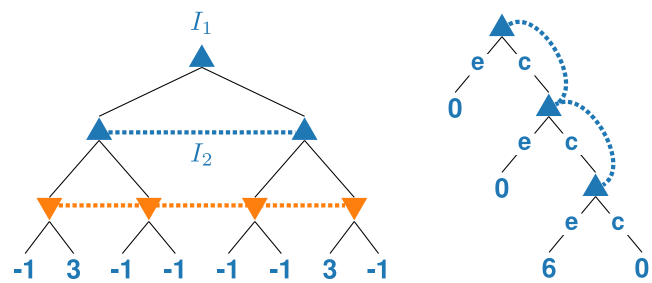
Imperfect-Recall Games: Equilibrium Concepts and Their Complexity*

Emanuel Tewolde[1,4] , **Brian Hu Zhang**[1] , **Caspar Oesterheld**[1,4] ,
Manolis Zampetakis[2] , **Tuomas Sandholm**[1,5,6,7] , **Paul W. Goldberg**[3] and **Vincent Conitzer**[1,3,4]
[1]Carnegie Mellon University
[2]Yale University
[3]University of Oxford
[4]Foundations of Cooperative AI Lab (FOCAL)
[5]Strategic Machine, Inc., [6]Strategy Robot, Inc., [7]Optimized Markets, Inc.
emanueltewolde@cmu.edu, bhzhang@cs.cmu.edu, oesterheld@cmu.edu, manolis.zampetakis@yale.edu,
sandholm@cs.cmu.edu, paul.goldberg@cs.ox.ac.uk, conitzer@cs.cmu.edu

Abstract

We investigate optimal decision making under imperfect recall, that is, when an agent forgets information it once held before. An example is the absentminded driver game, as well as team games in which the members have limited communication capabilities. In the framework of extensive-form games with imperfect recall, we analyze the computational complexities of finding equilibria in multiplayer settings across three different solution concepts: Nash, *multiselves* based on evidential decision theory (EDT), and multiselves based on causal decision theory (CDT). We are interested in both exact and approximate solution computation. As special cases, we consider (1) single-player games, (2) two-player zero-sum games and relationships to maximin values, and (3) games without exogenous stochasticity (chance nodes). We relate these problems to the complexity classes P, PPAD, PLS, Σ_2^{P}, $\exists\mathbb{R}$, and $\exists\forall\mathbb{R}$.



(a) Forgetful penalty shoot-out. This game has no Nash equilibrium.

(b) Extended absentminded driver.

Figure 1: Games with imperfect recall. P1's (▲) utility payoffs are labeled on each terminal node. If P2 (▼) is present, the game is zero sum. Infosets are joined by dotted lines.

1 Introduction

In game theory, it is common to restrict attention to games of *perfect recall*, that is, games in which no player ever forgets anything. At first, it seems that this assumption is even better motivated for AI agents than for human agents: humans forget things, but AI does not have to. However, we argue this view is mistaken: there are often reasons to design AI agents to forget, or to structure them so that they can be modeled as forgetful. Moreover, such forgetting-by-design follows predictable rules and is thereby easier to model formally than idiosyncratic human forgetting. Thus, games of imperfect recall are receiving renewed attention from AI researchers.

Imperfect recall is already being used for state-of-the-art *abstraction* algorithms for larger games of perfect recall [Waugh *et al.*, 2009; Ganzfried and Sandholm, 2014; Brown *et al.*, 2015]. The idea is that by forgetting unimportant aspects of the past, the AI can afford to conduct equilibrium-approximation computations with a game model that has a more refined abstraction of the present. Indeed, imperfect-recall abstractions were a key component in the first superhuman AIs in no-limit Texas hold'em poker [Brown and Sandholm, 2018; Brown and Sandholm, 2019].

Imperfect recall also naturally models settings in which forgetting is deliberate for other reasons, such as privacy of sensitive data [Conitzer, 2019; Zhang and Sandholm, 2022]. Conitzer provides the example of an AI driving assistant designed to intervene whenever the human car driver makes a significant error. In such instances, the AI must assess the overall skill level of the human driver, despite not being allowed to store information about the individual.

It can also model *teams* of agents with common goals and limited ability to communicate. Each team, represented by one agent with imperfect recall, is then striving for some notion of optimality among team members [von Stengel and Koller, 1997; Celli and Gatti, 2018; Emmons *et al.*, 2022; Zhang *et al.*, 2023]. Highly distributed agents are similarly well-described by imperfect recall: such an agent may take an action at one node based on information at that node, and then need to take another action at a second node without yet having learned yet what happened at the first node. Thus, effectively, the distributed agent has forgotten what it

*Published in the Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI 2024).

Multi-player

	Nash (D)	EDT (D)	CDT (S)
exact	$\exists\mathbb{R}$-hard and in $\exists\forall\mathbb{R}$ (Thms. 1 & 3)		—
1/exp	Σ_2^{P}-complete (Thms. 2 and 4)		PPAD-complete (Thm. 6)
1/poly			

Single-player

	Optimal (D)	EDT (S)	CDT (S)
exact	$\exists\mathbb{R}$-complete [Gimbert *et al.*, 2020]	—	—
1/exp	NP-complete [Koller and Megiddo, 1992; Tewolde *et al.*, 2023]	PLS-complete (Thm. 5*)	CLS-complete [Tewolde *et al.*, 2023]
1/poly		P (Cor. 22*)	P (Cor. 17)

Table 1: Summary of complexity results. New results from this paper are shown with a light green background. (S) stands for search problem, which is when we ask for a solution strategy profile. In multi-player, (D) stands for deciding whether such an equilibrium even exists. In single-player, Optimal (D) decides whether some target utility can be achieved. Citations are given for results found in the literature. All of the hardness results even hold for highly restricted game instances, such as, *e.g.*, for games with no chance nodes or two-player zero-sum games where one player has perfect recall. *: The number of actions per infoset is required to be constant for the membership result. '—': No results exist for these settings to our knowledge. Also note the technical complication that arises here from the fact that there exist single-player games in which every exact EDT or CDT equilibrium involves irrational values [Tewolde *et al.*, 2023].

knew before. Finally, a single agent can be instantiated multiple times in the same environment, where one copy does not know what another copy just knew [Conitzer and Oesterheld, 2023]. For example, we might want to test goal-oriented AI agents in simulation to ensure that they will later act in a trustworthy fashion in the real world [Kovařík *et al.*, 2023; Kovařík *et al.*, 2024]. Then, the AI agent will have to act in the real world without knowing how it acted in simulation.

Perfect recall is a common technical assumption in game theory because it implies many simplifying properties, such as polynomial-time solvability of single-player and two-player zero-sum settings [Koller and Megiddo, 1992]. In multi-player settings with *imperfect* recall, Nash equilibria may not exist anymore [Wichardt, 2008]; in fact, we show that deciding existence is computationally hard. To give an illustrative running example, consider a variation of Wichardt's game in Figure 1a, which we call the forgetful (soccer) penalty shoot-out. The shooter (P1) decides whether to shoot left or right, once before the whistle, and once again right before kicking the ball. At the second decision point, P1 has forgotten which direction they chose previously. P1 only succeeds in shooting in any direction if she chooses that direction at both decision points. Upon succeeding, it becomes a matching pennies game with the goalkeeper (P2) who chooses to jump left or right to block the ball. A similar analysis to the one of matching pennies implies that in a potential Nash equilibrium, none of the two players can play one side more often than the other. However, both players randomizing $50/50$ at each infoset is not a Nash equilibrium either: P1 is not best responding to P2 because she could instead deterministically shoot towards one side to avoid miscoordination with herself altogether which would achieve a payoff of 1 instead of 0.

Indeed, many of our intuitions fail for imperfect-recall games – to the point that a significant body of work in philosophy and game theory addresses conceptual questions about probabilistic reasoning and decision making in imperfect-recall games, such as in the Sleeping Beauty problem [Elga, 2000] or the absentminded driver game of Figure 1b [Piccione and Rubinstein, 1997]. From this literature, several distinct and coherent ways to approach games of imperfect recall have emerged. We will discuss these in detail in Section 4.

In this paper, we study the computational complexity of solving imperfect-recall extensive-form games. We focus on three solution concepts: (1) Nash equilibria where players play mutual best response strategies (or simply optimal strategies in single-player domains), (2) multiselves equilibria based on evidential decision theory, in which each infoset plays a best-response action to all other infosets and players, and (3) multiselves equilibria based on causal decision theory, in which each infoset plays a *Karush-Kuhn-Tucker (KKT)* point action for the current strategy profile. The latter two are relaxations of the first. Sections 2 and 4 cover preliminaries on imperfect-recall games and on multiselves equilibria, respectively. Sections 3 and 5 analyze the computation of Nash equilibria and of multiselves equilibria, respectively, in various setting. Our complexity results for these are summarized in Table 1. Last but not least, Section 6 shows that games with imperfect recall stay computationally equally hard even in the absence of exogenous stochasticity (*i.e.*, chance nodes).

2 Imperfect-Recall Games

We first define extensive-form games, allowing for imperfect recall. The concepts we use in doing so are standard; for more detail and background, see, *e.g.*, Fudenberg and Tirole [1991] and Piccione and Rubinstein [1997]. In this section, we follow the exposition of Tewolde *et al.* [2023], with the addition of introducing multi-player notation.

Definition 1. *An extensive-form* game with imperfect recall*, denoted by* Γ*, consists of:*

1. *A rooted tree, with nodes* $\mathcal{H}$ *and where the edges are labeled with* actions*. The game starts at the root node* h_0 *and finishes at a leaf node, also called* terminal node*. We denote the terminal nodes in* $\mathcal{H}$ *as* $\mathcal{Z}$ *and the set of actions available at a nonterminal node* $h \in \mathcal{H} \setminus \mathcal{Z}$ *as* A_h*.*
2. *A set of* $N+1$ *players* $\mathcal{N} \cup \{c\}$*, for* $N \in \mathbb{N}$*, and an assignment of nonterminal nodes to a player that shall choose an action at that node. Player* c *stands for* chance *and represents exogenous stochasticity that chooses an action. With* $\mathcal{H}^{(i)}$ *we denote all nodes associated to player* $i \in \mathcal{N}$*.*
3. *A fixed distribution* $\mathbb{P}^{(c)}(\cdot \mid h)$ *over* A_h *for each chance node* $h \in \mathcal{H}^{(c)}$*, with which an action is determined at* h*.*
4. *For each* $i \in \mathcal{N}$*, a* utility function $u^{(i)} : \mathcal{Z} \to \mathbb{R}$ *that specifies the payoff that player* i *receives from finishing the game at a terminal node.*

5. *For each $i \in \mathcal{N}$, a partition $\mathcal{H}^{(i)} = \sqcup_{I \in \mathcal{I}^{(i)}} I$ of player i's decision nodes into information sets (*infosets*). We require $A_h = A_{h'}$ for all nodes h, h' of the same infoset. Therefore, infoset I has a well-defined action set A_I.*

Imperfect Recall. Nodes of the same infoset are assumed to be indistinguishable to the player during the game even though the player is always aware of the full game structure. This may happen even in perfect-recall games due to *imperfect information*, that is, when it is unobservable to the player what another player (or chance) has played. This effect is present in Figure 1a for P2. In contrast, infoset I_2 of P1 exhibits *imperfect recall* because once arriving there, the player has forgotten information about the history of play that she once held when leaving I_1, namely whether she chose left or right back then. In Figure 1b, the player is unable to recall whether she has been in the same situation before or not. This phenomenon is a special kind of imperfect recall called *absentmindedness*. The *degree of absentmindedness* of an infoset shall be defined as the maximum number of nodes of the same game trajectory that belong to that infoset. In Figure 1b, it is 3. The *branching factor* of a game is the maximum number of actions at any infoset.

In contrast to that, games with *perfect* recall have every infoset reflect that the player remembers the sequence of infosets she visited and the actions she took. We note that any node $h \in \mathcal{H}$ uniquely corresponds to a history path hist(h) in the game tree, consisting of alternating nodes and actions from root h_0 to h. Let $\exp^{(i)}(h)$ be the experienced sequence of infosets visited and actions taken by player i on the path hist(h). Then, formally, a game has perfect recall if for all players $i \in \mathcal{N}$, all infosets $I \in \mathcal{I}^{(i)}$, and all nodes $h, h' \in I$, we have $\exp^{(i)}(h) = \exp^{(i)}(h')$.

Strategies. Let $\Delta(A_I)$ denote the set of probability distributions over the actions in A_I. These will also be referred to as *randomized actions*. A (behavioral) *strategy* $\mu^{(i)} : \mathcal{I}^{(i)} \to \sqcup_{I \in \mathcal{I}^{(i)}} \Delta(A_I)$ of a strategic player i assigns to each of her infosets I a probability distribution $\mu^{(i)}(\cdot \mid I) \in \Delta(A_I)$. Upon reaching I, the player draws an action randomly from $\mu^{(i)}(\cdot \mid I)$. A *pure* strategy maps deterministically[1] to $\sqcup_{I \in \mathcal{I}^{(i)}} A_I$. A strategy profile, or *profile*, $\mu = (\mu^{(i)})_{i \in \mathcal{N}}$ specifies a behavioral strategy for each player. We may write $\left(\mu^{(i)}, \mu^{(-i)}\right)$ to emphasize the influence of $i \in \mathcal{N}$ on μ. Denote the strategy set of player $i \in \mathcal{N}$ with $\mathcal{S}^{(i)}$, and the set of profiles with $\mathcal{S}$.

For a computational analysis, we identify a randomized action set $\Delta(A_I)$ with the simplex $\Delta^{|A_I|-1}$, where $\Delta^{n-1} := \{x \in \mathbb{R}^n \, : \, x_k \geq 0 \, \forall k \, , \sum_{k=1}^n x_k = 1\}$. Therefore, the strategy sets are Cartesian products of simplices:

$$\mathcal{S} \equiv \times_{i \in \mathcal{N}} \times_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1} \text{ and } \mathcal{S}^{(i)} \equiv \times_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1}.$$

Reach Probabilities and Utilities. Let $\mathbb{P}(\bar{h} \mid \mu, h)$ be the probability of reaching node $\bar{h} \in \mathcal{H}$ given that the current game state is at $h \in \mathcal{H}$ and that the players are playing profile μ. It evaluates as 0 if $h \notin \text{hist}(\bar{h})$, and as the product of probabilities of the actions on the path from h to $\bar{h}$ otherwise. The expected utility payoff of player $i \in \mathcal{N}$ at node $h \in \mathcal{H} \setminus \mathcal{Z}$ if profile μ is being followed henceforth is $U^{(i)}(\mu \mid h) := \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid \mu, h) \cdot u^{(i)}(z)$. We overload notation by defining $\mathbb{P}(h \mid \mu) := \mathbb{P}(h \mid \mu, h_0)$ for root h_0 of Γ, and by defining the function $U^{(i)}$ as $U^{(i)}(\mu) := U^{(i)}(\mu \mid h_0)$, mapping a profile μ to its expected utility from game start. In Figure 1b, this is $U^{(1)}(\mu) = 6c^2e$ – or, to follow our notation more precisely, $U^{(1)}(\mu) = 6\mu^{(1)}(c \mid I)^2 \mu^{(1)}(e \mid I)$.

Polynomials. Each summand $\mathbb{P}(z \mid \mu, h) \cdot u^{(i)}(z)$ in $U^{(i)}(\mu \mid h)$ is a monomial in μ times a scalar, and the expected utility function $U^{(i)}$ is a polynomial function in the profile μ. All these polynomials $U^{(i)}$ can be constructed in polynomial time (polytime) in the encoding size of Γ.

One might also ask how general those polynomial utility functions may be. Indeed, imperfect-recall games can be very expressive. We give a polytime construction in Appendix A.4 that, given a collection of N multivariate polynomials $p^{(i)} : \times_{i=1}^N \times_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}} \to \mathbb{R}$, yields an associated N-player game Γ with imperfect recall such that its expected utility functions satisfy $U^{(i)}(\mu) = p^{(i)}(\mu)$ on $\times_{i=1}^N \times_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}}$.

Approximate Solutions. The solution concepts we investigate will have a definition of the abstract form "Strategy μ is a *solution* if for all $y \in Y$ we have $f(\mu) \geq f_\mu(y)$" for some set Y of alternatives and some utility/objective functions f and f_μ. Then, we call a strategy μ an ϵ-solution if $\forall \, y \in Y : f(\mu) \geq f_\mu(y) - \epsilon$.

Computational Considerations. In this paper, we discuss *decision problems* and *search problems*. The former ask for a yes/no answer; the latter ask for a solution point. The input to these computational problems may be a game Γ, a precision parameter $\epsilon > 0$, and/or a target value t. Values in Γ, as well as ϵ and t are assumed to be rational. We assume that a game Γ is represented by its game tree structure, which has size $\Theta(|\mathcal{H}|)$, and by a binary encoding of its chance node probabilities and its utility payoffs. If there is a target t, then it shall be given in binary as well.

If there is no precision parameter ϵ, then we are dealing with problems involving *exact* solutions. In our settings, such problems are usually beyond NP because equilibria may require irrational probabilities and may therefore not be representable in finite bit length. In fact, Tewolde *et al.* [2023][Figure 6] give a simple single-player example in which the unique equilibrium takes on irrational values. That is, in part, why we will also be interested in approximations up to a small precision error $\epsilon > 0$. Here, we mean 'small' relative to the range of utility payoffs, which – by shifting and rescaling utilities – we can w.l.o.g. assume to be $[0, 1]$.

[1]Other work has also considered *mixed* strategies, that is, probability distributions over all pure strategies. In the presence of imperfect recall, mixed strategies are not realization-equivalent to behavioral strategies [Kuhn, 1953]. Mixed strategies require the agent to coordinate her actions across infosets (*e.g.*, access to a correlation device): For example, in contrast to our introductory discussion on the forgetful penalty shoot-out (Figure 1a), this game does admit a Nash equilibrium *in mixed strategies* since P1 can now choose to kick left twice in a row 50% of the time and to kick right twice in a row the other 50% of the time. As this would imply a form of memory, it does not fit the motivation of this paper.

Remark. *By default, $\epsilon > 0$ will be given in binary, in which case we require* inverse-exponential *(*1/exp*) precision.*

Here, the term 'inverse-exponential' indicates that $1/\epsilon$ can be exponentially larger than the tree size $|\mathcal{H}|$. Occasionally, we may instead require *inverse-polynomial* (1/poly) precision, which is when ϵ is given in unary, or require constant precision, which is when ϵ is fixed to a constant > 0. Naturally, 1/exp precision is hardest to achieve.

Complexity Classes. We give a brief overview of the complexity classes appearing in this paper, and refer to Appendix A.5 for references and more details. The subset relationships of the complexities classes we present here are believed to be strict. P describes the decision problems that can be solved in polytime. NP describes the decision problems that can be solved in non-deterministic polytime. Σ_2^{P} describes the decision problems that can be solved in non-deterministic polytime if given oracle access to an NP solver, such as a SAT oracle. We have $\mathsf{P} \subseteq \mathsf{NP} \subseteq \Sigma_2^{\mathsf{P}} \subseteq \mathsf{PSPACE}$. NP and Σ_2^{P} are classes for decision problems that can be formulated as one over discrete variables (w.l.o.g. Boolean variables). Their counterparts for real-valued decision problems are the *first-order-of-the-reals* classes $\exists\mathbb{R}$ and $\exists\forall\mathbb{R}$: A $\exists\mathbb{R}$ problem asks whether a sentence of the form $\exists x_1 \dots \exists x_n F(x_1, \dots, x_n)$ is true, where the x_i represent real-valued variables and F represents a quantifier-free formula of (in-)equalities of real polynomials in rational coefficients. $\exists\forall\mathbb{R}$ is defined analogously, except for sentences of the form $\exists x \in \mathbb{R}^{n_1} \forall y \in \mathbb{R}^{n_2} F(x, y)$. We have $\mathsf{NP} \subseteq \exists\mathbb{R} \subseteq \mathsf{PSPACE} \cap \exists\forall\mathbb{R}$.

The complexity classes FP and FNP are the search problem analogues of P and NP, and as such, essentially have the same complexity. The landscape between FP and FNP, however, is rich. Total NP search problems are those problems in FNP for which one knows that each problem instance admits a solution. The complexity classes in it can be characterized by the natural, but exponential-time method with which one can show that each problem instance admits a solution. For the class PPAD the method is that of a fixed point argument, as is the case, *e.g.*, for the existence of a Nash equilibrium. For the class PLS the method is that of a local optimization argument on a directed acyclic graph. For the class PLS the method is that of a CLS a local optimization argument on a bounded polyhedral (continuous) domain. We have $\mathsf{FP} \subseteq \mathsf{CLS} = \mathsf{PPAD} \cap \mathsf{PLS}$ and $\mathsf{PPAD}, \mathsf{PLS} \subseteq \mathsf{FNP}$.

3 Nash Equilibria and Optimal Play

In this section, we present our computational results for the classic and most important solution concept in game theory – the Nash equilibrium [Nash, 1950].

Definition 2. *A profile μ is said to be a* Nash equilibrium *(in behavioral strategies) for game Γ if for all player $i \in \mathcal{N}$, and all alternative strategies $\pi^{(i)} \in \mathcal{S}^{(i)}$, we have*

$$U^{(i)}(\mu^{(i)}, \mu^{(-i)}) \geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}).$$

In a Nash equilibrium, no player has any utility incentives to deviate unilaterally to another strategy. Nash showed that any finite perfect-recall game admits at least one Nash equilibrium. In contrast, some finite imperfect-recall games have no Nash equilibrium, as discussed in the introduction. If there is only a single player, however, finding a Nash equilibrium – *i.e.*, finding an *optimal* strategy – reduces to maximizing a polynomial utility function over a compact strategy space. Such a solution is guaranteed to exist, and its value is unique. Therefore, one may ask instead whether some target value t can be achieved in a given game. In Figure 1b, this would result in the $\exists\mathbb{R}$-sentence $\exists e, c : 6c^2 e \geq t \wedge c \geq 0 \wedge e \geq 0 \wedge c + e = 1$. This is an easier task than *finding* an optimal strategy. Nonetheless, we have:

Proposition 3 (Gimbert *et al.*, 2020)**.** *Deciding whether a single-player game with imperfect recall admits a strategy with value $\geq t$ is $\exists\mathbb{R}$-complete.*

For approximation, consider problem OPT-D that asks to distinguish between whether $\exists \mu \in S : U^{(1)}(\mu) \geq t$ and whether $\forall \mu \in S : U^{(1)}(\mu) \leq t - \epsilon$.

Proposition 4 (Koller and Megiddo, 1992; Tewolde *et al.*, 2023)**.** OPT-D *is* NP*-complete.*

Technically, Koller and Megiddo establish hardness for the *exact* decision problem. We shall merely add the observation that their proof also implies NP-hardness of the approximate problem; and via the PCP theorem [Håstad, 2001], even for a constant precision $\epsilon < 1/8$.

3.1 Two-Player Zero-Sum Games

A *two-player zero-sum (2p0s)* game is a two-player game where $U^{(2)} = -U^{(1)}$. In that case utilities can be given in terms of P1, and P2 simply minimizes that utility.

Koller and Megiddo [1992] prove Σ_2^{P}-completeness of deciding in 2p0s games with imperfect recall whether the max-min value in pure-strategy play exceeds some utility target $\geq t$. We will consider behavioral strategies instead.

Definition 5. *In a 2p0s game Γ, the (behavioral)* max-min value *and* min-max value *are defined as*

$$\underline{U} := \max_{\mu^{(1)} \in S^{(1)}} \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}),$$
$$\bar{U} := \min_{\mu^{(2)} \in S^{(2)}} \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}).$$

Gimbert *et al.* [2020] prove that deciding $\underline{U} \geq t$ is in $\exists\forall\mathbb{R}$ and is $\exists\mathbb{R}$-hard. For approximation, we know the following.

Lemma 6 (Zhang *et al.*, 2023)**.** *It is Σ_2^{P}-complete to distinguish $\underline{U} \geq 0$ from $\underline{U} \leq -\epsilon$ in 2p0s games with imperfect recall. Hardness holds even with no absentmindedness and* 1/poly *precision.*

To leverage this result in the subsequent sections, we will first show a tight connection between the existence of Nash equilibria in a 2p0s game Γ, and Γ's min-max and max-min values. Define the *duality gap* of Γ as the difference

$$\Delta := \bar{U} - \underline{U} \geq 0.$$

In Figure 1a the duality gap is $1 - 0 = 1$.

Proposition 7. *Let Γ be a 2p0s game with imperfect recall. If $\Delta \leq \epsilon$ then Γ admits an ϵ-Nash equilibrium. Conversely, if Γ admits an ϵ-Nash equilibrium, then $\Delta \leq 2\epsilon$.*

In particular, there is an equivalence between Nash equilibrium existence and vanishing duality gap. This result is not specific to behavioral strategies in imperfect-recall games; it holds for any family of strategies in any 2p0s game.

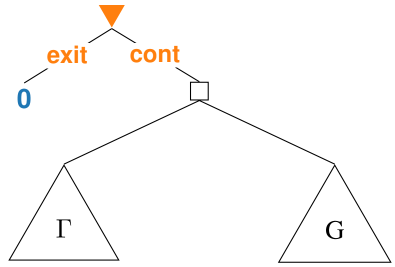


Figure 2: Game construction used to prove hardness of deciding equilibrium existence. We use boxes for chance nodes, at which chance plays uniformly at random. Γ is a placeholder game. G is a game with no equilibrium; Section 3.2 for example uses Figure 1a.

3.2 Deciding Nash Equilibrium Existence

We observe that the existence of a Nash equilibrium can be formulated as "there exists a profile μ such that for all other profiles π the condition of Definition 2 are satisfied for all $i \in \mathcal{N}$". This puts the exact and approximate decision problems in $\exists\forall\mathbb{R}$ and Σ_2^{P} respectively. For an intuitive idea of our upcoming hardness results, consider the game in Figure 2 where subgame G shall be that of Figure 1a and where subgame Γ is a game in which it is hard to decide what utility P1 can guarantee himself. Then a profile cannot be a Nash equilibrium if P2 is supposed to continue at the root node, because in that case G is reached with positive probability and the players cannot be in equilibrium in that subgame as we have discussed in the introduction. Note that exiting at the root node yields P2 a utility of 0, and best-responding to P1 in subgame G also yields P2 a utility of ≤ 0 (recall that P2 is the minimizer). Thus, for a profile to be a Nash equilibrium in the overall game, P2 must exit at the root node as a best response, which is the case exactly if P1 cannot achieve a utility of at least 0 in the subgame Γ. Using the problem instances of Proposition 3 for the subgame Γ, we obtain

Theorem 1. *Deciding if a game with imperfect recall admits a Nash equilibrium is $\exists\mathbb{R}$-hard and in $\exists\forall\mathbb{R}$. Hardness holds even for 2p0s games where one player has a degree of absentmindedness of* 4 *and the other player has perfect recall.*

Next, for the approximate case, we use the problem instances of Lemma 6 for the subgame Γ. Define NASH-D to ask to distinguish between whether an exact Nash equilibrium exists or whether no ϵ-Nash equilibrium exists.

Theorem 2. NASH-D *is Σ_2^{P}-complete. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

With Proposition 7, this immediately implies

Corollary 8. *It is Σ_2^{P}-complete to distinguish $\Delta = 0$ from $\Delta \geq \epsilon$ in 2p0s games. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

Later in this paper, Theorem 4 will imply another Σ_2^{P}-hardness for NASH-D but with different restrictions.

3.3 A Naïve Algorithm for Nash Equilibria

For game Γ, let $|\Gamma|$ denote its representation size and $m := \sum_{i \in \mathcal{N}} \sum_{I \in \mathcal{I}^{(i)}} |A_I|$ its the total number of pure actions.

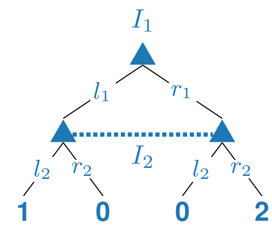


Figure 3: A single-player game with imperfect recall where miscoordinating actions with yourself is punished most.

Proposition 9. NASH-D *is solvable in time* $\mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\Big)$.

In fact, our algorithm *finds* an ϵ-Nash equilibrium whenever an exact Nash equilibrium exists. The idea is similar to that one of Lipton and Markakis [2004][Theorem 2] for multi-player normal-form games: Namely, we iteratively subdivide the strategy space, and repeatedly decide with first-order-of-the-reals solvers whether a Nash equilibrium exists in this smaller region. Those solvers also give rise to the exponential time dependence on m. In particular, the algorithm becomes polytime if m is bounded by a constant. This observation will aid us towards a PLS-membership proof in Theorem 5. Also note that such a bound on m will not restrict the size of the game tree since the degree of absentmindedness can still grow arbitrarily (cf. Figure 1b).

4 Introducing Multiselves Equilibria

Section 3 shows strong obstacles to finding Nash equilibria in games with imperfect recall. In light of these limitations, we relax the space of solutions and turn to the *multiselves* approach (cf. the agent-form [Kuhn, 1953]), which we review in this section. This approach argues that, whenever a player finds herself in an infoset, she has no influence over which actions she chooses at other infosets. Therefore, at a multiselves equilibrium μ, each player will play the best randomized action at each of their infosets, assuming that they themselves play according to μ at other infosets and assuming all other players also play according to μ.

Consider Figure 3. The optimal strategy is to play (r_1, r_2). This is also a multiselves equilibrium. However, (l_1, l_2) is also a multiselves equilibrium, because if the player is at the top-level infoset I_1 and assumes that she will follow left at the bottom-level infoset I_2, then it is best for her to go left now. On the other hand, if the player is at I_2 and assumes that she played left at I_1, then it is again best for her to play left now.

Multiselves equilibria can be arbitrarily bad in payoff in comparison to optimal strategies and Nash equilibria, as can be seen by replacing the payoff of 2 in Figure 3 with some $\lambda \to \infty$. This phenomenon is due to miscoordination across infosets, and it arises in the same manner across teams in team games: The corresponding normal-form game $\begin{pmatrix} \lambda, \lambda & 0, 0 \\ 0, 0 & 1, 1 \end{pmatrix}$ shows that Nash equilibria can be arbitrarily worse relative to Pareto-optimal profiles.

In games with absentmindedness it becomes controversial how to apply the multiselves idea. Specifically, how should a player reason about implications of a choice at the current decision point for her action choices at past and future decision points *within the same infoset*, and – as a consequence – compute incentives to deviate? That is, in considering deviating, will the player assume they would perform the same deviation at other nodes in the same infoset, or that the deviation is a one-time-only event? We will handle this question using two well-motivated[2] decision theories that correspond to these two cases: evidential decision theory and causal decision theory. We will see that Nash equilibria are multiselves equilibria under both decision theories.

This section is accompanied with an extensive Appendix C that – beyond proving the statements made in this section – also introduces some additional observations and lemmas needed for the development of our main results.

4.1 Evidential Decision Theory (EDT)

Suppose a game Γ is played with profile μ, and a player i arrives in one of her infosets $I \in \mathcal{I}^{(i)}$. EDT postulates that if that player deviates to a randomized action $\alpha \in \Delta(A_I)$ at the current node, then she will have also deviated to α whenever she arrived in I in the past, and that she will also deviate to α whenever she arrives in I again in the future. This is because EDT argues that the choice to play α now is evidence for the player playing the same α in the past and future.

We denote the behavioral strategy that results from an EDT deviation as $\mu^{(i)}_{I \mapsto \alpha}$. It plays according to $\mu^{(i)}$ at every infoset except for at $I \in \mathcal{I}^{(i)}$ where it plays according to $\alpha \in \Delta(A_I)$.

Definition 10. *We call μ an* EDT equilibrium *for game Γ if for all players $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all randomized actions $\alpha \in \Delta(A_I)$, we have*

$$U^{(i)}(\mu) \geq U^{(i)}(\mu^{(i)}_{I \mapsto \alpha}, \mu^{(-i)}).$$

In an EDT equilibrium, no player has an incentive to deviate at an infoset in an EDT fashion to another randomized action. This is because the right hand side of the inequality represents the expected *ex-ante* utility of such an EDT deviation. Section 4.4 gives an extensive discussion on the ex-ante perspective for multiselves equilibria. Regarding equilibrium computation, the following result is known:

Proposition 11 (Tewolde *et al.*, 2023)**.** *Unless* NP = ZPP*, finding an ϵ-EDT equilibrium in a single-player game for* 1/poly *precision is not in* P.

4.2 Causal Decision Theory (CDT)

Say, again, game Γ is played with profile μ, and a player i arrives in one of her infosets $I \in \mathcal{I}^{(i)}$. Then CDT postulates that the player can deviate to an action $\alpha \in \Delta(A_I)$ at the current node without violating that she has been playing according to $\mu^{(i)}$ at past arrivals in I, or that she will be playing according to $\mu^{(i)}$ at future arrivals in I. This is in addition to assuming that all other players follow $\mu^{(-i)}$ as usual. The intuition behind CDT is that the player's choice to deviate from $\mu^{(i)}$ at the current node does not *cause* any change in her behavior at any other node of the same infoset I.

Example 12. *Recall Figure 1b in which – as the story goes – the absentminded driver has to exit a highway at the second highway exit to find home. Say the player enters the game with μ = 'e' (exit), and upon arriving in the infoset, considers deviating to 'c' (continue) at this point of time. EDT then argues that the player will always continue on the highway and arrive at the third "0" payoff of the game. CDT, on the other hand, argues that the player will continue on the highway once – or more precisely, continue at the root node since that is the only decision node she could possibly be at given her strategy μ – and then exit the highway at its second exit.*

For node $h \in \mathcal{H}^{(i)}$ and pure action $a \in A_h$, let ha denote the child node reached if player i plays a at h. Consequently, $U^{(i)}(\mu \mid ha)$ is the expected utility of player i from being at h, playing a, and everyone following profile μ afterwards. When at an infoset $I \in \mathcal{I}^{(i)}$, the player does not know at which node of I she currently is. Therefore, when computing her utility incentives for a CDT-style deviation to a, she scales each node by the probability of reaching that node under profile μ. This yields utility incentives

$$\textstyle\sum_{h \in I} \mathbb{P}(h \mid \mu) \cdot U^{(i)}(\mu \mid ha).$$

to CDT-deviate to pure action a at infoset I. This value is known to be equal to the partial derivative $\nabla_{I,a} U^{(i)}(\mu)$ of utility function $U^{(i)}$ w.r.t. to action a of $I \in \mathcal{I}^{(i)}$ at point μ [Piccione and Rubinstein, 1997; Oesterheld and Conitzer, 2022]. Hence, we can formulate the following definition.

Definition 13. *Given a profile μ in game Γ, a player $i \in \mathcal{N}$ determines her (ex-ante) utility from CDT-deviating at infoset $I \in \mathcal{I}^{(i)}$ to randomized action $\alpha \in \Delta(A_I)$ as*

$$U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I) := U^{(i)}(\mu) + \textstyle\sum_{a \in A_I} (\alpha(a) - \mu(a \mid I)) \cdot \nabla_{I,a} U^{(i)}(\mu).$$

In other words, this is the first-order Taylor approximation of $U^{(i)}$ at μ in the subspace $\Delta(A_I)$. In Figure 8 of Appendix C, we illustrate on a simple game that the ex-ante CDT-utility – as a first-order approximation – may yield unreasonable utility payoffs for values α far away from $\mu(\cdot \mid I)$. Moreover, if $\alpha \neq \mu(\cdot \mid I)$, we observe that the resulting behavior of the deviating player cannot be captured by a *behavioral strategy* anymore that the player could have chosen from the beginning. That is because the player is now acting differently at different nodes of the same infoset.

Definition 14. *A profile μ is said to be a* CDT equilibrium *for game Γ if for all player $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all alternative randomized actions $\alpha \in \Delta(A_I)$, we have*

$$U^{(i)}(\mu) = U^{(i)}_{\text{CDT}}\big(\mu^{(i)}(\cdot \mid I) \bigm| \mu, I\big) \geq U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I).$$

Therefore, no player has any utility incentives to deviate at an infoset in a CDT fashion to another randomized action. CDT equilibria have received a more thorough treatment in the literature than EDT equilibria have.

[2] The debate around decision theories is related to the approach for belief formation (cf. the Sleeping Beauty problem [Elga, 2000]). Among other aspects, the literature has studied which combination of decision theories and belief formation avoid being Dutch-booked (money-pumped) [Piccione and Rubinstein, 1997; Briggs, 2010; Oesterheld and Conitzer, 2022].

Lemma 15 (Lambert *et al.*, 2019). *Any game Γ with imperfect recall admits a CDT equilibrium.*

Thus, we shall define CDT-S as the search problem that asks for an ϵ-CDT equilibrium in the game (which always exists). Let 1P-CDT-S be its restriction to single-player games.

Proposition 16 (Tewolde *et al.*, 2023).

1. *A profile μ is a CDT equilibrium of Γ if and only if for all player $i \in \mathcal{N}$, strategy $\mu^{(i)}$ is a KKT-point of*
$$\max_{\pi^{(i)} \in S^{(i)}} U^{(i)}(\pi^{(i)}, \mu^{(-i)}).$$
2. *The problem* 1P-CDT-S *is* CLS*-complete.*

The original formulation of Tewolde *et al.* was not given for the multi-player setting and the ex-ante perspective. The advantages of the latter are discussed in Section 4.4. Furthermore, we may also highlight a positive algorithmic implication which has not been stated before. It can be obtained analogously to [Fearnley *et al.*, 2023, Lemma C.4].

Corollary 17. 1P-CDT-S *for* 1/poly *precision is in* P.

4.3 Comparing the Solution Concepts

The three solution concepts form an inclusion hierarchy. This result is known for single-player settings and extends straightforwardly to multi-player settings.

Proposition 18 (Oesterheld and Conitzer, 2022). *A Nash equilibrium is an EDT equilibrium. An EDT equilibrium is a CDT equilibrium.*

This also implies that any single-player game admits both EDT and CDT equilibria since it admits an optimal strategy (= Nash equilibrium). In general, neither statement in Proposition 18 holds in reverse. Indeed, we have seen in Figure 3 that multiselves equilibria may not be the optimal strategy. Moreover, the strategy μ described in Example 12 forms a CDT equilibrium but not an EDT equilibrium (an EDT deviation to a uniformly randomized action achieves positive utility).

We will find in this paper that CDT equilibria are easier to compute than EDT equilibria. Indeed, Proposition 11 and Corollary 17 already serve as the first evidence towards such a separation. We can also find a hint towards such an insight by considering the easier problem of *verifying* whether a given profile could be an equilibrium. For CDT, this can be done in polytime: since $U^{(i)}_{\text{CDT}}$ is linear in α, we do not actually need to check Definition 14 for all $\alpha \in \Delta(A_I)$, but it suffices to only check it for *pure* actions $a \in A_I$. For EDT equilibria, on the other hand, there is no simple-to-check characterization: $U^{(i)}(\mu^{(i)}_{I\mapsto\cdot}, \mu^{(-i)})$ is a polynomial function over $\Delta(A_I)$, for which no easy verification method is known. At least, this is true in general. As for special cases, we have:

Remark 19. *Without absentmindedness, deviation incentives of EDT and of CDT coincide, and so do the equilibrium concepts. Hence, complexity results such as Proposition 16 and Theorem 6 will apply to EDT equilibria as well.*

Remark 20. *If each player has only one infoset in total, then the EDT equilibria coincide with the Nash equilibria.*

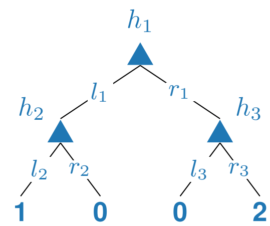


Figure 4: Differences of the ex-ante and de-se utility perspective explained on a perfect-recall variant of Figure 3. Again, the only optimal strategy takes the path $r_1 - r_3$. But what action can you choose at h_2?

4.4 On Utility Perspectives

Let us discuss the *ex-ante* and *de-se* perspectives on utilities, and why we chose the former. Consider the game in Figure 4, which is the *perfect-recall* version of Figure 3, and consider the strategy $\mu = (\epsilon l_1 + (1-\epsilon) r_1, r_2, r_3)$ for some small $\epsilon > 0$. In this case, how much does it matter what randomized action the player chooses at node h_2? In the *de-se* perspective, the player calculates her expected deviation gains for her current situation onwards. In our example, she would calculate an incentive of 1 to deviate to l_2 assuming she is already at h_2. In the *ex-ante* perspective, the player calculates her expected deviation gains on the ex-ante utility (from before the game started). In our example, she would calculate an incentive of ϵ to deviate to l_2 at h_2 since that node is rarely visited anyways.

Previous work in the literature has considered agents that maximize their de-se utilities, as in Strotz [1955] with the strategy of consistent planning or in Piccione and Rubinstein [1997]. This might fit well for human agents who are interested in the impacts of their actions on their *current* self. In this paper, however, we argue that for AI and team agents, the ex-ante perspective is more suitable. Indeed, such an agent should ground its optimization in the impact its actions has on the overall ex-ante utility; despite imperfect recall limiting the agent's decision or commitment powers to the current infoset (EDT) or decision node (CDT).

There are also technical advantages supporting the ex-ante perspective. At infosets that are never reached, the action choices do not affect the ex-ante utility. Under de-se reasoning, however, the agent would have to generate beliefs on the impossible event of being at that infoset. In order to make such beliefs well-defined, one has to pick one of many possible options for equilibrium refinement. In optimization terms, the de-se utility functions are fractions of polynomials with possible singularities on the boundary of the strategy set due to vanishing denominators. Tewolde *et al.*[2023, Theorem 2] circumvents this issue in their formulation of Proposition 16 by only considering games that come with universal lower bounds on the positive reach probabilities of all infosets. Unfortunately, many (simple) games such as Figure 4 do not satisfy this property.

5 Complexities of Multiselves Equilibria

In this section, we present our computational results for multiselves equilibria.

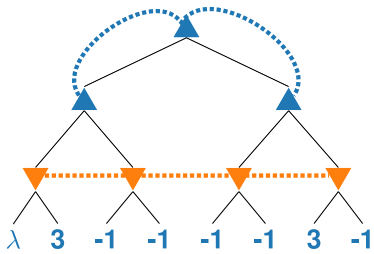


Figure 5: A variant of Figure 1a where P1 has one single infoset with absentmindedness. It is parametrized by the payoff $\lambda \in \mathbb{R}$ from P1 shooting left and P2 blocking left.

5.1 EDT Equilibria

Consider the (parametrized) *absentminded* penalty shoot-out in Figure 5. It shows that in multi-player settings, EDT equilibria may not exist. Absentmindedness is crucial for such an example due to Remark 19 and Lemma 15.

Lemma 21. *Figure 5 has an EDT equilibrium if and only if* $\lambda \geq 3$.

The next result establishes $\exists\mathbb{R}$-hardness again by similar arguments to Theorem 1. Except in this construction, we attach the single-player game Γ from Proposition 3 to the bottom left of Figure 5. Note here that by an appropriate payoff shift in Γ, we can w.l.o.g. assume the target t for Γ to be 3.

Theorem 3. *Deciding whether a game with imperfect recall admits an EDT equilibrium is* $\exists\mathbb{R}$*-hard and in* $\exists\forall\mathbb{R}$*. Hardness holds even for 2p0s games where one player has a degree of absentmindedness of* 4 *and the other player has perfect recall.*

Now consider problem EDT-D that asks to distinguish between whether an exact EDT equilibrium exists or whether no ϵ-EDT equilibrium exists.

Theorem 4. EDT-D *is* Σ_2^{P}*-complete. Hardness holds for* 1/poly *precision and 2p0s games with one infoset per player and a degree of absentmindedness of* 4.

The technically involved proof casts the game construction for Theorem 1 to a game where each player only has one infoset, in order to use Remark 20. For that, we cannot reduce from Lemma 6 this time, but we reduce directly from the Σ_2^{P}-complete problem $\exists\forall$3-DNF-SAT [Stockmeyer, 1976]. Moreover, we make use of the flexibility that EDT-utilities can represent arbitrary polynomial functions as long as they are only over a single simplex.

Next, we turn to the search problem. The algorithm of Proposition 9 can also find ϵ-EDT equilibria if we adjust for its equilibrium conditions. In single-player settings, however, we can do better since EDT equilibria are guaranteed to exist. Let 1P-EDT-S be the search problem that asks for an ϵ-EDT equilibrium. This problem was left open by Tewolde *et al.* [2023].

Theorem 5. 1P-EDT-S *is* PLS*-complete when the branching factor is constant. Hardness holds even when the branching factor and the degree of absentmindedness are* 2.

Before we touch on the proof idea, we shall highlight its contrast to Proposition 16 on the CLS-membership of 1P-CDT-S, since CLS is believed to be a proper subset of PLS (evidenced by conditional separations as discussed in Appendix A.5). Furthermore, we also get:

Corollary 22. 1P-EDT-S *for* 1/poly *precision is in* P *when the branching factor is constant.*

The proofs first establish that 1P-EDT-S is computationally equivalent to the search problem that takes a polynomial function p over a product of simplices, and asks for an approximate "Nash equilibrium point" of it. In the special case where the branching factor is 2, the domain becomes the hypercube $[0,1]^\ell$, and an ϵ-Nash equilibrium x is characterized by the property

$$\forall j \in [\ell]\, \forall y \in [0,1]: \quad p(x) \geq p(y, x_{-j}) - \epsilon.$$

We show that this problem is PLS-complete. This result may be of independent interest for the optimization literature.

The PLS-hardness follows from a reduction from the PLS-complete problem MaxCut/Flip [Schäffer and Yannakakis, 1991; Yannakakis, 2003]. For the positive algorithmic results of PLS and P membership respectively, we show that ϵ-best-response dynamics converges to an ϵ-EDT equilibrium. We run a similar method to Proposition 9 in order to compute an ϵ-best response randomized action of an infoset to the other infosets. This takes polytime if the number of actions per infoset (= branching factor) is bounded. Without this restriction, we run into the impossibility result of Proposition 11.

5.2 CDT Equilibria

How hard is CDT-S, now that we allow for many players? We can get PPAD-hardness straightforwardly because any normal-form game can be cast to extensive form, and because finding a Nash equilibrium in a normal-form game is PPAD-complete [Daskalakis *et al.*, 2009; Chen *et al.*, 2009]. Interestingly enough, we can also show PPAD-membership.

Theorem 6. CDT-S *is* PPAD*-complete. Hardness holds even for two-player perfect-recall games with one infoset per player and for* 1/poly *precision.*

For membership we investigate the existence proof of Lemma 15 by Lambert *et al.*. They first shows a connection to perfect-recall games with particular symmetries, and then give a Brouwer fixed point argument which resembles that of Nash's for symmetric games. However, the connection relies on a construction whose size blows up in the order of factorials, *i.e.*, super-polynomially. Therefore, we modify the fixed point argument to one that works directly on CDT utilities: In a game of imperfect recall, given a profile μ, define the advantage of a pure action a at infoset I of player i as

$$g_{I,a}^{(i)}(\mu) := U_{\mathrm{CDT}}^{(i)}(a \mid \mu, I) - U^{(i)}(\mu)\,.$$

Intuitively, if the advantage of an action a over the current randomized action $\mu^{(i)}(\cdot \mid I)$ is large, then the player should increase its probability of play. Therefore, we may define the Brouwer function to map any profile μ to profile π defined as

$$\pi^{(i)}(a \mid I) := \frac{\mu^{(i)}(a \mid I) + \max\{0, g_{I,a}^{(i)}(\mu)\}}{1 + \sum_{a' \in I} \max\{0, g_{I,a'}^{(i)}(\mu)\}}\,.$$

Then we show that this forms a valid a Brouwer function whose fixed points are indeed CDT equilibria of the underlying game, and that the Brouwer function and precision errors satisfy the computational requirements developed by Etessami and Yannakakis [2010] to imply PPAD-membership.

The PPAD-membership result is a positive algorithmic result: it shows that we can find CDT equilibria with fixed point solvers and path-following methods, just as it is the case with Nash equilibria in normal-form games. In particular, we shall highlight the stark contrast to Theorem 4. Finding a CDT equilibrium sits well within in the landscape of total NP search problems, whereas even deciding whether an EDT equilibrium exists is already on higher levels of the polynomial hierarchy, let alone finding one.

6 The Insignificance of Exogenous Stochasticity

As of now, the hardness results for single-player settings rely on the presence of chance nodes; see Propositions 3 and 4 and Theorem 5. In this section, we investigate the complexity of games without chance nodes. Of course, one might choose to add players to the game to simulate nature, even in games of perfect recall. However, adding players may add significantly to the computational complexity of the game, cf. P vs PPAD for Nash equilibria in single-player vs two-player settings under perfect recall, or Proposition 16 vs Theorem 6 for CDT equilibria under imperfect recall. Interestingly enough, we can show that in the presence of imperfect recall, chance nodes do not affect the complexity.

Theorem 7. *All computational hardness results in this paper for the three equilibrium concepts* {*Nash, EDT, CDT*} *still hold even when the game has no chance nodes. They hold together with previously possible restrictions (*e.g.*, on the branching factor), except that the restrictions on the number of infosets and the degree of absentmindedness increase by one and to* $\mathcal{O}(\log|\mathcal{H}|)$ *respectively.*

In other words, all exogenous stochasticity can be replaced by one infoset (of an arbitrary player, say P1) with absentmindedness, *i.e.*, replaced by uncertainty that arises from forgetting one's past actions in an identical situation. The proof first transforms the game Γ to an equivalent game $\tilde{\Gamma}$ that only has a single chance node h_c that is located at the root. Next, we replace h_c with an infoset I_c with absentmindedness. We illustrate in Figure 6 how to do it with a chance node that uniformly randomizes over two actions. The resulting game Γ' has the same number of players and strategy sets as Γ, except for the additional infoset I_c for P1. In equilibrium, the induced conditional probability distribution over the children of h_c in Γ and the nonterminal "children" of I_c in Γ' will be the same. Finally, there will be a polynomial relationship between the equilibrium precision errors in Γ and Γ'.

Next, recall OPT-D from Proposition 4 which asks whether an approximate target value can be achieved in a single-player game with imperfect recall. We improve on Theorem 7 in the specific problem OPT-D via an independent proof.

Proposition 23. OPT-D *is* NP*-hard, even for games with no chance nodes, one infoset, a degree of absentmindedness of* 2*, and* 1/poly *precision.*

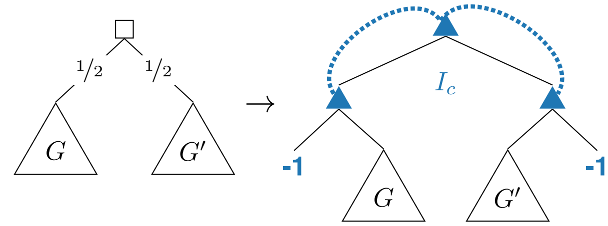


Figure 6: How to remove a chance node if it is located at the root. Starting with the game on the left, replace it with infoset I_c. Assuming w.l.o.g. that the subgames G and G' always yield positive payoffs, the player of I_c will want to randomize uniformly at I_c – independent of the play in G and G'. For another example with more chance node children, see Figure 10 in Appendix E.

Due to Remark 20, this hardness result also holds for deciding whether *all* EDT equilibria achieve an approximate target value. The proof reduces from the 2-MINSAT problem [Kohli *et al.*, 1994].

7 Conclusion

Historically, games of imperfect recall have received only limited attention, as it is not clear that they cleanly model any strategic interactions between humans. However, as we argued in the introduction, they are more practically significant in the context of AI agents. However, they also pose new challenges. Optimal decision making under imperfect recall is hard due to its close connections to polynomial optimization. This and previous work has shown this for the single-player setting. Moreover, it holds even more so in multi-player settings, where we established that even deciding whether a Nash equilibrium (*i.e.*, mutual best responses) exists is very hard. Therefore, we turned towards suitable relaxations that arose from the game theory and philosophy literature. We studied them, and their relationship to each other and to the Nash equilibrium concept, with a computational lens.

We find that CDT equilibria stay relatively easy to find, joining the complexity class of finding a Nash equilibrium in *perfect-recall* or *normal-form* games. This is because CDT defines the most local form of deviation, affecting only one decision node at a time. EDT equilibria show a more convoluted picture. In single-player settings, we relate it to polynomial local search via best-response dynamics. Furthermore, without absentmindedness, EDT and CDT equilibria coincide and hence become equally easy (Remark 19). *With* absentmindedness, on the other hand, the relevant decision problems for EDT equilibria (in single- or multi-player settings) tend to coincide in complexity with the analogous problems for Nash equilibria under *imperfect* recall.

One conclusion, however, has presented itself in all settings considered throughout this paper: (assuming well-accepted complexity assumptions), CDT equilibria are in general strictly easier to find and decide than EDT and Nash equilibria (Proposition 16 vs Theorem 5, Corollary 17 vs Proposition 11, and Theorem 6 vs Theorem 4). Does this imply that

CDT-based reasoning is more suitable for computationally-bounded agents?

Finally, the computational differences between EDT equilibria and Nash equilibria have yet to be properly understood, that is, the differences between global optimization of polynomials over a single simplex versus a product of simplices. We leave this open for future work, with a particular interest in the search complexities of these two equilibrium concepts.

Acknowledgments

We are grateful to Ioannis Anagnostides for the fruitful discussions during the development of this project, and to the anonymous reviewers for their valuable improvement suggestions for this paper. Emanuel Tewolde, Caspar Oesterheld, and Vincent Conitzer thank the Cooperative AI Foundation, Polaris Ventures (formerly the Center for Emerging Risk Research) and Jaan Tallinn's donor-advised fund at Founders Pledge for financial support. Caspar Oesterheld's work is also supported by an FLI PhD Fellowship. Paul Goldberg is supported by EPSRC Grant EP/X040461/1. The work of Prof. Sandholm's group is supported by the Vannevar Bush Faculty Fellowship ONR N00014-23-1-2876, National Science Foundation grants RI-2312342 and RI-1901403, ARO award W911NF2210266, and NIH award A240108S001. Brian Hu Zhang's work is supported in part by the CMU Computer Science Department Hans Berliner PhD Student Fellowship.

References

[Arora and Barak, 2009] Sanjeev Arora and Boaz Barak. *Computational Complexity - A Modern Approach*. Cambridge University Press, 2009.

[Briggs, 2010] Rachael Briggs. Putting a Value on Beauty. In *Oxford Studies in Epistemology: Volume 3*, pages 3–34. Oxford University Press, 2010.

[Brown and Sandholm, 2018] Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

[Brown and Sandholm, 2019] Noam Brown and Tuomas Sandholm. Superhuman AI for multiplayer poker. *Science*, 365(6456):885–890, 2019.

[Brown *et al.*, 2015] Noam Brown, Sam Ganzfried, and Tuomas Sandholm. Hierarchical Abstraction, Distributed Equilibrium Computation, and Post-Processing, with Application to a Champion No-Limit Texas Hold'em Agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS '15*, pages 7–15.

[Buresh-Oppenheim and Morioka, 2004] Joshua Buresh-Oppenheim and Tsuyoshi Morioka. Relativized NP search problems and propositional proof systems. In *Proceedings of the 19th IEEE Conference on Computational Complexity (CCC)*, pages 54–67.

[Buss and Johnson, 2012] Samuel R. Buss and Alan S. Johnson. Propositional proofs and reductions between NP search problems. *Annals of Pure and Applied Logic*, 163(9):1163–1182, 2012.

[Canny, 1988] John Canny. Some Algebraic and Geometric Computations in PSPACE. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 460–467.

[Celli and Gatti, 2018] Andrea Celli and Nicola Gatti. Computational Results for Extensive-Form Adversarial Team Games. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18)*, pages 965–972.

[Chen *et al.*, 2009] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the Complexity of Computing Two-Player Nash Equilibria. *J. ACM*, 56(3):14:1–14:57, 2009.

[Conitzer and Oesterheld, 2023] Vincent Conitzer and Caspar Oesterheld. Foundations of Cooperative AI. In *Thirty-Seventh AAAI Conference on Artificial Intelligence*, pages 15359–15367.

[Conitzer, 2019] Vincent Conitzer. Designing Preferences, Beliefs, and Identities for Artificial Intelligence. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, pages 9755–9759.

[Daskalakis and Papadimitriou, 2011] Constantinos Daskalakis and Christos Papadimitriou. Continuous Local Search. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 790–804.

[Daskalakis *et al.*, 2009] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *SIAM J. Comput.*, 39(1):195–259, 2009.

[Elga, 2000] Adam Elga. Self-locating belief and the Sleeping Beauty problem. *Analysis*, 60(2):143–147, 2000.

[Emmons *et al.*, 2022] Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, and Stuart Russell. For Learning in Symmetric Teams, Local Optima are Global Nash Equilibria. In *International Conference on Machine Learning, ICML 2022*, pages 5924–5943.

[Etessami and Yannakakis, 2010] Kousha Etessami and Mihalis Yannakakis. On the Complexity of Nash Equilibria and Other Fixed Points. *SIAM J. Comput.*, 39(6):2531–2597, 2010.

[Fearnley *et al.*, 2023] John Fearnley, Paul Goldberg, Alexandros Hollender, and Rahul Savani. The Complexity of Gradient Descent: CLS = PPAD $\cap$ PLS. *J. ACM*, 70(1):7:1–7:74, 2023.

[Fudenberg and Tirole, 1991] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, October 1991.

[Ganzfried and Sandholm, 2014] Sam Ganzfried and Tuomas Sandholm. Potential-Aware Imperfect-Recall Abstraction with Earth Mover's Distance in Imperfect-Information Games. In *Proceedings of the Twenty-Eighth*

AAAI Conference on Artificial Intelligence, pages 682–690.

[Gimbert *et al.*, 2020] Hugo Gimbert, Soumyajit Paul, and B. Srivathsan. A Bridge between Polynomial Optimization and Games with Imperfect Recall. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '20*, pages 456–464.

[Håstad, 2001] Johan Håstad. Some Optimal Inapproximability Results. *J. ACM*, 48(4):798–859, 2001.

[Johnson *et al.*, 1988] David S. Johnson, Christos H. Papadimitriou, and Mihalis Yannakakis. How Easy is Local Search? *Journal of Computer and System Sciences*, 37(1):79–100, 1988.

[Kohli *et al.*, 1994] Rajeev Kohli, Ramesh Krishnamurti, and Prakash Mirchandani. The Minimum Satisfiability Problem. *SIAM J. Discret. Math.*, 7(2):275–283, 1994.

[Koller and Megiddo, 1992] Daphne Koller and Nimrod Megiddo. The Complexity of Two-Person Zero-Sum Games in Extensive Form. *Games and Economic Behavior*, 4(4):528–552, 1992.

[Kovarík *et al.*, 2023] Vojtech Kovarík, Caspar Oesterheld, and Vincent Conitzer. Game Theory with Simulation of Other Players. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023*, pages 2800–2807.

[Kovarík *et al.*, 2024] Vojtech Kovarík, Caspar Oesterheld, and Vincent Conitzer. Recursive Joint Simulation in Games. *CoRR*, abs/2402.08128, 2024.

[Kuhn, 1953] Harold W. Kuhn. Extensive Games and the Problem of Information. In *Contributions to the Theory of Games (AM-28), Volume II*, chapter 11, pages 193–216. Princeton University Press, 1953.

[Lambert *et al.*, 2019] Nicolas S. Lambert, Adrian Marple, and Yoav Shoham. On equilibria in games with imperfect recall. *Games Econ. Behav.*, 113:164–185, 2019.

[Lipton and Markakis, 2004] Richard J. Lipton and Evangelos Markakis. Nash Equilibria via Polynomial Equations. In *LATIN 2004: Theoretical Informatics, 6th Latin American Symposium*, pages 413–422.

[Morioka, 2001] Tsuyoshi Morioka. Classification of Search Problems and Their Definability in Bounded Arithmetic. Master's thesis, University of Toronto, 2001.

[Nash, 1950] John F. Nash. Equilibrium Points in N-Person Games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[Nash, 1951] John Nash. Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295, 1951.

[Oesterheld and Conitzer, 2022] Caspar Oesterheld and Vincent Conitzer. Can *de se* choice be *ex ante* reasonable in games of imperfect recall? https://www.andrew.cmu.edu/user/coesterh/DeSeVsExAnte.pdf, 2022. Working paper. Accessed: 2022-12-14.

[Papadimitriou, 1994] Christos H. Papadimitriou. On the Complexity of the Parity Argument and Other Inefficient Proofs of Existence. *Journal of Computer and System Sciences*, 48(3):498–532, 1994.

[Piccione and Rubinstein, 1997] Michele Piccione and Ariel Rubinstein. On the Interpretation of Decision Problems with Imperfect Recall. *Games and Economic Behavior*, 20(1):3–24, 1997.

[Renegar, 1992] James Renegar. On the Computational Complexity and Geometry of the First-Order Theory of the Reals. Part I: Introduction. Preliminaries. The Geometry of Semi-Algebraic Sets. The Decision Problem for the Existential Theory of the Reals. *Journal of Symbolic Computation*, 13(3):255–299, 1992.

[Schaefer and Stefankovic, 2017] Marcus Schaefer and Daniel Stefankovic. Fixed Points, Nash Equilibria, and the Existential Theory of the Reals. *Theory Comput. Syst.*, 60(2):172–193, 2017.

[Schäffer and Yannakakis, 1991] Alejandro A. Schäffer and Mihalis Yannakakis. Simple Local Search Problems That are Hard to Solve. *SIAM J. Comput.*, 20(1):56–87, 1991.

[Shor, 1990] Peter W. Shor. Stretchability of Pseudolines is NP-Hard. In *Applied Geometry And Discrete Mathematics, Proceedings of a DIMACS Workshop*, pages 531–554.

[Stockmeyer, 1976] Larry J. Stockmeyer. The Polynomial-Time Hierarchy. *Theor. Comput. Sci.*, 3(1):1–22, 1976.

[Strotz, 1955] R. H. Strotz. Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies*, 23(3):165–180, 1955.

[Tewolde and Conitzer, 2024] Emanuel Tewolde and Vincent Conitzer. Game Transformations That Preserve Nash Equilibria or Best-Response Sets. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024*.

[Tewolde *et al.*, 2023] Emanuel Tewolde, Caspar Oesterheld, Vincent Conitzer, and Paul W. Goldberg. The Computational Complexity of Single-Player Imperfect-Recall Games. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023*, pages 2878–2887.

[von Stengel and Koller, 1997] Bernhard von Stengel and Daphne Koller. Team-Maxmin Equilibria. *Games and Economic Behavior*, 21(1-2):309–321, 1997.

[Waugh *et al.*, 2009] Kevin Waugh, Martin Zinkevich, Michael Johanson, Morgan Kan, David Schnizlein, and Michael H. Bowling. A Practical Use of Imperfect Recall. In *Eighth Symposium on Abstraction, Reformulation, and Approximation, SARA 2009*.

[Wichardt, 2008] Philipp C. Wichardt. Existence of Nash Equilibria in Finite Extensive Form Games with Imperfect Recall: A Counterexample. *Games Econ. Behav.*, 63(1):366–369, 2008.

[Yannakakis, 2003] Mihalis Yannakakis. *Computational Complexity*, pages 19–56. Princeton University Press, 2003.

[Zhang and Sandholm, 2022] Brian Hu Zhang and Tuomas Sandholm. Polynomial-Time Optimal Equilibria with a Mediator in Extensive-Form Games. In *Annual Conference on Neural Information Processing Systems, NeurIPS 2022*.

[Zhang *et al.*, 2023] Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. Team Belief DAG: Generalizing the Sequence Form to Team Games for Fast Computation of Correlated Team Max-Min Equilibria via Regret Minimization. In *International Conference on Machine Learning, ICML 2023*, pages 40996–41018.

A On Section 2

In this section, we expand on the technical background needed on games with imperfect recall to develop our results. Again, this section closely follows the notation of Tewolde *et al.* [2023] and simply extends their exposition to multi-player settings.

A.1 Notation

Players in $\mathcal{N}$ will be referred to as *strategic* players. Recall that we can identify their strategy sets with Cartesian products of simplices

$$\mathcal{S} \equiv \bigtimes_{i \in \mathcal{N}} \bigtimes_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1} \quad \text{and} \quad \mathcal{S}^{(i)} \equiv \bigtimes_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1} .$$

For that purpose, we have to fix an ordering of the infosets and actions in a considered game Γ. Denote the number of infosets of strategic player $i \in \mathcal{N}$ as $\ell^{(i)} := |\mathcal{I}^{(i)}|$, and fix an ordering $I_1^{(i)}, \dots, I_{\ell^{(i)}}^{(i)}$ of infosets in $\mathcal{I}^{(i)}$. Similarly, denote the number of actions at infoset $I_j^{(i)} \in \mathcal{I}^{(i)}$ as $m_j^{(i)} := |A_{I_j^{(i)}}|$, and fix an ordering $a_{j1}^{(i)}, \dots, a_{jm_j^{(i)}}^{(i)}$ of actions in $A_{I_j^{(i)}}$.

Then, a strategy $\mu \in \mathcal{S}$ is uniquely identified with the vector

$$\mu = \big(\mu_{jk}^{(i)}\big)_{i,j,k} \in \bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} \Delta^{m_j^{(i)}-1} \subset \bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}}$$

where $\mu_{jk}^{(i)} \in [0,1]$ is the probability $\mu^{(i)}\big(a_{jk}^{(i)} \mid I_j^{(i)}\big)$ that strategic player i assigns to action $a_{jk}^{(i)}$ at infoset $I_j^{(i)}$.

A.2 Representation of Polynomials

Since we draw some of our complexity results from known results in polynomial optimization, we first note that polynomials shall be represented in the Turing (bit) model, which will be described below. Say, we have a general polynomial function $p : \bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}} \to \mathbb{R}$ in variables $x = (x_{jk}^{(i)})_{i,j,k}$ and of total degree $d \in \mathbb{N}$. Let $\boldsymbol{m} := (m_j^{(i)})_{i,j=1}^{N,\ell^{(i)}}$ denote the associated vector of dimensions. Then, the relevant standard monomial basis to p is $\big\{ \prod_{i,j,k=1}^{N,\ell^{(i)},m_j^{(i)}} (x_{jk}^{(i)})^{D_{ijk}} \big\}_{D \in \text{MB}(d,\boldsymbol{m})}$. Here, each vector D indicates one unique way to distribute the total degree d to the variables, and $\text{MB}(d,\boldsymbol{m})$ denotes the collection of them. That is

$$\text{MB}(d,\boldsymbol{m}) := \{ D = (D_{ijk})_{ijk} \in \bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} (\mathbb{N} \cup \{0\})^{m_j^{(i)}} : \sum_{i,j,k=1}^{N,\ell^{(i)},m_j^{(i)}} D_{ijk} \leq d \} .$$

By abuse of notation, let $x^D := \prod_{i,j,k} (x_{jk}^{(i)})^{D_{ijk}}$. Then, polynomial p can be uniquely represented as $p(x) = \sum_{D \in \text{MB}(d,\boldsymbol{m})} \lambda_D \cdot x^D$ where each $\lambda_D \in \mathbb{Q}$ for computational considerations. Finally, any general polynomials p in this paper are assumed to be represented as a binary encoding of these values $(\ell^{(i)})_{i=1}^N, \boldsymbol{m}$ and coefficients $(\lambda_D)_{D \in \text{MB}(d,\boldsymbol{m})}$.

A.3 Lipschitz Constants

As polynomial functions, utility functions $U^{(i)}$ are Lipschitz continuous over strategy space S. We will sometimes need access to these Lipschitz constants. Tewolde *et al.* [2023][Lemma 22] describe how, given a polynomial function $p : \mathbb{R}^n \to \mathbb{R}$ in the Turing (bit) model, one can obtain a Lipschitz constant L_∞ of p over the hypercube $[0,1]^n$ w.r.t. the infinity norm within polytime.

One possible Lipschitz constant is the maximum gradient norm over the hypercube

$$\max_{x \in [0,1]^n} \{ ||\nabla p(x)||_\infty \} = \max_{x \in [0,1]^n} \max_{j \in [n]} |\nabla_j \, p(x)| = \max_{j \in [n]} \max_{x \in [0,1]^n} |\nabla_j \, p(x)| .$$

Consider a dimension $j \in [n]$ and suppose polynomial $\nabla_j \, p(x)$ has monomial coefficients $(\lambda_D)_D$. Since all variables $x_{j'}$ are bounded by 1, we get

$$\max_{x \in [0,1]^n} |\nabla_j \, p(x)| \leq \max_{x \in [0,1]^n} |\sum_D \lambda_D \cdot x^D| \leq \sum_{\lambda_D} |\lambda_D| =: L_j ,$$

which is polytime computable. Hence, we can set

$$L_\infty := \max\{1, \max_{j \in [n]} L_j\} .$$

Now say, we start with a game Γ with imperfect recall. Observe that the strategy space S is a subset of one high-dimensional hypercube $\bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} [0,1]^{m_j^{(i)}}$. By the method described above, we can get a Lipschitz constant $L_\infty^{(i)}$ of each player's utility function $U^{(i)}$, and a Lipschitz constant $L_\infty^{(i,j,k)}$ of each partial derivative $\nabla_{jk} \, U^{(i)}$. Summarize them to one Lipschitz constant

$$L_\infty := \max\{1, \max_{i \in [N]} L_\infty^{(i)}, \max_{i \in [N], j \in [\ell^{(i)}], k \in [m_j^{(i)}]} L_\infty^{(i,j,k)}\}$$

for the game Γ.

Furthermore, if the utility payoffs in Γ are bounded, say in $[-2,2]$, then the coefficients λ_D will be (at worst) a monomial coefficient from $U^{(i)}$ times two monomial degrees from $U^{(i)}$ (due to second derivative). These values are bounded for the by the maximum absolute utility value 2 and by the squared number of nodes $|\mathcal{H}|^2$ respectively. Moreover, there will be at most as many monomials as there are terminal nodes in Γ, which itself is bounded by $|\mathcal{H}|$ again. Hence, each $L_j \leq 2|\mathcal{H}|^3$, and thus $L_\infty = \mathsf{poly}(|\mathcal{H}|)$.

A.4 From Polynomials to Imperfect-Recall Games

Given a set of polynomials $\boldsymbol{p} = (p^{(1)}, \dots, p^{(N)}) : \bigtimes_{i \in [N]} \bigtimes_{j \in [\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}} \to \mathbb{R}^N$, we can construct an associated N-player game with imperfect recall Γ such that Γ's its expected utility functions satisfy $U^{(i)}(\mu) = p^{(i)}(\mu)$ on $\bigtimes_{i=1}^N \bigtimes_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}}$.

Denote $\text{supp}(\boldsymbol{p}) := \{D \in \text{MB}(d,\boldsymbol{m}) : \lambda_D^{(i)} \neq 0 \text{ for some } i \in [N]\}$). The constructed game Γ shall have N players and an infoset $I_j^{(i)}$ for each $i \in [N]$ and $j \in [\ell^{(i)}]$.

At $I_j^{(i)}$, there shall $m_j^{(i)}$ action choices. The game tree will have a depth of up to $d+1$. The root h_0 will be a chance node that has one outgoing edge to a subtree T_D for each monomial index $D \in \operatorname{supp}(\boldsymbol{p})$. An outgoing edge is drawn uniformly at random. Let us build T_D associated to D, which, in turn, is associated to monomial $\prod_{i,j,k}(x_{jk}^{(i)})^{D_{ijk}}$. Let $\operatorname{supp}^{\text{ms}}(D)$ be a lexicographically ordered version of the multiset which contains D_{ij} many copies of element (i,j,k) if $D_{ijk} > 0$. Then, going through the list $\operatorname{supp}^{\text{ms}}(D)$ means that we will encounter a variable $x_{jk}^{(i)}$ that degree D_{ijk} exactly D_{ijk}-many times back to back. Therefore, starting with the edge that goes into T_D, do the following loop:

1. Take the next element (i,j,k) of $\operatorname{supp}^{\text{ms}}(D)$.
2. Add a nonterminal node h to the current edge and assign h to info set $I_j^{(i)}$ of player i.
3. Create $m_j^{(i)}$ outgoing edges from h, one for each action at $I_j^{(i)}$.
4. At the end of edges $\neq k$, add a terminal node with utility payoff 0 for all $i \in [N]$.
5. Go to the $k-th$ edge.

Lastly, once we are through with $\operatorname{supp}^{\text{ms}}(D)$, add a final terminal node z_D to the current edge, with utility payoff $\lambda_D^{(1)} \cdot |\operatorname{supp}(\boldsymbol{p})|, \dots, \lambda_D^{(N)} \cdot |\operatorname{supp}(\boldsymbol{p})|$. With this procedure, subtree T_D will have depth $|\operatorname{supp}^{\text{ms}}(D)| = ||D||_1$.

In this reduction, we have that any point $x \in \bigtimes_{i\in[N]} \bigtimes_{j\in[\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}}$ that is also in $\bigtimes_{i\in[N]} \bigtimes_{j\in[\ell^{(i)}]} \Delta^{m_j^{(i)}-1}$ naturally comprises a strategy in Γ that selects the k-th action at $I_j^{(i)}$ with probability $x_{jk}^{(i)}$. Moreover, each expected utility function $U^{(i)}$ of player i in Γ satisfies for such a point x:

$$\begin{aligned}
U^{(i)}(x) &= \sum_{z\in\mathcal{Z}} \mathbb{P}(z \mid x) \cdot u^{(i)}(z) \\
&= \sum_{D\in\operatorname{supp}(\boldsymbol{p})} \mathbb{P}(z_D \mid x) \cdot \lambda_D^{(i)} \cdot |\operatorname{supp}(\boldsymbol{p})| \\
&= \sum_{D\in\operatorname{supp}(\boldsymbol{p})} \Big[\Big(\frac{1}{|\operatorname{supp}(\boldsymbol{p})|} \cdot \prod_{D_{ijk}>0} (x_{jk}^{(i)})^{D_{ijk}} \Big) \\
&\qquad \cdot \lambda_D^{(i)} \cdot |\operatorname{supp}(\boldsymbol{p})| \Big] \\
&= \sum_{D\in\mathrm{MB}(d,\boldsymbol{m})} \lambda_D \cdot \prod_{i,j} (x_{jk}^{(i)})^{D_{ijk}} \\
&= p(x)\,.
\end{aligned}$$

This identity extends to the whole space $\bigtimes_{i\in[N]} \bigtimes_{j\in[\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}}$. Finally, the construction of Γ takes polytime in the encoding size of $p^{(1)}, \dots, p^{(N)}$.

A.5 Further Comments on the Complexity Classes

Decision Problems P, NP, and Σ_2^{P} are parts of the lower levels of the polynomial hierarchy. For a more detailed treatment we refer to [Arora and Barak, 2009][Section 5]. The first order of the reals, its subclasses $\exists\forall\mathbb{R}$ and the existential theory of the reals $\exists\mathbb{R}$, as well as algorithms to solve those decision problems are discussed in [Renegar, 1992; Schaefer and Stefankovic, 2017]. The chain NP $\subseteq \exists\mathbb{R} \subseteq$ PSPACE$\cap\exists\forall\mathbb{R}$ is due to [Shor, 1990; Canny, 1988], and it ties into the previous discussion that the solutions of an $\exists\mathbb{R}$ sentence may take on irrational solution values.

Let us discuss the running time of deciding a $\exists\forall\mathbb{R}$-sentence $\exists x \in \mathbb{R}^{n_1} \forall y \in \mathbb{R}^{n_2} F(x,y)$, using the standard variable notation that may overlap with our variable notation in this paper. In its standard form, quantifier-free formula F is assumed to be a Boolean formula P where the atomic predicates can be of the form $g_i \Delta_i 0$. Here, g_i is a polynomial function in integer coefficients and $\Delta_i \in \{>, \geq, =, \neq, \leq, <\}$. Let n_1 and n_2 be the number of variables of the $\exists$ and $\forall$ quantifiers, a be the length of P, m be the number of atomic predicates $g_i \Delta_i 0$, d be an upper bound on the degree of polynomials g_i, and L be the bit length to represent the coefficients in the g_i. Then, Renegar [1992] gives an algorithm $\mathcal{A}$ that takes time $poly(L,a) \cdot (md)^{\mathcal{O}(n_1 \cdot n_2)}$ to decide whether such a sentence is true or not. In order for this to be polynomial time, we may require the number of variables n_1 and n_2 to be constant. For deciding an $\exists\mathbb{R}$-sentence, just set $n_2 = 1$ in the above running time bound.

Total NP search problems Decision problems ask for a yes/no answer. Search problems can ask for more sophisticated answers, usually they ask directly for solutions (if existent) with which we can verify a "yes" instances of the associated decision problem. The complexity classes FP and FNP are the search problem analogues of P and NP. We have P = NP if and only if FP = FNP. However, the landscape between FP and FNP has been characterized better than the landscape between P and NP. More specifically, there is a special interest in search problems for which one knows that each problem instance admits a solution (the landscape of total NP search problems). Within that, we will be interested in the three complexity classes CLS, PLS, PPAD, which can be characterized by the method with which one can show that each problem instance admits a solution. For the class PPAD ("Polynomial Parity Arguments on Directed graphs", Papadimitriou, 1994), that is if one can show that the existence of a solution can be proven by a fixed point argument. This is the case for example for the existence of an approximate Nash equilibrium [Nash, 1951; Daskalakis *et al.*, 2009]. For the class PLS ("Polynomial Local Search", Johnson *et al.*, 1988), that is if one can show that the existence of a solution can be proven by a local optimization argument on a directed acyclic graph. Since we will prove PLS-membership directly, we will give a precise definition further below. For the class CLS ("Continuous Local Search", Daskalakis and Papadimitriou, 2011), that is if one can show that the existence of a solution can be proven by a local optimization argument on a bounded polyhedral (continuous) domain. Alternatively, CLS can be characterized as the intersection of PPAD and PLS [Fearnley *et al.*, 2023].

In Section 5.1, we discuss the differences in complexity of finding an approximate EDT vs CDT equilibrium in single-player settings, and relate it to PLS versus CLS. As of yet, CLS is believed to be a proper subclass of PLS, a belief supported by separation oracles (a kind of conditional separation) for PPAD and PLS [Buresh-Oppenheim and Morioka, 2004; Buss and Johnson, 2012; Morioka, 2001]. An unconditional separation of these classes would imply P $\subsetneq$ NP.

Definition of PLS A *local search problem* is given by a set of instances $\mathcal{J}$. For every instance $J \in \mathcal{J}$ we are given a finite set of feasible solutions $\mathcal{S}(J)$, an objective function $c : \mathcal{S}(J) \to \mathbb{Q}$, and for every feasible solution $s \in \mathcal{S}(J)$ a *neighborhood* $N(s, J) \subseteq \mathcal{S}(J)$. Given an instance J, the goal is to compute a *locally optimal solution* s^*, that is, a solution that does not have a strictly better neighbor in terms of the objective value.

Definition 24 (Johnson *et al.* [1988])**.** *The complexity class polynomial local search (*PLS*) consists of all local search problems that admit a polytime algorithm for each of the following tasks:*

1. *Given an instance J, compute an initial feasible solution;*
2. *Given an instance J and a feasible solution $s \in \mathcal{S}(J)$, compute the objective value $C(s)$;*
3. *Given an instance J and a feasible solution $s \in \mathcal{S}(J)$, determine if s is a local optimum, or otherwise determine a feasible solution $s \in N(s, J)$ with a strictly higher objective value.*

B On Section 3

In this section, we prove the results in Section 3. To that end, we restate results taken from the main body, and give new numbers to results presented first in this appendix.

Proposition (Restatement of Proposition 7)**.** *Let Γ be a 2p0s game with imperfect recall. If $\Delta \leq \epsilon$ then Γ admits an ϵ-Nash equilibrium. Conversely, if Γ admits an ϵ-Nash equilibrium, then $\Delta \leq 2\epsilon$.*

Proof.

1. Suppose $\Delta \leq \epsilon$. Let

$$\pi^{(1)} = \operatorname*{argmax}_{\mu^{(1)} \in S^{(1)}} \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}),$$

and

$$\pi^{(2)} = \operatorname*{argmin}_{\mu^{(2)} \in S^{(2)}} \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}).$$

Then we can show that $(\pi^{(1)}, \pi^{(2)})$ is an ϵ-Nash equilibrium. Indeed, we have

$$\begin{aligned} \underline{U} = \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\pi^{(1)}, \mu^{(2)}) &\leq U^{(1)}(\pi^{(1)}, \pi^{(2)}) \\ \leq \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \pi^{(2)}) &= \bar{U}. \end{aligned}$$

Thus, using $\bar{U} - \underline{U} = \Delta \leq \epsilon$, we obtain the ϵ-Nash equilibrium conditions

$$U^{(1)}(\pi^{(1)}, \pi^{(2)}) \geq \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \pi^{(2)}) - \epsilon,$$

and

$$U^{(1)}(\pi^{(1)}, \pi^{(2)}) \leq \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\pi^{(1)}, \mu^{(2)}) + \epsilon.$$

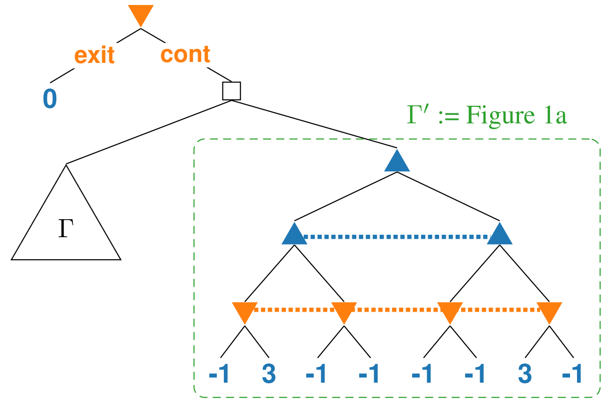


Figure 7: Game used in Lemma 25 and the proofs of Theorems 1 and 2. Γ is a placeholder game. We use boxes for chance nodes, at which chance plays uniformly at random.

2. Suppose μ^* is an ϵ-Nash equilibrium. Then

$$\begin{aligned} \Delta = &\min_{\mu^{(2)} \in S^{(2)}} \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}) \\ &- \max_{\mu^{(1)} \in S^{(1)}} \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}) \\ \leq &\max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, (\mu^*)^{(2)}) \\ &- \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}((\mu^*)^{(1)}, \mu^{(2)}) \\ = &\max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, (\mu^*)^{(2)}) - U^{(1)}(\mu^*) \\ &+ U^{(1)}(\mu^*) - \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}((\mu^*)^{(1)}, \mu^{(2)}) \\ \leq &\epsilon + \epsilon = 2\epsilon. \end{aligned}$$

□

Next, consider the game of Figure 7 where Γ is a 2p0s game. Let $\underline{U}$ be the max-min value in the subgame Γ. Moreover, let $v := \min_{z \in \mathcal{Z}} u^{(1)}(z) \leq -1$ be the minimum terminal payoff in Figure 7.

Lemma 25. *If max-min value $\underline{U} \geq 0$ in Γ, then the game in Figure 7 has an exact Nash equilibrium. If $\underline{U} \leq -\epsilon$ in Γ for $\epsilon > 0$ sufficiently small ($\epsilon \leq -\frac{1}{8v}$), then Figure 7 has no $\epsilon/2$-Nash equilibrium.*

Proof. First we note that the max-min value of the forgetful penalty shoot-out subgame Γ' (cf. Figure 1a) is 0, realized for example by P1 randomizing 50/50 at her infosets, and P2 going left deterministically. The min-max value in Γ', on the other hand, is 1, realized for example by P2 randomizing 50/50, and P1 going left deterministically. Therefore, $\Delta = 1$, and by Proposition 7, there is no 1/4-Nash equilibrium in Γ'.

The proof idea is that given a strategy of P1, P2 would be interested in continuing at the root node if and only if he can achieve negative utility after that. This comes down to if he can achieve negative utility in Γ because in Γ', there is always a best response with which he can achieve non-positive utility. Suppose that is the case and, thus, P2 is continuing at the root node with sufficiently high probability. Then the profile is already guaranteed to not be an (approximate) Nash equilibrium because the players will never be in a $1/4$-Nash equilibrium in the subgame Γ'.

1. Suppose $\underline{U} \geq 0$ in Γ. Consider the profile where P2 exits 100% of the time, and where P1 plays her max-min strategies in Γ and Γ', and P2 plays anything in Γ and Γ'. This makes an exact Nash equilibrium: The only play that has an effect on the ex-ante utility is P2 exiting at his first node. And this is optimal for P2, because upon continuing he would receive a loss of ≥ 0 instead (recall the 2p0s assumption).

2. Suppose $\underline{U} \leq -\epsilon$ for some $0 < \epsilon < -\frac{1}{8v}$. For the sake of a contradiction, assume Figure 7 also has an $\epsilon^2/4$-Nash equilibrium π. Let ρ be the probability with which P2 continues at the root in profile π. Then observe that P2 can deviate to the strategy $\mu^{(2)}$ that deterministically continues at the root and that plays a best response to $\pi^{(1)}$ in Γ and Γ'. Therefore, since π is an $\epsilon^2/4$-Nash equilibrium, we get

$$\begin{aligned}\rho v &\leq U^{(1)}(\pi) \leq U^{(1)}(\pi^{(1)}, \mu^{(2)}) + \frac{\epsilon^2}{4} \\ &\leq \frac{1}{2}\underline{U} + \frac{1}{2}0 + \frac{1}{4}\epsilon \leq \frac{1}{2}\underline{U} + \frac{1}{4}\cdot(-\underline{U}) = \frac{1}{4}\underline{U} < 0\,.\end{aligned}$$

Since $v < 0$, we get $\rho > 0$ and hence $\rho \geq \frac{1}{4v}\underline{U}$.

Again, π is an $\epsilon^2/4$-Nash equilibrium. Thus, in particular, no player has an incentive of more than $\epsilon^2/4$ to deviate to another strategy in the game Γ'. Hence $\pi|_{\Gamma'}$ would make an approximate Nash equilibrium in Γ' (considered as its own game) with rescaled approximation error

$$\begin{aligned}&\frac{1}{\rho\cdot\frac{1}{2}}\cdot\frac{\epsilon^2}{4} \leq \frac{\epsilon^2}{\frac{1}{2v}\underline{U}} = \frac{2\cdot(-v)\cdot\epsilon^2}{-\underline{U}} \leq \frac{2\cdot(-v)\cdot\epsilon^2}{\epsilon} \\ &= -2v\epsilon \leq -2v\frac{-1}{8v} = \frac{1}{4}\,.\end{aligned}$$

This contradicts that Γ' has no $1/4$-Nash equilibrium as discussed in the beginning of the proof.

□

Theorem (Restatement of Theorem 1)**.** *Deciding if a game with imperfect recall admits a Nash equilibrium is $\exists\mathbb{R}$-hard and in $\exists\forall\mathbb{R}$. Hardness holds even for 2p0s games where one player has a degree of absentmindedness of 4 and the other player has perfect recall.*

Proof. $\exists\forall\mathbb{R}$-membership follows because we can formulate the question of Nash equilibrium existence as the sentence

$$\begin{aligned}\exists\mu\forall\pi : \mu \in S \wedge \Bigg[&\pi \notin S \vee \\ &\wedge_{i\in[N]}\left(U^{(i)}(\mu) \geq U^{(i)}(\pi^{(i)}, \mu^{(-i)})\right)\Bigg]\,.\end{aligned}$$

$\exists\mathbb{R}$-hardness follows from a reduction from Proposition 3. Let (Γ, t) be an instance of that decision problem. W.l.o.g. we can assume $t = 0$ (otherwise first shift the payoffs in Γ by $1 - t$). Insert Γ into Figure 7 by letting P1 play in that subgame. Payoffs of P2 in that subgame shall simply be the negative of the payoffs of P1. Call this new game construction G. This is a polytime construction. Moreover, asking whether a utility of 0 can be achieved in original Γ is equivalent to asking whether $\underline{U} \geq 0$ in subgame Γ of G.

The equivalence of those decision problems follows by Lemma 25: If $\underline{U} \geq 0$ then G has a Nash equilibrium. If G has a Nash equilibrium, then it also has an $\epsilon/2$-Nash equilibrium for arbitrary small $\epsilon > 0$. Hence, $\underline{U} \geq 0$.

About the hardness restrictions: Gimbert *et al.* [2020] show hardness of Proposition 3 even for degree 4 polynomials, that is, games with degree of absentmindedness 4. Moreover, G is a 2p0s game in which P2 has perfect recall. □

Theorem (Restatement of Theorem 2)**.** NASH-D *is Σ_2^{P}-complete. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

Proof. Σ_2^{P}-membership: Given an instance (Γ, ϵ), we can guess a profile μ and verify in non-deterministic polytime whether it is an ϵ-Nash equilibrium. Namely, guess μ to have action probabilities with values with denominators $\leq \frac{2L_\infty}{\epsilon}$, where L_∞ is obtained as in Appendix A.3. Note that this is a polysized guess in (Γ, ϵ). Then, to verify, guess another profile π in the same way. Finally, check for each player $i \in \mathcal{N}$ whether $U^{(i)}(\mu^{(i)}, \mu^{(-i)}) \geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - \epsilon$. If so, then this serves as a verification that Γ has an ϵ-Nash equilibrium, and therefore, an exact Nash equilibrium.

This works, because if Γ has an exact Nash equilibrium μ^*, then this method is able to find an ϵ-Nash equilibrium. Namely, μ^* will be at most $\frac{\epsilon}{2L_\infty}$ away (in $||\cdot||_\infty$) from a profile μ that could have been guessed by the method above. And this profile will satisfy for all player $i \in \mathcal{N}$, and all alternative strategies $\pi^{(i)} \in \mathcal{S}^{(i)}$

$$\begin{aligned}U^{(i)}(\mu) &= U^{(i)}(\mu) - U^{(i)}(\mu^*) + U^{(i)}(\mu^*) \\ &\geq U^{(i)}(\mu^*) - L_\infty||\mu - \mu^*||_\infty \\ &\geq U^{(i)}(\pi^{(i)}, (\mu^*)^{(-i)}) - L_\infty\frac{\epsilon}{2L_\infty} \\ &\quad - U^{(i)}(\pi^{(i)}, \mu^{(-i)}) + U^{(i)}(\pi^{(i)}, \mu^{(-i)}) \\ &\geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - \frac{\epsilon}{2} \\ &\quad - L_\infty\left|\left|\left(\pi^{(i)}, (\mu^*)^{(-i)}\right) - \left(\pi^{(i)}, \mu^{(-i)}\right)\right|\right|_\infty \\ &\geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - \frac{\epsilon}{2} - L_\infty\frac{\epsilon}{2L_\infty} \\ &\geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - \epsilon\,.\end{aligned}$$

Σ_2^{P}-hardness: We reduce from Lemma 6, and the idea is analogous to the proof of Theorem 1. Given an instance (Γ, ϵ) for it, insert Γ in Figure 7 and call the construction G. Let $\underline{U}$ be the min-max value in Γ, and v be the minimum terminal payoff in G. Set $\epsilon' := \frac{1}{2}\min\{\epsilon, -\frac{1}{8v}\}$. Then, by Lemma 25, we have for corresponding NASH-D instance (Γ, ϵ'): If $\underline{U} \geq 0$, then G has an exact Nash equilibrium, which will be correctly identified as such by a NASH-D solver. If $\underline{U} \leq -\epsilon$, then also $\underline{U} \leq -\min\{\epsilon, -\frac{1}{8v}\}$. Hence, G has no ϵ'-Nash equilibrium, which will be correctly identified as such by a NASH-D solver.

The restrictions on the hardness result follow directly from Lemma 6 or Theorem 2. □

Corollary (Restatement of Corollary 8)**.** *It is* Σ_2^{P}*-complete to distinguish* $\Delta = 0$ *from* $\Delta \geq \epsilon$ *in 2p0s games. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

Proof. Reduce from Theorem 2 using Proposition 7. $\square$

Proposition (Restatement of Proposition 9)**.** NASH-D *is solvable in time* $\mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\Big)$.

Proof. Let (Γ, ϵ) be an instance of NASH-D. Let $m := \sum_{i \in [N]} \sum_{j \in [\ell^{(i)}]} m_j^{(i)}$ be the total number of pure actions in the game. This will be the number of variables n_1 and n_2 in the resulting $\exists\forall\mathbb{R}$-sentences (recall Appendix A.5). By abuse of notation, let $S(\mu)$ be the system of linear (in-)equalities in a profile μ that describes whether μ lies in the profile set, that is, the conjunctions of

$$\mu_{jk}^{(i)} \geq 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]$$
$$\text{and} \sum_{k=1}^{m_j^{(i)}} \mu_{jk}^{(i)} = 1 \quad \forall i \in [N], \forall j \in [\ell^{(i)}]\,.$$

Notice that the profile set S lies in the standard hypercube $[0,1]^m$ which can be described as the (conjuction) system $B(\mu)$ of linear (in-)equalities

$$\mu_{jk}^{(i)} \geq 0 := y_{jk}^{(i)} \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]\,,$$
$$\text{and } \mu_{jk}^{(i)} \leq 1 := z_{jk}^{(i)} \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]\,.$$

We will make use and adjust the values $y_{jk}^{(i)}$ and $z_{jk}^{(i)}$ later on.

First, we decide the sentence whether there exists $(\exists)$ $\mu \in \mathbb{R}^m$ such that for all $(\forall)$ $\pi \in \mathbb{R}^m$ we have

$$\begin{aligned} &S(\mu) \wedge B(\mu) \\ &\wedge \Big[\neg S(\pi) \vee \wedge_{i \in [N]} \Big(U^{(i)}(\mu) \geq U^{(i)}(\pi^{(i)}, \mu^{(-i)})\Big)\Big]\,. \end{aligned} \tag{1}$$

If this is false, then we can return that no (ϵ-)Nash equilibrium exists in Γ (since NASH-D is a promise problem). If the sentence is true, then we can work on finding an ϵ-Nash equilibrium. We do it by shrinking the region of profile space S that we may consider further and further, until any point of that region is an ϵ-Nash equilibrium. Namely, compute a Lipschitz constant L_∞ of Γ as described in Appendix A.3, and run the subdivision method in Algorithm 1.

Algorithm 1 Subdivison Search for a Nash Equilibrium

while diam $\geq \frac{\epsilon}{2L_\infty}$ **do**
  for $i \in [N], j \in [\ell^{(i)}], k \in [m_j^{(i)}]$ **do**
    if $\exists\mu\forall\pi :\ (1) \wedge \mu_{jk}^{(i)} \leq \frac{y_{jk}^{(i)} + z_{jk}^{(i)}}{2}$ **then**
      $z_{jk}^{(i)} \leftarrow \frac{y_{jk}^{(i)} + z_{jk}^{(i)}}{2}$
    else
      $y_{jk}^{(i)} \leftarrow \frac{y_{jk}^{(i)} + z_{jk}^{(i)}}{2}$
    end if
    Update B accordingly
  end for
  diam $\leftarrow$ diam/2
end while

After each for loop, the box B shrinks by $1/2$ along each dimension, while making sure that the the sentence to (1) remains true. Therefore, once the while loop terminates, there (still) is a profile $\hat{\mu} \in B$ that is an exact Nash equilibrium for Γ. Select any point μ that satisfies the linear (in-)equality system $S(\mu) \wedge B(\mu)$. Then, due to termination condition, we have $||\mu - \hat{\mu}||_\infty < \frac{\epsilon}{2L_\infty}$. All in all, we can therefore derive by analogous reasoning to the Σ_2^{P}-membership proof of Theorem 2 that μ is an ϵ-Nash equilibrium of Γ.

Running time: Let us assume oracle access to an $\exists\forall\mathbb{R}$ solver for a second. The diameter of B w.r.t. the infinity norm starts at diam $= 1$, and it halves once after each finished for loop. Any for loop takes $\mathcal{O}(m)$ time. This makes the subdivison algorithm take $\mathcal{O}\Big(m \cdot \big(\log L_\infty + \log\frac{1}{\epsilon}\big)\Big)$ time, which is polytime in instance (Γ, ϵ). Therefore, the bounds $y_{jk}^{(i)}$ and $z_{jk}^{(i)}$ remain polysized. On the other hand, observe that the maximal degree of the polynomial functions in (1) is bounded by the maximal tree depth in Γ, which is in turn bounded by $|\mathcal{H}|$. Hence, by the discussion in Appendix A.5, each $\exists\forall\mathbb{R}$-sentence can be decided in running time

$$\begin{aligned} &\mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}\Big) \cdot \Big(\mathcal{O}(m) \cdot |\mathcal{H}|\Big)^{\mathcal{O}(m^2)} \\ &= \mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\Big)\,. \end{aligned}$$

Finally, solving a linear (in-)equality system to get the point μ takes $\mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}\Big)$-time. This gives the overall running time bound

$$\begin{aligned} &\mathcal{O}\Big(m \cdot \big(\log L_\infty + \log\frac{1}{\epsilon}\big)\Big) \cdot \mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\Big) \\ &= \mathsf{poly}\Big(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\Big)\,. \end{aligned}$$

$\square$

C On Section 4

In this section, we expand on the technical background needed on multiselves equilibria and prove the claims made in Section 4. To that end, we restate results taken from the main body, and give new numbers to results presented first in this appendix.

C.1 On CDT Utilities and Equilibria

CDT Utilities and Derivatives

Lemma 26 (Piccione and Rubinstein; Oesterheld and Conitzer, 1997; 2022)**.** *For all player $i \in \mathcal{N}$, infosets $I \in \mathcal{I}^{(i)}$,* pure *actions $a \in A_I$, and strategy $\mu \in S$, we have*

$$\nabla_{I,a} U^{(i)}(\mu) = \sum_{h \in I} \mathbb{P}(h \mid \mu) \cdot U^{(i)}(\mu \mid ha)\,,$$

where the l.h.s. denotes the partial derivative of utility function $U^{(i)}$ w.r.t. to action a of $I \in \mathcal{I}^{(i)}$ at point μ.

Proof. Take a player $i \in \mathcal{N}$. Recall that $U^{(i)}$ is a polynomial function over strategy set $S \subset \bigtimes_{i' \in [N]} \bigtimes_{j' \in [\ell^{(i')}]} \mathbb{R}^{m_{j'}^{(i')}}$. Fix an infoset $j \in [\ell^{(i)}]$ and pure action $k \in [m_j^{(i)}]$. Let $e_{jk}^{(i)}$ denote the unit vector in direction of dimension (i,j,k) of S. The partial derivative of $U^{(i)}$ in that dimension and at a strategy μ is then defined as

$$\nabla_{jk} U^{(i)}|_\mu := \lim_{\epsilon \to 0} \frac{1}{\epsilon} \cdot \Big(U^{(i)}(\mu + \epsilon \cdot e_{jk}^{(i)}) - U^{(i)}(\mu) \Big)\,.$$

Note that $x := \mu + \epsilon \cdot e_{jk}^{(i)}$ is not a profile since its action values at infoset I_j sum up to $1+\epsilon$. Nonetheless, utility $U^{(i)}$ and reach probability $\mathbb{P}$ as polynomials are still well defined there. Thus

$$U^{(i)}(x) = \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid x) \cdot u^{(i)}(z) = (\dagger)\,.$$

Here, product $\mathbb{P}(z \mid x)$ is equal to the product $\mathbb{P}(z \mid \mu)$, except that factor $\mu_{jk}^{(i)}$ is replaced by factor $\mu_{jk}^{(i)} + \epsilon$. This factor occurs as often in that product as a_k needs to be played in the history of z. Multiply out this product and sort the resulting sum by their order in ϵ:

$$\begin{aligned}
(\dagger) &= \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid \mu) \cdot u^{(i)}(z) \\
&\quad + \sum_{z \in \mathcal{Z}} \Big(u^{(i)}(z) \cdot \sum_{h \in \mathrm{hist}(z) \cap I_j} \mathbb{P}(h \mid \mu) \cdot \epsilon \cdot \mathbb{P}(z \mid \mu, ha_k) \Big) \\
&\quad + \mathcal{O}(\epsilon^2) \\
&= U^{(i)}(\mu) + (\star) + \mathcal{O}(\epsilon^2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\nabla_{jk} U^{(i)}|_\mu &= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \cdot \Big(U^{(i)}(\mu) + (\star) + \mathcal{O}(\epsilon^2) - U^{(i)}(\mu) \Big) \\
&= \lim_{\epsilon \to 0} \frac{(\star)}{\epsilon} \\
&= \sum_{z \in \mathcal{Z}} \sum_{h \in I_j} u^{(i)}(z) \cdot \mathbb{P}(h \mid \mu) \cdot \mathbb{P}(z \mid \mu, ha_k) \\
&= \sum_{h \in I_j} \mathbb{P}(h \mid \mu) \cdot \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid \mu, ha_k) \cdot u^{(i)}(z) \\
&= \sum_{h \in I_j} \mathbb{P}(h \mid \mu) \cdot U^{(i)}(\mu \mid ha_k)\,.
\end{aligned}$$

□

Alternative Characterizations

Remark 27. *The CDT utility of a player $i \in \mathcal{N}$ in Γ at infoset $I \in \mathcal{I}^{(i)}$ from randomized action $\alpha \in \Delta(A_I)$ under profile μ satisfies the linearity property*

$$U_{\text{CDT}}^{(i)}(\alpha \mid \mu, I) = \sum_{a \in A_I} \alpha(a) \cdot U_{\text{CDT}}^{(i)}(a \mid \mu, I)\,.$$

Proof. We have

$$\begin{aligned}
&\sum_{a \in A_I} \alpha(a) \cdot U_{\text{CDT}}^{(i)}(a \mid \mu, I) \\
&= \sum_{a \in A_I} \alpha(a) \cdot \Big(U^{(i)}(\mu) + \nabla_{I,a} U^{(i)}(\mu) \\
&\qquad - \sum_{a' \in A_I} \mu(a' \mid I) \cdot \nabla_{I,a'} U^{(i)}(\mu) \Big) \\
&= \sum_{a \in A_I} \alpha(a) \cdot \Big(U^{(i)}(\mu) - \sum_{a' \in A_I} \mu(a' \mid I) \cdot \nabla_{I,a'} U^{(i)}(\mu) \Big) \\
&\qquad + \sum_{a \in A_I} \alpha(a) \cdot \nabla_{I,a} U^{(i)}(\mu) \\
&= U^{(i)}(\mu) - \sum_{a' \in A_I} \mu(a' \mid I) \cdot \nabla_{I,a'} U^{(i)}(\mu) \\
&\qquad + \sum_{a \in A_I} \alpha(a) \cdot \nabla_{I,a} U^{(i)}(\mu) \\
&= U_{\text{CDT}}^{(i)}(\alpha \mid \mu, I)\,.
\end{aligned}$$

□

Lemma 28. *A profile μ is a CDT equilibrium for game Γ if and only if for all player $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all alternative* pure *actions $a \in A_I$, we have*

$$U_{\text{CDT}}^{(i)}(a \mid \mu, I) \geq \max_{a' \in A_I} U_{\text{CDT}}^{(i)}(a' \mid \mu, I)\,.$$

Proof. Using Remark 27, we have for all $i \in [N]$ and $j \in [\ell^{(i)}]$

$$\begin{aligned}
U_{\text{CDT}}^{(i)}(\mu_{j\cdot}^{(i)} \mid \mu, I_j) &= \sum_{k \in [m_j^{(i)}]} \mu_{jk}^{(i)} \cdot U_{\text{CDT}}^{(i)}(a_k \mid \mu, I_j) \\
&= \sum_{k \in \mathrm{supp}(\mu_{j\cdot}^{(i)})} \mu_{jk}^{(i)} \cdot U_{\text{CDT}}^{(i)}(a_k \mid \mu, I_j)\,,
\end{aligned}$$

and analogously for randomized action α instead of $\mu_{j\cdot}^{(i)}$. This allows us to see that randomized action $\mu_{j\cdot}^{(i)}$ is not optimal in $\Delta(A_{I_j})$ if and only if it does not solely randomize over optimal pure actions. □

Unreasonable CDT Utilities In Distance

As a first-order Taylor approximation of $U^{(i)}$, the ex-ante CDT-utility may yield unreasonable utility payoffs for values α far away from $\mu(\cdot \mid I)$. Consider Figure 8. Say, the player enters the game with the strategy μ that always continues at I, the player arrives at I, and considers a deviation to action α that deterministically exits now. Then

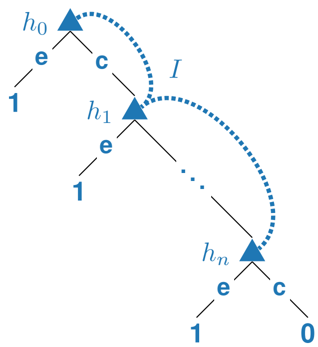


Figure 8: A single-player single-infoset game with n nodes. Exiting at any of them gives the player a utility of 1.

$$
\begin{aligned}
U^{(1)}_{\text{CDT}}(\alpha \mid \mu, I) &= 0 + \nabla_{I,\text{exit}} U^{(1)}(\mu) - \nabla_{I,\text{continue}} U^{(1)}(\mu) \\
&= \sum_{h_i \in I} \mathbb{P}(h_i \mid \mu) \cdot U^{(1)}(\mu \mid h_i \circ \text{exit}) - \\
&\quad \sum_{h_i \in I} \mathbb{P}(h_i \mid \mu) \cdot U^{(1)}(\mu \mid h_i \circ \text{continue}) \\
&= n \cdot 1 - n \cdot 0 = n \,.
\end{aligned}
$$

This is unreasonable to expect because the game only has payoffs between 0 and 1. However, we can give the following bound on CDT utilities:

$$
\begin{aligned}
&|U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I)| \\
&= |U^{(i)}(\mu) + \sum_{a \in A_I} (\alpha(a) - \mu(a \mid I)) \cdot \nabla_{I,a} U^{(i)}(\mu)| \\
&\leq |U^{(i)}(\mu)| + ||\alpha - \mu(\cdot \mid I)||_\infty ||\nabla_{I,\cdot} U^{(i)}(\mu)||_\infty \qquad (2) \\
&\overset{(*)}{\leq} |U^{(i)}(\mu)| + 1 \cdot |\mathcal{H}||U^{(i)}(\mu)| \\
&\leq 2|\mathcal{H}| \max_{z \in \mathcal{Z}} |u^{(i)}(z)| \,,
\end{aligned}
$$

where in $(*)$ we used Lemma 26.

Computational Results Inserting Definition 13 into Definition 14 immediately yields

Remark 29. *A profile μ is an ϵ-CDT equilibrium ($\epsilon \geq 0$) for game Γ if and only if for all player $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all alternative randomized actions $\alpha \in \Delta(A_I)$, we have*

$$
\sum_{a \in A_I} \mu(a \mid I) \cdot \nabla_{I,a} U^{(i)}(\mu) \geq \sum_{a \in A_I} \alpha(a) \cdot \nabla_{I,a} U^{(i)}(\mu) - \epsilon \,.
$$

Proposition (Restatement of Proposition 16; Tewolde *et al.*, 2023)**.**

1. *A profile μ is a CDT equilibrium of Γ if and only if for all player $i \in \mathcal{N}$, strategy $\mu^{(i)}$ is a KKT-point of*
$$\max_{\pi^{(i)} \in S^{(i)}} U^{(i)}(\pi^{(i)}, \mu^{(-i)}).$$
2. *Problem* 1P-CDT-S *is* CLS*-complete.*

Proof. We refer to Tewolde *et al.* [2023][Theorem 1 and 2] for these results. We shall use characterization Remark 29 which is analogous to their Definition 9, except that ours is in the ex-ante perspective, so make use of Lemma 26. Tewolde *et al.* prove the KKT correspondence for single-player settings. Our proof works analogously because the single-player result then implies the multi-player result: A profile μ is a CDT-equilibrium of Γ if and only if for each player $i \in \mathcal{N}$, their strategy $\mu^{(i)}$ is a CDT-equilibrium of the single-player version of Γ where all players $i' \neq i$ play fixed the strategy $\mu^{(i')}$.

We highlight that all KKT conditions together for a point $\mu \in \times_{i \in [N]} \times_{j \in [\ell^{(i)}]} \mathbb{R}^{m_j^{(i)}}$ become that there exist KKT multipliers $\{\tau^{(i)}_{jk} \in \mathbb{R}\}_{i,j,k=1}^{N,\ell^{(i)},m_j^{(i)}}$ and $\{\kappa^{(i)}_j \in \mathbb{R}\}_{i,j=1}^{N,\ell^{(i)}}$ such that

$$
\begin{aligned}
&\mu^{(i)}_{jk} \geq 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m^{(i)}_j] \\
&\sum_{k=1}^{m_j^{(i)}} \mu^{(i)}_{jk} = 1 \quad \forall i \in [N], \forall j \in [\ell^{(i)}] \\
&\tau^{(i)}_{jk} \geq 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m^{(i)}_j] \\
&\tau^{(i)}_{jk} = 0 \quad \text{or} \quad \mu^{(i)}_{jk} = 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m^{(i)}_j] \\
&\nabla_{jk} U^{(i)}(\mu) + \tau^{(i)}_{jk} - \kappa^{(i)}_j = 0 \, \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m^{(i)}_j]
\end{aligned}
\qquad (3)
$$

□

Next, we remark that Lemma 28 motivates another notion of approximate CDT equilibrium: A profile μ is said to be an ϵ-well-supported CDT equilibrium for a game if it satisfies inequality Lemma 28 up to ϵ relaxation on the r.h.s..

This approximation concept is polynomially precision-related to ϵ-CDT equilibria, and it has a close connection to approximate KKT points.

Lemma 30 (Tewolde *et al.*, 2023)**.** *Let μ be a profile of a game Γ with imperfect recall. Then:*

1. *If μ is an ϵ-well-supported CDT equilibrium of Γ then it is also an ϵ-CDT equilibrium of Γ*
2. *If μ is an ϵ CDT equilibrium of Γ then we can compute a $(3L_\infty|\mathcal{H}|\sqrt{\epsilon})$-well-supported CDT equilibrium of Γ, where L_∞ is the Lipschitz constant obtained as in Appendix A.3.*
3. *If μ is an ϵ-well-supported CDT equilibrium of Γ then it is an ϵ-KKT point to Proposition 16.1, that is, it satisfies KKT conditions (3), except the last one that is replaced by*
$$|\nabla_{jk} U^{(i)}(\mu) + \tau^{(i)}_{jk} - \kappa^{(i)}_j| \leq \epsilon \,.$$

Proof. Analogous Tewolde *et al.* [2023][Section H.3]. □

Corollary (Restatement of Corollary 17)**.** 1P-CDT-S *for* 1/poly *precision is in* P.

Proof. Let the reward range of the game be in $[0, 1]$. Then the utility function $U^{(1)}$ and its gradient are both L_∞-Lipschitz continuous for $L_\infty = \mathsf{poly}(|\Gamma|)$ (Appendix A.3). Thus, it suffices to run projected gradient descent for $\mathsf{poly}(|\Gamma|, 1/\epsilon)$ steps to obtain an ϵ-KKT point, which, in return, makes an ϵ-CLS equilibrium of Γ. A formal analysis of can be found, for example, in [Fearnley *et al.*, 2023, Lemma C.4].) $\square$

C.2 On Comparing the Solution Concepts

Proposition (Restatement of Proposition 18; Oesterheld and Conitzer, 2022)**.** *A Nash equilibrium is an EDT equilibrium. An EDT equilibrium is a CDT equilibrium.*

Proof. If μ is a Nash equilibrium, then every player $i \in \mathcal{N}$ plays their optimal strategy in $S^{(i)}$ in response to the profile of the others. In particular, for each of her infosets I_j, she plays the optimal randomized action in $\Delta^{m_j^{(i)}-1}$ in response to her own strategy at other infosets $j' \neq j$ and to the profile of the others. Therefore, μ is an EDT-equilibrium.

If μ is an EDT-equilibrium, then for every player $i \in \mathcal{N}$ and each of her infosets I_j, we have that $\mu_{j\cdot}^{(i)}$ is the global optimum of

$$\max_{\alpha \in \Delta^{m_j^{(i)}-1}} U^{(i)}(\mu^{(i)}_{I_j \mapsto \alpha}, \mu^{(-i)}) \, .$$

In particular, it will be a KKT point of this maximization problem. These KKT conditions, grouped together for all infosets I_j, coincide with with the KKT conditions of Proposition 16. Therefore, μ is a CDT-equilibrium. $\square$

Remark (Restatement of Remark 19)**.** *Without absentmindedness, deviation incentives of EDT and of CDT coincide, and so do the equilibrium concepts. Hence, complexity results such as Proposition 16 and Theorem 6 will apply to EDT equilibria as well.*

Proof. If there is no absentmindedness in a game Γ, then each player enters each of her infosets at most once within the same game play. Therefore, utility function $U^{(i)}(\mu) = U^{(i)}(\mu^{(i)}_{I \mapsto \mu(\cdot|I)}, \mu^{(-i)})$ is linear in an action probability $\mu(a \in I)$. In particular, the first order approximation $U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I)$ of $U^{(i)}(\mu^{(i)}_{I \mapsto \alpha}, \mu^{(-i)})$ becomes exact. Hence, EDT utilities and CDT utilities coincide, yielding coinciding equilibrium sets.

$\square$

Remark (Restatement of Remark 20)**.** *If each player has only one infoset in total, then the EDT equilibria coincide with the Nash equilibria.*

Proof. If a player $i \in \mathcal{N}$ has only one infoset I, then $S^{(i)} = \Delta(A_I)$ as well as $U^{(i)}(\mu^{(i)}_{I \mapsto \alpha}, \mu^{(-i)}) = U^{(i)}(\alpha, \mu^{(-i)})$ for $\alpha \in \Delta(A_I)$. Hence, the Definitions 2 and 10 coincide. $\square$

D On Section 5

In this section, we prove the results in Section 5. To that end, we restate results taken from the main body, and give new numbers to definitions and results presented first in this appendix.

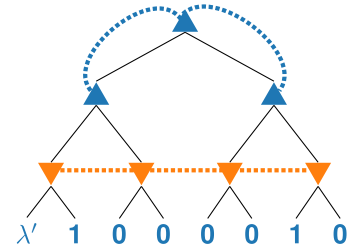


Figure 9: As in Figure 5 but with payoffs first shifted by 1 and then scaled with $1/4$, such that $\lambda' \equiv (\lambda + 1)/4$.

D.1 On the Existence of EDT Equilibria

Recall the absentminded penalty shoot-out in Figure 5, and also consider the variation of Figure 9 in which its payoffs are shifted and scaled.

Lemma 31. *In the games of Figure 5 for value λ and Figure 9 for value $\lambda' = (\lambda+1)/4$, we have coinciding EDT and Nash equilibrium sets.*

Proof. First note that EDT and Nash equilibria coincide within each game due to Remark 20. Thus we may turn our attention to Nash equilibria. The Nash equilibrium sets of the two games coincide because they are positive affine transforms of each other [Tewolde and Conitzer, 2024]: P1 maximizes her utility function $U^{(1)}(\cdot, \mu^{(-1)})$. But its maximum stays the same even if one adds constant 1 and scales it afterwards with positive $1/4$. Hence, P1's best response sets stay the same in both games. Analogously for P2 whose utility function was shifted by -1 and scaled by positive $1/4$. In particular, any Nash equilibria in one game (if existent at all) will remain Nash equilibria in the other game. $\square$

Lemma (Restatement of Lemma 21)**.** *Figure 5 has an EDT equilibrium if and only if $\lambda \geq 3$.*

Proof. Again, the first equivalence follows from Remark 20.

For the second equivalence note that due to Lemma 31, Figure 5 has a Nash equilibrium for some λ if and only if Figure 9 has a Nash equilibrium for $\lambda' = (\lambda+1)/4$. Hence, we can show instead that Lemma 31 has a Nash equilibrium if and only if $\lambda' \geq 1$.

Suppose $\lambda' \geq 1$ in Figure 9. Then P2 has a (weakly) dominant strategy of playing right. P1 best responds to that with playing left. This profile forms a Nash equilibrium.

Suppose $\lambda' < 1$ in Figure 9. For the sake of contradiction, suppose furthermore that the game has a Nash equilibrium μ^*.

First, note that this game has the following best response cycle: P1 plays left $\implies$ P2 plays left $\implies$ P1 plays right $\implies$ P2 plays right $\implies$ P1 plays left. Therefore, μ^* cannot contain a pure strategy because this would lead to a contradiction due to the best response cycle.

Therefore, μ^* is fully randomized. Let $0 < s^*, t^* < 1$ be the probabilities with which P1 and P2 play left in μ^* respectively. Note that the strategy spaces of P1 and P2 are fully determined by the probability put on left, which is the interval $[0, 1]$. As a Nash equilibrium of the game, s^* and t^* have to simultaneously maximize $U^{(1)}(\cdot, t^*)$ and $U^{(2)}(s^*, \cdot)$ over $[0, 1]$ respectively. Since s^* and t^* both lie in the interior, they must therefore, in particular, be stationary points. We obtain

$$\begin{aligned} U^{(1)}(s, t^*) &= s^2 \cdot \big(t^*\lambda' + (1 - t^*)\big) + (1 - s)^2 t^* \\ &= s^2 \cdot \big(t^*\lambda' + 1\big) - 2st^* + t^* \,. \end{aligned}$$

Hence for P1's perspective

$$\begin{aligned} 0 &= \frac{d}{ds} U^{(1)}(\cdot, t^*)\Big|_{s=s^*} = \Big(2s \cdot \big(t^*\lambda' + 1\big) - 2t^*\Big)\Big|_{s=s^*} \\ &= 2s^* \cdot \big(t^*\lambda' + 1\big) - 2t^* \,. \end{aligned}$$

In particular, $t^*\lambda' + 1 \neq 0$, because otherwise would imply $t^* = 0$. Therefore,

$$s^* = \frac{t^*}{t^*\lambda' + 1} \,. \tag{4}$$

We can now deduce that actually $t^*\lambda' + 1 > 0$ is the case because otherwise

$$t^*\lambda' + 1 < 0 \implies 0 > \frac{t^*}{t^*\lambda' + 1} = s^* \implies \text{Contradiction.}$$

Finally, we can deduce that s^* could not have been a best response to t^* because $s = 0$ performs better than s^*. We get $U^{(1)}(0, t^*) = t^*$ and therefore

$$\begin{aligned} &U^{(1)}(0, t^*) - U^{(1)}(s^*, t^*) \\ &= t^* - (s^*)^2 \cdot \big(t^*\lambda' + 1\big) + 2s^*t^* - t^* \\ &\overset{(4)}{=} -\frac{(t^*)^2}{t^*\lambda' + 1} + \frac{2(t^*)^2}{t^*\lambda' + 1} = \frac{(t^*)^2}{t^*\lambda' + 1} > 0 \,. \end{aligned}$$

Hence, (s^*, t^*) could not have been a Nash equilibrium. ☐

Theorem (Restatement of Theorem 3)**.** *Deciding whether a game with imperfect recall admits an EDT equilibrium is* $\exists\mathbb{R}$*-hard and in* $\exists\forall\mathbb{R}$*. Hardness holds even for 2p0s games where one player has a degree of absentmindedness of* 4 *and the other player has perfect recall.*

Proof. $\exists\forall\mathbb{R}$-membership follows because we can formulate the question of EDT equilibrium existence as the sentence

$$\exists \mu \forall \pi : \mu \in S \wedge \Bigg[\pi \notin S \vee \bigwedge\nolimits_{i \in [N], j \in [\ell^{(i)}]} U^{(i)}(\mu) \geq U^{(i)}(\mu^{(i)}_{I^{(i)}_j \mapsto \pi^{(i)}_{j \cdot}}, \mu^{(-i)}) \Bigg] \,.$$

$\exists\mathbb{R}$-hardness follows from a reduction from Proposition 3. Let (Γ, t) be an instance of that decision problem. W.l.o.g. we can assume $t = 3$ (otherwise first shift the payoffs in Γ by $1 - t$). Create a new game G by attaching Γ to the game G' of Figure 5, by replacing the terminal node on the bottom left of G' with action history (left, left, left) with the root node of Γ. Let P1 play the subgame Γ in G and receive its payoffs. Payoffs of P2 in that subgame shall simply be the negative of the payoffs of P1. This is a polytime construction. Let us show equivalence of the decision problems.

Suppose a utility ≥ 3 can be achieved in Γ. Then, there is also an optimal strategy π for Γ with utility ≥ 3. By Proposition 18, this is also an EDT equilibrium of Γ. Then, the profile ((left, π), right) makes an EDT equilibrium in G: P1 cannot improve at her first (and only relevant) infoset because P2 is going right, and P2 cannot improve at his infoset because he would receive a utility (loss) ≥ 3 upon going left.

Suppose G has an EDT equilibrium μ. Let λ be the de-se utility that μ achieves in subgame Γ. Observe that the infosets of subgame Γ in G are separated from the other infosets in G. In particular, P1 knows when she is in subgame Γ of G, and P2 does not get to act in that subgame. Therefore, μ restricted to the first infoset of each player must make an EDT equilibrium of Figure 5 for that value λ. By Lemma 21, we obtain $\lambda > 3$. Hence, a utility of ≥ 3 can be achieved in Γ.

About the hardness restrictions: Gimbert *et al.* [2020] show hardness of Proposition 3 even for degree 4 polynomials, that is, games with degree of absentmindedness 4. Moreover, G is a 2p0s game in which P2 has perfect recall. ☐

Theorem (Restatement of Theorem 4)**.** EDT-D *is* Σ_2^{P}*-complete. Hardness holds for* 1/poly *precision and 2p0s games with one infoset per player and a degree of absentmindedness of* 4*.*

Proof. Σ_2^{P}-membership: EDT-D is the special case of NASH-D in which each infoset is played by a new additional player. Thus membership follows from Theorem 2.

Σ_2^{P}-hardness: We reduce from the Σ_2^{P}-complete problem $\exists\forall$3-DNF-SAT [Stockmeyer, 1976][Section 4], which is the following problem: given a 3-DNF formula $\phi(x, y)$ with $k-2$ clauses where $x \in \{0,1\}^{m-1}$ and $y \in \{0,1\}^{n-1}$, decide whether $\exists x \, \forall y \; \phi(x, y)$. We consider the standard multilinear form of a DNF formula $\phi : \mathbb{R}^{m-1} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by replacing $\wedge$ with multiplication, $\vee$ with addition, and $\neg z$ with $1 - z$, so that, for $x \in \{0,1\}^{m-1}$ and $y \in \{0,1\}^{n-1}$, $\phi(x, y)$ is the number of clauses satisfied by (x, y). In particular, formula $\phi(x, y)$ is satisfied if and only if $\phi(x, y) \geq 1$.

Construction (Part 1) First, we add variables x_m and y_n and the two clauses $x_m \wedge \neg y_n$ and $\neg x_m \wedge y_n$ to ϕ, to get a formula

$$\begin{aligned} &\phi' : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \\ &(x, y) \mapsto \phi(x, y) + x_m(1 - y_n) + (1 - x_m)y_n . \end{aligned}$$

Note that ϕ' is also a 3-DNF formula with $m + n$ variables and k clauses, and ϕ' (as a SAT formula) is $\exists\forall$-satisfiable if and only if ϕ is. That is because y_n can always be set to x_m in order to not satisfy the last two added clauses. The sole purpose of these two clauses will be, in our game construction later on, to ensure that if there is no equilibrium, there is also no ϵ-equilibrium. Next, construct a 2p0s game as follows. P1 and P2 each have one infoset of $m+1$ and $n+2$ actions. For randomized action profile $(x, y) \in \Delta(m+1) \times \Delta(n+2) =: X \times Y$, we set P1's utility function as

$$\begin{aligned} U^{(1)}(x, y) := (1 - y_{n+2}) \Big(\phi'(m x_{1:m}, n y_{1:n}) - \frac{1}{2} \Big) \\ - L\psi_m(x) + L\psi_n(y) \,. \end{aligned}$$

Here, $x_{1:m}$ denotes indexing, that is, the subvector of x comprising of the first m elements, and analogous for $y_{1:n}$. For simplicity of notation, we will hereafter write $\tilde{x} = mx_{1:m}$ and $\tilde{y} = ny_{1:n}$. Moreover, ψ_m is defined as

$$\psi_m(x) := \sum_{i=1}^{m} (mx_i)^2(1 - mx_i)^2,$$

and ψ_n is defined analogously. Lastly, L is a large number to be picked later. Note that we have $\psi_m \geq 0$, and equality if and only if $\tilde{x}_i \in \{0, 1\}$ for all $i \in \{1, \dots, m\}$. Also note that $U^{(1)}$ has polynomial degree at most 4.

Intuition With this utility function, the two players play the multilinear form of ϕ' against each other, determined by a rescaling of randomization $x_{1:m}$ and $ny_{1:n}$. In order to make those rescalings practicall boolean, we harshly punish with $L \cdot \psi$ any values in $(x_{1:m}, ny_{1:n})$ that do not rescale to close to boolean. Then, $\left(\phi'(mx_{1:m}, ny_{1:n}) - \frac{1}{2}\right)$ will be strictly positive / negative depending on if ϕ' is satisfied or not. P2 has an additional action alternative e_{n+2} that also allows him to not play this game ϕ' in the first place if he – as the minimizer – does not see how to satisfy ϕ'. Finally, actions e_{m+1} of P1 and e_{n+1} of P2 have the sole purpose to fill up the vector sum to 1 in order to represent an element of the simplex, i.e., a probability vector.

Construction (Part 2) Use Appendix A.4 to construct a 2p0s imperfect-recall game Γ whose utility function is $U^{(1)}$ (that is, $U^{(2)} = -U^{(1)}$). In particular, Γ has a degree of absentmindedness at most 4 and one infoset for each player. Due to the latter, we can speak of an ϵ-equilibrium which entails both, ϵ-EDT-equilibria and ϵ-nash equilibria. Next, set

$$R := k \cdot (\max\{m, n\})^3$$
$$\epsilon := \left(\frac{1}{28k}\right)^2$$
$$L := \frac{8R}{\epsilon}.$$

Then the construction of corresponding EDT-D instance (Γ, ϵ) takes polytime and ϵ is of 1/poly precision. The remaining goal is to proof that

Claim 1: if ϕ' is $\exists\forall$-satisfiable, then Γ admits an exact equilibrium.

Claim 2: if Γ admits an ϵ-equilibrium, then from it, we can construct an assignment $\tilde{x} \in \{0, 1\}^m$ that shows $\exists\forall$-satisfiability of ϕ'.

Those two claims together imply that Γ either admits an exact equilibrium or no ϵ-equilibrium. As a conclusion, ϕ will be a "yes" (and resp. "no") instance of $\exists\forall$3-DNF-SAT if and only if ϕ' is if and only if the corresponding EDT-D instance (Γ, ϵ) has an exact equilibrium (and resp. has no ϵ-equilibrium equilibrium), which is the decision question of EDT-D. This completes the reduction.

Claim 1: Suppose ϕ' is $\exists\forall$-satisfiable, that is, there exists $\tilde{x} \in \{0, 1\}^m$ for which $\phi'(\tilde{x}, \tilde{y}) \geq 1$ for all $\tilde{y} \in \{0, 1\}^n$. We claim that in this case the profile (x^*, y^*) where $x^* := (\tilde{x}/m, 1 - \|\tilde{x}/m\|_1)$ and $y^* := e_{n+2}$ (i.e., always play pure action $n + 2$) is an exact equilibrium. Indeed, we have $U^{(1)}(x^*, y^*) = 0$,

$$\max_{x \in X} U^{(1)}(x, y^*) = \max_{x \in X} -L\psi_m(x) = 0$$

and

$$\begin{aligned}
&\min_{y \in Y} U^{(1)}(x^*, y) \\
&= \min_{y \in Y} \left((1 - y_{n+2}) \left(\phi'(\tilde{x}, \tilde{y}) - \frac{1}{2} \right) + L\psi_n(y) \right) \\
&\geq \min_{y \in Y} (1 - y_{n+2}) \cdot \frac{1}{2} + 0 \geq 0.
\end{aligned}$$

Hence, neither player can unilaterally improve on the outcome of (x^*, y^*). Before we get to Claim 2, we will need two lemma-like observations.

(I) Best Response and Integrality We say that x is δ-*integral* for $0 < \delta < 1/2$ if $\tilde{x}_i \in [0, \delta] \cup [1 - \delta, 1 + \delta]$ for every $1 \leq i \leq m$. We define it analogously for y where we now have to check for $1 \leq j \leq n$. We claim that if x is an ϵ-best response to y (resp. y is an ϵ-best response to x), then x (resp. y) is $\sqrt{\epsilon}$-integral (note that $\sqrt{\epsilon} \leq 1/28 < 1/2$). We show its contraposition. First, observe that ϕ' has k terms of degree at most 3, and that $\tilde{x}_i, \tilde{y}_j \in [0, \max\{m, n\}]$ for all i, j. Hence, parameter R was chosen large enough such that we have for all $(x, y) \in X \times Y$ that $0 \leq \phi'(\tilde{x}, \tilde{y}) \leq R$. For the contraposition, suppose x is not $\sqrt{\epsilon}$-integral, that is, there is $i \in [m]$ such that $\tilde{x}_i \notin [0, \sqrt{\epsilon}] \cup [1 - \sqrt{\epsilon}, 1 + \sqrt{\epsilon}]$. Then

$$\psi_m(x) \geq \tilde{x}_i^2(1 - \tilde{x}_i)^2 \geq \left(\frac{1}{2}\right)^2 (\sqrt{\epsilon})^2 = \epsilon/4$$

where the second inequality is based on the fact that $\tilde{x}_i$ must have at least $1/2$ distance from either 0 or 1 (or both). Hence, by how we set L, we have

$$\begin{aligned}
U^{(1)}(x, y) &\leq 1 \cdot (R - 1/2) - L\psi_m(x) + L\psi_n(y) \\
&\leq R - \frac{8R}{\epsilon} \cdot \frac{\epsilon}{4} + L\psi_n(y) \\
&= R - 2R + L\psi_n(y) \\
&\leq L\psi_n(y) - 1.
\end{aligned}$$

But then P1 at least $1/2$ utility units of incentives to deviate to her pure action e_{m+1} because

$$\begin{aligned}
U^{(1)}(e_{m+1}, y) &\geq 1 \cdot (0 - 1/2) - L \cdot 0 + L\psi_n(y) \\
&= L\psi_n(y) - 1/2.
\end{aligned}$$

In particular, x couldn't have been an ϵ-best response to y since we set $\epsilon < 1/2$. Almost analogous reasoning yields that if y is an ϵ-best response to x, then y is $\sqrt{\epsilon}$-integral. The only difference is that we derive

$$U^{(1)}(x, y) \geq 1 \cdot (0 - 1/2) - L\psi_m(x) + 2R \geq -L\psi_m(x) + 3/4$$

and

$$U^{(1)}(x, e_{n+2}) \leq 0 \cdot (\dots) - L\psi_m(x) + 0 = -L\psi_m(x).$$

(II) Integrality and Value Approximation Suppose (x', y') are $\sqrt{\epsilon}$-integral, and let $[\cdot]$ denote element-wise rounding. Recall that we denote $\tilde{x}' = m \cdot x'_{1:m}$ and analogous for $\tilde{y}'$. Then $\tilde{x}'$ and $[\tilde{x}']$ are at most $\sqrt{\epsilon}$ distant from each other in each entry i; and same for $\tilde{y}'$ and $[\tilde{y}']$. Hence, using how we set ϵ, we can derive that

$$\begin{aligned}|\phi'(\tilde{x}', \tilde{y}') - \phi'([\tilde{x}'], [\tilde{y}'])| &\leq k\big((1+\sqrt{\epsilon})^3 - 1^3\big) \\ &\overset{\sqrt{\epsilon}<1}{\leq} k(1 + 3\sqrt{\epsilon} + 3\sqrt{\epsilon} + \sqrt{\epsilon} - 1) \\ &= 7k\sqrt{\epsilon} = 1/4\,.\end{aligned}$$

The first inequality comes from the fact that that ϕ' has k terms with unit coefficients and degree at most 3, so that the maximal difference in a term is if $\tilde{x}'_i$ (resp. $\tilde{y}'_j$) is $1 + \sqrt{\epsilon}$ and $[\tilde{x}'_i]$ (resp. $[\tilde{y}'_j]$) is 1. The second inequality uses the Binomial Theorem.

Starting on Claim 2 Suppose (x^*, y^*) is an ϵ-equilibrium. By paragraph (I), both x^* and y^* are $\sqrt{\epsilon}$-integral. Thus, we can set $\tilde{x}_i := [\tilde{x}^*_i]$ and $\tilde{y}_j := [\tilde{y}^*_j]$ for $i \in [m]$ and $j \in [n]$, and obtain $\tilde{x} \in \{0,1\}^m$ and $\tilde{y} \in \{0,1\}^n$. That is, $(\tilde{x}, \tilde{y})$ are Boolean assignments for ϕ'. We will show that for all possible assignments $y \in \{0,1\}^n$, we have $\phi'(\tilde{x}, y) \geq 1$. To that end, we will first show an intermediary fact.

Observation $\phi'(\tilde{x}, \tilde{y}) \geq 1$: For the sake of contraposition, suppose $\phi'(\tilde{x}, \tilde{y}) < 1$, that is, since $(\tilde{x}, \tilde{y})$ are exactly integral, $\phi'(\tilde{x}, \tilde{y}) = 0$. Then consider P2's alternative strategy $\hat{y} = (\tilde{y}/n, 1 - ||\tilde{y}/n||_1, 0)$. Since y^* is an ϵ-best response to x^*, we derive

$$\begin{aligned}&-\frac{3}{4}(1 - y^*_{n+2}) - L\psi_m(x^*) \\ &= (1 - y^*_{n+2})(\phi'(\tilde{x}, \tilde{y}) - \frac{3}{4}) - L\psi_m(x^*) \\ &\overset{\text{(II)}}{\leq} (1 - y^*_{n+2})(\phi'(\tilde{x}^*, \tilde{y}^*) + \frac{1}{4} - \frac{3}{4}) - L\psi_m(x^*) + L\psi_n(y^*) \\ &= U^{(1)}(x^*, y^*) \leq U^{(1)}(x^*, \hat{y}) + \epsilon \\ &= 1 \cdot (\phi'(\tilde{x}^*, \tilde{y}) - \frac{1}{2}) - L\psi_m(x^*) + 0 + \epsilon \\ &\overset{\text{(II)}}{\leq} \phi'(\tilde{x}, \tilde{y}) + \frac{1}{4} - \frac{1}{2} - L\psi_m(x^*) + \frac{1}{8} \\ &= 0 - \frac{1}{8} - L\psi_m(x^*)\,.\end{aligned}$$

This simplifies to $-\frac{3}{4}(1 - y^*_{n+2}) \leq -\frac{1}{8}$, that is,

$$1 - y^*_{n+2} \geq \frac{1}{6}\,. \tag{5}$$

This allows us to show that x^* could not have been an ϵ-best response to y^*. Consider P1's alternative assignment $\tilde{z} = (\tilde{x}_{1:m-1}, 1 - \tilde{x}_m) \in \{0,1\}^m$, which is $\tilde{x}$ except for the last bit flipped. Since we started with $\phi'(\tilde{x}, \tilde{y}) = 0$, we in particular have $\tilde{x}_m(1 - \tilde{y}_n) + (1 - \tilde{x}_m)\tilde{y}_n = 0$. But then $\tilde{z}_m(1 - \tilde{y}_n) + (1 - \tilde{z}_m)\tilde{y}_n = 1$, and thus $\phi'(\tilde{z}, \tilde{y}) = 1$. Moreover, we also have versus $\tilde{y}^*$ that

$$\begin{aligned}\phi'(\tilde{z}, \tilde{y}^*) &\overset{\text{(II)}}{\geq} \phi'(\tilde{z}, \tilde{y}) - 1/4 = 3/4 + 0 = 3/4 + \phi'(\tilde{x}, \tilde{y}) \\ &\overset{\text{(II)}}{\geq} 3/4 + \phi'(\tilde{x}^*, \tilde{y}^*) - 1/4 = \phi'(\tilde{x}^*, \tilde{y}^*) + 1/2\end{aligned}$$

Finally, consider P1's alternative strategy $z := (\tilde{z}/m, 1 - ||\tilde{z}/m||_1)$. Then

$$\begin{aligned}&U^{(1)}(z, y^*) = (1 - y^*_{n+2})(\phi'(\tilde{z}, \tilde{y}^*) - \frac{1}{2}) - 0 + L\psi_n(y^*) \\ &\geq (1 - y^*_{n+2})(\phi'(\tilde{x}^*, \tilde{y}^*) + \frac{1}{2} - \frac{1}{2}) - L\psi_m(x^*) + L\psi_n(y^*) \\ &= (1 - y^*_{n+2}) \cdot \frac{1}{2} + U^{(1)}(x^*, y^*) \overset{(5)}{\geq} \frac{1}{6} \cdot \frac{1}{2} + U^{(1)}(x^*, y^*) \\ &\overset{\frac{1}{12}>\epsilon}{>} U^{(1)}(x^*, y^*) + \epsilon\,.\end{aligned}$$

Therefore, x^* could not have been an ϵ-best response to y^*.

Completing Claim 2 By above observation, we have $\phi'(\tilde{x}, \tilde{y}) \geq 1$. Let $y \in \{0,1\}^n$ be any possible assignment in the 3-DNF formula $\phi'(\tilde{x}, \cdot)$. Define $\hat{y} := (y/n, 1 - ||y/n||_1, 0)$. Then, since y^* is an ϵ-best response to x^*, we can derive

$$\begin{aligned}&\phi'(\tilde{x}, y) - \frac{1}{4} - L\psi_m(x^*) \overset{\text{(II)}}{\geq} \phi'(\tilde{x}^*, y) - \frac{1}{4} - \frac{1}{4} - L\psi_m(x^*) \\ &= U^{(1)}(x^*, \hat{y}) \geq U^{(1)}(x^*, y^*) - \epsilon \\ &= (1 - y^*_{n+2})(\phi'(\tilde{x}^*, \tilde{y}^*) - \frac{1}{2}) - L\psi_m(x^*) + L\psi_n(y^*) - \epsilon \\ &\geq (1 - y^*_{n+2})(\phi'(\tilde{x}, \tilde{y}) - \frac{1}{4} - \frac{1}{2}) - L\psi_m(x^*) + L\psi_n(y^*) - \epsilon \\ &\geq (1 - y^*_{n+2})(1 - \frac{3}{4}) - L\psi_m(x^*) - \epsilon \\ &\geq -L\psi_m(x^*) - \epsilon\,.\end{aligned}$$

But this implies

$$\phi'(\tilde{x}, y) \geq \frac{1}{4} - \epsilon > 0\,.$$

Since $\phi'(\tilde{x}, y)$ is an integer, we have $\phi'(\tilde{x}, y) \geq 1$, that is, $\phi'(\tilde{x}, y)$ is satisfied. Note that $y \in \{0,1\}^n$ was chosen arbitrarily, hence, we conclude that ϕ' is $\exists\forall$-satisfiable (with $\tilde{x}$). $\square$

D.2 On the Search of EDT Equilibria in Single-Player

Theorem (Restatement of Theorem 5)**.** 1P-EDT-S *is* PLS-*complete when the branching factor is constant. Hardness holds even when the branching factor and the degree of absentmindedness are* 2*.*

We will prove this result over multiple steps. First, we consider the corresponding polynomial optimization problem to 1P-EDT-S.

Definition 32. *An instance of the search problem* NE-POLY-S *consists of*

1. *integers ℓ and $(m_j)_{j\in[\ell]}$ in binary, determining simplices $S_j := \Delta^{m_j - 1}$*
2. *a polynomial function $p : \bigtimes_{j\in[\ell]} \mathbb{R}^{m_j} \to \mathbb{R}$ in the Turing (bit) model, and*
3. *a precision parameter $\epsilon > 0$ in binary.*

A solution consists of a point $x \in S := \bigtimes_{j\in[\ell]} S_j$ such that for all $j \in [\ell]$ and $y_j \in S_j$, we have $p(x) \geq p(y_j, x_{-j}) - \epsilon$.

Lemma 33. 1P-EDT-S *is computationally equivalent to* NE-POLY-S.

Proof. This follows straightforwardly from the connection of imperfect-recall games and polynomial optimization described in Section 2 and appendix A.4. It only requires the realization that condition $p(x) \geq p(y_j, x_{-j}) - \epsilon$ corresponds to $U^{(1)}(\mu^{(1)}) \geq U^{(1)}(\mu^{(1)}_{I \mapsto \alpha}) - \epsilon$ in this connection. □

For PLS-hardness, it is enough to work on the the hypercube as a domain.

Definition 34. *An instance of the search problem* CUBE-NE-POLY-S *consists of*

1. *integer ℓ in binary, determining hypercube $[0,1]^\ell$*
2. *a polynomial function $p : \mathbb{R}^\ell \to \mathbb{R}$ in the Turing (bit) model, and*
3. *a precision parameter $\epsilon > 0$ in binary.*

A solution consists of a point $x \in [0,1]^\ell$ such that for all $j \in [\ell]$ and $y_j \in [0,1]$, we have $p(x) \geq p(y_j, x_{-j}) - \epsilon$.

Lemma 35. CUBE-NE-POLY-S *reduces to* NE-POLY-S.

Proof. Take an instance $J = (\ell, p : \mathbb{R}^\ell \to \mathbb{R}, \epsilon)$ of CUBE-NE-POLY-S. Define the corresponding NE-POLY-S instance as $\hat{J} = (\ell, (m_j)_j, \hat{p}, \epsilon)$, where $\forall j \in [\ell] : m_j := 2$ and

$$\begin{aligned} \hat{p} : \bigtimes_{j=1}^{\ell} \mathbb{R}^2 &\to \mathbb{R} \\ \Big((x_{j1}, x_{j2}) \Big)_{j=1}^{\ell} &\mapsto p(x_{11}, x_{21}, \ldots, x_{\ell 1}) \,. \end{aligned}$$

Then, if $\hat{x}^*$ is an ϵ-Nash equilibrium of $\hat{J}$, then so will be $(\hat{x}^*_{j1})_{j \in [\mathcal{I}]}$ for J. □

Next, we show that CUBE-NE-POLY-S is PLS-hard. For that, we introduce the PLS-complete problem MAXCUT/FLIP.

Let $G = (V, E, w)$ be an undirected graph, $w : E \to \mathbb{N}$ be positive edge weights, and $V = A \sqcup B$ be a vertex partition. Then, the cut of $A \sqcup B$ is defined as all the edges in between A and B:

$$\begin{aligned} &E \cap (A, B) := \\ &\{\{u, v\} = e \in E : u \in A \wedge v \in B \text{ or } u \in B \wedge v \in A\} \,. \end{aligned}$$

Its weight is defined as $w(A, B) := \sum_{e \in E \cap (A,B)} w(e)$. The FLIP neighbourhood of partition $A \sqcup B$ is the set of partitions that can be obtained from (A, B) by just moving one vertex from one part to the other:

$$\begin{aligned} &\text{FLIP}(A, B) := \\ &\Big\{ (A \cup \{b\}) \sqcup (B \setminus \{b\}) \Big\}_{b \in B} \cup \Big\{ (A \setminus \{a\}) \sqcup (B \cup \{a\}) \Big\}_{a \in A} \,. \end{aligned}$$

Definition 36. *An instance of the search problem* MAXCUT/FLIP *consists of an undirected graph $G = (V, E, w)$ with weights $w : E \to \mathbb{N}$. A solution consists of a partition $V = A \sqcup B$ that has maximal cut weight among its* FLIP *neighbourhood.*

For problems involving weighted graphs $G = (V, E, w)$, we are interested in their computational complexities in terms of $|V|$, $|E|$, and a binary encoding of all weight values.

Lemma 37 (Yannakakis [2003],Schäffer and Yannakakis [1991])**.** MAXCUT/FLIP *is* PLS*-complete.*

This allows us to proof PLS-hardness of our problems of interest.

Lemma 38. MAXCUT/FLIP *reduces to* CUBE-NE-POLY-S.

Corollary 39. CUBE-NE-POLY-S, NE-POLY-S, *and* 1P-EDT-S *are* PLS*-hard. Hardness holds even when the branching factor ($\max_j m_j$) and the degree of the polynomial / absentmindedness are* 2*.*

Proof of Lemma 38. Let $G = (V, E, w)$ be an instance of MAXCUT/FLIP. First, we create the associated CUBE-NE-POLY-S instance. Let $\ell = |V|$ such that each vertex $v \in V$ is associated to an entry x_v in $x \in S = [0,1]^\ell$. We can define for point $x \in S$ and vertices $t, v \in V$ the function

$$d_{t,v}(x) := x_t(1 - x_v) + (1 - x_t)x_v$$

which is maximized if one of the values x_t and x_v is 0 and the other is 1; corresponding to t and v belonging to different partitions. Set $W = \sum_{e \in E} w(e)$, $W' := 2(W + 1)$ and

$$p(x) = W' \sum_{v \in V} (\frac{1}{2} - x_v)^2 + \sum_{\{t,v\} \in E} w(t, v) \cdot d_{t,v}(x) \,.$$

The first summand has a large weight W' and forces any solution x^* to have values x^*_v far away from $\frac{1}{2}$. We can get the Lipschitz constant $L_\infty = 15W$ for p over S by the method described in Appendix A.3. Set $\epsilon = 1/(2L_\infty + 2) < \frac{1}{2}$.

Let x^* be a solution to this CUBE-NE-POLY-S instance. Then we claim: (1) L_∞ is actually a Lipschitz constant of p over S, (2) We have $\forall v : x^*_v \leq \epsilon \vee x^*_v \geq 1 - \epsilon$. Define $z^* \in \{0,1\}^\ell$ as $z^*_v = 0$ if $x^*_v \leq \epsilon$ and as $z^*_v = 1$ if $x^*_v \geq 1 - \epsilon$. Then, (3) partition $V = \{v \in V : z^*_v = 0\} \sqcup \{v \in V : z^*_v = 1\}$ is a solution to the original MAXCUT/FLIP instance.

Claim (1): We have for $u \in V$

$$\begin{aligned} \nabla_u \, p(x) &= -2W'(\frac{1}{2} - x_u) + \sum_{\{u,v\} \in E} w(u, v) \cdot (1 - 2x_v) \\ &= -W' + 2W'x_u + W - 2 \sum_{\{u,v\} \in E} w(u, v) x_v \end{aligned}$$

Using $W \geq 1$ and $W' \leq W$, these polynomial coefficients yield Lipschitz constant

$$W' + 2W' + W + 2W \leq 15W =: L_\infty$$

for p over the hypercube.

Claim (2): We start with x^* being an ϵ-Nash equilibrium. Suppose vertex $u \in V$ has $x_u^* \leq \frac{1}{2}$. Then

$$\begin{aligned}
\epsilon &\geq p(0, x_{-u}^*) - p(x^*) \\
&= W'(\frac{1}{2} - 0)^2 - W'(\frac{1}{2} - x_u^*)^2 \\
&\quad + \sum_{\{u,v\} \in E} w(u,v) \cdot d_{u,v}(0, x_{-u}^*) \\
&\quad - \sum_{\{u,v\} \in E} w(u,v) \cdot d_{u,v}(x^*) \\
&= W'\big(\frac{1}{4} - (\frac{1}{2} - x_u^*)^2\big) + \sum_{\{u,v\} \in E} w(u,v) \cdot (x_v^* - d_{u,v}(x^*)) \\
&= (\dagger)
\end{aligned}$$

Note that with $x_u^* \leq \frac{1}{2}$, we have

$$\frac{1}{4} - (\frac{1}{2} - x_u^*)^2 = (1 - x_u^*) x_u^* \geq \frac{1}{2} x_u^*$$

and

$$x_v^* - d_{u,v}(x^*) = x_v^* - x_u^* + x_u^* x_v^* - x_v^* + x_u^* x_v^* \geq -x_u^* \,.$$

Therefore, recalling that u was a fixed vertex, we can continue with

$$\begin{aligned}
(\dagger) &\geq W' \cdot \frac{1}{2} x_u^* + \sum_{\{u,v\} \in E} w(u,v) \cdot (-x_u^*) \\
&= x_u^* \Big(\frac{1}{2} W' - \sum_{\{u,v\} \in E} w(u,v) \Big) \geq x_u^* \Big(W + 1 - W \Big) \\
&= x_u^*
\end{aligned}$$

On the other hand, suppose that vertex $u \in V$ has $x_u^* > \frac{1}{2}$. Then

$$\begin{aligned}
\epsilon &\geq p(1, x_{-u}^*) - p(x^*) \\
&= W'\big(\frac{1}{4} - (\frac{1}{2} - x_u^*)^2\big) \\
&\quad + \sum_{\{u,v\} \in E} w(u,v) \cdot (1 - x_v^* - d_{u,v}(x^*)) \\
&\overset{(*)}{\geq} W' \cdot \frac{1}{2}(1 - x_u^*) + \sum_{\{u,v\} \in E} w(u,v) \cdot (-(1 - x_u^*)) \\
&\geq (1 - x_u^*) \Big(\frac{1}{2} W' - \sum_{\{u,v\} \in E} w(u,v) \Big) \\
&\geq 1 - x_u^*
\end{aligned}$$

which implies $x_u^* \geq 1 - \epsilon$. In $(*)$, we used

$$\begin{aligned}
1 - x_v^* - d_{u,v}(x^*) &= 1 - x_u^* - 2x_v^* + 2x_u^* x_v^* \\
&= (1 - x_u^*)(1 - 2x_v^*) \geq -(1 - x_u^*) \,.
\end{aligned}$$

Claim (3): Any point $z =\in \{0,1\}^\ell$ induces a partition

$$V = A(z) \sqcup B(z) := \{v \in V : z_v = 0\} \sqcup \{v \in V : z_v = 1\} \,.$$

Moreover, for any such point z and vertices $t, v \in V$, we have

$$d_{t,v}(z) = \begin{cases} 0 & \text{if } z_t, z_v \in A(z) \text{ or } z_t, z_v \in B(z) \\ 1 & \text{else} \end{cases} .$$

Therefore, the cut weight associated to point $z =\in \{0,1\}^\ell$ has a relationship to polynomial p in the form of

$$p(z) = W' \cdot \ell \cdot \frac{1}{4} + w\Big(A(z), B(z)\Big) \,. \tag{6}$$

Now define z^* as

$$z_v^* := \begin{cases} 0 & \text{if } x_v \leq \epsilon \\ 1 & \text{if } x_v \geq 1 - \epsilon \end{cases} .$$

Let us now show that its induced partition is a solution to the original MAXCUT/FLIP instance. Consider a vertex $u \in V$ that wants to change the part of the partition it is in. The new partition is induced by the point $(1 - z_u^*, z_{-u}^*)$. Using that p is L_∞-Lipschitz and that x^* is an ϵ-Nash equilibirum that is also ϵ-close to z^*, we get

$$\begin{aligned}
& w\Big(A(1 - z_u^*, z_{-u}^*), B(1 - z_u^*, z_{-u}^*)\Big) - w\Big(A(z^*), B(z^*)\Big) \\
&\overset{(6)}{=} p(1 - z_u^*, z_{-u}^*) - p(z^*) \\
&= p(1 - z_u^*, z_{-u}^*) - p(1 - z_u^*, x_{-u}^*) + p(1 - z_u^*, x_{-u}^*) \\
&\quad - p(x^*) + p(x^*) - p(z^*) \\
&\leq L_\infty ||(1 - z_u^*, z_{-u}^*) - (1 - z_u^*, x_{-u}^*)||_\infty \\
&\quad + \epsilon + L_\infty ||x^* - z^*||_\infty \\
&\leq L_\infty \epsilon + \epsilon + L_\infty \epsilon = \epsilon(2L_\infty + 1) \\
&< 1
\end{aligned}$$

by the choice of ϵ. Recall that edge weights are integers, and hence, also the weight of a cut. Therefore, the inequality chain above started with an integer that was shown to be strictly less than 1 at the end. We get

$$w\Big(A(1 - z_u^*, z_{-u}^*), B(1 - z_u^*, z_{-u}^*)\Big) \leq w\Big(A(z^*), B(z^*)\Big) \,,$$

proving that if vertex u changes the part of the partition it is in, then the cut weight does not increase. Since $u \in V$ was arbitrary, we have shown that partition $V = A(z^*) \sqcup B(z^*)$ has maximal weight among its FLIP neighbourhood. □

Proof of Corollary 39. Follows from Lemmata 33, 35 and 37. For the hardness restrictions, note that we started with a degree two polynomial and a hypercube. This is associated to a game tree of depth 3, with a depth of absentmindedness of at most 2, and with a number of actions per infoset of 2. □

Next, we show PLS-membership.

Lemma 40. 1P-EDT-S *and* NE-POLY-S *when the branching factor is constant is in* PLS.

Proof. By Lemma 33, it suffices to show this for 1P-EDT-S. We show it by giving a best response dynamics that can be run between the infosets in order to find an ϵ-EDT equilibrium. So let (Γ, ϵ) be a 1P-EDT-S instances.

Computing an ϵ-best response: We will now describe a method that, given a profile μ for Γ and an infoset I_j, computes an $\epsilon/2$-best response $\alpha \in \Delta^{m_j^{(1)}-1} =: S_j$ of that infoset to strategy $\mu_{-j}^{(1)}$ at other infosets. The method is similar to the one described in the proof of Proposition 9. This time, however, instead of working on the whole profile set S, we only work on the randomized action simplex S_j. We also initialize the hypercube as $B_j := [0,1]^{m_j^{(1)}}$, and describe it with bounds $(y_k, z_k)_k$. Then, the sentences we will have to solve are whether there exists ($\exists$) $\alpha \in \mathbb{R}^{m_j^{(1)}}$ such that

$$S_j(\alpha) \wedge B_j(\alpha) \wedge U^{(1)}(\alpha, \mu_{-j}^{(1)}) \geq t\,, \tag{7}$$

where $t \in \mathbb{Q}$ is a target value.

As a preprocessing step, we shall first approximate the maximal utility value $u^* \in \mathbb{R}$ achievable with an exact best response. To that regard, initialize

$$\underline{u} := \min_{z \in \mathcal{Z}} u^{(1)}(z) - 1 \text{ and } \bar{u} := \min_{z \in \mathcal{Z}} u^{(1)}(z) + 1\,.$$

Then $\underline{u} < u^* < \bar{u}$. Therefore, sentence (7) is true and false for values $t = \underline{u}$ and $t = \bar{u}$ respectively. Hence, we can do binary search on $\mathbb{R}$ to pinpoint u^* by updating the lower and upper bounds $\underline{u}$ and $\bar{u}$ accordingly such that $\underline{u} < u^* < \bar{u}$ says satisfied and until $|\bar{u} - \underline{u}| < \epsilon/4$. Then, in particular, $\hat{u} := \underline{u}$ satisfies $|u^* - \hat{u}| < \epsilon/4$.

Algorithm 2 Subdivison Search for a Best Response

while diam $\geq \frac{\epsilon}{4L_\infty}$ **do**
  for $k \in [m_j^{(i)}]$ **do**
    if $\exists \alpha :$ (7) $\wedge\ \alpha_k \leq \frac{y_k + z_k}{2}$ **then**
      $z_k \leftarrow \frac{y_k + z_k}{2}$
    else
      $y_k \leftarrow \frac{y_k + z_k}{2}$
    end if
    Update B_j accordingly
  end for
  diam $\leftarrow$ diam/2
end while

Next, we run Algorithm 2 where (7) is always invoked for value $t = \hat{u}$. Upon termination, select any point α that satisfies the linear (in-)equality system $S_j(\alpha) \wedge B_j(\alpha)$. If α^* is the point that satisfies (7), then due to termination condition, we have $||\alpha - \alpha^*||_\infty < \frac{\epsilon}{4L_\infty}$. This yields

$$\begin{aligned}
&U^{(1)}(\alpha, \mu_{-j}^{(1)}) \\
&= U^{(1)}(\alpha, \mu_{-j}^{(1)}) - U^{(1)}(\alpha^*, \mu_{-j}^{(1)}) + U^{(1)}(\alpha^*, \mu_{-j}^{(1)}) \\
&= U^{(1)}(\alpha^*, \mu_{-j}^{(1)}) - |U^{(1)}(\alpha, \mu_{-j}^{(1)}) - U^{(1)}(\alpha^*, \mu_{-j}^{(1)})| \\
&\geq \hat{u} - L_\infty \cdot ||(\alpha, \mu_{-j}^{(1)}) - (\alpha^*, \mu_{-j}^{(1)})||_\infty \\
&\geq u^* - \epsilon/4 - L_\infty \cdot \frac{\epsilon}{4L_\infty} = u^* - \epsilon/2\,.
\end{aligned}$$

Finally, the running time analysis works analogous to the one in the proof of Proposition 9, except that the number of variables "m" now is $m_j^{(1)}$. Since it is constant by assumption, we get polytime computability of such an ϵ-best response α.

ϵ-Best Response Dynamics: The best response dynamics can start at any profile $\mu \in S$, e.g., at the one that plays the first action of each infoset deterministically. The neighbourhood of an iterate $\mu \in S$ shall be all profiles of the form $(\alpha, \mu_{-j}^{(1)})$ where $j \in [\ell^{(1)}]$ and α is an $\epsilon/2$-best response to $\mu_{-j}^{(1)}$. This neighbourhood can be determined within polytime. Finally, we can evaluate the utility $U^{(1)}(\pi)$ of any iterate $\pi = \mu$ or any neighbour $\pi = (\alpha, \mu_{-j}^{(1)})$ within polytime. If there is a neighbour $(\alpha, \mu_{-j}^{(1)})$ with utility $U^{(1)}(\mu) \leq U^{(1)}(\alpha, \mu_{-j}^{(1)}) - \epsilon/2$, then take that neighbour as the next iterate, otherwise, terminate and return μ.

Let π be returned by this algorithm. Then π is an ϵ-EDT equilibrium of Γ: Take any infoset I_j. Let $(\alpha, \pi_{-j}^{(1)})$ be the neighbour associated to that infoset and u^* be the maximal utility value achievable from that infoset with an exact best response. Then, for any alternative randomized action $\alpha' \in \Delta^{m_j^{(1)}-1}$, we have

$$\begin{aligned}
U^{(1)}(\pi) &\geq U^{(1)}(\alpha, \pi_{-j}^{(1)}) - \epsilon/2 \\
&\geq u^* - \epsilon/2 - \epsilon/2 = u^* - \epsilon \\
&\geq U^{(1)}(\alpha', \pi_{-j}^{(1)}) - \epsilon/2\,.
\end{aligned}$$

This concludes the proof. □

Corollary (Restatement of Corollary 22)**.** 1P-EDT-S *for* 1/poly *precision is in* P *when the branching factor is constant.*

Proof. Let the reward range of the game be in $[0,1]$ and the desired approximation error be ϵ. Take the best response method described in the previous proof. When the branching factor is constant, it iteratively computes and transitions to a $\epsilon/2$-best response in time $\mathsf{poly}(|\Gamma|, \log(1/\epsilon))$. But this process can update the strategy at most $\mathcal{O}(1/\epsilon)$ times. Thus, if ϵ is of 1/poly size, then the method runs in time $\mathsf{poly}(|\Gamma|, 1/\epsilon)$ overall. □

D.3 On the Search of CDT Equilibria

Theorem (Restatement of Theorem 6)**.** CDT-S *is* PPAD-*complete. Hardness holds even for two-player perfect-recall games with one infoset per player and for* 1/poly *precision.*

We prove this in parts.

Lemma 41. CDT-S *is* PPAD-*hard, even for two-player perfect-recall games with one infoset per player and for* 1/poly *precision.*

Proof. For PPAD-hardness, we can use a well-known reduction from normal-form games to perfect-recall extensive-form games. In particular, we reduce from the PPAD-complete problem of computing an approximate Nash equilibrium, for inverse-polynomial precision, of a two-player normal-form game [Chen *et al.*, 2009]. The representation of such a game is the number of pure actions of each player, and all utility payoffs encoded in binary.

Starting from such an instance (G, ϵ), we can a corresponding instance (Γ, ϵ) of CDT-EQ as follows: Assign the root node of Γ to P1, with an infoset $I_1^{(1)}$, and the action set that P1 has in G. Each child of the root node shall be assigned to P2, grouped together to one infoset $I_1^{(2)}$, and with the action set that P2 has in G. Finally, each node of depth 3 is a terminal node, with utility payoffs equal to what the players would have received in G if they played the same action there that lead to this terminal node. This is a polytime construction.

Then, Γ has perfect recall, hence no absentmindedness, and it has one infoset per player. By Remarks 19 and 20, ϵ-CDT equilibria of Γ equal its ϵ-Nash equilibria which, in return, equal the ϵ-Nash equilibria in G. $\square$

Next, we show PPAD-membership of CDT-S by leveraging the general tool of Etessami and Yannakakis [2010][Section 2.3] to show that a fixed point problem is in PPAD. First, we define the fixed point function that is similar in structure to the one given by Nash [1951]. Given a game Γ with imperfect recall and a profile μ for it, define the advantage of a pure action $k \in [m_j^{(i)}]$ at infoset $j \in [\ell^{(i)}]$ of player $i \in [N]$ as

$$g_{jk}^{(i)}(\mu) := U_{\text{CDT}}^{(i)}(a_k \mid \mu, I_j) - U^{(i)}(\mu)\,.$$

If these advantages are at most ϵ for all i, j, k at μ, then μ is an ϵ-CDT equilibrium by Definition 14 and Remark 27. The fixed point mapping F we will consider is one that sends profile μ of Γ to profile $F(\mu)$ of Γ with entries

$$F(\mu)_{jk}^{(i)} := \frac{\mu_{jk}^{(i)} + \max\{0, g_{jk}^{(i)}(\mu)\}}{1 + \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\}}\,. \tag{8}$$

Intuitively, from μ to $F(\mu)$ it increases the probabilities of those actions that have a positive advantage over the currently played randomized action $\mu_{j\cdot}^{(i)}$ at $\Delta(A_{I_j})$.

Lemma 42. *Let Γ be a game with imperfect recall. Then mapping F as defined above maps the profile set S onto itself. Moreover, a profile μ is an exact CDT equilibrium of Γ if and only if it is a fixed point of F, that is, $F(\mu) = \mu$.*

Proof. If $\mu \geq 0$, then also $F(\mu) \geq 0$. Moreover, for all player i and infosets j, we have that if $\sum_{k=1}^{m_j^{(i)}} \mu_{jk}^{(i)}$ then also

$$\sum_{k=1}^{m_j^{(i)}} F(\mu)_{jk}^{(i)} = \frac{\sum_{k=1}^{m_j^{(i)}} \mu_{jk}^{(i)} + \sum_{k=1}^{m_j^{(i)}} \max\{0, g_{jk}^{(i)}(\mu)\}}{1 + \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\}} = 1\,.$$

Hence, $F(\mu)$ is a profile of Γ if μ is.

Next, suppose μ is an exact CDT equilibrium of Γ. Then, by definition, we have for all i, j, k that the advantage $g_{jk}^{(i)}(\mu)$ is non-positive, hence $F(\mu)_{jk}^{(i)} = \frac{\mu_{jk}^{(i)}+0}{1+0} = \mu_{jk}^{(i)}$, yielding $F(\mu) = \mu$.

Last but not least, suppose $F(\mu) = \mu$. Let us show that any player i plays optimally at any infoset I_j with randomized action $\mu_{j\cdot}^{(i)}$, by showing that no pure action a_k does better at I_j, using Remark 27. We show this by proving that advantage $g_{jk}^{(i)}(\mu)$ is non-positive for each action a_k. For the sake of contradiction, suppose that the subset $\mathcal{K} := \{k \in [m_j^{(i)}] : g_{jk}^{(i)}(\mu) > 0\}$ of actions with positive advantage is non-empty. Then observe that $\sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\} > 0$. Thus, all actions $\tilde{k} \notin \mathcal{K}$ with non-positive advantage satisfy

$$\mu_{j\tilde{k}}^{(i)} = F(\mu)_{j\tilde{k}}^{(i)} = \frac{\mu_{j\tilde{k}}^{(i)} + 0}{1 + \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\}}\,,$$

which implies $\mu_{j\tilde{k}}^{(i)} = 0$. Hence, $\text{supp}(\mu_{j\cdot}^{(i)}) \subseteq \mathcal{K}$. Knowing this, we can derive the contradiction

$$\begin{aligned}
U_{\text{CDT}}^{(i)}(\mu_{j\cdot}^{(i)} \mid \mu, I_j) &= \sum_{k \in [m_j^{(i)}]} \mu_{jk}^{(i)} \cdot U_{\text{CDT}}^{(i)}(a_k \mid \mu, I_j) \\
&= \sum_{k \in \text{supp}(\mu_{j\cdot}^{(i)})} \mu_{jk}^{(i)} \cdot U_{\text{CDT}}^{(i)}(a_k \mid \mu, I_j) \\
&= \sum_{k \in \text{supp}(\mu_{j\cdot}^{(i)})} \mu_{jk}^{(i)} \cdot \left(g_{jk}^{(i)}(\mu) + U_{\text{CDT}}^{(i)}(\mu_{j\cdot}^{(i)} \mid \mu, I_j)\right) \\
&> \sum_{k \in \text{supp}(\mu_{j\cdot}^{(i)})} \mu_{jk}^{(i)} \cdot U_{\text{CDT}}^{(i)}(\mu_{j\cdot}^{(i)} \mid \mu, I_j) \\
&= U_{\text{CDT}}^{(i)}(\mu_{j\cdot}^{(i)} \mid \mu, I_j)\,.
\end{aligned}$$

Thus, $\mathcal{K}$ must have been empty. Since this holds for each player and infoset, μ must have been an exact CDT equilibrium for Γ. $\square$

Next, we show that F is Lipschitz continuous for a moderately sized Lipschitz constant in terms of the game instance Γ.

Lemma 43. *Given a game Γ with imperfect recall, its mapping F from (8) is Lipschitz continuous on the profile set S with Lipschitz constant $L_F := 11|\mathcal{H}|^2 L_\infty$, where L_∞ is the Lipschitz constant of Γ as described in Appendix A.3.*

Proof. We follow the proof outline of [Daskalakis and Papadimitriou, 2011][Lemma 3.4], and show that if profiles μ and π have distance $||\mu - \pi||_\infty \leq \epsilon$, then $||F(\mu) - F(\pi)||_\infty \leq 11|\mathcal{H}|^2 L_\infty \epsilon$.

First, consider $h_{jk}^{(i)} = U_{\text{CDT}}^{(i)}(a_k \mid \cdot, I_j)$ as a function in profile $\mu' \in S$, for a given i, j, k.

$h_{jk}^{(i)}$ is Lipschitz continuous: We show this with the profiles μ and π above. We have

$$
\begin{aligned}
&|h_{jk}^{(i)}(\mu) - h_{jk}^{(i)}(\pi)| \\
&= |U^{(i)}(\mu) + \nabla_{jk} U^{(i)}(\mu) - \sum_{k' \in [m_j^{(i)}]} \mu_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\mu) \\
&\quad - U^{(i)}(\pi) - \nabla_{jk} U^{(i)}(\pi) + \sum_{k' \in [m_j^{(i)}]} \pi_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\pi)| \\
&\leq |U^{(i)}(\mu) - U^{(i)}(\pi)| + |\nabla_{jk} U^{(i)}(\mu) - \nabla_{jk} U^{(i)}(\pi)| \\
&\quad + |\sum_{k' \in [m_j^{(i)}]} \mu_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\mu) - \pi_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\pi)| \\
&\leq L_\infty \epsilon + L_\infty \epsilon \\
&\quad + \sum_{k' \in [m_j^{(i)}]} |\mu_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\mu) - \pi_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\mu) \\
&\quad + \pi_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\mu) - \pi_{jk'}^{(i)} \cdot \nabla_{jk'} U^{(i)}(\pi)| \\
&\leq 2L_\infty \epsilon + \sum_{k' \in [m_j^{(i)}]} |\nabla_{jk'} U^{(i)}(\mu)| \epsilon + |\pi_{jk'}^{(i)}| L_\infty \epsilon \\
&\leq 2L_\infty \epsilon + \sum_{k' \in [m_j^{(i)}]} (L_\infty \epsilon + 1 \cdot L_\infty \epsilon) \\
&\leq 2L_\infty \epsilon + 2|\mathcal{H}| L_\infty \epsilon = 4|\mathcal{H}| L_\infty \epsilon \,.
\end{aligned}
$$

$g_{jk}^{(i)}$ is Lipschitz continuous:

$$
\begin{aligned}
|g_{jk}^{(i)}(\mu) - g_{jk}^{(i)}(\pi)| &= |h_{jk}^{(i)}(\mu) - U^{(i)}(\mu) - h_{jk}^{(i)}(\pi) + U^{(i)}(\pi)| \\
&\leq |h_{jk}^{(i)}(\mu) - h_{jk}^{(i)}(\pi)| + |U^{(i)}(\mu) - U^{(i)}(\pi)| \\
&\leq 4|\mathcal{H}| L_\infty \epsilon + L_\infty \epsilon \leq 5|\mathcal{H}| L_\infty \epsilon \,.
\end{aligned}
$$

Therefore, by case distinction, we also get

$$
\begin{aligned}
&|\max\{0, g_{jk}^{(i)}(\mu)\} - \max\{0, g_{jk}^{(i)}(\pi)\}| \\
&\leq |g_{jk}^{(i)}(\mu) - g_{jk}^{(i)}(\pi)| \leq 5|\mathcal{H}| L_\infty \epsilon \,.
\end{aligned}
$$

F is Lipschitz continuous: We show for each entry index i, j, k

$$
\begin{aligned}
&|F(\mu)_{jk}^{(i)} - F(\pi)_{jk}^{(i)}| \\
&\overset{(*)}{\leq} |\mu_{jk}^{(i)} - \pi_{jk}^{(i)}| + |\max\{0, g_{jk}^{(i)}(\mu)\} - \max\{0, g_{jk}^{(i)}(\pi)\}| \\
&\quad |\sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\} - \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\pi)\}| \\
&\leq \epsilon + 5|\mathcal{H}| L_\infty \epsilon + |\mathcal{H}| \cdot 5|\mathcal{H}| L_\infty \epsilon \\
&\leq 11|\mathcal{H}|^2 L_\infty \epsilon \,,
\end{aligned}
$$

where in $(*)$ we used [Daskalakis and Papadimitriou, 2011][Lemma 3.6]. Thus,

$$
||F(\mu) - F(\pi)||_\infty \leq 11|\mathcal{H}|^2 L_\infty \epsilon \,. \qquad \square
$$

We call $\mu \in S$ an ϵ-fixed point of F if $||F(\mu) - \mu||_\infty < \epsilon$.

Lemma 44. *It is in* PPAD *to find an ϵ-fixed point of associated mapping F to a game Γ with imperfect recall.*

Proof. We invoke [Etessami and Yannakakis, 2010][Proposition 2.2 (2)] for this, for which we need that mapping F is *polynomially continuous* and *polynomially computable*. The former follows from Lemma 43. The latter follows because the profile set S is easy to describe and F is polytime computable. $\square$

For the next (and last) result, note that the value

$$
\theta := \max \left\{ 1 \,,\, 3|\mathcal{H}| \cdot \max_{z \in \mathcal{Z}, i \in \mathcal{N}} |u^{(i)}(z)| \right\}
$$

serves as an upper bound on values $g_{jk}^{(i)}(\mu)$. (We need the factor $|\mathcal{H}|$ because of (2)).

Lemma 45. *For any game Γ with imperfect recall, if μ is an ϵ-fixed point ($\epsilon < \frac{1}{4}$) of its associated mapping F, then μ is an ϵ'-CDT equilibrium of Γ, where $\epsilon' := 2\theta|\mathcal{H}|^{3/2}\sqrt{\epsilon}$ and θ is defined as above.*

Proof. We follow the proof outline of [Etessami and Yannakakis, 2010][Proposition 2.3]. Let μ be an ϵ-fixed point of F. Then we show that for all indices i, j, k, we have $\max\{0, g_{jk}^{(i)}(\mu)\} \leq \epsilon'$. This then implies that μ is an ϵ'-CDT equilibrium by Definition 14 and Remark 27.

Take any player i and infoset j. Then $||F(\mu) - \mu||_\infty < \epsilon$ implies for any action $k \in [m_j^{(i)}]$:

$$
\begin{aligned}
&|\max\{0, g_{jk}^{(i)}(\mu)\} - \mu_{jk}^{(i)} \cdot \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\}| \\
&\leq \epsilon \cdot \left(1 + \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\}\right) \leq \epsilon(1 + |\mathcal{H}| \cdot \theta) \\
&\leq 2\theta|\mathcal{H}|\epsilon \,.
\end{aligned} \tag{9}
$$

Next, define $\mathcal{K}^+ := \{k \in [m_j^{(i)}] : g_{jk}^{(i)}(\mu) > 0\}$ and $\mathcal{K}^- := \{k \in [m_j^{(i)}] : g_{jk}^{(i)}(\mu) < 0\}$.

Case 1: There is an index $\tilde{k} \in \mathcal{K}^-$ with $\mu_{j\tilde{k}}^{(i)} \geq \sqrt{\frac{\epsilon}{|\mathcal{H}|}}$, then for any action $k \in [m_j^{(i)}]$:

$$
\begin{aligned}
&\max\{0, g_{jk}^{(i)}(\mu)\} \leq \sqrt{\frac{|\mathcal{H}|}{\epsilon}} \cdot \sqrt{\frac{\epsilon}{|\mathcal{H}|}} \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\} \\
&\leq \sqrt{\frac{|\mathcal{H}|}{\epsilon}} \cdot \mu_{j\tilde{k}}^{(i)} \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\} \\
&\leq \sqrt{\frac{|\mathcal{H}|}{\epsilon}} \cdot \left| 0 - \mu_{j\tilde{k}}^{(i)} \sum_{k'=1}^{m_j^{(i)}} \max\{0, g_{jk'}^{(i)}(\mu)\} \right|
\end{aligned}
$$

$$\overset{\tilde{k}\in\mathcal{K}^-}{=} \sqrt{\frac{|\mathcal{H}|}{\epsilon}} \cdot \left| \max\{0, g^{(i)}_{j\tilde{k}}(\mu)\} - \mu^{(i)}_{j\tilde{k}} \sum_{k'=1}^{m^{(i)}_j} \max\{0, g^{(i)}_{jk'}(\mu)\} \right|$$
$$\overset{(9)}{\leq} \sqrt{\frac{|\mathcal{H}|}{\epsilon}} \cdot 2\theta|\mathcal{H}|\epsilon = 2\theta|\mathcal{H}|^{3/2}\sqrt{\epsilon} = \epsilon' \,.$$

Case 2: For all indices $\tilde{k} \in \mathcal{K}^-$, we have $\mu^{(i)}_{j\tilde{k}} < \sqrt{\frac{\epsilon}{|\mathcal{H}|}}$. We first have to observe that

$$\begin{aligned}
&\sum_{k'=1}^{m^{(i)}_j} \mu^{(i)}_{jk'} \cdot \max\{0, g^{(i)}_{jk'}(\mu)\} = \sum_{k'\in\mathcal{K}^+} \mu^{(i)}_{jk'} \cdot g^{(i)}_{jk'}(\mu) \\
&= \sum_{k'=1}^{m^{(i)}_j} \mu^{(i)}_{jk'} \cdot g^{(i)}_{jk'}(\mu) - 0 - \sum_{\tilde{k}\in\mathcal{K}^-} \mu^{(i)}_{j\tilde{k}} \cdot g^{(i)}_{j\tilde{k}}(\mu) \\
&= \sum_{k'=1}^{m^{(i)}_j} \mu^{(i)}_{jk'} \cdot U^{(i)}_{\text{CDT}}(a_k \mid \mu, I_j) - \sum_{k''=1}^{m^{(i)}_j} \mu^{(i)}_{jk''} \cdot U^{(i)}(\mu) \\
&\quad - \sum_{\tilde{k}\in\mathcal{K}^-} \mu^{(i)}_{j\tilde{k}} \cdot g^{(i)}_{j\tilde{k}}(\mu) \\
&= U^{(i)}_{\text{CDT}}(\mu^{(i)}_{j\cdot} \mid \mu, I_j) - U^{(i)}(\mu) - \sum_{\tilde{k}\in\mathcal{K}^-} \mu^{(i)}_{j\tilde{k}} \cdot g^{(i)}_{j\tilde{k}}(\mu) \\
&= \sum_{\tilde{k}\in\mathcal{K}^-} \mu^{(i)}_{j\tilde{k}} \cdot (-g^{(i)}_{j\tilde{k}}(\mu)) \leq |\mathcal{H}|\sqrt{\frac{\epsilon}{|\mathcal{H}|}} \cdot \theta \\
&\leq \theta\sqrt{|\mathcal{H}|}\sqrt{\epsilon} \,.
\end{aligned}$$

Next, set $k^* = \text{argmax}_{k\in\mathcal{K}^+}\, g^{(i)}_{jk}(\mu)$.

Case 2.1: Probability $\mu^{(i)}_{jk^*} \geq \frac{1}{2|\mathcal{H}|}$. Then for any action $k \in [m^{(i)}_j]$:

$$\begin{aligned}
&\max\{0, g^{(i)}_{jk}(\mu)\} \leq \max\{0, g^{(i)}_{jk^*}(\mu)\} \\
&= \frac{2|\mathcal{H}|}{2|\mathcal{H}|} \max\{0, g^{(i)}_{jk^*}(\mu)\} \leq 2|\mathcal{H}|\mu^{(i)}_{jk^*} \max\{0, g^{(i)}_{jk^*}(\mu)\} \\
&\leq 2|\mathcal{H}| \sum_{k'=1}^{m^{(i)}_j} \mu^{(i)}_{jk'} \max\{0, g^{(i)}_{jk'}(\mu)\} \\
&\leq 2|\mathcal{H}| \cdot \theta\sqrt{|\mathcal{H}|}\sqrt{\epsilon} = 2\theta|\mathcal{H}|^{3/2}\sqrt{\epsilon} = \epsilon' \,.
\end{aligned}$$

Case 2.2: Probability $\mu^{(i)}_{jk^*} < \frac{1}{2|\mathcal{H}|}$. Then for any action $k \in [m^{(i)}_j]$:

$$\begin{aligned}
&\max\{0, g^{(i)}_{jk}(\mu)\} \leq \max\{0, g^{(i)}_{jk^*}(\mu)\} \\
&= 2 \cdot \Big(\max\{0, g^{(i)}_{jk^*}(\mu)\} - \frac{1}{2} \max\{0, g^{(i)}_{jk^*}(\mu)\} \Big) \\
&= 2 \cdot \Big(\max\{0, g^{(i)}_{jk^*}(\mu)\} - \frac{1}{2|\mathcal{H}|} |\mathcal{H}| \cdot \max\{0, g^{(i)}_{jk^*}(\mu)\} \Big) \\
&\leq 2 \cdot \Big(\max\{0, g^{(i)}_{jk^*}(\mu)\} - \frac{1}{2|\mathcal{H}|} \sum_{k'=1}^{m^{(i)}_j} \max\{0, g^{(i)}_{jk^*}(\mu)\} \\
&\leq 2 \cdot \Big(\max\{0, g^{(i)}_{jk^*}(\mu)\} - \mu^{(i)}_{jk^*} \sum_{k'=1}^{m^{(i)}_j} \max\{0, g^{(i)}_{jk'}(\mu)\} \Big) \\
&\leq 2 \cdot \Big| \max\{0, g^{(i)}_{jk^*}(\mu)\} - \mu^{(i)}_{jk^*} \sum_{k'=1}^{m^{(i)}_j} \max\{0, g^{(i)}_{jk'}(\mu)\} \Big| \\
&\overset{(9)}{\leq} 2 \cdot 2\theta|\mathcal{H}|\epsilon \leq 2\theta|\mathcal{H}|^{3/2}\sqrt{\epsilon} \cdot 2\sqrt{\epsilon} \\
&\leq 2\theta|\mathcal{H}|^{3/2}\sqrt{\epsilon} = \epsilon' \,,
\end{aligned}$$

where we used at the last line that $\epsilon \leq \frac{1}{4}$.

This covers all cases, and hence μ is an ϵ'-CDT equilibrium of Γ. □

Proposition 46. CDT-S *is in* PPAD.

Proof. Given an instance (Γ, ϵ) of CDT-S, construct mapping F as in (8) and $\delta := \Big(\frac{\epsilon}{2\theta|\mathcal{H}|^{3/2}}\Big)^2$. Use Lemma 44. Then, a δ-fixed point of F makes an $2\theta|\mathcal{H}|^{3/2}\sqrt{\delta} = \epsilon$-CDT equilibrium of Γ by Lemma 45. □

E On Section 6

In this section, we prove the results in Section 6. To that end, we restate results taken from the main body, and give new numbers to results presented first in this appendix.

E.1 On General Chance Node Removal

Here, we will show:

Theorem (Restatement of Theorem 7)**.** *All computational hardness results in this paper for the three equilibrium concepts* {*Nash, EDT, CDT*} *still hold even when the game has no chance nodes. They hold together with previously possible restrictions* (e.g.*, on the branching factor), except that the restrictions on the number of infosets and the degree of absentmindedness increase by one and to* $\mathcal{O}(\log|\mathcal{H}|)$ *respectively.*

This result immediately follows from the following construction. For this subsection, let *equilibrium* be any of the three equilibrium concepts {Nash, EDT, CDT}.

Theorem 8. *For an N-player game Γ with imperfect recall, we can create an N-player game Γ' with imperfect recall and without chance nodes such that*

1. *Γ' has the same info and action sets as Γ except an arbitrary player (PL1) has one additional infoset I_c with absentmindedness which will induce randomness,*
2. *there is a randomized action α for I_c such that for exact equilibrium sets E and E' of Γ and Γ', we have identity $E' = E \times \{\alpha\}$,*
3. *there is an polynomial relationship between precision errors of approximate equilibria in Γ' and the approximate equilibria they induce in Γ.*

We prove Theorem 8 in parts.

First, we show how to transition to a game with a single, polysized chance node h_c at the root.

Lemma 47. *A game Γ with imperfect recall can be polytime reduced to a game Γ' with imperfect recall such that Γ' has the same strategy sets and utility functions as Γ, and Γ' has only one chance node which is placed at the root. That chance node randomizes uniformly over 2^t actions for some integer t in $\mathcal{O}(\log |\mathcal{H}|)$ where $|\mathcal{H}|$ is the number of nodes in Γ.*

In particular, Γ and Γ' have the same ϵ-equilibria.

Proof. Let Γ be a game with imperfect recall. Get its utility functions $U^{(1)}, \ldots, U^{(N)}$. Use Appendix A.4 to create a game $\tilde{\Gamma}$ with imperfect recall out of it, that has utility functions $U^{(1)}, \ldots, U^{(N)}$. Both of these steps take polytime. Notably, $\tilde{\Gamma}$ has only one chance node h_0 at the root, and that one is randomizing uniformly over a number of outgoing actions r that is equal to the number of monomials with nonzero coefficients in the functions $U^{(1)}, \ldots, U^{(N)}$. This is bounded by the number of terminal nodes $|\mathcal{Z}|$ in Γ, which is bounded by the number of nodes $|\mathcal{H}|$ in Γ. Next, we pad the number of outgoing edges at h_0 in $\tilde{\Gamma}$ to $2^{\lceil \log(r) \rceil}$ to obtain the final game Γ': Add $2^{\lceil \log(r) \rceil} - r$ many actions to h_0, each leading to a terminal node with utility 0 for all player. Next, make the probability distribution at h_0 uniform over these $2^{\lceil \log(r) \rceil}$ actions, and rescale the payoffs at terminal nodes that were in $\tilde{\Gamma}$ before this padding action by $2^{\lceil \log(r) \rceil}/r$. This padding procedure is polytime (we added at most r additional actions at the root) and it ensures that the new game Γ' has the same utility functions as $\tilde{\Gamma}$ which has the same utility function as original game Γ. Hence, Γ' and Γ have the same ϵ-equilibria since those are defined in terms of the strategy sets and ex-ante utility functions. Last but not least, Γ' has only one chance node at the root which randomizes uniformly over $2^{\lceil \log(r) \rceil}$ actions where $\lceil \log(r) \rceil = \mathcal{O}(\log |\mathcal{H}|)$. $\square$

Proposition 48. *Let Γ be a game with imperfect recall that has only one chance node at the root which randomizes uniformly over 2^t actions for some $t \in \mathbb{N}$. Then Γ can be polytime reduced to a game Γ' with imperfect recall such that*

1. *Γ' has no chance nodes*
2. *exact equilibria of Γ correspond 1-1 to exact equilibria of Γ'*
3. *δ-equilibria of Γ' give rise to ϵ-equilibria of Γ, where we (might) set*

$$\delta = \min\{\frac{1}{4}, \frac{\epsilon}{2^t + t \cdot L_\infty}\}$$

using a Lipschitz constant L_∞ as described in Appendix A.3.

Proof. Let Γ be a game with imperfect recall with one chance node h_0 at the root with the above description. We assume w.l.o.g. that the payoffs in Γ are ≥ 1 (otherwise first shift the payoffs in Γ by $1 - \min_{z \in \mathcal{Z}, i \in \mathcal{N}} u^{(i)}(z)$). Replace h_0 with one big infoset I_c with 2 actions l (left) and r (right) and a degree of absentmindedness of $2t$. It is irrelevant which player is assigned to I_c, so let it be P1. Figure 6 gives an example for $t = 1$ and Figure 10 gives an example for $t = 3$.

The Construction:
Formally, we replace h_0 by a tree $\mathcal{T}$ of terminal nodes and nodes in I_c as follows. $\mathcal{T}$ will have a depth of $2t + 1$. Start at the root h'_0, assign it to I_c, and call its depth level 0. Create two outgoing edges l and r to nodes h'_1 and h'_2. Assign those nodes to I_c as well. Make nodes $h'_1 l$ and $h'_2 r$ terminal nodes with payoff -1 to all players. Assign nodes $h'_1 r$ and $h'_2 l$ to I_c. Next, we create depth levels 2 and 3 to $2t - 3$ and $2t - 2$ by induction. Let node h' be in I_c and at an even depth level. Assign its two children $h'l$ and $h'r$, as well as their respective child $h'lr$ and $h'rl$, to I_c. The nodes $h'll$ and $h'rr$ shall be terminal nodes with payoff 0 to all players. Finally, depth levels $2t - 1$ and $2t$ shall be created in the same way, except that the nodes at level $2t$ that would have been assigned to I_c are now instead being replaced by the children of the original chance node h_0 (order of replacement is irrelevant). This works out number-wise because there are exactly $2^{2t/2} = 2^t$ many nonterminal nodes at depth level $2t$.

How Γ' looks like:
If $S = \times_{i \in \mathcal{N}} S^{(i)}$ is the strategy set of Γ, then

$$S' = S^{(1)} \times \Delta(\{l, r\}) \times \bigtimes_{i \in \mathcal{N} \setminus \{1\}} S^{(i)} = S \times [0, 1]$$

is the strategy set of Γ, where the interval $[0, 1]$ stands for the probability that P1 assigns to playing left l at I_c. Take the utility function of any player $i \in \mathcal{N}$ and write it as

$$U^{(i)}(\mu) = \sum_{a \in [2^t]} \frac{1}{2^t} U^{(i)}(\mu \mid h_0 a)$$

for strategy $\mu \in S$.

Observe that each children $h_0 a$ of h_0 in the corresponding game Γ' has an action history with exactly t appearances of l and t appearances of r. Therefore, they are all reached with equal probability $l^t (1 - l)^t$, where by abuse of notation we use $l \in [0, 1]$ for the probability put by P1 on action l at I_c. Moreover, after the constructed tree $\mathcal{T}$ at the beginning, infoset I_c does not occur again. Next, recall that all terminal nodes in $\mathcal{T}$, except for the two on depth level 2, have a payoff of 0. Hence, if $V^{(i)}$ denotes the utility functions in Γ', we can rewrite them as

$$\begin{aligned} V^{(i)}(\mu, l) &= -l^2 - (1 - l)^2 + \sum_{a \in [2^t]} l^t (1 - l)^t U^{(i)}(\mu \mid h_0 a) \\ &= -l^2 - (1 - l)^2 + 2^t \big(l(1 - l)\big)^t U^{(i)}(\mu) \\ &= -\frac{1}{2} - 2(l - \frac{1}{2})^2 + 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t U^{(i)}(\mu) \end{aligned}$$

for strategy $(\mu, l) \in S'$. Note that $0 \leq l \leq 1$ implies $0 \leq (l - \frac{1}{2})^2 \leq \frac{1}{4}$ and thus the factor in front of $U^{(i)}(\mu)$ is always non-negative in a valid profile (μ, l). Moreover, using that utility $U^{(i)}$ is positive, we observe that any profile (μ, l) for $l \neq \frac{1}{2}$ is strictly dominated by profile $(\mu, \frac{1}{2})$. In fact, the best response set of P1 to $\mu^{(-1)}$ in Γ' is the best response set to $\mu^{(-1)}$ in Γ and playing $\frac{1}{2}$ at I_c. This is because function $V^{(1)}(\cdot, \mu^{(-1)}, \frac{1}{2})$ is just a positive factor scaling and subsequent constant shift of $U^{(1)}(\cdot, \mu^{(-1)})$. Analogous reasoning yields that the best response set of player $i \neq 1$ in Γ'

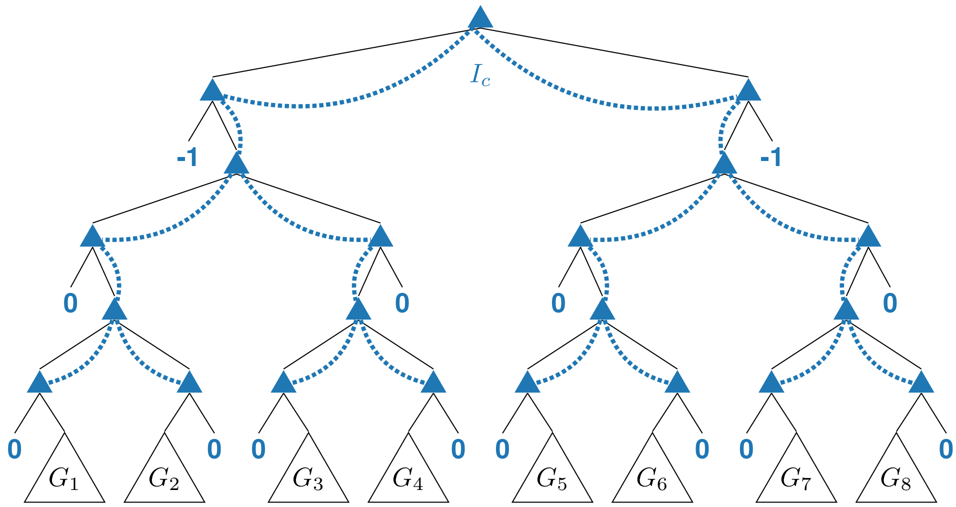


Figure 10: Another example for the construction of Proposition 48, this time the chance node at the root randomizes uniformly over 8 actions into subgames $G_1, \dots, G_8$.

to $(\mu^{(-i)}, l)$ for $0 \neq l \neq 1$ is equal her best response set in Γ to $\mu^{(-i)}$.

Exact Nash and EDT:
Suppose μ is an exact Nash equilibrium or EDT equilibrium of Γ. Then by above reasoning, profile $(\mu, \frac{1}{2})$ makes an exact Nash equilibrium or, respectively, EDT equilibrium in Γ'.

Approximate Nash:
We will now show that given (Γ, ϵ), we can set $\delta > 0$ sufficiently small but of size $\mathsf{poly}(\epsilon, |\Gamma|)$, such that if (μ, l) is a δ-Nash equilibrium in Γ', then μ is an ϵ-Nash equilibrium in Γ.

First, we bound how far away l can be from $\frac{1}{2}$. A δ-Nash equilibrium in particularly satisfies the δ-EDT equilibrium condition that (μ, l) does not perform more than δ worse than $(\mu, \frac{1}{2})$ for P1. Thus, using that utility $U^{(i)}$ is positive, we get

$$\begin{aligned}
\delta &\geq V^{(1)}(\mu, \frac{1}{2}) - V^{(1)}(\mu, l) \\
&= 2(l - \frac{1}{2})^2 + 2^t U^{(1)}(\mu) \Big[\frac{1}{4^t} - \Big(\frac{1}{4} - (l - \frac{1}{2})^2 \Big)^t \Big] \\
&\geq 2(l - \frac{1}{2})^2 + 2^t U^{(1)}(\mu) \cdot 0 \,,
\end{aligned}$$

which implies that l must satisfy

$$(l - \frac{1}{2})^2 \leq \delta / 2 \,. \tag{10}$$

In particular, we will choose $\delta \leq \frac{1}{4}$, and have $0 \neq l \neq 1$.

Next, we show that μ is an ϵ-Nash equilibrium of Γ. Consider any deviation strategy $\pi^{(i)}$ of player $i \in \mathcal{N}$ in Γ. Then, we get

$$\begin{aligned}
&U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - U^{(i)}(\mu) \\
&= \frac{2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2 \Big)^t}{2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2 \Big)^t} (U^{(i)}(\pi^{(i)}, \mu^{(-i)}) - U^{(i)}(\mu)) \\
&= \frac{V^{(i)}(\pi^{(i)}, \mu^{(-i)}, l) - V^{(i)}(\mu, l)}{2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2 \Big)^t} \\
&\overset{(*)}{\leq} \frac{\delta}{2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2 \Big)^t} \\
&\overset{(10)}{\leq} \frac{\delta}{2^t \Big(\frac{1}{4} - \frac{\delta}{2} \Big)^t} \\
&\overset{(\dagger)}{\leq} \frac{\delta}{2^t \Big(\frac{1}{4} - \frac{1}{8} \Big)^t} \\
&= \frac{\delta}{2^t \frac{1}{4^t}} \\
&\overset{(\star)}{\leq} \epsilon \,,
\end{aligned}$$

where we use in $(*)$ that (μ, l) is a δ-Nash equilibrium in Γ', in $(\dagger)$ that we will choose $\delta \leq \frac{1}{4}$, and in $(\star)$ that we will choose $\delta \leq \frac{1}{2^t} \epsilon$. Hence, $\mu^{(i)}$ is an ϵ-best response of player i to $\mu^{(-i)}$ in Γ. All in all, if we set $\delta := \min\{\frac{1}{4}, \frac{1}{2^t} \epsilon\}$, then any δ-Nash equilibrium (μ, l) in Γ' induces μ to be an ϵ-Nash equilibrium in Γ.

Approximate EDT:
Analogous reasoning as in approximate Nash. One merely has to consider each infoset I in Γ and only deviations $\mu^{(i)}_{I \mapsto \alpha}$.

Exact CDT:

The KKT characterization Proposition 16 for game Γ states that profile μ is an exact CDT equilibrium of Γ if and only if there exist KKT multipliers $\{\tau_{jk}^{(i)} \in \mathbb{R}\}_{i,j,k=1}^{\mathcal{N},\ell^{(i)},m_j^{(i)}}$ and $\{\kappa_j^{(i)} \in \mathbb{R}\}_{i,j=1}^{\mathcal{N},\ell^{(i)}}$ such that

$$\mu_{jk}^{(i)} \geq 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]$$
$$\sum_{k=1}^{m_j^{(i)}} \mu_{jk}^{(i)} = 1 \quad \forall i \in [N], \forall j \in [\ell^{(i)}]$$
$$\tau_{jk}^{(i)} \geq 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]$$
$$\tau_{jk}^{(i)} = 0 \quad \text{or} \quad \mu_{jk}^{(i)} = 0 \quad \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}]$$
$$\nabla_{jk} U^{(i)}(\mu) + \tau_{jk}^{(i)} - \kappa_j^{(i)} = 0 \, \forall i \in [N], \forall j \in [\ell^{(i)}], \forall k \in [m_j^{(i)}] \tag{11}$$

The KKT characterization for a profile (μ, l) in game Γ' is the same, except that we replace $\nabla_{jk} U^{(i)}(\mu)$ in (11) with $\nabla_{jk} V^{(i)}(\mu, l)$, and that we need additional multipliers τ_l^- and τ_l^+ such that

$$l \geq 0 \quad \text{and} \quad l \leq 1$$
$$\tau_l^- \geq 0 \quad \text{and} \quad \tau_l^+ \geq 0$$
$$\tau_l^- = 0 \quad \text{or} \quad l = 0$$
$$\tau_l^+ = 0 \quad \text{or} \quad l = 1$$
$$\nabla_l V^{(i)}(\mu, l) + \tau_l^- - \tau_l^+ = 0 \, . \tag{12}$$

We first observe that

$$\begin{aligned}
&\nabla_l V^{(i)}(\mu, l) \\
&= 2(1-2l) + 2^t(1-2l)t\Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^{t-1} U^{(i)}(\mu) \\
&= (1-2l)\bigg(2 + 2^t t\Big(\frac{1}{4} - (l-\frac{1}{2})^2\Big)^{t-1} U^{(i)}(\mu)\bigg) .
\end{aligned}$$

Here, using that utility $U^{(i)}$ is positive, the second factor (the big bracket) will always be positive. Hence, after another look at the KKT conditions, boundary points $l = 0$ and $l = 1$ cannot satisfy the KKT conditions in Γ' no matter the choice of $\tau_l^- \geq 0$ or, respectively, $\tau_l^+ \geq 0$. In the interior $(0, 1)$, the KKT conditions on l reduce to stationary condition $\nabla_l V^{(i)}(\mu, l) = 0$ which is only satisfied at $l = \frac{1}{2}$.

Next, we observe that for indices i, j, k, we have

$$\nabla_{jk} V^{(i)}(\mu, l) = 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \nabla_{jk} U^{(i)}(\mu) \, ,$$

which, for $0 < l < 1$ implies that $\nabla_{jk} V^{(i)}(\mu, l)$ is simply a positive rescaling of $U^{(i)}(\mu)$.

Therefore, all in all, we get the equivalence that (1) a point (μ, l) satisfies the KKT conditions of Γ' for multipliers $\{\tau_{jk}^{(i)} \in \mathbb{R}\}_{i,j,k=1}^{\mathcal{N},\ell^{(i)},m_j^{(i)}}$, $\{\kappa_j^{(i)} \in \mathbb{R}\}_{i,j=1}^{\mathcal{N},\ell^{(i)}}$, τ_l^-, and τ_l^+ if and only if (2) $l = 1/2$, $\tau_l^- = 0 = \tau_l^+$, and μ satisfies the KKT conditions of Γ for multipliers $\{2^t \tau_{jk}^{(i)} \in \mathbb{R}\}_{i,j,k=1}^{\mathcal{N},\ell^{(i)},m_j^{(i)}}$ and $\{2^t \kappa_j^{(i)} \in \mathbb{R}\}_{i,j=1}^{\mathcal{N},\ell^{(i)}}$.

Approximate CDT:

We will now show that given (Γ, ϵ), we can set $\tilde{\delta} > 0$ sufficiently small but of size $\mathsf{poly}(\epsilon, |\Gamma|)$, such that if $(\tilde{\mu}, \tilde{l})$ is a $\tilde{\delta}$-CDT equilibrium in Γ', then we can compute a profile μ' from it in polytime such that μ is an ϵ-CDT equilibrium in Γ.

First, we use Lemma 30 to transition from the $\tilde{\delta}$-CDT equilibrium in Γ' to a δ-well-supported CDT equilibrium (μ, l) in Γ', where $\delta = 3L_\infty |\mathcal{H}| \sqrt{\tilde{\delta}}$ and L_∞ is chosen as in Appendix A.3. Next, we note that an approximate well-supported CDT equilibrium (for precision ϵ in Γ or δ in Γ') satisfies the exact KKT conditions above except that equality (11) is replaced by

$$|\nabla_{jk} U^{(i)}(\mu) + \tau_{jk}^{(i)} - \kappa_j^{(i)}| \leq \epsilon$$

and that equality (12) is replaced by

$$|\nabla_l V^{(i)}(\mu, l) + \tau_l^- - \tau_l^+| \leq \delta \, .$$

So let $\{\tau_{jk}^{(i)} \in \mathbb{R}\}_{i,j,k=1}^{\mathcal{N},\ell^{(i)},m_j^{(i)}}$, $\{\kappa_j^{(i)} \in \mathbb{R}\}_{i,j=1}^{\mathcal{N},\ell^{(i)}}$, τ_l^-, and τ_l^+ be those KKT multipliers for (μ, l) in Γ. Then, we will show that multiplier $\{2^t \tau_{jk}^{(i)} \in \mathbb{R}\}_{i,j,k=1}^{\mathcal{N},\ell^{(i)},m_j^{(i)}}$ and $\{2^t \kappa_j^{(i)} \in \mathbb{R}\}_{i,j=1}^{\mathcal{N},\ell^{(i)}}$ show that μ is an ϵ-well-supported CDT equilibrium for Γ.

First, we bound how far away l can be from $\frac{1}{2}$. Analogous to the exact case, we will implicitly choose $\delta \leq 1$ and therefore, boundary points $l = 0$ and $l = 1$ cannot be the case in a δ-well-supported CDT equilibrium. Hence, $\tau_l^- = 0 = \tau_l^+$ and the above inequality simplifies to

$$\begin{aligned}
\delta &\geq |\nabla_l V^{(i)}(\mu, l)| \\
&= |1 - 2l| \cdot \Big| 2 + 2^t t\Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^{t-1} U^{(i)}(\mu) \Big| \\
&\geq |1 - 2l| \cdot |2|
\end{aligned}$$

that is

$$(l - \frac{1}{2})^2 \leq \delta^2 / 4 \, . \tag{13}$$

Next, we bound

$$\begin{aligned}
0 = \frac{1}{2^t} - 2^t(\frac{1}{4} - 0)^t &\leq \frac{1}{2^t} - 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \\
&\overset{(13)}{\leq} \frac{1}{2^t} - 2^t(\frac{1}{4} - \frac{\delta^2}{4})^t = \frac{1}{2^t} - 2^t \frac{1}{4^t}(1 - \delta^2)^t \\
&\overset{(\star)}{\leq} \frac{1}{2^t} - \frac{1}{2^t}(1 - t\delta^2) = \frac{t}{2^t}\delta^2
\end{aligned}$$

where in $(\star)$ we used Bernoulli's inequality. Hence

$$\Big| \frac{1}{2^t} - 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \Big| \leq \frac{t}{2^t}\delta^2 \, . \tag{14}$$

Finally, we have all the tools to conclude that μ is an ϵ-well-supported CDT equilibrium of Γ: Clearly, the domain and complementary conditions are still satisfied since those are the same in Γ and Γ', and because we simply rescaled the

τ multiplier by a positive constant. Next, fix any coordinate (i, j, k). Then, we get

$$
\begin{aligned}
&|\nabla_{jk} U^{(i)}(\mu) + 2^t \tau_{jk}^{(i)} - 2^t \kappa_j^{(i)}| \\
&\leq 2^t |\frac{1}{2^t} \nabla_{jk} U^{(i)}(\mu) + \tau_{jk}^{(i)} - \kappa_j^{(i)}| \\
&\leq 2^t \cdot \Big| \frac{1}{2^t} \nabla_{jk} U^{(i)}(\mu) - 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \nabla_{jk} U^{(i)}(\mu) \\
&\qquad + 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \nabla_{jk} U^{(i)}(\mu) + \tau_{jk}^{(i)} - \kappa_j^{(i)} \Big| \\
&\leq 2^t \cdot \Big| \frac{1}{2^t} \nabla_{jk} U^{(i)}(\mu) - 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \nabla_{jk} U^{(i)}(\mu) \Big| \\
&\qquad + 2^t \cdot \Big| 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \nabla_{jk} U^{(i)}(\mu) + \tau_{jk}^{(i)} - \kappa_j^{(i)} \Big| \\
&\leq 2^t \Big| \nabla_{jk} U^{(i)}(\mu) \Big| \cdot \Big| \frac{1}{2^t} - 2^t \Big(\frac{1}{4} - (l - \frac{1}{2})^2\Big)^t \Big| \\
&\qquad + 2^t \cdot |\nabla_{jk} V^{(i)}(\mu, l) + \tau_{jk}^{(i)} - \kappa_j^{(i)}| \\
&\overset{(14),(*)}{\leq} 2^t \Big| \nabla_{jk} U^{(i)}(\mu) \Big| \cdot \frac{t}{2^t} \delta^2 + 2^t \delta \\
&\leq \delta \Big(2^t + t\delta \cdot \max_{\mu \in S} ||\nabla U^{(i)}(\mu)||_\infty \Big) \\
&\overset{(\dagger)}{\leq} \delta \Big(2^t + t \cdot \max_{\mu \in S} L_\infty \Big) \\
&\overset{(\star)}{\leq} \epsilon \,,
\end{aligned}
$$

where we use in $(*)$ that (μ, l) is a δ-well-supported CDT equilibrium in Γ', in $(\dagger)$ that we will implicitly choose $\delta \leq 1$, and in $(\star)$ that we will implicitly choose $\delta \leq \frac{1}{2^t + t \cdot L_\infty} \epsilon$. All in all, if we set

$$
\tilde{\delta} := \left(\frac{\min\{1, \frac{\epsilon}{2^t + t \cdot L_\infty}\}}{3 L_\infty |\mathcal{H}|} \right)^2 ,
$$

then any $\tilde{\delta}$-CDT equilibrium $(\tilde{\mu}, \tilde{l})$ in Γ' gives rise to a $\min\{1, \frac{\epsilon}{2^t + t \cdot L_\infty}\}$-well-supported equilibrium (μ, l) in Γ', which in turn induces μ to be an ϵ-well-supported CDT equilibrium in Γ and therefore, by Lemma 30, an ϵ-CDT equilibrium in Γ.

□

We conclude with the main result of this section.

Proof of Theorems 7 and 8. Starting with a game Γ with utility payoffs in the range of $[0, 2]$, a precision parameter $\epsilon \geq 0$ and an equilibrium concept *equilibrium* in {Nash, EDT, CDT}, apply Lemma 47 and then Proposition 48 on Γ to get a game Γ'. Then Γ' was constructed in polytime, and it has no chance nodes. It also has the same strategy set and game tree structure as Γ, except for one additional infoset I_c at the beginning. I_c has a degree of absentmindedness that is bounded by $2 \cdot \lceil \log |\mathcal{Z}| \rceil = \mathcal{O}(\log |\mathcal{H}|)$.

For exact computational problems ($\epsilon = 0$) we have that μ is an equilibrium of Γ if and only if $(\mu, \frac{1}{2})$ is an equilibrium in Γ'. For approximate computational problems ($\epsilon > 0$), we still have the correspondence with exact equilibria, but we can also choose $\delta > 0$ as in Proposition 48 such that δ-equilibria in Γ' will be or give rise to ϵ-equilibria of Γ. Note that $2^t = \text{poly}(|\mathcal{H}|)$ and that because of bounded utility payoffs in Γ, Lipschitz constant L_∞ will be of size poly$(|\mathcal{H}|)$ as well. Thus, if ϵ was of 1/poly or 1/exp precision (in $|\Gamma$), then δ will continue to be so. □

E.2 On Single-Player Games without Chance Nodes

Recall that a Boolean formula ϕ is in conjunctive normal form (CNF) if it is a conjunction of a collection of m clauses $C_1', \dots, C_m'$ each of which is a disjuntion of literals $\{x_i, \neg x_i\}_i$. The problem MINSAT takes a Boolean formula ϕ in CNF together with an integer threshold $0 \leq s^* \leq m$ as an instance, and asks whether there is a truth assignment for the variables in ϕ that satisfies at most s^* clauses in ϕ. The problem 2-MINSAT restricts MINSAT to those instances where each clause C_j' of ϕ contains no more than two literals.

Lemma 49 (Kohli *et al.* [1994])**.** *2-MINSAT is NP-complete.*

We will consider a variant of 2-MINSAT. First, suppose a clause C_j' uses the same variable x in both of its literals. Then, it is either always satisfied $(x \vee \neg x)$ in which case it can be removed. Otherwise, it reduces to a singleton clause in which case it can be padded with a dummy variable y that is only used in that clause. There are only at most m such paddings needed, and for the minimization procedure it is sufficient to consider only those truth assignments that set y to be false. Hence, with linear blowup we can assume w.l.o.g. that the 2-MINSAT instances solely consist of clauses that use two distinct variables.

Next, observe that the negation $\neg\phi$ – after distributing the negation into the clauses – is a disjunction of the collection of clauses $C_1, \dots, C_m$ where C_j is a conjunction of the negations of the literals in C_j'. Moreover, a truth assignment π satisfies M clauses in ψ if and only if it satisfies (exactly the other) $m - M$ clauses in $\neg\phi$. Putting both of these together, we obtain from Lemma 49:

Corollary 50. *The following problem 2-DNF-MAXSAT is NP-complete: Given a threshold DNF formula ϕ which uses exactly two distinct variables in each of its m clauses, and given an integer threshold $0 \leq s^* \leq m$, does there exist a truth assignment for the variables in ϕ that satisfies at least s^* clauses in ϕ?*

We get to the main result of this section.

Proposition (Restatement of Proposition 23)**.** *OPT-D is NP-hard, even for games with no chance nodes, one infoset, a degree of absentmindedness of 2, and 1/poly precision.*

Proof. We reduce from 2-DNF-MAXSAT. Let (ϕ, s^*) be one of its instances, that is, ϕ is a collection of clauses $C_1, \dots, C_m$ over variables $x_1, \dots, x_n$, where each clause is a conjunction of 2 literals of distinct variables. Construct a single-player game Γ with imperfect recall from it as follows. It has one infoset I with $2n$ actions $\{t_1, f_1, t_2, f_2, \dots, t_n, f_n\}$, where taking action t_i or f_i will correspond to setting x_i to true or false respectively in a corresponding truth assignment. Root h_0 belongs to I, each of its $2n$ children belong to I, and each of their respective $2n$ children are terminal nodes. Hence, there are $4n^2$ terminal nodes in Γ. Each terminal node z has a history that corresponds to setting some variable x_i to truth value v, and setting some (possibly other) variable $x_{i'}$ to truth value w, where $v, w \in \{t, f\}$. The utility payoff at such z shall be as follows:

- If $i = i'$, then z yields a penalty payoff of $u(z) = -B$, where $B = (16mn^2)^3 \in \mathbb{N}$ is a sufficiently large value (but still polynomially large). We will later see that because of this penalty, the player will try to maximize the probability that the case $i \neq i'$ happens. That is, the player will be incentivized to allocate, for each i, approximately $1/n$ probability to the actions t_i and f_i together.

- If $i \neq i'$, then $x_i \neq x_{i'}$, hence the partial truth value assignment might already satisfy some clauses C_j of ϕ. Let $\mathcal{C}(x_i \mapsto v, x_{i'} \mapsto w) \in \{0, 1, \ldots, m\}$ be the number of such satisfied clauses. For example, a terminal node z with history (f_5, t_3) satisfies all the occurrences of the clauses $x_3 \wedge \neg x_5$ and $\neg x_5 \wedge x_3$ in formula ϕ (and no other clauses). Define the payoff $u(z)$ to be $\mathcal{C}(x_i \mapsto v, x_{i'} \mapsto w)$.

Finally, choose target value $t^* := -B\frac{1}{n} + 2\frac{1}{n^2}s^*$ and precision $\epsilon := 2\frac{1}{n^2} \cdot \frac{1}{4}$. This whole construction takes poly-time and ϵ indeed makes a 1/poly precision.

We claim that

Claim 1: if there is a truth assignment that satisfies at least s^* clauses of ϕ, then there is also a strategy of Γ with utility at least t^*,

Claim 2: if Γ has a strategy μ with utility $\geq t^* - \epsilon$, then from it, we can construct an assignment ψ that satisfies at least s^* clauses of ϕ, and hence, Γ admits a strategy with utility at least t^*.

Those two claims imply that one can either achieve utility t^* in Γ, or one cannot achieve $t^* - \epsilon$. In particular, (ϕ, s^*) will be a "yes" (and resp. "no") instance of 2-DNF-MAXSAT if and only if the corresponding (Γ, t^*, ϵ) is a "yes" (and resp. "no") instance of OPT-D. This concludes the reduction.

Utility in Γ: First, we characterize the utility function U of the single player in Γ. In general, a strategy μ contains action probabilities $\mu(x_i \mapsto f)$ and $\mu(x_i \mapsto t)$ on actions f_i and t_i respectively. Any such strategy can instead be described by values $p_i = \mu(x_i \mapsto f) + \mu(x_i \mapsto t) \in [0, 1]$, which are the probabilities with which variables x_i are chosen under μ, and values $\alpha_i = \frac{\mu(x_i \mapsto f)}{p_i} \in [0, 1]$, which are the fractions of times with which variable x_i – if chosen – is set to false. If $p_i = 0$, then we can set α_i to an arbitrary value in $[0, 1]$ instead. Since μ is a strategy, we have $\sum_i p_i = 1$. We get for any strategy μ the identity

$$
\begin{aligned}
U(\mu) = &\sum_{i\in[n]} (-B) \cdot \Big[\mu(x_i \mapsto f) \cdot \mu(x_i \mapsto f) \\
&+ \mu(x_i \mapsto f) \cdot \mu(x_i \mapsto t) + \mu(x_i \mapsto t) \cdot \mu(x_i \mapsto f) \\
&+ \mu(x_i \mapsto t) \cdot \mu(x_i \mapsto t) + \Big] \\
&+ \sum_{i\in[n]} \sum_{i' \neq i} \Big[\\
&\mu(x_i \mapsto f) \cdot \mu(x_{i'} \mapsto f) \cdot \mathcal{C}(x_i \mapsto f, x_{i'} \mapsto f) \\
&+ \mu(x_i \mapsto f) \cdot \mu(x_{i'} \mapsto t) \cdot \mathcal{C}(x_i \mapsto f, x_{i'} \mapsto t) \\
&+ \mu(x_i \mapsto t) \cdot \mu(x_{i'} \mapsto f) \cdot \mathcal{C}(x_i \mapsto t, x_{i'} \mapsto f) \\
&+ \mu(x_i \mapsto t) \cdot \mu(x_{i'} \mapsto t) \cdot \mathcal{C}(x_i \mapsto t, x_{i'} \mapsto t)\Big] \\
&= -B \sum_{i\in[n]} p_i^2 + 2 \sum_{i\in[n]} \sum_{i'>i} p_i p_{i'} \Big[\\
&\alpha_i \alpha_{i'} \cdot \mathcal{C}(x_i \mapsto f, x_{i'} \mapsto f) \\
&+ \alpha_i (1 - \alpha_{i'}) \cdot \mathcal{C}(x_i \mapsto f, x_{i'} \mapsto t) \\
&+ (1 - \alpha_i) \alpha_{i'} \cdot \mathcal{C}(x_i \mapsto t, x_{i'} \mapsto f) \\
&+ (1 - \alpha_i)(1 - \alpha_{i'}) \cdot \mathcal{C}(x_i \mapsto t, x_{i'} \mapsto t)\Big] \\
&= -B \sum_{i\in[n]} p_i^2 + 2V(\mu) \,,
\end{aligned}
$$

where V stands for the second (big double) sum. We will later use the fact that

$$
0 \leq V(\pi) \leq \sum_{i\in[n]} \sum_{i'>i} 1 \cdot (m + m + m + m) \leq 4mn^2 \,. \tag{15}
$$

Claim 1 Suppose ψ is a truth value assignment for variables $x_1, \ldots, x_n$. Let $s(\psi)$ be the number of clauses ψ satisfies in ϕ. Define ψ's associated strategy in Γ as

$$
\mu_\psi(x_i \mapsto v) = \begin{cases} 1/n & \text{if } \psi(x_i) = v \\ 0 & \text{if } \psi(x_i) = \neg v \end{cases}
$$

for all $i \in [n]$ and $v \in \{f, t\}$. Then observe that $p_i(\mu_\psi) = 1/n$, and $\alpha_i(\mu_\psi) = \psi(x_i)$, and

$$
\begin{aligned}
V(\mu_\psi) &= \sum_{i\in[n]} \sum_{i'>i'} \frac{1}{n^2} \mathcal{C}(x_i \mapsto \psi(x_i), x_{i'} \mapsto \psi(x_{i'})) \\
&= \frac{1}{n^2} s(\psi) \,.
\end{aligned} \tag{16}
$$

Hence,

$$
U(\mu_\psi) = -B\frac{1}{n} + 2\frac{1}{n^2} s(\psi) = t^* + 2\frac{1}{n^2}(s(\psi) - s^*)
$$

Therefore, overall, if there is truth value assignment ψ for ϕ with $s(\psi) \geq s^*$, then μ_ψ will achieve a utility of at least t^*.

Observation 1 for Claim 2 First, we show that in an (exactly) optimal strategy π for Γ, the probabilities p_i have distance at most

$$
\delta := \sqrt{16mn^2/B} = \frac{1}{16mn^2} < 1 \tag{17}
$$

from value $\frac{1}{n}$. That is because by optimality, it in particular performs better than a strategy μ defined as follows: Give it the same distribution α, and almost the same distribution q as p in π. The only difference is that for $i^* \in \operatorname{argmax}_i p_i$ and $i_* \in \operatorname{argmin}_i p_i$, we define $q_{i^*} = (p_{i^*} + p_{i_*})/2 = q_{i_*}$ instead. Then

$$0 \le U(\pi) - U(\mu) = -B \sum_{i\in[n]} p_i^2 + 2V(\pi) + B \sum_{i\in[n]} q_i^2 - 2V(\mu)$$
$$\overset{(15)}{\le} -B \sum_{i\in[n]} (p_i^2 - q_i^2) + 2 \cdot 4mn^2 - 0$$
$$= -B\Big(p_{i^*}^2 + p_{i_*}^2 - 2 \cdot (p_{i^*} + p_{i_*})^2/4\Big) + 8mn^2$$
$$= -B\Big(p_{i^*}^2/2 + -p_{i^*}p_{i_*} + p_{i_*}^2/2\Big) + 8mn^2$$
$$= -\frac{1}{2}B(p_{i^*} - p_{i_*})^2 + 8mn^2 ,$$

which implies $(p_{i^*} - p_{i_*})^2 \le 16mn^2/B$, and hence,

$$|p_{i^*} - p_{i_*}| \le \sqrt{16mn^2/B} = \delta .$$

Note that i^* and i_* were chosen as extreme values, and thus,

$$1/n = 1/n \sum_{i\in[n]} p_i \in [1/n \sum_{i\in[n]} p_{i_*}, 1/n \sum_{i\in[n]} p_{i^*}] = [p_{i_*}, p_{i^*}] ,$$

where this interval has a length of at most δ. Putting the last two derivations together, we obtain for any $i \in [n]$:

$$p_i \in [p_{i_*}, p_{i^*}] \subset [1/n - \delta, 1/n + \delta] .$$

As another consequence, we want to observe for later that for $i, i' \in [n]$, we have

$$\begin{aligned} p_i p_{i'} &\le (1/n + \delta)^2 = 1/n^2 + 2\delta/n + \delta^2 \\ &\le 1/n^2 + 2\delta + \delta = 1/n^2 + 3\delta . \end{aligned} \tag{18}$$

Observation 2 for Claim 2 Next, we shall argue that $-B \sum_{i\in[n]} p_i^2$ is maximized at $p_i = 1/n \, \forall i$. Recall that for any strategy μ we have $p \in \Delta^{n-1}$. Hence, minimizing the term above is equivalent to $\max_{p\in\Delta^{n-1}} ||p||_2^2$. This is a uniformly convex function over a convex, compact polytope, hence it attains its global minimum in the relative interior of the simplex (i.e. the inequality constraints are slack). A global minimum also satisfies the KKT conditions, and the KKT conditions for a relative interior point become that $p \in \Delta^{n-1}$ and that $p = \frac{1}{2}\kappa\mathbb{1}$ for some $\kappa \in \mathbb{R}$ and the vector $\mathbb{1}$ that consists of 1's in each entry. This condition is only satisfied at $p^* = \frac{1}{n}\mathbb{1}$. In particular, for all $p \in \Delta^{n-1}$, we therefore obtain

$$-B \sum_{i\in[n]} p_i^2 \le -B \sum_{i\in[n]} \left(\frac{1}{n}\right)^2 = -B\frac{1}{n} . \tag{19}$$

Claim 2 Now suppose Γ has a strategy μ with utility $\ge t^* - \epsilon$. Then, an optimal strategy π' also achieves $U(\pi') \ge t^* - \epsilon$. Let us create another optimal strategy π from π' that satisfies $\alpha \in \{0,1\}^n$, that is, its distribution α to truth and false values make a proper truth value assignment of variables $x_1, \dots, x_n$. To that end, note that $U(\pi')$ is linear in α'_i of π' for any given values p and α'_{-i}, and hence it is maximized at the boundary $\alpha'_i = 0$ or $\alpha'_i = 1$. Therefore, starting from π', we can iterative over $i = 1, \dots, n$ and set α_i to one of these extreme values without decreasing the utility value. Denote the resulting strategy with π. It is also optimal for Γ and its $\alpha \in \{0,1\}^n$. For that strategy, we can derive

$$-B\frac{1}{n} + 2\frac{1}{n^2}(s^* - \frac{1}{4}) = t^* - \epsilon \le U(\pi')$$
$$= U(\pi) = -B \sum_{i\in[n]} p_i^2 + 2V(\pi) \overset{(19)}{\le} -B\frac{1}{n} + 2V(\pi)$$
$$\overset{(16)}{=} -B\frac{1}{n} + 2 \sum_{i\in[n]} \sum_{i'>i'} p_i p_{i'} \mathcal{C}(x_i \mapsto \alpha_i, x_{i'} \mapsto \alpha_{i'})$$
$$\overset{(18)}{=} -B\frac{1}{n} + 2 \sum_{i\in[n]} \sum_{i'>i'} (1/n^2 + 3\delta) \cdot \mathcal{C}(x_i \mapsto \alpha_i, x_{i'} \mapsto \alpha_{i'})$$
$$\overset{(16)}{=} -B\frac{1}{n} + 2 \cdot (1/n^2 + 3\delta) \cdot s(\alpha) .$$

Rearranging yields

$$s^* \le \frac{1}{4} + s(\alpha) + s(\alpha) \cdot 3\delta n^2 \le s(\alpha) + \frac{1}{4} + \delta \cdot 3mn^2$$
$$\overset{(17)}{=} s(\alpha) + \frac{1}{4} + \frac{1}{16mn^2} \cdot 3mn^2 \le s(\alpha) + \frac{1}{2} .$$

Since both s^* and $s(\alpha)$ are integers, this can only be the case if $s(\alpha) \ge s^*$, that is, there is a truth value assignment that satisfies at least s^* clauses of ϕ. Using Claim 1 for $\psi = \alpha$, we obtain that there must also be a strategy in Γ that achieves a utility of at least t^* in Γ. ☐

Corollary 51. *It is* NP*-hard to distinguish between whether all EDT equilibria μ in a single-player game have an utility $U^{(1)}(\mu) \ge t$ from whether there is an EDT equilibrium μ that satisfies $U^{(1)}(\mu) \le t - \epsilon$. Hardness holds even for games with no chance nodes, one infoset, a degree of absentmindedness of* 2*, and* 1/poly *precision.*

Proof. This follows from Proposition 23, and from the fact that in single-infoset games all EDT equilibria are optimal strategies due to Remark 20. ☐