Sleeping Beauty Loses Her Magic Psychic Powers*

John Cremin†

Aix-Marseille University, CNRS, AMSE

First version: May 25, 2026
Latest version: June 25, 2026

Abstract

An agent, *Sleeping Beauty*, in a game with *self-locating uncertainty* (i.e. one play of the game visits the same information set multiple times, as in the paradox of the absent-minded driver) must select a behavioural strategy that is *self-ratifying*: a best-response to the belief that her other instances do likewise. When there are multiple such fixed points, the standard treatment of Aumann et al. (1997) assumes that all instances of the agent can simply coordinate. I drop this assumption (supposing that Sleeping Beauty '*lose her magic psychic powers*'), and study said agent iteratively reasoning her way to an equilibrium selection instead. Which strategy other instances select can be seen as *ambiguous*, so I model Beauty's choice via a *response function* that encodes her response to ambiguity, and a *procedure* that describes in what manner she iterates the application of this response function. I consider two response functions, '*Bayesian*' and EU-maximin, and two procedures, *replacement* and *accumulation*, and characterise in which cases the iterative reasoning converges. For either response function, *accumulation* always converges, but *replacement* does so if and only if there are no cycles of length weakly greater than two on a particular finite functional digraph I call the *support digraph*.

Keywords: Absent-Minded Driver, Self-Locating Uncertainty, Imperfect Recall, Decision Instability, Ambiguity
JEL Codes: C72, C73, D81, D83

*The Sleeping Beauty problem is in fact an 'SLU' game *without* decision instability; it has a unique interim fixed point regardless of the awakening structure. It is the paradox of the absent-minded driver which can exhibit multiple fixed points. However, 'The Absent-Minded Driver Loses His Magic Psychic Powers' is too clunky and dreary a title, so I ask the reader to tolerate this wrinkle. Readers do not need prior awareness of either the Sleeping Beauty problem, the paradox of the absent-minded driver, self-locating uncertainty, or decision instability to understand this paper. Bear with me!

†Email: john-walter-edward.cremin@univ-amu.fr. The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference: ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University – A*MIDEX.

1 Introduction

Distressed at the news that the evil fairy is back in town, Sleeping Beauty has decided to drown her sorrows at the pub, downing 10 pints of lager (she is, by English standards, not a big drinker). Unwisely, she then decides to drive home, though fully aware that in her drink-addled state of absent-mindedness she will not be able to distinguish between the three intersections between her and chez Beauty (see Figure 1). Upon arriving at any of these intersections, and of course not knowing at which one she finds herself, she considers with what probability she should continue if she believes that all her other temporal parts continue with probability q. She then reasons that if her optimal continue-probability is not also q, then it does not make sense as a solution: it is not *self-ratifying* or *consistent*. To see this, notice that she expects that her other temporal parts will engage in exactly the same reasoning upon arriving at an intersection as she does. If instead of believing her other temporal parts continue with a fixed-point probability, she believes other temporal parts continue with probability q to which her best-response is $p \neq q$, then she should believe that her other temporal parts will form exactly the same beliefs and draw exactly the same conclusion. Thus they will continue with probability p, and not q. Only a fixed point avoids this issue. In this instance, however, she notices another problem. The particular pay-off values associated with the different outcomes are such that there are *multiple* such fixed points.

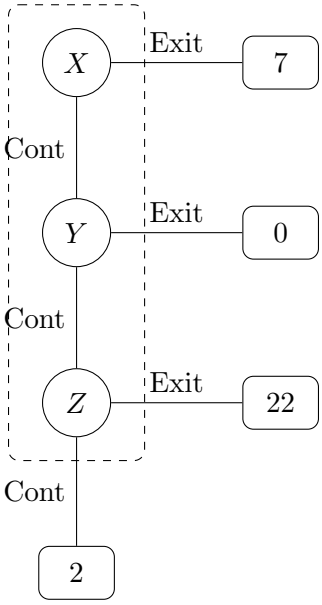


Figure 1: The absent-minded driver with three intersections and payoffs $(7, 0, 22, 2)$. All three decision nodes X, Y, Z lie in a single information set (dashed box). This game is reproduced from Figure 3 of Aumann et al. (1997), where they also note the three 'action-optimal decisions' or interim fixed points are 0, 7/30 and 1/2.

Aumann, Hart and Perry (who are on a lads' holiday in Sutton Coldfield, where this is all taking place) shout from the sidelines that she should simply have already decided on which fixed point to coordinate before setting out, but this is unhelpful in the circumstances.[1] Beauty forgot to do so, and must deal with the consequences. On a normal day, she would simply use her magic psychic powers to intuit on which fixed point her other temporal parts had settled, but having alas recently lost these, must instead reason her way to a solution.

[1] "*...he had better coordinate his actions; this coordination can take place only before he starts out— at the planning stage... (If he chose... nothing at all, then at the action stage he will have some hard thinking to do.)*" This paper considers exactly the hard thinking in question.

As a Bayesian, with no particular reason to consider either of the three fixed points more plausible than the others, she could apply the principal of insufficient reason in predicting the behaviour of each other temporal part, and adopt the belief that they are choosing each with a probability of 1/3. Since these fixed points are 0, 7/30 and 1/2, this implies a belief that each other temporal part will choose to continue with probability 11/45 which in turn implies that she should continue, since the expected utilities of continue and exit are (roughly) 6.5 and 6.4 respectively.[2] Having decided to continue though, she bears in mind once more that the solution must be self-ratifying in order to be valid: her other temporal parts will have reasoned identically. She now believes all temporal parts will have processed the original three fixed points and decided to certainly continue, but her best response to this is to exit for an expected utility of roughly 9.7 against the 2 that continuing would provide. Finally however, the chain of reasoning has landed at an interim fixed point: the best-response to continuing with probability 0 is to follow suit. All temporal parts choose to exit.

This, of course, was just one example game. The purpose of this paper is to study whether or not such iterative reasoning generally resolves the problem of multiplicity in finite-action games with *self-locating uncertainty*. This is the uncertainty agents face when a single play (terminal history) of a game visits the same information set multiple times. When one considers formalising this sort of reasoning, it emerges that there are a few different approaches one could adopt, which are, I argue, appropriate to different situations. Given this, I choose to model agents facing multiple *interim fixed points* as responding to them with a given *procedure* and *response function*. The latter captures how the agent responds to ambiguity, as the presence of multiple fixed points is arguably an ambiguous situation. The former reflects to what extent the agent maintains candidate probabilities produced by earlier stages of reasoning. I consider two candidate procedures, one of which is intended to respond to situations of absent-mindedness (where there is only one agent at the information set), with the other more appropriate for situations in which multiple agents share the same information set. When studying Bayesian social learning in the presence of temporal ignorance, as I do in some of my other research, this second procedure will be the correct one to use, whereas in models with absent-mindedness the first one is better. The response functions reflect the two responses I consider to ambiguity: Bayesianism with the principle of insufficient reason, and expected utility maximin à la Gilboa and Schmeidler (1989).

The first contribution of this paper is to formalise this problem, which I do in Section 2. Whether or not the iterative procedure under consideration 'converges', that is, eventually produces a unique strategy that is self-ratifying, is the outcome of interest. I can answer this thanks to a reduction of Beauty's reasoning to a finite directed graph. Although her deliberation ranges over probability distributions (at any stage of Procedure 1 her candidate set is the simplex $\Delta(T)$ of all strategies supported on some action set T), her response to it is obtained by evaluating her best response to the mean of a uniform belief over the candidates. The response therefore depends only on T and not on any other feature of the candidate set, so the whole process collapses to a map g on the finite family $\mathcal{T}$ of possible supports (actions played with positive probability in a given strategy). This is the *support digraph* Γ: one vertex for each support (element of the power set of actions), with a single outgoing edge $T \to g(T)$. Because the response is uniquely pinned down, every vertex has exactly one successor, so Γ is a functional digraph, and the ordinary best-response relation between pure actions, G, is simply Γ restricted to singleton supports.

Procedure 1, which is intended for single agent games, does not necessarily converge. It discards 'earlier' candidates (those from earlier rounds of reasoning) and each round posits that other instances of Beauty are uniformly randomising over the best-responses identified in the last round. From the first round onwards, the state is a single support over a subset of the action-space, and each step simply applies g to identify a new set of actions between which Beauty is indifferent. The procedure converges exactly

[2] I am assuming here that there is no reason to assume *other* temporal parts are somehow managing to perfectly coordinate with magic psychic powers, though the instance in question does not manage to. This implies their choices are independent.

when this orbit reaches a fixed point of g, a support that reproduces itself under the response and is therefore self-ratifying. Since Γ is a function on finitely many vertices, an orbit has only two possible fates: it reaches a fixed point, or it falls into a cycle and repeats forever. A deterministic walk on a finite vertex set must eventually revisit a vertex, and the first such revisit closes a cycle, so no third possibility exists. Procedure 1 therefore converges from a given starting point exactly when that orbit avoids a cycle of length at least two, and from every starting point exactly when Γ contains no such cycle at all. (In generic binary-action games such as the absent-minded driver, this Γ-criterion is equivalent to requiring the acyclicity of G, as noted in Corollary 3.3.) A converging orbit visits distinct vertices until it stops, since a repeat would be a cycle, and there are only $2^m - 1$ vertices, so it reaches its fixed point within $2^m - 2$ steps.

Procedure 2, intended for the case of several agents sharing the information set, behaves differently, and better: it always converges, whichever response function Beauty adopts. This is the good news for the multi-agent setting, where coordinating in advance on a single fixed point is least plausible and the symmetry that would let a single agent trust her copies to reason identically is weakest. Accumulation never discards a candidate; each round upon deciding she is indifferent between a given set of actions, Beauty adds all mixtures over this set of actions to the set of strategies she thinks other instances may be playing. She stops once her response to the accumulated pile is already contained in something she has gathered, so that adding it would change nothing. Any round at which this does not terminate must therefore contribute a genuinely new support, and there are only $2^m - 1$ supports to contribute, so she is logically must stop within that many rounds. The strategy to which it converges to need not be an interim fixed point of the game: the terminal play is an average of the mixtures Beauty has visited, or the uniform mixture over her final response, and it is self-ratifying only at the coarse level of supports, not necessarily a best response to itself.

Finally, none of this depends on whether Beauty reasons as a Bayesian with the Principle of Insufficient Reason or alternatively following the EU-maximin prescription. The convergence arguments use only what the two rules share: for Procedure 1, that g is a function on a finite vertex set, and for Procedure 2, that each response is again a support, so that accumulation exhausts a finite family. They can differ in the terminal play they produce, but whether a procedure converges is the same under both.

The remainder of the paper is organised as follows. Section 2 sets up the model: single-information-set games with self-locating uncertainty, the interim fixed-point set MPP, the two response functions (the PIR Bayesian rule and the maximin rule), and the two procedures (replacement and accumulation). Section 3 develops the convergence theory in full for the PIR Bayesian responder, introducing the support digraph Γ and characterising when each procedure converges. Section 4 then carries the analysis over to the maximin responder; because the two rules agree on singleton candidate sets they generate the same best-response digraph G, so the convergence conditions of Section 3 transfer without change, and only the orbits through multi-element vertices and the terminal play of Procedure 2 differ. Section 5 situates the paper within the wider literature, and Section 6 concludes. Proofs and calculations omitted from the main text are collected in the appendices.

2 Model

I study *single-information-set games with self-locating uncertainty*. These are extensive-form games which feature only one information set, ι, containing multiple decision nodes $h_1, \ldots, h_n$. In contrast to the standard restriction in game theory that each information set belongs to only one agent, in my other articles on such games (*Sleeping Beauty's Dismal Day Out* and *Sleeping Newcomb*) I allow multiple agents to exist at the same information set. Formally, I will allow this here too, though it is unimportant for the analysis. One could impose that there is only one agent, *Sleeping Beauty*, whose distinct temporal parts awake at the single information set of the game as many times as there are decision nodes on the realised

play of the game. Whilst it changes which of the two procedures introduced below seems most natural, one can apply either to either case. I have Procedure 1 (Replacement) in mind for the single-player reading and Procedure 2 (Accumulation) for the multi-agent reading. The game has a finite state space Θ, and at each node Beauty chooses an action from a finite set $A(h) = \{a_1, \ldots, a_m\}$, $m \geq 2$. She does not know which node she occupies; this is the self-locating uncertainty. As there is only one information set, her (behavioural) strategy for the game consists of a single probability distribution over available actions $\sigma \in \Delta(A(h))$.

Beauty's payoff depends on her action, the state, and the actions of her other temporal parts (the other instances of herself at other nodes in ι). Denote this payoff as $U(q, \sigma)$, conditional on being at ι, when she plays $q \in \Delta(A(h))$ and believes each other instance's action is drawn independently from behavioural strategy (this is just a distribution over actions, since there is only one information set) σ. A behavioural strategy, σ, that is a best-response to the belief that all 'other' temporal parts are also playing it is an *interim fixed point*:

Definition 2.1 (Interim Fixed Points)**.** The set of *interim fixed points* is

$$\text{MPP}(\iota) \;=\; \big\{\sigma \in \Delta(A(h)) : \sigma \in \underset{q \in \Delta(A(h))}{\arg\max}\, U(q, \sigma)\big\}.$$

That is, $\sigma \in \text{MPP}(\iota)$ if σ is a best response to the belief that all other instances play σ. I write MPP when ι is clear from context.

MPP stands for 'magic psychic powers': each fixed point provides a solution to the game if each temporal part of Beauty can psychically coordinate with all others. If we are considering a single-agent game and there is some planning stage, these become more plausible solutions to the game, as committing to a strategy before arriving at the information set at all is possible.[3] Aumann et al. (1997) call the elements of *MPP action-optimal probabilities* and Piccione and Rubinstein (1997) use *modified multi-selves equilibria*. When $|\text{MPP}| = 1$, the interim problem has a unique solution, and coordination on this obvious focal-point is easy. I am interested in the case $|\text{MPP}| > 1$, where Beauty faces genuine ambiguity. To model such an agent faced with multiple fixed points, I consider the use of a *response function* R, applied to a set of candidate behavioural strategies. Beauty uses her response function to identify, for any such candidate set of strategies, the subset of pure actions over which she would be indifferent given that set and the belief that her other temporal parts are playing one of those candidate strategies. Definitions 2.2 and 2.3 give the two response functions considered in this paper.

Definition 2.2 (Bayesian Response Under Insufficient Reason)**.** Given a candidate set $S \subseteq \Delta(A(h))$ equipped with the *uniform reference measure* μ_S, let $\sigma_S := \mathbb{E}_{\sigma \sim \mu_S}[\sigma]$ denote the marginal mean strategy under μ_S. The *Bayesian response under the Principle of Insufficient Reason* (PIR) is the subset of pure actions

$$R^{\text{PIR}}(S) \;=\; \underset{a \in A(h)}{\arg\max}\, U(a, \bar{\sigma}_S) \;\subseteq\; A(h).$$

The uniform reference measure μ_S is the principle of insufficient reason applied to S, defined

[3] In their original article, Piccione and Rubinstein (1997), one solution concept they consider involves agents who are able to control the actions of other instances at other decision nodes. Aumann et al. (1997) consider this silly, and involving some 'mysterious psychic process'. It seems similarly strange to assume an ability to coordinate in general, hence I use 'MPP' (Magic Psychic Powers - the process in question) to describe equilibria at which the agents are assumed to be able to do so nonetheless.

componentwise:

(i) if S is finite, μ_S assigns mass $1/|S|$ to each element, so $\bar{\sigma}_S = \frac{1}{|S|}\sum_{\sigma\in S}\sigma$;

(ii) if $S = \Delta(T)$ for some non-empty $T \subseteq A(h)$, μ_S is the uniform (Lebesgue) probability measure on the simplex $\Delta(T)$ (I write μ_T as shorthand), and $\bar{\sigma}_{\Delta(T)} = \sigma_T$, the uniform mixture $\frac{1}{|T|}\sum_{a\in T} e_a$ over T;

(iii) if S is a finite union of pieces of types (i) and (ii), μ_S is the equal-weight mixture over those pieces, and $\bar{\sigma}_S$ is the corresponding equal-weight mixture of the piecewise means.

$R^{\mathrm{PIR}}(S)$ is set-valued: when Beauty is indifferent between several pure actions it returns the maximal indifference set, and otherwise it is a singleton.

An immediate Bayesian response to this ambiguity might be to resort to the *Principle of Insufficient Reason*. Consider any other instance of herself, Beauty might suppose that since there is no specific reason to pick any particular candidate within S, she should act as if said instance is randomising uniformly between them. Exactly what this implies depends on the cardinality of S, as is set out in points (i), (ii), and (iii). An advantage of this is that when we define procedures according to which we iteratively add new elements to our set S, the new Bayesian belief over the new set will still be well-defined as a uniform distribution.

Since Bayesianism is the standard approach to uncertainty in game theory, I consider this the principle response to multiplicity. However, the most well-known other response to ambiguity in the literature is that EU maximin rule of Gilboa and Schmeidler (1989), which applied here can be expressed:

Definition 2.3 (Maximin Response)**.** Given a candidate set $S \subseteq \Delta(A(h))$, the *maximin response* is the subset of pure actions

$$R^{\max}(S) \;=\; \arg\max_{a\in A(h)} \inf_{\sigma\in S} U(a,\sigma) \;\subseteq\; A(h).$$

Beauty assumes that nature picks the worst-case element of S for whatever pure action she chooses, and best-responds to this worst case. As with PIR, $R^{\max}(S)$ is set-valued: when several pure actions tie at the maximin value it returns the maximal indifference set.

For both response rules, $R(S)$ is uniquely determined by S: each is the argmax of a real-valued function over the finite set $A(h)$, hence the (unique) maximal subset of pure actions tied at the optimum. Beauty therefore faces no ambiguity in computing the action subset over which she is indifferent. Any residual ambiguity concerns only the distribution with which she will mix over those actions; this is the question she resolves according to one of the procedures I set out below. This is the second of the two big elements I must set out to complete my description of Beauty's reasoning: the *iterative procedures*. Applying the response function to the set MPP produces a maximal indifference set of pure actions, $R(\mathrm{MPP}) \subseteq A(h)$, but this still leaves Beauty's choice underdetermined: she must commit to a specific distribution over those actions, and even once she has done so, the resulting strategy may not be a best-response to the candidate set. Given a response function R (either R^{PIR} or $R^{\max}$), I present two candidate iterative procedures, and study whether they converge to a strategy that is a best-response to the candidate set. Both procedures invoke the support of a strategy and the family of supports realised in a candidate set: for any $\sigma \in \Delta(A(h))$, write $\mathrm{supp}(\sigma) = \{a \in A(h) : \sigma(a) > 0\}$ for its support, and for any $S \subseteq \Delta(A(h))$ write $\mathrm{Supp}(S) = \{\mathrm{supp}(\sigma) : \sigma \in S\} \subseteq 2^{A(h)} \setminus \{\emptyset\}$ for the family of supports realised in S.

Procedure 1: Replacement

1. Compute $S_0 = \text{MPP}$.
2. If $|S_0| = 1$, play this strategy. Stop.
3. Compute $R(S_0) \subseteq A(h)$.
   (a) If $R(S_0) \in \text{Supp}(S_0)$ and a unique $s^* \in S_0$ has $\text{supp}(s^*) = R(S_0)$: play s^*. Stop.
   (b) Otherwise, set $T_1 = R(S_0)$ and $S_1 = \Delta(T_1)$.
4. For $k \geq 1$: $S_k = \Delta(T_k)$. Compute $R(S_k) \subseteq A(h)$.
   (a) If $R(S_k) = T_k$: the candidate support is fixed under R. Play the terminal action $\bar{\sigma}_K$ (Definition 2.4). Stop.
   (b) Otherwise, set $T_{k+1} = R(S_k)$ and $S_{k+1} = \Delta(T_{k+1})$. Repeat.

Procedure 2: Accumulation

1. Compute $S_0 = \text{MPP}$.
2. If $|S_0| = 1$, play this strategy. Stop.
3. Compute $R(S_0) \subseteq A(h)$.
   (a) If $R(S_0) \in \text{Supp}(S_0)$ and a unique $s^* \in S_0$ has $\text{supp}(s^*) = R(S_0)$: play s^*. Stop.
   (b) Otherwise, set $T_1 = R(S_0)$ and $S_1 = \Delta(T_1)$.
4. For $k \geq 1$: $S_k = \Delta(T_1) \cup \cdots \cup \Delta(T_k)$. Compute $R(S_k) \subseteq A(h)$.
   (a) If $R(S_k) \subseteq T_i$ for some $i \in \{1, \ldots, k\}$ (equivalently, $\Delta(R(S_k)) \subseteq S_k$): the candidate set is self-ratifying under accumulation. Play the terminal action $\bar{\sigma}_K$ (Definition 2.4). Stop.
   (b) Otherwise, set $T_{k+1} = R(S_k)$ and $S_{k+1} = S_k \cup \Delta(T_{k+1})$. Repeat.

Under Procedure 1, the candidate set is *replaced* at each stage by the simplex $\Delta(R(S_k))$ of strategies supported on Beauty's current response: she discards prior candidates and treats every distribution over her current indifference set as a separate (and equally plausible) candidate. Under Procedure 2, Beauty *accumulates* her successive response simplices: the candidate set at step k is the union $\Delta(T_1) \cup \cdots \cup \Delta(T_k)$ of every response simplex she has visited. The original MPP elements σ_p have support contained in some T_i generically (their supports are exactly the ones R surfaces, beginning with $T_1 = R(\text{MPP})$), so the simplex pieces capture them as candidates automatically. Accumulation thus formalises the idea that Beauty does not dismiss the strategies her own reasoning has surfaced at previous stages, even as she expands the candidate set. Why introduce two procedures rather than one? I intend Procedure 1 to be the canonical approach in the single-agent reading: if only one absent-minded agent awakens at the information set, she can be confident that any other instance of herself awakening at said information set engages in an identical line of reasoning, so if after k rounds she replaces S_k with S_{k+1}, she can be confident that those other instances also perform that step. Discarding the strategies in S_k is then obvious, and Replacement is the natural formalisation.

In the multi-agent reading the symmetry that licenses this confidence weakens. With several distinct agents at the same information set there may not be common knowledge of infinite reasoning depth: even a fully rational Bayesian, uncertain about how many rounds her counterparts have iterated, should hedge over the strategies surfaced at every round visited rather than commit only to the support reached at her own current depth. That hedge is exactly the union $\Delta(T_1) \cup \cdots \cup \Delta(T_k)$ of Procedure 2.[4] Accumulation is therefore not a concession to bounded rationality but the Bayesian response to a strictly weaker common-

[4]This paper's framework aligns directly with the cognitive-hierarchy literature (Camerer et al. (2004) and earlier Stahl and Wilson (1995)), in which each opponent is independently drawn from a distribution over reasoning depths. Procedure 2's accumulation here, with uniform weighting across visited rounds, corresponds to a uniform prior over depths $\{1, \ldots, k\}$; a Poisson-weighted refinement would track CHC more directly but is not pursued here. The structural results survive most reasonable choices of weights: the termination bound $K \leq 2^m - 1$ is a purely combinatorial count of distinct supports in the downward closure and uses no measure at all, and what declining weights would shift is the specific PIR terminal play (a weighted average $\sum_i w_i \sigma_{T_i}$ in place of the equal-weight $\frac{1}{K} \sum_i \sigma_{T_i}$), not the qualitative fact of termination.

knowledge assumption, in the spirit of the level-k literature's treatment of heterogeneous reasoning depths. A reader who is unconvinced, who takes Procedure 1 to be appropriate even in the multi-agent case, will treat the Procedure 1 results as the principal ones; I take the reasoning-depth-uncertainty motivation to be strong enough that the Procedure 2 results earn their place alongside. Of course, this uncertainty over reasoning depth actually helps here: the procedure motivated by the weaker common-knowledge assumption is also the one that always terminates.

The procedure boxes themselves are response-rule-neutral: every step has the same form under R^{PIR} and $R^{\max}$. The supports $T_1, T_2, \ldots$ that the orbit visits will generically differ between the two rules (since R^{PIR} and $R^{\max}$ disagree at multi-element vertices of the support digraph, introduced in Section 3), but the box's structural dynamics do not. The terminal action $\bar{\sigma}_K$ at the stop condition does depend on which response rule the agent is using, however, and is therefore specified separately:

Definition 2.4 (Terminal play)**.** At the terminating step K of a procedure, with accumulated supports $T_1, \ldots, T_K$ and final candidate set S_K, Beauty's terminal action $\bar{\sigma}_K \in \Delta(A(h))$ depends on her response rule:

(i) Under R^{PIR}, $\bar{\sigma}_K$ is the mean of the PIR reference measure μ_K on S_K:

$$\bar{\sigma}_K^{\mathrm{PIR}} \;=\; \mathbb{E}_{p\sim\mu_K}[p] \;=\; \begin{cases} \sigma_{T_K} & \text{(Procedure 1)},\\ \frac{1}{K}\sum_{i=1}^{K}\sigma_{T_i} & \text{(Procedure 2)}. \end{cases}$$

Here $\sigma_T := \frac{1}{|T|}\sum_{a\in T} e_a$ denotes the uniform mixture over T, and the reference measure is $\mu_K = \mu_{T_K}$ for Procedure 1 and $\mu_K = \frac{1}{K}\sum_{i=1}^{K}\mu_{T_i}$ for Procedure 2 (Definition 2.2). Beauty's terminal play under PIR is therefore the action-level marginal of her belief about an arbitrary copy's strategy.

(ii) Under $R^{\max}$, $\bar{\sigma}_K$ is the uniform mixture over the final response set:

$$\bar{\sigma}_K^{\max} \;=\; \sigma_{R^{\max}(S_K)} \;=\; \frac{1}{|R^{\max}(S_K)|}\sum_{a\in R^{\max}(S_K)} e_a.$$

For Procedure 1 this equals σ_{T_K}, since the stop condition forces $R^{\max}(S_K) = T_K$. For Procedure 2 it is the uniform mixture over $R^{\max}(S_K)$ itself, which is generally not the same point in $\Delta(A(h))$ as the PIR formula in (i).

Three technical points about the procedure boxes are worth flagging. Firstly, Step 2 is generically subsumed by Step 3(a): for an interim fixed point σ^* one has $R(\{\sigma^*\}) = \mathrm{supp}(\sigma^*)$, so Step 3(a) fires with σ^* as the unique match. Step 2 is retained only to cover the non-generic case where some action outside $\mathrm{supp}(\sigma^*)$ ties at the best response, so that the procedure still terminates at the unique interim fixed point in that edge case. Secondly, whilst for Procedure 1 the two response rules' terminal plays coincide (both equal σ_{T_K}), for Procedure 2 they generically differ, since PIR averages over the whole accumulation history while maximin reads off only the final response. This asymmetry is intrinsic to the two rules: PIR carries a reference measure that registers Beauty's full reasoning history, while maximin evaluates only the worst case against the current candidate set. Thirdly, the uniform-mixture specification of $\bar{\sigma}_K^{\max}$ in Definition 2.4(ii) is not a substantive one: all pure actions in $R^{\max}(S_K)$ are tied at the maximin value, so how exactly Beauty randomises over them does not matter. Choosing the uniform mixture simply lets the procedure formally produce a single terminal action.

Example. This example walks through the procedures on the 4-intersection AMD with payoffs (100, 132, 0, 200, 100), whose interim fixed-point set is

$$\text{MPP} \;=\; \{\sigma_{0.2},\ \sigma_{0.5},\ \sigma_{0.8}\}\,.$$

The detailed calculations underlying the PIR responses below are gathered in Appendix B; here I present only the algorithmic steps.

Procedure 1. Starting from $S_0 = \text{MPP}$, the mean is $\bar{\sigma}_{S_0} = (\sigma_{0.2}+\sigma_{0.5}+\sigma_{0.8})/3 = \sigma_{1/2}$, at which Beauty is indifferent between continuing and exiting, so $R^{\text{PIR}}(S_0) = \text{BR}(\sigma_{1/2}) = \{\text{Cont}, \text{Exit}\}$. This response lies in $\text{Supp}(S_0)$, but three elements of S_0 share it, so Step 3(a) fails by non-unique match and Step 3(b) fires: $T_1 = \{\text{Cont}, \text{Exit}\}$, $S_1 = \Delta(\{\text{Cont}, \text{Exit}\})$. The mean of S_1 is again $\sigma_{1/2}$, whose response is once more $\{\text{Cont}, \text{Exit}\} = T_1$, so Step 4(a) fires and the procedure converges in one iteration to $\bar{\sigma}_1 = \sigma_{T_1} = \sigma_{1/2}$, the uniform mixture over $\{\text{Cont}, \text{Exit}\}$. The terminal play coincides with the middle element of MPP because $\sigma_{1/2}$ is itself an interim fixed point, a non-generic characteristic of these payoffs (Proposition 3.2): for generic payoffs the support $\{\text{Cont}, \text{Exit}\}$ would *not* reproduce itself.

Procedure 2 runs identically through the formation of $T_1 = \{\text{Cont}, \text{Exit}\}$ and $S_1 = \Delta(\{\text{Cont}, \text{Exit}\})$. Its stop condition asks whether $R(S_1) \subseteq T_i$ for some accumulated T_i; since $R(S_1) = \text{BR}(\sigma_{1/2}) = \{\text{Cont}, \text{Exit}\} \subseteq T_1$, Step 4(a) fires immediately and Beauty plays $\bar{\sigma}_1 = \sigma_{T_1} = \sigma_{1/2}$. The two procedures coincide here because both come to rest at the same self-reproducing multi-element support $\{\text{Cont}, \text{Exit}\}$; they generically diverge, as the AHP game below shows.

Applying Procedure 1 under PIR to the Figure 1 game of the introduction reproduces that walkthrough step-for-step: $\bar{\sigma}_{\text{MPP}} = \sigma_{11/45}$ and $\text{BR}(\sigma_{11/45}) = \{\text{Cont}\}$, so $T_1 = \{\text{Cont}\}$; then $\text{BR}(\sigma_{\text{Cont}}) = \{\text{Exit}\}$, so $T_2 = \{\text{Exit}\}$; then $\text{BR}(\sigma_{\text{Exit}}) = \{\text{Exit}\} = T_2$ and Step 4(a) fires with terminal play σ_{Exit} (Appendix B.2 for the details). Procedure 2 differs on the same game: its accumulated $S_2 = \{\sigma_{\text{Cont}}, \sigma_{\text{Exit}}\}$ has mean $\sigma_{1/2}$ at which the response is tied, and after one more accumulation round it terminates at $\sigma_{1/2}$ rather than σ_{Exit}.

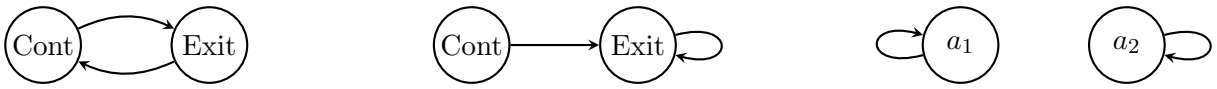
3 PIR Bayesian Results

To state the convergence results I first need two directed graphs that encode the structure of Beauty's best responses. The *best-response digraph* G takes the individual actions as its vertices and records, for each action, the best response to the belief that every other instance plays it. The *support digraph* Γ instead takes the *sets* of actions as its vertices and records Beauty's response to the uniform mixture over each set. Whether either procedure converges will turn out to be a combinatorial property of these two graphs, so I set them up before turning to the results themselves. Let us begin by defining $\text{BR}(a) = R(\{e_a\}) \subseteq A(h)$ as the *set* of best responses to the singleton belief that all other instances play a.

Definition 3.1 (Best-Response Digraph)**.** The *best-response digraph* $G = (A(h), E)$ has vertex set $A(h)$ and a directed edge $a_i \to a_j$ whenever $a_j \in \text{BR}(a_i)$. Each vertex has out-degree at least 1.

A *fixed point* of BR is a vertex with a self-loop: $a \in \text{BR}(a)$. A *cycle* of length $k \geq 2$ is a sequence $a_{i_1} \to a_{i_2} \to \cdots \to a_{i_k} \to a_{i_1}$ of distinct actions. In the binary case ($m = 2$, actions $\{\text{Cont}, \text{Exit}\}$), the only possible cycle is the 2-cycle Cont $\to$ Exit $\to$ Cont, which occurs when $\text{Exit} \in \text{BR}(\text{Cont})$ and $\text{Cont} \in \text{BR}(\text{Exit})$. The absent-minded driver exhibits exactly this 2-cycle.

The best-response digraph G records the procedures' dynamics on singleton-support candidate sets. The procedures' candidate set can also carry strategies with multi-element supports (the full simplex $\Delta(T)$



(a) 4-intersection AMD $(100, 132, 0, 200, 100)$: 2-cycle.

(b) AHP 3-intersection AMD $(7, 0, 22, 2)$: BR-fixed point at Exit.

(c) Symmetric coordination game (Appendix C): two BR-fixed points.

Figure 2: Best-response digraphs G for three example games: the 4-intersection AMD of Section 2 (left), the AHP 3-intersection AMD of the introduction (centre), and the symmetric 2-action coordination game of Appendix C (right), which I include to produce examples of digraphs here and in Figure 3.

once Beauty's response correspondence returns a multi-element T), and the response correspondence on those richer candidate sets is what governs convergence in general. The relevant object is:

Definition 3.2 (PIR Support Digraph)**.** The *PIR support digraph* $\Gamma = (\mathcal{T}, F)$ has vertex set $\mathcal{T} = 2^{A(h)} \setminus \{\emptyset\}$, the family of non-empty subsets of $A(h)$, and an edge from each $T \in \mathcal{T}$ to $g(T) := R^{\mathrm{PIR}}(\Delta(T))$, the best response to the uniform mixture over T:

$$g(T) \;=\; \arg\max_{a \in A(h)} U(a, \sigma_T) \;=\; \mathrm{BR}(\sigma_T),$$

where $\sigma_T = \frac{1}{|T|} \sum_{a \in T} e_a$ is the uniform mixture over T.

By the uniqueness of R, Γ is a *functional* digraph: each vertex has out-degree exactly 1. It contains the best-response digraph G as a sub-digraph: the singleton supports $\{\{a\} : a \in A(h)\}$ inherit the edges of G exactly, because $\sigma_{\{a\}} = e_a$ is the pure strategy, so $g(\{a\}) = \mathrm{BR}(\sigma_{\{a\}}) = \mathrm{BR}(a)$. However, Γ has $2^m - 1$ vertices in total, and its multi-element-support vertices and edges have no analogue in G. The maximin responder induces a different support digraph $\Gamma^{\max}$, which agrees with Γ on the singleton supports but differs at multi-element vertices; it is introduced in Section 4.

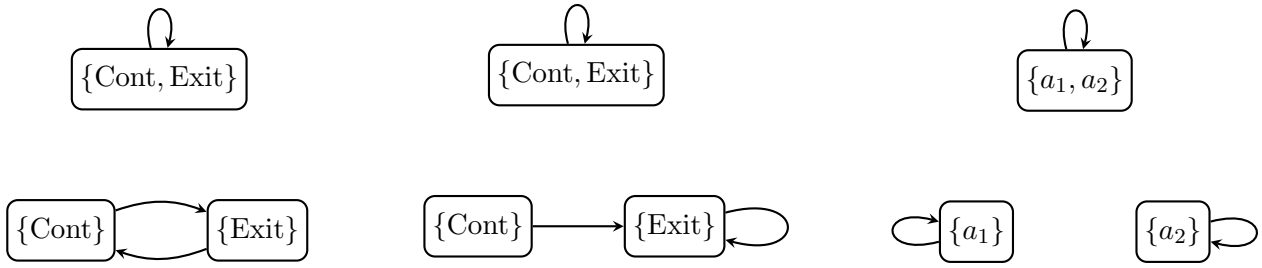
A *fixed point* of Γ is a vertex with a self-loop: $g(T) = T$. A singleton fixed point $\{a\}$ is exactly a BR self-loop in G. A multi-element fixed point (a T with $|T| \geq 2$ and $g(T) = T$) corresponds to the uniform mixture σ_T being itself an interim fixed point with support exactly T: $\mathrm{BR}(\sigma_T) = T$. A *cycle of length* $k \geq 2$ in Γ is a sequence of distinct supports $T_1 \to T_2 \to \cdots \to T_k \to T_1$. A cycle in G at the action level $a_{i_1} \to \cdots \to a_{i_k} \to a_{i_1}$ projects to a cycle in Γ at the singleton level $\{a_{i_1}\} \to \cdots \to \{a_{i_k}\} \to \{a_{i_1}\}$; not all Γ-cycles arise this way, but as Proposition 3.2 below makes precise, for generic AMD games the singleton-projected ones are the only cycles to worry about.

3.1 Convergence of the Procedures

The procedures' state from step 1 onward consists of either a single simplex $\Delta(T)$ (Procedure 1, replacement) or a growing union of simplices $\bigcup_i \Delta(T_i)$ (Procedure 2, accumulation), and each step transitions according to Γ. The convergence question therefore becomes a question about the structure of Γ.

3.1.1 Procedure 1: Replacement

Procedure 1 maintains $S_k = \Delta(T_k)$ for all $k \geq 1$: each replacement step sets $T_{k+1} = R(S_k) = g(T_k)$ and $S_{k+1} = \Delta(T_{k+1})$ (and, under R^{PIR}, equips S_{k+1} with the uniform reference measure $\mu_{T_{k+1}}$ of



(a) 4-int AMD: $g(\{\text{Cont}, \text{Exit}\}) = \{\text{Cont}, \text{Exit}\}$; multi-element fixed point at $\sigma_{1/2}$.

(b) AHP: $g(\{\text{Cont}, \text{Exit}\}) = \{\text{Cont}, \text{Exit}\}$; multi-element fixed point at $\sigma_{1/2}$.

(c) Symmetric coordination: $\{a_1, a_2\}$ is a multi-element Γ-fixed point.

Figure 3: Support digraphs Γ for three example games: the 4-intersection AMD of Section 2 (left), the AHP 3-intersection AMD of the introduction (centre), and the symmetric 2-action coordination game of Appendix C (right).

Definition 2.2). Letting $T_1 = R(\text{MPP})$, the post-step-0 candidate sequence is the orbit of g on Γ starting at T_1, and the procedure terminates exactly when this orbit reaches a fixed point of g.

Theorem 3.1 (Procedure 1: Convergence)**.** Procedure 1 converges from initial candidate set S_0 if and only if either

(i) $|S_0| = 1$, or $R(S_0) \in \text{Supp}(S_0)$ has a unique match in S_0 (the procedure terminates at step 0), or

(ii) the orbit T_1, $g(T_1)$, $g^2(T_1), \ldots$ of g on Γ, starting at $T_1 = R(S_0)$, reaches a fixed point of g.

Procedure 1 converges from every initial S_0 if and only if Γ contains no cycles of length ≥ 2. If convergence occurs, it does so in at most $|\mathcal{T}| - 1 = 2^m - 2$ steps.

Proof. (i) is immediate from steps 2 and 3(a) of the procedure. For (ii), the candidate set after step 0 is $\Delta(T_1)$ with $T_1 = R(S_0)$. At step $k \geq 1$, $S_k = \Delta(T_k)$, and step 4(a) fires iff $R(S_k) = T_k$, i.e. T_k is a fixed point of g. If reached, the procedure plays $\bar{\sigma}_k = \sigma_{T_k}$, the uniform mixture over T_k (Definition 2.4(i)). Otherwise $T_{k+1} = g(T_k)$ and the procedure continues. Since $\mathcal{T}$ is finite and g is a function, the orbit must eventually either reach a fixed point or enter a cycle of length ≥ 2; convergence holds in the former case and fails in the latter. The acyclicity condition follows because every $T_1 \in \mathcal{T}$ is reachable from some choice of S_0. □

We can say considerably more once we restrict to the absent-minded driver games that feature so prominently in this article. An AMD game is binary-action, so its support digraph Γ has a single multi-element vertex $T^* = \{\text{Cont}, \text{Exit}\}$, and whether Γ does anything that the best-response digraph G does not comes down to that one vertex: whether T^* is a fixed point or lies on a cycle. The following proposition shows that, outside a knife-edge set of payoffs, it is neither, so the convergence test collapses onto G.

Proposition 3.2 (Generic AMD reduction)**.** For an n-intersection AMD with payoffs in an open dense subset of $\mathbb{R}^{n+1}$, the support digraph Γ has no fixed points of cardinality ≥ 2 and no multi-element cycles: every T with $|T| \geq 2$ satisfies $g(T) \neq T$, and the only cycles of Γ live in its singleton-support sub-digraph G.

Proof. See Appendix D. □

Corollary 3.3 (For generic AMD, Procedure 1 convergence reduces to G-acyclicity)**.** For payoffs in the open dense subset identified in Proposition 3.2, Procedure 1 converges from every initial S_0 if and only if the best-response digraph G contains no cycles of length ≥ 2.

Proof. By Proposition 3.2, the only fixed points of Γ are singleton self-loops (i.e. BR-fixed points in G), and the only cycles of length ≥ 2 are the singleton-projected images of G-cycles. Theorem 3.1 then equates universal convergence with G-acyclicity. □

Theorem 3.1 reduces convergence to the absence of cycles of length ≥ 2 in Γ, but Γ has $2^m - 1$ vertices, so testing this directly is exponential in m. This is what makes the generic-AMD reduction useful: for generic payoffs, Corollary 3.3 allows us to check the much smaller best-response digraph G instead.

3.1.2 Procedure 2: Accumulation

Procedure 2's candidate set at step k is $S_k = \Delta(T_1) \cup \cdots \cup \Delta(T_k)$, and under R^{PIR} this is equipped with the reference measure $\mu_k = \frac{1}{k}\sum_{i=1}^{k} \mu_{T_i}$ (Definition 2.2). The accumulation never removes strategies from the candidate set: $S_k \subseteq S_{k+1}$. By the stop condition (step 4(a)), termination occurs once $R(S_k) \subseteq T_i$ for some accumulated T_i, so that adding it would not enlarge the candidate set. Whereas Theorem 3.1 gave us that Procedure 1 need not necessarily terminate, but can cycle forever, this next result establishes that Procedure 2 will always do so.

Theorem 3.4 (Procedure 2: Always Terminates)**.** In any single-information-set SLU game with $|A(h)| = m$, Procedure 2 terminates in at most $2^m - 1$ accumulation steps. Under R^{PIR}, the terminal play at step K is

$$\bar{\sigma}_K \;=\; \frac{1}{K}\sum_{i=1}^{K} \sigma_{T_i}\,,$$

the equal-weight average of the uniform mixtures over the accumulated supports $T_1, \ldots, T_K$ (Definition 2.4(i)).

Proof. At step k, the support family associated to S_k is the downward closure $\mathcal{D}_k := \{T \subseteq A(h) : \emptyset \neq T \subseteq T_i \text{ for some } i \leq k\}$, since each $\Delta(T_i)$ contributes every non-empty sub-support of T_i as the support of some element. If the procedure does not stop at step k, then $R(S_k) = T_{k+1}$ satisfies $T_{k+1} \not\subseteq T_i$ for any $i \leq k$ (else the stop condition would fire), and T_{k+1} is a new accumulated label.

The downward closure $\mathcal{D}_k$ is monotone in k: $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$, and the new label T_{k+1} adds at least itself (which was not in $\mathcal{D}_k$). Since $\mathcal{D}_k \subseteq \mathcal{T} = 2^{A(h)} \setminus \{\emptyset\}$ has at most $2^m - 1$ elements, accumulation can grow $\mathcal{D}_k$ strictly at most $2^m - 1$ times before $\mathcal{D}_k = \mathcal{T}$, at which point any $R(S_k) \in \mathcal{T}$ satisfies $R(S_k) \subseteq T_i$ for the appropriate i and the stop condition fires. The terminal play formula in the statement follows from Definition 2.4(i): under PIR, $\bar{\sigma}_K = \mathbb{E}_{p\sim\mu_K}[p]$ with $\mu_K = \frac{1}{K}\sum_i \mu_{T_i}$, and the mean of μ_{T_i} on $\Delta(T_i)$ is the uniform mixture σ_{T_i}. □

Procedure 2's terminal play $\bar{\sigma}_K$ does not in general have support equal to any Γ-fixed-point support: termination is a property of the accumulation operation (no new strategies added) rather than of g at any single vertex, so Procedure 2's output need not be an interim fixed point of the underlying game.

4 Maximin Responses

The convergence analysis of Section 3 was developed for the PIR Bayesian responder. The maximin alternative (Definition 2.3) replaces the uniform-mixture evaluation with the worst-case minimum over S_k. At multi-element vertices the two rules pull apart: $g^{\mathrm{PIR}}(T) = \mathrm{BR}(\sigma_T)$ is the best response to the uniform mixture, while $g^{\max}(T) = R^{\max}(\Delta(T))$ involves minimisation over $\Delta(T)$. They agree on singletons:

Proposition 4.1 (Singleton Agreement)**.** For any strategy $\sigma \in \Delta(A(h))$, $R^{\max}(\{\sigma\}) = R^{\mathrm{PIR}}(\{\sigma\}) = \arg\max_{a \in A(h)} U(a, \sigma)$.

Proof. When $S = \{\sigma\}$, both $\min_{\sigma' \in S} U(a, \sigma') = U(a, \sigma)$ and $\frac{1}{|S|} \sum_{\sigma' \in S} U(a, \sigma') = U(a, \sigma)$. Both reduce to the standard pure-action best response to σ. □

The two rules therefore generate the same singleton-support sub-digraph G, and the support digraphs Γ^{PIR} and $\Gamma^{\max}$ differ only at multi-element vertices. The convergence theorems carry over without change:

Theorem 4.2 (Maximin convergence)**.** Theorems 3.1 and 3.4 hold verbatim for convergence and termination with R^{PIR} replaced by $R^{\max}$ and Γ replaced by $\Gamma^{\max}$: Procedure 1 converges from every S_0 iff $\Gamma^{\max}$ has no cycles of length ≥ 2, and Procedure 2 terminates in at most $2^m - 1$ accumulation steps. The terminal action is given by Definition 2.4(ii).

Proof. The Procedure 1 proof depended only on g being a function on a finite vertex set; the Procedure 2 proof used only that $R(S_k) \in \mathcal{T}$. The terminal action is then read off from Definition 2.4(ii). □

The orbits on Γ^{PIR} and $\Gamma^{\max}$ differ at multi-element vertices of the support digraph, so the accumulated support sequences and the response sets visited generally diverge between the two rules. They also differ in the terminal action played by Procedure 2: under PIR Beauty plays the history-averaged mixture $\frac{1}{K} \sum_i \sigma_{T_i}$, the mean of her reference measure on S_K; under maximin she plays the uniform mixture $\sigma_{R^{\max}(S_K)}$ over the final response, which depends only on $R^{\max}(S_K)$ and not on the rest of the support history. At Procedure 1's stop condition, by contrast, this maximin terminal action $\sigma_{R^{\max}(S_K)}$ coincides with the PIR terminal action σ_{T_K}, so the two rules part company only in Procedure 2's terminal play. The convergence condition for Procedure 1, in turn, is response-rule-independent in the regimes worked out here. For generic AMD payoffs this rule-independence sharpens further: the codimension argument of Proposition 3.2 adapts directly to the maximin rule, with uniform-mixture evaluations replaced by minima of rational functions that remain analytic in the payoffs outside a measure-zero subvariety, so Procedure 1 convergence again reduces to G-acyclicity, exactly as for PIR (Corollary 3.3).

5 Literature Review

This paper follows on, of course, from Piccione and Rubinstein (1997) in which the paradox of the absent-minded driver was first set out, and Aumann et al. (1997), in which they noted the possible multiplicity of solutions in passing (the point of the article was to argue in favour of the now standard analysis of it). This initial discussion then inspired a vast literature on the *Sleeping Beauty Problem*, which is a distinct thought experiment that isolates the epistemological paradox (as opposed to the decision theoretic one) revolving around one's beliefs about the state of the world and one's own self-location (which decision node are you at?). Winkler (2017) provides a useful literature review and summary of this problem, but its particular relevance in economics, I believe, is that it illustrates how self-locating uncertainty interacts curiously with learning problems. Whenever agents are forming beliefs about some state of the world, but their number of 'awakenings' (decision nodes within a given information set) itself

depends on said state of the world, the '*thirder*' beliefs the consensus supports are counter-intuitive. When this number of awakenings depends also on their choice of action (as in the absent-minded driver) the multiplicity confronted in this article can emerge as an additional problem.

A particularly promising direction to pursue in economic theory, I believe, is the study of learning under *temporal ignorance*, and indeed this is what much of my other research is currently geared at doing. The present article serves a clear function here, in that particularly in models of social learning (for which Procedure 2 is intended) in which the number of awakenings depends on the action taken by different instances at the information set in question, it shows how to resolve the multiplicity of interim fixed points that may arise. If, as I am, one is convinced that agents simply managing to coordinate on a given solution in a game with many agents is implausible, this is necessary. The primary relation between *Sleeping Beauty Loses Her Magic Psychic Powers* is this one: it has been written to fill this particular gap.

Beyond the existing literature on self-locating uncertainty, there is also a philosophical literature on *unstable decision problems*. Decision instability is precisely the property of a decision problem that the choice of one action is itself evidence that another ought instead be chosen. Where above I noted that the Sleeping Beauty problem isolates the epistemological paradox of the absent-minded driver from the decision theoretic one: this is the nature of said decision-theoretic paradox. One of the most famous decision problems considered in this literature is that of *Death in Damascus* (first considered by Gibbard and Harper (1978)). An agent learns that Death will come for him tomorrow in whichever city he is in, and must decide whether to stay in Damascus or flee to Aleppo. If he stays, he expects Death to come to Damascus; but then he should flee. If he flees, he expects Death to come to Aleppo; but then he should stay. The agent's deliberation oscillates: any resolution of his indecision undermines itself. This is of course reminiscent of the paradox of the absent-minded driver, and the conclusion in this literature that self-ratifying solutions are the only viable ones foreshadows that same conclusion of Piccione and Rubinstein. In Death in Damascus, a causal decision theorist recommends uniform randomisation as the optimal resolution, and in the absent-minded driver continuing with probability 2/3 wins the day (Harper, 1986). Joyce (2012) provides an analysis, developing on Arntzenius (2008) and Skyrms (1990), of what he calls *regret-based instability*, of which the convergence of the present paper is reminiscent. For one thing, he consider applying Skyrms' model of deliberation to unstable problems, and since this is also a reasoning process that iterates through steps on a finite set of acts just as in this paper.

Although reminiscent of this paper however, it does not consider the same question, and the similarities are more superficial than substantial. Instead, Joyce considers problems in which every action induces regret and simply argues that the premise that non-ratifiable actions are irrational must be dropped when evaluating the rationality of an action in such problems. Here of course, I accept this premise in dealing with problems with a multiplicity of self-ratifying solutions and ask whether the agent can reason her way to one of them. The iterated reasoning is attempting to solve a different problem, and beyond involving iteration the two processes are not related.

6 Conclusion

I have studied what happens when an agent in a game with self-locating uncertainty 'loses her magic psychic powers', that is, when she acknowledges the multiplicity of interim fixed points and attempts to reason her way to a definite choice through iterative best-responding, rather than managing to coordinate perfectly with her other temporal parts by assumption. The agent's response function R takes a candidate set of strategies S and returns the maximal subset of pure actions over which she is indifferent given S. The procedures iterate this construction, either replacing (Procedure 1) or accumulating (Procedure 2) candidate strategies, until the candidate set is self-ratifying. Whether or not these procedures eventually terminate is the main question of the article, and whilst Procedure 2 always converges, Procedure 1

convergence is more complicated.

This is governed by the *support digraph* Γ, a functional digraph on the family of non-empty subsets of feasible actions $A(h)$, in which the edges gives 'best responses' (under specific conditions) to mixes over the subset of actions represented by a given node. Procedure 1 converges from every initial candidate set if and only if Γ contains no cycles of length ≥ 2. The best-response digraph G on individual actions sits inside Γ as the singleton-support sub-digraph, and for generic absent-minded-driver games it is sufficient to consider this smaller subgraph instead: Procedure 1 convergence reduces to acyclicity of G alone. Hence, the decision instability problem that randomisation solves in the absent-minded driver is resurrected via fixed point multiplicity if one follows Procedure 1.

The results are robust to one's choice of the two response rules considered: PIR and maximin agree on singleton candidate sets, and the convergence conditions for both procedures carry over to the maximin responder. The two rules differ in which intermediate supports the orbit visits and in the terminal action selected by Procedure 2, but not in whether the procedure converges.

A Interim Fixed Point Multiplicity in the Absent-Minded Driver

In this appendix I consider absent-minded driver games in a little more detail, noting when one can guarantee a unique interim fixed point. In an n-intersection absent-minded driver with payoffs $(\alpha_1, \ldots, \alpha_{n+1})$ (α_k utility upon exiting at the k-th intersection and α_{n+1} for continuing all the way through), the agent's belief about her node position is $\pi(X_k \mid \sigma_p) = p^{k-1}/D_n(p)$, with $D_n(p) = 1 + p + \cdots + p^{n-1}$. The interim indifference condition $U(\text{Cont}, \sigma_p) = U(\text{Exit}, \sigma_p)$ reduces to

$$P_n(p) \;:=\; \sum_{k=0}^{n-1} (k+1)\,\beta_{k+1}\,p^k \;=\; 0, \qquad \beta_j := \alpha_{j+1} - \alpha_j. \tag{A.1}$$

The interior interim fixed points are exactly the roots of P_n in $(0,1)$. The corner σ_{Exit} is in MPP iff $\alpha_1 \geq \alpha_2$; the corner σ_{Cont} is in MPP iff $\alpha_{n+1} \geq \frac{1}{n}\sum_{k=1}^{n} \alpha_k$. If an AMD game satisfies the conditions $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ and $\alpha_{n+1} < \frac{1}{n}\sum_{k=1}^{n} \alpha_k$, let us say it satisfies 'monotone payoffs' (i.e. exiting later is always better than exiting early, but continuing off at the end gives a worse payoff than the average exit payoff).

Under monotone payoffs, $\beta_1, \ldots, \beta_{n-1}$ are all positive and β_n is negative. The polynomial $P_n(p)$ in equation (A.1) therefore has coefficient sequence

$$\beta_1,\ 2\beta_2,\ 3\beta_3,\ \ldots,\ (n-1)\beta_{n-1},\ n\beta_n$$

with the first $n-1$ entries positive and the last entry negative, so there must be exactly one sign change.

Proposition A.1 (Monotone-payoff AMDs have a unique interim fixed point)**.** For every $n \geq 2$ and every monotone-payoff n-intersection AMD,

$$|\text{MPP}| \;=\; 1.$$

The unique element of MPP is an interior fixed point σ_{p^*} for some $p^* \in (0,1)$.

Proof. By Descartes' rule of signs, the number of positive real roots of P_n, counting repeated roots once for each repetition, equals the number of sign changes in the coefficient sequence minus a non-negative even integer (necessarily 0 here). With exactly one sign change, P_n has at most one positive real root.

The corner conditions show $\sigma_{\text{Exit}} \notin \text{MPP}$ (because $\alpha_1 < \alpha_2$ violates $\alpha_1 \geq \alpha_2$) and $\sigma_{\text{Cont}} \notin \text{MPP}$ (because $\alpha_{n+1} < \frac{1}{n}\sum_k \alpha_k$ violates the corner condition). So all interim fixed points must be interior, i.e. roots of P_n in $(0,1)$.

It remains to verify P_n has *exactly* one root in $(0,1)$. Observe $P_n(0) = \beta_1 > 0$ and

$$P_n(1) \;=\; \sum_{k=0}^{n-1}(k+1)\beta_{k+1} \;=\; \sum_{k=1}^{n} k\,\beta_k \;=\; n\,\alpha_{n+1} - \sum_{k=1}^{n}\alpha_k \;<\; 0$$

by the monotone-payoff condition. By the intermediate value theorem there is at least one root in $(0,1)$; by Descartes' rule there is at most one positive real root. Hence exactly one root in $(0,1)$, which is also the unique positive real root. □

Multiplicity in MPP therefore requires breaking the monotone-payoff structure. There are two ways to do so, and each enlarges the possible number of sign changes in P_n.

- **Sleeping in the car is pretty comfortable:** $\alpha_k > \alpha_{k+1}$ for some $k < n$. The coefficient β_k becomes negative, introducing additional sign changes. AHP's 3-intersection $(7, 0, 22, 2)$ has $\alpha_1 = 7 > 0 = \alpha_2$, giving $\beta_1 = -7,\ \beta_2 = 22,\ \beta_3 = -20$: two sign changes, hence up to two positive roots. The actual roots are $\{7/30,\ 1/2\}$.
- **The motel isn't so bad:** $(\alpha_{n+1} \geq \frac{1}{n}\sum \alpha_k)$. The corner σ_{Cont} becomes an interim fixed point. This doesn't directly change the sign-change count of P_n, but it can add a pure-strategy fixed point to MPP alongside any interior roots.

The 4-intersection AMD $(100, 132, 0, 200, 100)$ used as a running counter-example elsewhere in this paper is non-monotone in a more dramatic way: $\alpha_3 = 0$ sits in a valley between $\alpha_2 = 132$ and $\alpha_4 = 200$. The coefficients are $\beta_1 = 32,\ \beta_2 = -132,\ \beta_3 = 200,\ \beta_4 = -100$, giving the polynomial-sign-change sequence

$$+,\ -,\ +,\ -$$

with *three* sign changes. By Descartes, there are up to three positive real roots; and we already saw all three lie in $(0,1)$, namely $p \in \{0.2,\ 0.5,\ 0.8\}$.

The sign of P_n at the corners has a separate interpretation in terms of the best-response digraph G (on the action set $\{\text{Cont}, \text{Exit}\}$):

- $P_n(0) > 0 \iff \text{Cont} \in \text{BR}(\sigma_{\text{Exit}})$ (edge Exit → Cont in G).
- $P_n(1) < 0 \iff \text{Exit} \in \text{BR}(\sigma_{\text{Cont}})$ (edge Cont → Exit in G).

A 2-cycle in G requires both conditions: $P_n(0) > 0$ *and* $P_n(1) < 0$. In a monotone-payoff AMD, both hold by construction (because $\beta_1 > 0$ and the corner condition gives $P_n(1) < 0$). So the 2-cycle in G is a built-in feature of monotone AMD.

B Computations for the Examples

All PIR responses are point-evaluations of BR at the mean of the relevant candidate set. The computations below evaluate the indifference polynomial $P_n(p)$ and the rational function $V_{\text{Cont}}(\sigma_p) - V_{\text{Exit}}(\sigma_p) = P_n(p)/D_n(p)$ (Appendix A) at the relevant mean values of p.

B.1 4-Intersection AMD with Payoffs $(100, 132, 0, 200, 100)$

For these payoffs, Appendix A gives the indifference polynomial

$$P_4(p) \;=\; 32 - 264p + 600p^2 - 400p^3 \;=\; -400(p-0.2)(p-0.5)(p-0.8)$$

and MPP $= \{\sigma_{0.2}, \sigma_{0.5}, \sigma_{0.8}\}$.

Step 3: The mean of MPP is $\bar{\sigma}_{\text{MPP}} = \sigma_{(0.2+0.5+0.8)/3} = \sigma_{1/2}$. At this strategy,

$$V_{\text{Cont}}(\sigma_{1/2}) - V_{\text{Exit}}(\sigma_{1/2}) \;=\; \frac{P_4(1/2)}{D_4(1/2)} \;=\; \frac{0}{15/8} \;=\; 0,$$

since $P_4(1/2) = 32 - 132 + 150 - 50 = 0$ (the factor $(p - 1/2)$ in the factorisation of P_4). The two pure actions tie at the PIR response, hence $R^{\text{PIR}}(\text{MPP}) = \{\text{Cont}, \text{Exit}\}$.

Step 4 with $k = 1$. The uniform mixture over $\{\text{Cont}, \text{Exit}\}$ is also $\sigma_{1/2}$, so the same evaluation applies: $R(S_1) = \text{BR}(\sigma_{1/2}) = \{\text{Cont}, \text{Exit}\} = T_1$. Step 4(a) fires.

The fact that the mean of MPP and the uniform mixture over $\{\text{Cont}, \text{Exit}\}$ coincide at $\sigma_{1/2}$ is what makes Procedure 1 terminate on this AMD: $\sigma_{1/2}$ is a tied point of the BR map, and is itself an interim fixed point. This is the non-generic situation excluded from Proposition 3.2's open dense set U by the hyperplane $H_{1/2} = \{P_n(1/2) = 0\}$. Off this hyperplane the multi-element vertex T^* would map to one of the singletons ($P_4(0) = 32 > 0$ and $P_4(1) = -8 < 0$, the two edges of the G-cycle), and Procedure 1 would enter the G-cycle and fail to converge.

B.2 AHP 3-Intersection AMD with Payoffs $(7, 0, 22, 2)$

For these payoffs (introduced informally in Section 1), Appendix A gives the indifference polynomial $P_3(p) = -7 + 44p - 60p^2$ with positive roots $7/30$ and $1/2$, and the corner σ_{Exit} also satisfies its fixed-point condition. So MPP $= \{\sigma_{\text{Exit}},\ \sigma_{7/30},\ \sigma_{1/2}\}$.

Step 3: The mean is $\bar{\sigma}_{\text{MPP}} = \sigma_{(0+7/30+1/2)/3} = \sigma_{11/45}$. At this strategy,

$$V_{\text{Cont}}(\sigma_{11/45}) - V_{\text{Exit}}(\sigma_{11/45}) \;=\; \frac{P_3(11/45)}{D_3(11/45)}.$$

$P_3(11/45) \approx -7 + 10.76 - 3.58 = 0.18$ and $D_3(11/45) \approx 1.30$, so the difference is $\approx 0.14 > 0$ and $R^{\text{PIR}}(\text{MPP}) = \text{BR}(\sigma_{11/45}) = \{\text{Cont}\}$. Since $\{\text{Cont}\} \notin \text{Supp}(\text{MPP}) = \{\{\text{Exit}\}, \{\text{Cont}, \text{Exit}\}\}$, Step 3(a) fails and Step 3(b) fires: $T_1 = \{\text{Cont}\}$.

Step 4 (Procedure 1). With $k = 1$: $\text{BR}(\sigma_{\text{Cont}}) = \{\text{Exit}\}$ ($P_3(1) = -23 < 0$), so $R(S_1) = \{\text{Exit}\} \neq T_1$, Step 4(b) fires, $T_2 = \{\text{Exit}\}$. With $k = 2$: $\text{BR}(\sigma_{\text{Exit}}) = \{\text{Exit}\}$ ($P_3(0) = -7 < 0$), so $R(S_2) = T_2$, Step 4(a) fires. Terminal play $\bar{\sigma}_2 = \sigma_{\text{Exit}}$; Procedure 1 terminates in three iterations.

Step 4 (Procedure 2). With $k = 2$: the accumulated $S_2 = \{\sigma_{\text{Cont}}, \sigma_{\text{Exit}}\}$ has mean $\sigma_{1/2}$, and $P_3(1/2) = 0$, so $\text{BR}(\sigma_{1/2}) = \{\text{Cont}, \text{Exit}\}$. Since $\{\text{Cont}, \text{Exit}\} \not\subseteq T_1, T_2$, Step 4(b) fires: $T_3 = \{\text{Cont}, \text{Exit}\}$, $S_3 = \Delta(\{\text{Cont}, \text{Exit}\})$. With $k = 3$: the mean of S_3 is still $\sigma_{1/2}$, so $\text{BR}(\sigma_{1/2}) = \{\text{Cont}, \text{Exit}\} \subseteq T_3$ and Step 4(a) fires. Terminal play $\bar{\sigma}_3 = \frac{1}{3}(\sigma_{\text{Cont}} + \sigma_{\text{Exit}} + \sigma_{1/2}) = \sigma_{1/2}$, differing from Procedure 1's σ_{Exit}.

C The Symmetric Coordination Game

This appendix describes the simplest game in which the support digraph Γ has a *robust* multi-element fixed point, one that survives generic (symmetry-preserving) perturbation, in contrast to the knife-edge multi-element fixed points of the AMD examples above (Figures 3a and 3b), which vanish under generic perturbation by Proposition 3.2. It is the symmetric 2-action coordination game whose digraphs appear as the rightmost panels of Figures 2 and 3.

The game. A single-information-set SLU game with two decision nodes X_1, X_2 in one information set and action set $A(h) = \{a_1, a_2\}$. The agent receives payoff 1 if her play coincides at both nodes and 0 otherwise:

$$U(q, \sigma) \;=\; q(a_1)\,\sigma(a_1) \;+\; q(a_2)\,\sigma(a_2) \;=\; \sum_a q(a)\,V_a(\sigma),$$

with $V_{a_i}(\sigma) = \sigma(a_i)$. Note that $V_a(\sigma)$ is *linear* in σ, in contrast to AMD-style games, where the location prior makes $V_a(\sigma_p)$ a rational function of p.

Interim fixed points. The agent is indifferent over a_1, a_2 iff $\sigma(a_1) = \sigma(a_2) = 1/2$; otherwise her unique best response is the action with larger σ-weight. So

$$\mathrm{MPP} \;=\; \{\sigma_{a_1},\ \sigma_{a_2},\ \sigma_{1/2}\}.$$

Best-response and support digraphs. The best response to σ_{a_i} is a_i itself, so G consists of two self-loops with no cross-edges (Figure 2c). For the support digraph, the transition at the multi-element vertex is the best response to the uniform mixture $\sigma_{1/2}$ over $\{a_1, a_2\}$. There $V_{a_1}(\sigma_{1/2}) = V_{a_2}(\sigma_{1/2}) = \frac{1}{2}$, so the two actions are tied and $g(\{a_1, a_2\}) = \{a_1, a_2\}$: the multi-element vertex is a Γ-fixed point (Figure 3c). Since V_a is linear, the symmetric payoffs are already tied at the uniform mixture, and perturbations of the payoffs that preserve the symmetry under the transposition $(a_1\, a_2)$ keep this fixed point in place.

Procedure 1. Starting from $S_0 = \mathrm{MPP}$, the PIR average $R^{\mathrm{PIR}}(\mathrm{MPP}) = \{a_1, a_2\}$ (by symmetry, each summand contributes equally), and $\sigma_{1/2}$ is the unique element of MPP with support $\{a_1, a_2\}$. Step 3(a) therefore fires immediately: the agent plays $\sigma_{1/2}$.

Procedure 2. The two procedures coincide on this game. The shared steps 0-3 terminate before any accumulation occurs: exactly as for Procedure 1, $R^{\mathrm{PIR}}(\mathrm{MPP}) = \{a_1, a_2\}$ matches the unique element $\sigma_{1/2}$ of MPP carrying that support, so Step 3(a) fires at step 0 and the agent plays $\sigma_{1/2}$. This is the opposite of the AMD examples, where the response $\{\mathrm{Cont}, \mathrm{Exit}\}$ is shared by several elements of MPP, Step 3(a) fails on the non-unique match, and the two procedures diverge once accumulation begins.

Robustness of the interior fixed point: The selection of the interior strategy $\sigma_{1/2}$ is special to the symmetric configuration: if we break the symmetry by an arbitrary small perturbation of the payoffs, the interior interim fixed point survives but the multi-element Γ-fixed point does not. The transition $g(\{a_1, a_2\}) = \mathrm{BR}(\sigma_{1/2})$ evaluates the best response at the uniform mixture, not at the displaced indifference point σ_{x^*}; once $x^* \neq 1/2$ one action strictly wins at $\sigma_{1/2}$, so $g(\{a_1, a_2\})$ collapses to a singleton and the multi-element vertex is no longer a fixed point of Γ. Procedure 1 then converges to one of the two pure fixed points, and the interior interim fixed point, though it still exists, is no longer selected. This is the same reference-point phenomenon as in the generic AMD: Γ samples best responses only at the uniform mixtures σ_T, so an interior interim fixed point that does not lie at such a mixture is invisible to the procedure. Symmetry matters here precisely because it is what forces the interior fixed point onto the uniform mixture $\sigma_{1/2}$ that Γ tests; absent that alignment, the procedure tracks the reference point rather than the interior fixed point.

D Omitted Proofs

Proof of Proposition 3.2. Since the AMD action set is $A(h) = \{\mathrm{Cont}, \mathrm{Exit}\}$, the support family is $\mathcal{T} = \{\{\mathrm{Cont}\}, \{\mathrm{Exit}\}, T^*\}$ with $T^* := \{\mathrm{Cont}, \mathrm{Exit}\}$, and T^* is the unique multi-element vertex of Γ. Identify $\Delta(T^*)$ with $[0, 1]$ via $p := \sigma(\mathrm{Cont})$; the uniform mixture over T^* is $\sigma_{1/2}$, corresponding to $p = 1/2$.

I prove the statement in four steps. Step 1 simply reproduces useful notation from Appendix A, which will give us an indifference condition (in a convenient form). Step 2 defines the candidate open dense subset (U) on which we will prove the proposition statement holds. Step 3 then proves the first part of it, that there are no fixed points of cardinality ≥ 2. Step 4 completes the proof by showing the second part: there are no multi-element cycles.

Step 1: A convenient expression of the utility difference. Fix payoffs $\alpha = (\alpha_1, \ldots, \alpha_{n+1}) \in \mathbb{R}^{n+1}$ and set $\beta_j := \alpha_{j+1} - \alpha_j$ for $j = 1, \ldots, n$. By the computation recorded in equation (A.1) of Appendix A,

$$V_{\text{Cont}}(\sigma_p) - V_{\text{Exit}}(\sigma_p) \;=\; \frac{P_n(p)}{D_n(p)} \qquad \text{for all } p \in [0,1], \tag{D.1}$$

where $D_n(p) := 1 + p + \cdots + p^{n-1}$ is strictly positive on $[0,1]$ and $P_n(p) := \sum_{k=0}^{n-1}(k+1)\beta_{k+1}p^k$ has coefficients that are linear functions of α.

Step 2: Defining a candidate open dense subset. Define

$$I_{1/2}(\alpha) \;:=\; P_n(1/2), \qquad I_{\text{Cont}}(\alpha) \;:=\; P_n(1) \;=\; \sum_{k=1}^{n} k\,\beta_k, \qquad I_{\text{Exit}}(\alpha) \;:=\; P_n(0) \;=\; \beta_1.$$

Each is a linear functional on $\mathbb{R}^{n+1}$: a fixed linear combination of $\beta_1, \ldots, \beta_n$, and the β_j are themselves linear in α. By the rank-nullity theorem, the zero set of a functional on $\mathbb{R}^{n+1}$ is a proper linear hyperplane exactly when the functional is linear and not identically zero. Linearity having been noted, it suffices to observe that none of the three vanishes at every payoff vector, which is immediate:

$$I_{\text{Exit}}(\alpha) = \alpha_2 - \alpha_1, \qquad I_{\text{Cont}}(\alpha) = n\alpha_{n+1} - \sum_{k=1}^{n}\alpha_k, \qquad I_{1/2}(\alpha) = \sum_{k=0}^{n-1}(k+1)\,\beta_{k+1}/2^k.$$

The zero set $H_\cdot := I_\cdot^{-1}(0)$ of each is therefore a proper linear hyperplane in $\mathbb{R}^{n+1}$: closed, nowhere dense, and of Lebesgue measure zero. The set

$$U \;:=\; \mathbb{R}^{n+1} \setminus (H_{1/2} \cup H_{\text{Cont}} \cup H_{\text{Exit}})$$

is the complement of a finite union of proper hyperplanes in a Euclidean space, hence open and dense. If we prove that there are no fixed points of cardinality ≥ 2 and no multi-element cycles for payoffs in U, the proposition is proved. In Step 3 I prove the first part, and in Step 4 the second.

Step 3: For $\alpha \in U$, T^* is not a fixed point of g. Since $\alpha \notin H_{1/2}$, $P_n(1/2) \neq 0$, so by (D.1) evaluated at $p = 1/2$,

$$V_{\text{Cont}}(\sigma_{1/2}) - V_{\text{Exit}}(\sigma_{1/2}) \;=\; \frac{P_n(1/2)}{D_n(1/2)} \;\neq\; 0.$$

One of Cont, Exit strictly dominates the other at the uniform mixture $\sigma_{1/2}$. Hence $g(T^*) = \text{BR}(\sigma_{T^*}) = \text{BR}(\sigma_{1/2})$ is a singleton in $\mathcal{T}$, and in particular $g(T^*) \neq T^*$.

Step 4: For $\alpha \in U$, no cycle of Γ of length ≥ 2 contains T^*. In a functional digraph each vertex has out-degree exactly one, so every cycle is simple. Suppose for contradiction that C is a simple cycle of length ≥ 2 containing T^*. Since T^* is the unique multi-element vertex of Γ and C visits each of its vertices once, C has the form

$$T^* \;\to\; \{a_1\} \;\to\; \{a_2\} \;\to\; \cdots \;\to\; \{a_{k-1}\} \;\to\; T^*$$

for some $k \geq 2$ with each $a_i \in \{\text{Cont}, \text{Exit}\}$. The final edge requires $g(\{a_{k-1}\}) = T^*$, that is, $R(\{\sigma_{a_{k-1}}\}) = \{\text{Cont}, \text{Exit}\}$. Since $R(\{\sigma_a\}) = \arg\max_{b \in A(h)} V_b(\sigma_a)$ is set-valued only when there is an exact tie, this forces $V_{\text{Cont}}(\sigma_{a_{k-1}}) = V_{\text{Exit}}(\sigma_{a_{k-1}})$. By (D.1), evaluated at $p = 1$ if $a_{k-1} = \text{Cont}$ and at $p = 0$ if

$a_{k-1} = \text{Exit}$, and using $D_n(1) = n > 0$ and $D_n(0) = 1 > 0$, this tie is equivalent to $I_{a_{k-1}}(\alpha) = 0$. But $\alpha \notin H_{\text{Cont}} \cup H_{\text{Exit}}$ excludes both possibilities $a_{k-1} \in \{\text{Cont}, \text{Exit}\}$, contradicting the supposition.

Conclusion. For every $\alpha \in U$ the unique multi-element vertex T^* is neither a fixed point of g (Step 3) nor on any cycle of length ≥ 2 in Γ (Step 4). Hence every cycle of Γ avoids T^* and so lies entirely in the singleton-support sub-digraph G. The set U is open and dense in $\mathbb{R}^{n+1}$, as claimed. □

References

Arntzenius, F. (2008). No regrets, or: Edith piaf revamps decision theory. *Erkenntnis*, 68(2):277–297.

Aumann, R. J., Hart, S., and Perry, M. (1997). The absent-minded driver. *Games and Economic Behavior*, 20(1):102–116.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898.

Gibbard, A. and Harper, W. L. (1978). Counterfactuals and two kinds of expected utility. pages 125–162.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Harper, W. L. (1986). Mixed strategies and ratifiability in causal decision theory. *Erkenntnis*, 24(1):25–36.

Joyce, J. M. (2012). Regret and instability in causal decision theory. In *Synthese*, volume 187, pages 123–145.

Piccione, M. and Rubinstein, A. (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3–24.

Skyrms, B. (1990). *The Dynamics of Rational Deliberation*. Harvard University Press, Cambridge, MA.

Stahl, D. O. and Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254.

Winkler, P. (2017). The sleeping beauty controversy. *The American Mathematical Monthly*, 124(7):579–587.