game G with $n \geq 2$ players, each of these having at most $\alpha \geq 2$ actions and $\beta \geq 2$ types, along with a k-additive mass function with s focal elements, each of them is a set of n-tuples of types – so $s \leq \beta^{kn}$.

Proof 13 (Proposition 1 – Spatial complexity of a Bel game). *It holds that:*

- *G contains n utility tables of size $(\alpha\beta)^n$ (one for each agent, assigning an utility value to a strategy profile and a type configuration).*
- *The size of m is bounded by kns (each of the s focal elements contains at most k n-tuples of types).*

Thus, $\mathrm{Size}(G) \in O(n(\alpha\beta)^n + kns)$.

Proof 14 (Proposition 4 – Complexity of the direct transform). *The direct transform $\tilde{G}$ of G has exactly s local games (one for each focal element B), in which players are possible pairs (i, θ_i) such that $\exists \theta' \in B, \theta'_i = \theta_i$. There may be kn such pairs, so the corresponding local game is described by at most kn matrices of α^{kn} cells, hence a spatial cost for the representation of $\tilde{G}$ in $O(skn\alpha^{kn})$. Recall that $s \leq \beta^{kn}$ and $kn \leq n^k$, it holds that* $\mathrm{Size}(\tilde{G})$ *is bounded by $kn(\alpha\beta)^{kn} \leq n^k(\alpha\beta)^{kn}$, i.e.,* $\mathrm{Size}(\tilde{G}) \in O(\mathrm{Size}(G)^k)$.

To instantiate those matrices, one has to compute each of the utility values as from Definition 19:

- *First, for each of the $n\beta$ pairs (i, θ_i), compute $\mathrm{Pl}(E_{\theta_i})$ by a single loop over m's focal set in which $n\beta$ tests are made, so it involves $O(skn\beta)$ operations.*
- *Then, for each of the s focal elements B (i.e., a local game e), for each of the kn possible corresponding pairs (i, θ_i) and for each of the α^{kn} possible local strategy profiles $\tilde{\sigma}_e$, set $\tilde{u}^e_{(i,\theta_i)}(\tilde{\sigma}_e) := m(B) \times f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{\sigma}}}(B \cap E_{\theta_i}) / \mathrm{Pl}(E_{\theta_i})$, where $f^{\mathrm{XEU}}_{\tilde{v}_i^{\tilde{sigma}}}(B \cap E_{\theta_i})$ involves k operations.*

Thus we have $\mathrm{Time}(\tilde{G}) \in O(skn\beta + sk^2n\alpha^{kn}) = O(skn(\beta + k\alpha^{kn}))$*. Since* m *is* k*-additive,* $s \leq \beta^{kn}$*, so* $\mathrm{Time}(\tilde{G})$ *is bounded by* $kn\beta^{kn}(\beta + k\alpha^{kn}) \in O(k^2n\beta(\alpha\beta)^{kn}) \subseteq O(\beta kn^k(\alpha\beta)^{kn})$*; i.e.,* $\mathrm{Time}(\tilde{G}) \in O(\beta k \times \mathrm{Size}(G)^k)$*.*

Note that since the normalization doesn't change the equilibria of $\tilde{G}$*, usually the first loop is not necessary and the complexity becomes* $\mathrm{Time}(\tilde{G}) \in O(k \times \mathrm{Size}(G)^k)$*.*

Proof 15 (Proposition 6 – Complexity of the conditioned transform). *Let* $s' = |\mathcal{S}_\cup|$ *be the total number of focal elements, after all conditioning, and* $k' = \max_{B\in\mathcal{S}_\cup} |B|$ *their maximal size. The conditioned transform* $\tilde{G}$ *of* F *has exactly* s' *local games, which involve at most* $k'n$ *players each, thus they are described by at most* $k'n$ *matrices of* $\alpha^{k'n}$ *cells, hence a spatial cost for the representation of* $\tilde{G}$ *in* $O\big(k'ns'\alpha^{k'n}\big)$*.*

If the conditioning is one of the Dempster, Strong or Weak ones, it holds that $k' \leq k$ *and* $s' \leq \beta^{kn}$ *(from Lemma 1, Lemma 2 or Lemma 3), so the bound becomes* $kn(\alpha\beta)^{kn}$*; i.e.,* $\mathrm{Size}(\tilde{G}) \in O\big(\mathrm{Size}(G)^k\big)$*.*

On the contrary, if the Fagin-Halpern conditioning is used, we can just bound $k' \in O(n\beta)$ *and* $s' \in O(2^{\beta^n})$ *(by Lemma 4), and thus get a spatial complexity* $\mathrm{Size}(\tilde{G}) \in O\big(n^2\beta 2^{\beta^n}\alpha^{n^2\beta}\big)$*.*

To construct those local utility matrices, one has to compute each of the utility values:

- *First, for each of the* $n\beta$ *pairs* (i, θ_i)*, compute* $m_{|\theta_i}$*, according to the chosen conditioning, say it costs* T_{cond}*.*
- *Then, for each of the* s' *local games, for each of its* $k'n$ *players and for each of the* $\alpha^{k'n}$ *possible local strategy profiles* $\tilde{\sigma}_e$*, set* $\tilde{u}^e_{(i,\theta_i)}(\tilde{\sigma}_e) := m_{|\theta_i}(B_e) \times f^{XEU}_{\tilde{v}_i^{\tilde{\sigma}}}(B_e)$*, where* $f^{XEU}_{\tilde{v}_i^{\tilde{\sigma}}}(B_e)$ *involves at most* k' *operations.*

The first loop costs $n\beta T_{cond}$ *operations, the second one costs* $s'k'^2n\alpha^{k'n}$ *operations.*

If the conditioning is one of the Dempster, Strong or Weak ones, it holds that $k' \leq k$*,* $s' \leq \beta^{kn}$ *and* $T_{cond} \in O(kn)$ *(from Lemma 1, Lemma 2 or Lemma*

3), so the bound becomes $k^2n(\alpha\beta)^{kn}$; i.e., $\mathrm{Time}(\tilde{G}) \in O\big(k \cdot \mathrm{Size}(G)^k\big)$.

On the contrary, if the Fagin-Halpern conditioning is used, we can just bound $k' \in O(n\beta)$, $s' \in O(2^{\beta^n})$. [33]'s algorithm is used to compute masses in the first loop – $T_{cond} \in O(2^{\beta^n})$ – thus we get a temporal complexity $\mathrm{Time}(\tilde{G}) \in O(n\beta 2^{\beta^n} + n^3\beta^2 2^{\beta^n}\alpha^{n^2\beta}) = O(n^3\beta^2 2^{\beta^n}\alpha^{n^2\beta})$.

Proof 16 (Proposition 8 – Complexity of the TBM transform).
The TBM transform $\tilde{G}$ of G has at most $|\Theta| = \beta^n$ local games (one for each possible type configuration θ), in which players are the n possible pairs (i, θ_i). So each local game is described by n matrices of α^n cells, hence a spatial cost for the representation of $\tilde{G}$ in $O\big(n(\alpha\beta)^n\big)$ – that is, $\mathrm{Size}(\tilde{G}) \in O\big(\mathrm{Size}(G)\big)$. *To construct these local utility matrices, one has to compute each of the utility values. First precompute* $\mathrm{Pl}(E_{\theta_i})$ *for each type of each agent, it requires $O(skn\beta)$ operations (see Proof 14). Then, for each agent $i \in N$, we compute in a single loop the β conditioned mass functions $m^{Pign}_{|\theta_i}$:*

- *For each $\theta \in \Theta$, initialize $m^{Pign}_{|\theta_i}(\theta) := 0$. It costs $O(\beta^n)$ operations.*
- *For each focal element $B \in \mathcal{S}_m$, and for each $\theta \in B$, add $\frac{m(B)}{|B \cap E_{\theta_i}| \times \mathrm{Pl}(E_{\theta_i})}$ to $m^{Pign}_{|\theta_i}(\theta)$. It costs $O(sk^2)$ operations.*

Thus, computing all $m^{Pign}_{|\theta_i}$'s value (for each i and θ_i) requires $O\big(n\beta^n + nsk^2\big)$ operations. Finally, for each of the β^n local games, for each of the n local players (i, θ_i) and each of the α^n action profiles, set the local utility in $O(1)$ – it requires $O\big(n(\alpha\beta)^n\big)$ operations.

Finally, the full transform requires $O\big(skn\beta + n(\alpha\beta)^n + sk^2n\big)$ operations. Note that since the normalization doesn't change the equilibria of $\tilde{G}$, it is usually not needed to compute $\mathrm{Pl}(E_{\theta_i})$*'s values, then the complexity becomes* $\mathrm{Time}(\tilde{G}) \in O\big(n(\alpha\beta)^n + sk^2n\big) = O\big(k \times \mathrm{Size}(G)\big)$.

References

[1] O. Morgenstern, J. Von Neumann, Theory of Games and Economic Behavior, Princeton University Press, 1953.

[2] J. Nash, Non-Cooperative Games, Annals of Mathematics (1951) 286–295.

[3] J. C. Harsanyi, Games With Incomplete Information Played by "Bayesian" Players, I–III Part I. The Basic Model, Management Science 14 (3) (1967) 159–182.

[4] D. Ellsberg, Risk, Ambiguity, and the Savage Axioms, The Quarterly Journal of Economics (1961) 643–669.

[5] R. Von Mises, Probability, statistics, and truth, Courier Corporation, 1981.

[6] A. P. Dempster, Upper and Lower Probabilities Induced by a Multivalued Mapping, The Annals of Mathematical Statistics 38 (1967) 325–339.

[7] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[8] G. Choquet, Theory of Capacities, in: Annales de l'institut Fourier, Vol. 5, 1954, pp. 131–295.

[9] I. Gilboa, D. Schmeidler, Maxmin Expected Utility with Non-Unique Prior, Journal of Mathematical Economics 18 (2) (1989) 141–153.

[10] D. Schmeidler, Integral Representation without Additivity, Proceedings of the American Mathematical Society 97 (2) (1986) 255–261.

[11] P. Smets, R. Kennes, The Transferable Belief Model, Artificial Intelligence 66 (2) (1994) 191–234.

[12] J.-Y. Jaffray, Linear Utility Theory for Belief Functions, Operations Research Letters 8 (2) (1989) 107–112.

[13] J. T. Howson Jr, R. W. Rosenthal, Bayesian Equilibria of Finite Two-Person Games with Incomplete Information, Management Science 21 (3) (1974) 313–315.

[14] E. Yanovskaya, Equilibrium Points in Polymatrix Games, Lithuanian Mathematical Journal 8 (1968) 381–384.

[15] C. H. Papadimitriou, T. Roughgarden, Computing Correlated Equilibria in Multi-Player Games, Journal of the ACM 55 (3) (2008) 1–29.

[16] M. Wahbi, K. N. Brown, A Distributed Asynchronous Solver for Nash Equilibria in Hypergraphical Games, in: Frontiers in Artificial Intelligence and Applications, Vol. 285, 2016, pp. 1291–1299.

[17] A. Chapman, A. Farinelli, E. M. de Cote, A. Rogers, N. Jennings, A Distributed Algorithm for Optimising over Pure Strategy Nash Equilibria, in: Proc. of the 24th AAAI Conference on Artificial Intelligence, Vol. 24, 2010, pp. 749–755.

[18] D. Dubois, H. Prade, On the Unicity of Dempster Rule of Combination, International Journal of Intelligent Systems 1 (2) (1986) 133–142.

[19] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman & Hall, 1991.

[20] T. Denoeux, Maximum likelihood estimation from uncertain data in the belief function framework, IEEE Transactions on knowledge and data engineering 25 (1) (2011) 119–130.

[21] R. Martin, False confidence, non-additive beliefs, and valid statistical inference, International Journal of Approximate Reasoning 113 (2019) 39–73.

[22] T. Denœux, Logistic regression, neural networks and dempster–shafer theory: A new perspective, Knowledge-Based Systems 176 (2019) 54–67.

[23] T. Denoeux, Decision-making with belief functions: a review, International Journal of Approximate Reasoning 109 (2019) 87–110.

[24] I. Gilboa, Expected Utility with Purely Subjective Non-Additive Probabilities, Journal of Mathematical Economics 16 (1) (1987) 65–88.

[25] D. Schmeidler, Subjective Probability and Expected Utility without Additivity, Econometrica: Journal of the Econometric Society (1989) 571–587.

[26] I. Gilboa, D. Schmeidler, Additive Representations of Non-Additive Measures and the Choquet Integral, Annals of Operations Research 52 (1) (1994) 43–65.

[27] T. Denoeux, P. P. Shenoy, An Interval-Valued Utility Theory for Decision Making with Dempster-Shafer Belief Functions, International Journal of Approximate Reasoning 124 (2020) 194–216.

[28] L. S. Shapley, Notes on the N-person Game–II: Characteristic-point Solutions of the Four-person Game, Rand Corporation, 1951.

[29] J. Y. Halpern, R. Fagin, Two Views of Belief: Belief as Generalized Probability and Belief as Evidence, Artificial intelligence 54 (3) (1992) 275–317.

[30] J.-Y. Jaffray, Bayesian Updating and Belief Functions, IEEE Transactions on Systems, Man, and Cybernetics 22 (5) (1992) 1144–1152.

[31] D. Dubois, T. Denoeux, Conditioning in Dempster-Shafer Theory: Prediction vs. Revision, in: Proc. of the 2nd International Conference on Belief Functions, Springer, 2012, pp. 385–392.

[32] R. Fagin, J. Y. Halpern, A New Approach to Updating Beliefs, in: Proc. of the 6th conference on Uncertainty in Artificial Intelligence (UAI), 1990, pp. 347–374.

[33] L. G. Polpitiya, K. Premaratne, M. N. Murthi, S. J. Murrell, D. Sarkar, Efficient computation of conditionals in the dempster-shafer belief theoretic framework, IEEE transactions on cybernetics (2020).

[34] R. Gong, X.-L. Meng, Judicious judgment meets unsettling updating: Dilation, sure loss and simpson's paradox, Statistical Science 36 (2) (2021) 169–190.

[35] P. Smets, Jeffrey's Rule of Conditioning Generalized to Belief Functions, in: Uncertainty in artificial intelligence, Elsevier, 1993, pp. 500–505.

[36] B. Planchet, Credibility and Conditioning, Journal of Theoretical Probability 2 (3) (1989) 289–299.

[37] M. J. Kearns, M. L. Littman, S. P. Singh, Graphical Models for Game Theory, in: J. S. Breese, D. Koller (Eds.), Proceedings of the 17th conference in Uncertainty in Artificial Intelligence, 2001, pp. 253–260.

[38] R. B. Myerson, Game Theory, Harvard university press, 2013.

[39] J.-Y. Jaffray, Linear Utility Theory and Belief Functions: a Discussion, in: Progress in decision, utility and risk theory, Springer, 1991, pp. 221–229.

[40] P. Miranda, M. Grabisch, P. Gil, Dominance of Capacities by k-Additive Belief Functions, European Journal of Operational Research 175 (2) (2006) 912–930.

[41] M. Grabisch, Upper Approximation of Non-additive Measures by k-additive Measures—The Case of Belief Functions., in: Proc. of the 1st International Symposium on Imprecise Probabilities and Their Applications (ISIPTA), 1999, pp. 158–164.

[42] D. Dubois, H. Prade, Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge, in: Qualitative and quantitative practical reasoning, Springer, 1997, pp. 96–107.

[43] J. Quiggin, A Theory of Anticipated Utility, Journal of economic behavior & organization 3 (4) (1982) 323–343.

[44] The Coq Development Team, The Coq Proof Assistant: Reference Manual: version 8.15 (2022).
URL `https://coq.github.io/doc/V8.15.0/refman/`