DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

A Note on Cursed Sequential Equilibrium and Sequential Cursed Equilibrium

Meng-Jhang Fong
California Institute of Technology

Po-Hsuan Lin
California Institute of Technology

Thomas R. Palfrey
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 1467
April 2023

A Note on Cursed Sequential Equilibrium and Sequential Cursed Equilibrium*

Meng-Jhang Fong† Po-Hsuan Lin‡ Thomas R. Palfrey§

April 11, 2023

Abstract

In this short note, we compare the cursed sequential equilibrium (CSE) by Fong et al. (2023) and the sequential cursed equilibrium (SCE) by Cohen and Li (2023). We identify eight main differences between CSE and SCE with respect to the following features:

(1) the family of applicable games,

(2) the number of free parameters,

(3) the belief updating process,

(4) the treatment of public histories,

(5) effects in games of complete information,

(6) violations of subgame perfection and sequential rationality,

(7) re-labeling of actions, and

(8) effects in one-stage simultaneous-move games.

JEL Classification Numbers: C72, D83

Keywords: Multi-stage Games, Private Information, Cursed Equilibrium

*We are grateful to Shengwu Li and Shani Cohen for multiple email correspondences in December 2022 that helped to clarify some of the differences between the CSE and SCE approaches to the generalization of cursed equilibrium for dynamic games. An initial version of the SCE working paper was posted on December 13, 2022, and this note is written based on the third version, Cohen and Li (2023), dated February 27, 2023.

†Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125 USA. mjfong@caltech.edu

‡Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125 USA. plin@caltech.edu

§Corresponding Author: Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125 USA. trp@hss.caltech.edu. Fax: +16263958967 Phone: +16263954088

1 Introduction

In this note, we compare two recently proposed theories: the cursed sequential equilibrium (CSE) of Fong et al. (2023); and the sequential cursed equilibrium (SCE) of Cohen and Li (2023). Both approaches generalize the cursed equilibrium (CE) by Eyster and Rabin (2005) to dynamic games, yet in different ways.

Following the specification of CE, CSE extends the concept of cursedness—the failure to fully account for the correlation between the types and actions of other players—to multi-stage games with publicly observed actions. After each stage of the game, each player in the game updates their current beliefs about the profile of other players' types in a cursed way, i.e., based on the *average* behavioral strategies of the other players rather than the type conditional behavioral strategies of the other players. As in CE, CSE is generalized to χ-CSE, where the parameter $\chi \in [0, 1]$ indicates the degree of cursedness, with $\chi = 0$ corresponding to standard fully rational equilibrium behavior (i.e., updating based on the type conditional behavior strategies of the other players), and the case of $\chi = 1$ is *fully cursed* (i.e., updating based on the *average* behavioral strategies of the other players). Values of $\chi \in (0, 1)$ correspond to a mixture of these two extremes, defined similarly to CE, but with respect to behavioral strategies rather than mixed strategies.

The SCE extension of CE to dynamic games differs from χ-CSE in a number of ways, and we will illustrate through a series of examples some of these differences. Two differences are immediate and are due to a major difference in basic approach. First, while χ-CSE is developed for a subclass of games in extensive form with perfect recall, SCE is developed for a larger class: extensive form games with perfect recall. In such games, the "stages" and "public histories" do not generally exist or are not well-defined. As a result, instead of applying the cursed updating stage-by-stage, a player's cursedness is defined in terms of coarsening the partition of other players' information sets. In this sense, the difference between CSE and SCE reminds one of the differences between the CE and Analogy Based Expectations (ABEE) approach by Jehiel (2005) and Jehiel and Koessler (2008), where the latter approach is based on the bundling of nodes at which other players move into analogy classes, which formally is a coarsening of information sets. Second, SCE is generalized by introducing *two* free parameters, $(\chi_S, \psi_S) \in [0, 1]^2$.[1] With two parameters, (χ_S, ψ_S)-SCE can distinguish inferences (and degree of neglect) based on past observed actions by other players and the inferences (and degree of neglect) based on simultaneous or future (hypothetical) strategies of the other players.

In addition to the family of applicable games and the dimension of parameter space, this

[1]To avoid confusion, we will henceforth add the subscript "S" to the parameters of SCE.

note identifies and illustrates six additional differences between CSE and SCE. To formally illustrate these technical differences, in this note, we will focus on the class of multi-stage games with observed actions, using the framework of Fudenberg and Tirole (1991). We next introduce the framework and the two solution concepts in Section 2. After that, we will discuss and illustrate the rest of the six differences between CSE and SCE in Section 3 which are organized as follows:

(1) the belief updating process (Section 3.1),

(2) the treatment of public histories (Section 3.2),

(3) effects in games of complete information (Section 3.3),

(4) violations of subgame perfection and sequential rationality (Section 3.4),

(5) the effect of re-labeling actions (Section 3.5), and

(6) effects in one-stage simultaneous-move games (Section 3.6).

2 Preliminary

2.1 Multi-Stage Games with Observed Actions

We first briefly summarize the structure of multi-stage games with observed actions. Let N denote the set of n players. Each player $i \in N$ has a *type* θ_i drawn from a finite set Θ_i. Let $\theta \in \Theta \equiv \times_{i=1}^{n}\Theta_i$ be the type profile and $\theta_{-i} \in \Theta_{-i} \equiv \times_{j\neq i}\Theta_j$ be the type profile without player i. All players share a common (full support) prior distribution $\mathcal{F}(\cdot) : \Theta \to (0,1)$. Therefore, for every player i, the belief of other players' types conditional on his own type is

$$\mathcal{F}(\theta_{-i}|\theta_i) = \frac{\mathcal{F}(\theta_{-i},\theta_i)}{\sum_{\theta'_{-i}\in\Theta_{-i}}\mathcal{F}(\theta'_{-i},\theta_i)}.$$

At the beginning, players observe their own types, but not the other players' types.

The game is played in stages $t = 1, 2, \ldots, T$. In each stage, players simultaneously choose actions, which will be revealed at the end of the stage. Let $\mathcal{H}^{t-1}$ be the set of all possible *public* histories[2] at stage t, where $\mathcal{H}^0 = \{h_\emptyset\}$ and $\mathcal{H}^T$ is the set of terminal histories. Let

[2]Technically, each public history is not a *history* in general extensive form games as a public history only specifies the past action profiles but not the realized type profile. In the following, we will use the term *public history* when referring to the past action profiles, and use the term *history* when referring to the past action profiles with a type profile.

$\mathcal{H} = \cup_{t=0}^{T}\mathcal{H}^t$ be the set of all possible public histories of the game, and $\mathcal{H}\backslash\mathcal{H}^T$ be the set of non-terminal public histories.

For every player i, the available information at stage t is in $\Theta_i \times \mathcal{H}^{t-1}$. Therefore, player i's information sets can be specified as $\mathcal{I}_i \in \mathcal{Q}_i = \{(\theta, h) : h \in \mathcal{H}\backslash\mathcal{H}^T, \theta_i \in \Theta_i\}$. That is, a type θ_i player i's information set at the public history h^t can be defined as $\bigcup_{\theta_{-i}\in\Theta_{-i}}(\theta_i, \theta_{-i}, h^t)$. With a slight abuse of notation, it will be simply denoted as (θ_i, h^t). At each public history, we assume that the feasible set of actions for every player is independent of their type and use $A_i(h^{t-1})$ to denote the feasible set of actions for player i at the public history h^{t-1}. Let $A_i = \times_{h\in\mathcal{H}\backslash\mathcal{H}^T} A_i(h)$ denote player i's feasible actions in all public histories of the game and $A = A_1 \times \cdots \times A_n$. In addition, we assume A_i is finite for all $i \in N$ and $|A_i(h)| \geq 1$ for all $i \in N$ and any $h \in \mathcal{H}\backslash\mathcal{H}^T$.

A behavioral strategy for player i is a function $\sigma_i : \mathcal{Q}_i \to \Delta(A_i)$ satisfying $\sigma_i(\theta_i, h^{t-1}) \in \Delta(A_i(h^{t-1}))$. Furthermore, we use $\sigma_i(a_i^t|\theta_i, h^{t-1})$ to denote the probability player i chooses $a_i^t \in A_i(h^{t-1})$. We use $a^t = (a_1^t, \ldots, a_n^t) \in \times_{i=1}^{n} A_i(h^{t-1}) \equiv A(h^{t-1})$ to denote the action profile at stage t and a_{-i}^t to denote the action profile at stage t without player i. If a^t is the action profile realized at stage t, then $h^t = (h^{t-1}, a^t)$. For the sake of simplicity, we define a partial order $\prec$ on $\mathcal{H}$ such that $h \prec h'$ represents that h is an earlier public history than h'. Finally, each player i has a payoff function $u_i : \Theta \times \mathcal{H}^T \to \mathbb{R}$, and we let $u = (u_1, \ldots, u_n)$ be the profile of payoff functions. A multi-stage game with observed actions, Γ, is defined by the tuple $\Gamma = \langle T, A, N, \mathcal{H}, \Theta, \mathcal{F}, u\rangle$.

2.2 CSE in Multi-Stage Games

Consider an assessment (μ, σ), where μ is a belief system and σ is a behavioral strategy profile. The belief system specifies, for each player, a conditional distribution over the set of type profiles conditional on each *public history*. Following the spirit of the cursed equilibrium, for player i at the public history h^{t-1}, we define the *average behavioral strategy profile of the other players* as:

$$\bar{\sigma}_{-i}(a_{-i}^t|\theta_i, h^{t-1}) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a_{-i}^t|\theta_{-i}, h^{t-1})$$

for any $i \in N$, $\theta_i \in \Theta_i$ and $h^{t-1} \in \mathcal{H}^{t-1}$.

A χ-CSE is parameterized by a single parameter $\chi \in [0, 1]$. Instead of thinking the other players are using σ_{-i}, a χ-cursed type θ_i player i would believe they are using a χ-weighted average of the average behavioral strategy and the true behavioral strategy:

$$\sigma^{\chi}_{-i}(a^t_{-i}|\theta_{-i}, \theta_i, h^{t-1}) = \chi\bar{\sigma}_{-i}(a^t_{-i}|\theta_i, h^{t-1}) + (1-\chi)\sigma_{-i}(a^t_{-i}|\theta_{-i}, h^{t-1}).$$

The beliefs of player i about the type profile θ_{-i} are updated via Bayes' rule in χ-CSE, whenever possible, assuming others are using the χ-cursed behavioral strategy rather than the true behavioral strategy. Specifically, an assessment satisfies the *χ-cursed Bayes' rule* if the belief system is derived from Bayes' rule while perceiving others are using σ^{χ}_{-i} rather than σ_{-i}:

Definition 1 (χ-cursed Bayes' rule, Definition 1 of Fong et al., 2023)**.** *An assessment (μ, σ) satisfies χ-cursed Bayes' rule if the following rule is applied to update the posterior beliefs whenever $\sum_{\theta'_{-i}\in\Theta_{-i}} \mu_i(\theta'_{-i}|\theta_i, h^{t-1})\sigma^{\chi}_{-i}(a^t_{-i}|\theta'_{-i}, \theta_i, h^{t-1}) > 0$:*

$$\mu_i(\theta_{-i}|\theta_i, h^t) = \frac{\mu_i(\theta_{-i}|\theta_i, h^{t-1})\sigma^{\chi}_{-i}(a^t_{-i}|\theta_{-i}, \theta_i, h^{t-1})}{\sum_{\theta'_{-i}\in\Theta_{-i}} \mu_i(\theta'_{-i}|\theta_i, h^{t-1})\sigma^{\chi}_{-i}(a^t_{-i}|\theta'_{-i}, \theta_i, h^{t-1})}.$$

Finally, χ-CSE places a consistency restriction, analogous to consistent assessments in sequential equilibrium, on how χ-cursed beliefs are updated off the equilibrium path:

Definition 2 (χ-consistency, Definition 2 of Fong et al., 2023)**.** *Let Ψ^{χ} be the set of assessments (μ, σ) such that σ is a totally mixed behavioral strategy profile and μ is derived from σ using χ-cursed Bayes' rule. (μ, σ) satisfies χ-consistency if there is a sequence of assessments $\{(\mu_k, \sigma_k)\} \subseteq \Psi^{\chi}$ such that $\lim_{k\to\infty}(\mu_k, \sigma_k) = (\mu, \sigma)$.*

For any $i \in N$, $\chi \in [0, 1]$, σ, and $\theta \in \Theta$, let $\rho^{\chi}_i(h^T|h^t, \theta, \sigma^{\chi}_{-i}, \sigma_i)$ be player i's perceived conditional realization probability of terminal history $h^T \in \mathcal{H}^T$ at the public history $h^t \in \mathcal{H}\backslash\mathcal{H}^T$ if the type profile is θ and player i uses the behavioral strategy σ_i whereas perceives other players' using the cursed behavioral strategy σ^{χ}_{-i}. At every non-terminal public history h^t, a χ-cursed player in χ-CSE will use χ-cursed Bayes' rule (Definition 1) to derive the posterior belief about the other players' types. Accordingly, a type θ_i player i's conditional expected payoff at the public history h^t is given by:

$$\mathbb{E}u_i(\sigma|\theta_i, h^t) = \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{h^T\in\mathcal{H}^T} \mu_i(\theta_{-i}|\theta_i, h^t)\rho^{\chi}_i(h^T|h^t, \theta, \sigma^{\chi}_{-i}, \sigma_i)u_i(\theta_i, \theta_{-i}, h^T).$$

Definition 3 (χ-cursed sequential equilibrium, Definition 3 of Fong et al., 2023)**.** *An assessment (μ^*, σ^*) is a χ-cursed sequential equilibrium if it satisfies χ-consistency and $\sigma^*_i(\theta_i, h^t)$ maximizes $\mathbb{E}u_i(\sigma^*|\theta_i, h^t)$ for all i, θ_i, $h^t \in \mathcal{H}\backslash\mathcal{H}^T$.*

2.3 SCE in Multi-Stage Games

Some additional notation is required to define sequential cursed equilibrium (SCE) in a multi-stage game. First, let λ denote nature. In the framework of multi-stage games with observed actions, λ only moves once in stage 0, i.e., at the initial history, denoted $h_\emptyset$. The action set for λ is the set of states of nature, i.e., the set of profiles of player types, $\Theta = \times_{i=1}^{n}\Theta_i$. The totally mixed strategy of nature, denoted by $\sigma_\lambda(h_\emptyset) \in \Delta(\Theta)$, is common knowledge to all players $i \in N$ and exogenously given by $\mathcal{F}$. Note that nature's information set is singleton. Also, note that after nature's action (i.e., type profile) is realized, θ_i will be observed only by player i but not by the other players. That is, each player's type is his own private information.

Second, let $\mu[\sigma]$ denote the probability measure on $\Theta \times \mathcal{H}^T$ that is induced by $\sigma \equiv (\sigma_1, ..., \sigma_n, \sigma_\lambda)$. Following Cohen and Li (2023), we allow $\mu[\sigma]$ to denote the probability measure on $\Theta \times \mathcal{H}^t$ for some $t < T$ by viewing $(\theta', (a^1, ..., a^t)) \in \Theta \times \mathcal{H}^t$ as

$$\{(\theta, h^T) \in \Theta \times \mathcal{H}^T : \theta = \theta', (a^1, ..., a^t) \prec h^T\}.$$

Third, let $\mathcal{P}$ denote the *coarsest (valid) partition* of the non-terminal histories $\Theta \times \mathcal{H} \backslash \mathcal{H}^T$,[3] and $P(\mathcal{I}_i)$ denotes $P \in \mathcal{P}$ such that $\mathcal{I}_i \subseteq P$. Note that a partition $\mathcal{P}'$ is coarser than $\mathcal{P}''$ if, given any $P'' \in \mathcal{P}''$, there exists a cell of $\mathcal{P}'$ that contains P''.

Finally, given $\mathcal{I}_j = (\theta_j, h^t)$, let $\sigma_i^{\mathcal{I}_j} : \mathcal{Q}_i^{\mathcal{I}_j} \to \Delta(A_i)$ denote a *partial strategy* for player i at $\mathcal{I}_j$, where $\mathcal{Q}_i^{\mathcal{I}_j} = \{(\theta', h') : h' \in \mathcal{H} \backslash \mathcal{H}^T, \theta'_i \in \Theta_i,\ h^t \preceq h' \text{ or } h' \preceq h^t\}$. In other words, $\sigma_i^{\mathcal{I}_j}$ is a function from information sets containing public histories compatible with h^t to distributions over player i's feasible actions.

To define a SCE, Cohen and Li (2023) in their Section 2.4 define three types of conjectures about opponents' and nature's strategies: the *sequentially cursed conjecture* ($\tilde{\sigma}$), the *typically cursed conjecture* ($\check{\sigma}$), and the *Bayesian conjecture* ($\hat{\sigma}$).

Consider a non-terminal information set $\mathcal{I}_i = (\theta_i, h^t)$ where $h^t = (a^1, ..., a^t)$. Given a totally mixed behavioral strategy σ, Cohen and Li (2023) define the three types of conjectures of player i about player $j \neq i$ at information set $\mathcal{I}_i$ as follows:[4]

1. Sequentially cursed conjecture: $\tilde{\sigma}_j^{\mathcal{I}_i}(\cdot)$ is said to be player i's *sequentially cursed con-*

[3] A partition of the non-terminal histories is invalid if it defines information sets that violate the assumption of perfect recall.

[4] For general behavioral strategy profiles, SCE imposes a consistency requirement similar to Kreps and Wilson (1982) as it requires the limits of all three types of conjectures exist.

jecture about j's behavioral strategy at $\mathcal{I}_i$ if, for all $\mathcal{I}_j = (\theta_j, h^{t'})$ and all $a_j \in A_j(h^{t'})$,

$$\tilde{\sigma}_j^{\mathcal{I}_i}(a_j|\, \mathcal{I}_j) \equiv \mu[\sigma](\{((\theta,h),a_j) : (\theta,h) \in P(\mathcal{I}_j)\}|\, \mathcal{I}_i \cap P(\mathcal{I}_j)).$$

2. Typically cursed conjecture: $\check{\sigma}_j^{\mathcal{I}_i}(\cdot)$ is said to be player i's *typically cursed conjecture* about j's behavioral strategy at $\mathcal{I}_i$ if, for all $\mathcal{I}_j = (\theta_j, h^{t'})$ and all $a_j \in A(h^{t'})$,

$$\check{\sigma}_j^{\mathcal{I}_i}(a_j|\, \mathcal{I}_j) \equiv \mu[\sigma](\{((\theta,h),a_j) : (\theta,h) \in P(\mathcal{I}_j)\}|\, (\theta_i, \cdot) \cap P(\mathcal{I}_j))$$

where $(\theta_i, \cdot) \equiv \{(\theta_i, \theta_{-i}, h) : h \in \mathcal{H} \backslash \mathcal{H}^T, \theta_{-i} \in \Theta_{-i}\}$.

3. Bayesian conjecture: $\hat{\sigma}_j^{\mathcal{I}_i}(\cdot)$ is said to be player i's *Bayesian conjecture* about j's behavioral strategy at $\mathcal{I}_i$ if, for all $\mathcal{I}_j = (\theta_j, h^{t'})$ and all $a_j \in A(h^{t'})$,

$$\hat{\sigma}_j^{\mathcal{I}_i}(a_j|\, \mathcal{I}_j) \equiv \mu[\sigma](\{((\theta,h),a_j) : (\theta,h) \in \mathcal{I}_j\}|\, \mathcal{I}_i \cap \mathcal{I}_j).$$

A (χ_S, ψ_S)-SCE has two parameters, $(\chi_S, \psi_S) \in [0,1]^2$. At every information set $\mathcal{I}_i$, a (χ_S, ψ_S)-sequentially cursed player i would perceive that the game has been and will be played by the other players according to his sequentially cursed, typically cursed, and Bayesian conjectures with probability $\chi_S \psi_S$, $\chi_S(1-\psi_S)$, and $1-\chi_S$, respectively. Based on such perceptions, a (χ_S, ψ_S)-sequential cursed equilibrium is defined as follows:

Definition 4 ((χ_S, ψ_S)-sequential cursed equilibrium, Section 2.4 of Cohen and Li, 2023)**.** *A strategy profile σ^* is said to be a (χ_S, ψ_S)-sequential cursed equilibrium if, given any $i \in N$ and $\mathcal{I}_i = (\theta_i, h^t)$ with $h^t = (a^1, ..., a^t)$, there exists a partial strategy $\sigma_i^{\mathcal{I}_i}$ such that $\sigma_i^{\mathcal{I}_i}(\mathcal{I}_i) = \sigma_i^*(\mathcal{I}_i)$ and $\sigma_i^{\mathcal{I}_i}$ maximizes i's expected utility under the belief that θ_{-i} is realized with probability $\mu[\bar{\sigma}^{\mathcal{I}_i}](\theta_{-i}, \theta_i, h^t|\, \mathcal{I}_i)$ and that the game will proceed according to $\bar{\sigma}_{-i}^{\mathcal{I}_i}$, where*

1. $\bar{\sigma}_i^{\mathcal{I}_i}(a^{t'}|\theta_i, h^{t'-1}) = 1$ *for all* $t' \leq t$
2. $\bar{\sigma}_{-i}^{\mathcal{I}_i} = \tilde{\sigma}_{-i}^{\mathcal{I}_i}$ *with probability* $\chi_S \psi_S$
3. $\bar{\sigma}_{-i}^{\mathcal{I}_i} = \check{\sigma}_{-i}^{\mathcal{I}_i}$ *with probability* $\chi_S(1-\psi_S)$
4. $\bar{\sigma}_{-i}^{\mathcal{I}_i} = \hat{\sigma}_{-i}^{\mathcal{I}_i}$ *with probability* $(1-\chi_S)$

When $\chi_S = 0$, a (χ_S, ψ_S)-SCE reduces to a sequential equilibrium. A sequential cursed equilibrium refers to a boundary case when $\chi_S = \psi_S = 1$ in a (χ_S, ψ_S)-SCE.

3 Differences Between CSE and SCE

3.1 Difference in the Belief Updating Dynamics

One key difference between the two equilibrium concepts is in the belief updating process about nature's strategy. In the framework of CSE, players update their beliefs about the other players' type profile *at each public history* by χ-cursed Bayes' rule, which is stated in Claim 1. In the case of totally mixed strategies, the updating process is very simple to characterize:

Claim 1 (Lemma 1 in Fong et al., 2023)**.** *For any χ-cursed sequential equilibrium (μ, σ) where σ is a totally mixed behavioral strategy profile, any $i \in N$, any non-terminal history $h^t = (h^{t-1}, a^t) \in \mathcal{H} \backslash \mathcal{H}^T$ and any $\theta \in \Theta$,*

$$\mu_i(\theta_{-i}|\theta_i, h^t) = \chi\mu_i(\theta_{-i}|\theta_i, h^{t-1}) + (1-\chi)\left[\frac{\mu_i(\theta_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a^t_{-i}|\theta_{-i}, h^{t-1})}{\sum_{\theta'_{-i}} \mu_i(\theta'_{-i}|\theta_i, h^{t-1})\sigma_{-i}(a^t_{-i}|\theta'_{-i}, h^{t-1})}\right]. \quad (1)$$

Proof. See the proof of Lemma 1 in Fong et al. (2023). □

On the other hand, SCE is defined as a mixture of three different kinds of conjectures. The mathematical object that corresponds to the belief of SCE is the *conjecture about nature's strategy*. In the following, we will characterize the evolution of this conjecture. In general, the characterization of the belief updating process about nature's strategy can be quite complicated. To simplify the comparison between CSE and SCE, it is instructive to first examine the comparison in games where the coarsest valid partition, $\mathcal{P}$, accords with public histories, in the sense that $\mathcal{P}$ is measurable with respect to public histories. Later examples in this note will explore some of the implications if this condition is not satisfied.

Definition 5 (Public History Consistency (PHC))**.** *The coarsest valid partition, $\mathcal{P}$, satisfies* Public History Consistency *if, for all $t < T, i \in N, \theta_i \in \Theta_i, h^t \in \mathcal{H}^t, \hat{h}^t \in \mathcal{H}^t$:*

$$h^t \neq \hat{h}^t \Rightarrow P(\theta_i, h^t) \neq P(\theta_i, \hat{h}^t).$$

Following Definition 5, we can explicitly characterize $P(\mathcal{I}_i)$ under PHC, as demonstrated in Lemma 1:

Lemma 1. *The coarsest valid partition, $\mathcal{P}$, satisfies* Public History Consistency *if and only if, for all $t < T$, $i \in N$, $\theta_i \in \Theta_i$, and $h^t \in \mathcal{H}^t$, we have $P(\theta_i, h^t) = \bigcup_{\theta \in \Theta}(\theta, h^t)$.*

Proof.

($\Longrightarrow$):

We prove by contrapositive. Suppose that there exists some non-terminal information set (θ_i, h^t) such that $P(\theta_i, h^t) \neq \bigcup_{\theta\in\Theta}(\theta, h^t)$, then by the definition of the coarsest partition, we have $P(\theta_i, h^t) \supsetneq \bigcup_{\theta\in\Theta}(\theta, h^t)$. However, in games of perfect recall, this implies that there is some $P \in \mathcal{P}$ and $\hat{h}^t \in \mathcal{H}^t$ such that $h^t \neq \hat{h}^t$ but $(\theta_i, h^t) \subset P$ and $(\theta_i, \hat{h}^t) \subset P$, which contradicts Definition 5.

($\Longleftarrow$):

Definition 5 follows immediately from Lemma 1. □

A key implication of PHC is that no matter how cursed the players are, they know that everyone's behavioral strategy depends on the realized public history. In other words, this condition shuts down one source of cursedness: players fully understand that other players' future actions are conditional on current and past actions. Other implications of the PHC condition will be discussed in later sections. Of course, it is still the case that at any history the players are cursed in the sense of neglecting the dependence of other players' current and future actions on the type profile (i.e., nature's initial move).[5]

Fix any totally mixed behavioral strategy profile σ and any information set $\mathcal{I}_i = (\theta_i, h^t)$ where $h^t = (h^{t-1}, a^t)$. Under SCE, player i's belief about other players' type profile (Nature's initial move) is the weighted average of the conjectures about nature's strategy. Specifically, for any $(\chi_S, \psi_S) \in [0,1]^2$, the belief can be derived by

$$\mu_i^{(\chi_S,\psi_S)}(\theta_{-i}|\theta_i, h^t) \equiv \chi_S\psi_S \underbrace{\tilde{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset)}_{\textit{sequentially cursed}} + \chi_S(1-\psi_S)\underbrace{\check{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset)}_{\textit{typically cursed}} + (1-\chi_S)\underbrace{\hat{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset)}_{\textit{Bayesian}}.$$

If PHC is satisfied, one can relatively easily characterize the evolution of the (χ_S, ψ_S)-SCE conjecture about nature. This is done next in Claim 2.

Claim 2. *Under PHC, for any (χ_S, ψ_S)-sequential cursed equilibrium where σ is a totally mixed behavioral strategy profile, and any non-terminal information set (θ_i, h^t) where $h^t = (a^1, \ldots, a^t)$, the conjecture about nature's strategy is:*

$$\mu_i^{(\chi_S,\psi_S)}(\theta_{-i}|\theta_i, h^t) = \chi_S(1-\psi_S)\mathcal{F}(\theta_{-i}|\theta_i) + [1-\chi_S(1-\psi_S)]\mu_i^*(\theta_{-i}|\theta_i, h^t) \tag{2}$$

[5] PHC may seem like a weak condition, but for most games the coarsest valid partition in SCE violates the condition. For example, it will be violated even in very simple games of perfect information, where there is no initial move by nature, and no simultaneous moves, as we illustrate later in this note. One alternative to avoid this would be to modify the definition of SCE so that the coarsest valid partition, $\mathcal{P}$, is *required* to satisfy PHC, leading to a different equilibrium concept, which could be called *Public Sequential Cursed Equilibrium (PSCE)*. We thank Shengwu Li for suggesting this possibility. For multistage games with *public* histories, PHC seems like a plausible requirement for valid coarsening.

where

$$\mu_i^*(\theta_{-i}|\theta_i, h^t) = \frac{\mathcal{F}\left(\theta_{-i}|\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j, h^{l-1}\right)\right]}{\sum_{\theta_i'\in\Theta_{-i}}\mathcal{F}\left(\theta_{-i}'|\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j', h^{l-1}\right)\right]}.$$

Proof. To obtain the belief of (χ_S, ψ_S)-SCE under PHC, we need to derive the sequentially cursed conjecture, typically cursed conjecture and the Bayesian conjecture of nature separately.

1. Player i's **Bayesian conjecture** about nature's strategy is:

$$\begin{aligned}\hat{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset) &= \mu[\sigma](h_\emptyset, \theta|(\theta_i, h^t)\cap h_\emptyset)\\ &= \frac{\mathcal{F}\left(\theta_{-i},\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j, h^{l-1}\right)\right]}{\sum_{\theta_i'\in\Theta_{-i}}\mathcal{F}\left(\theta_{-i}',\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j', h^{l-1}\right)\right]}\\ &= \frac{\mathcal{F}\left(\theta_{-i}|\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j, h^{l-1}\right)\right]}{\sum_{\theta_i'\in\Theta_{-i}}\mathcal{F}\left(\theta_{-i}'|\theta_i\right)\left[\prod_{j\neq i}\prod_{1\leq l\leq t}\sigma_j\left(a_j^l|\theta_j', h^{l-1}\right)\right]} \equiv \mu_i^*(\theta_{-i}|\theta_i, h^t).\end{aligned}$$

2. Player i's **sequentially cursed conjecture** about nature's strategy is:

$$\tilde{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset) = \mu[\sigma](h_\emptyset, \theta|(\theta_i, h^t)\cap P(h_\emptyset)) = \mu_i^*(\theta_{-i}|\theta_i, h^t),$$

   which coincides with the Bayesian conjecture. Note that $P(h_\emptyset) = h_\emptyset$ since the nature's information set is singleton.

3. Player i's **typically cursed conjecture** about nature's strategy is:

$$\check{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset) = \mu[\sigma](h_\emptyset, \theta|(\theta_i, \cdot)\cap h_\emptyset) = \frac{\mathcal{F}(\theta_{-i},\theta_i)}{\mathcal{F}(\theta_i)} = \mathcal{F}(\theta_{-i}|\theta_i),$$

   which is just the common prior conditional on player i's type.

Combined these three conjectures, we can obtain that

$$\begin{aligned}\mu_i^{(\chi_S,\psi_S)}(\theta_{-i}|\theta_i, h^t) =&\ \chi_S\psi_S\tilde{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset) + \chi_S(1-\psi_S)\check{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset) + (1-\chi_S)\hat{\sigma}_\lambda^{(\theta_i,h^t)}(\theta|h_\emptyset)\\ =&\ \chi_S(1-\psi_S)\mathcal{F}(\theta_{-i}|\theta_i) + [1-\chi_S(1-\psi_S)]\mu_i^*(\theta_{-i}|\theta_i, h^t).\end{aligned}$$

This completes the characterization of the evolution of the conjecture. □

From the proof of Claim 2, we can find that under the typically cursed conjecture, players do not update their beliefs about others' types while under the sequentially cursed and the Bayesian conjectures, players will update their beliefs correctly—in the sense that their conjecture about nature's strategy coincides with the Bayesian posterior. By contrast, under the framework of CSE, players update their beliefs via Bayes' rule while having incorrect perceptions about other players' behavioral strategies. Therefore, as characterized by Claim 1, in χ-CSE the posterior belief about others' types at stage t is a weighted average between the posterior belief at stage $t-1$ and the Bayesian posterior.

The results from Claim 1 and Claim 2 sharply contrast the difference in the belief updating process of the two theories. Under the framework of (χ_S, ψ_S)-SCE, a player perceives the others' strategies by weighting over cursed and actual strategies formed under different partitions of *entire game trees*. As a result, the posterior belief induced by a (χ_S, ψ_S)-SCE at every stage is a weighted average over two extreme cases—the prior belief (i.e., no belief updating) and the Bayesian posterior belief implied by the actual strategy profile (i.e., correct Bayesian belief updating). Alternatively, under the framework of χ-CSE, a player perceives the others' strategies by weighting between cursed (average) and actual strategies *at each stage*. The player would update his belief by applying Bayes rule to his belief in the previous stage, but with an incorrect perception about the behavioral strategy used in that stage. As a result, the posterior belief induced by a χ-CSE at stage t is a weighted average over two cases—the belief at stage $t-1$, and the Bayesian posterior belief implied jointly by that belief and actual behavioral strategy at $t-1$.

Another implication of Claim 1 and Claim 2 concerns the anchoring of the belief evolution to the original prior (nature's move). Under the framework of (χ_S, ψ_S)-SCE, a player always puts $\chi_S(1-\psi_S)$ weight on his prior belief about the joint distribution of types when forming the posterior belief. In this sense, the impact of prior on posterior is persistent and independent of which period the player is currently in. By contrast, under the framework of χ-CSE, a player at stage t puts a χ weight on his belief in previous stage, which has typically evolved over $t-1$ stages. The impact of prior on posterior belief thus diminishes over time compared to the posterior beliefs formed in the latest periods. The following example highlights this difference in the belief updating processes and sheds light on how they reflect distinct views about (cursed) learning behavior in dynamic games.

Before diving into the illustrative example, in Remark 1, we derive a player's conjectures about other players' strategies in past and future events under PHC. This remark is useful for understanding the mechanism of the belief updating in SCE and useful for the calculations in the illustrative example.

Remark 1. *Fix any non-terminal information set $\mathcal{I}_i = (\theta_i, h^t)$ such that there are public histories $h^{t'}, h^{t''}$ satisfying $h^{t'} \prec h^t \preceq h^{t''}$. Let $h^{t''} \equiv (a^1, \ldots, a^{t'}, \ldots, a^t, \ldots, a^{t''})$. Following* *the definition of SCE, under PHC, we can find that player i's conjectures about type θ_j player j's partial strategy at the past public history $h^{t'}$ are:*

1. *The* ***sequentially cursed conjecture*** *coincides with the* ***Bayesian conjecture****:*

$$\tilde{\sigma}_j^{\mathcal{I}_i}(\tilde{a}_j^{t'+1}|\theta_j, h^{t'}) = \hat{\sigma}_j^{\mathcal{I}_i}(\tilde{a}_j^{t'+1}|\theta_j, h^{t'}) = \begin{cases} 1 & \text{if } \tilde{a}_j^{t'+1} = a_j^{t'+1} \\ 0 & \text{if } \tilde{a}_j^{t'+1} \neq a_j^{t'+1} \end{cases}.$$

2. *The* ***typically cursed conjecture*** *is the average behavioral strategy:*

$$\check{\sigma}_j^{\mathcal{I}_i}(a_j^{t'+1}|\theta_j, h^{t'}) = \sum_{\theta'_j \in \Theta_j} \mu_i^*(\theta'_j|\theta_i, h^{t'})\sigma_j(a_j^{t'+1}|\theta'_j, h^{t'}).$$

On the other hand, player i's conjectures about type θ_j player j's partial strategy at the future public history $h^{t''}$ are:

1. *The* ***sequentially cursed conjecture*** *coincides with the* ***typically cursed conjecture****:*

$$\tilde{\sigma}_j^{\mathcal{I}_i}(a_j^{t''+1}|\theta_j, h^{t''}) = \check{\sigma}_j^{\mathcal{I}_i}(a_j^{t''+1}|\theta_j, h^{t''}) = \sum_{\theta'_j \in \Theta_j} \mu_i^*(\theta'_j|\theta_i, h^{t''})\sigma_j(a_j^{t''+1}|\theta'_j, h^{t''}).$$

2. *The* ***Bayesian conjecture*** *coincides with the true behavioral strategy:*

$$\hat{\sigma}_j^{\mathcal{I}_i}(\tilde{a}_j^{t''+1}|\theta_j, h^{t''}) = \sigma_j(\tilde{a}_j^{t''+1}|\theta_j, h^{t''}).$$

In summary, a player who is sequentially cursed accurately perceives another player's past behavior by assigning a probability of 1 to the actions that have been played, but assumes that the other player will use an average strategy (averaged based on μ^*) going forward. On the other hand, a player who is typically cursed perceives both the past and future strategies of another player in the same way, assuming that the other player uses the average strategy.

Illustrative Example[6]

Consider the following game where there is one broadcaster (B) and two other players. We denote the set of players as $N = \{B, 1, 2\}$. There are two possible states—good (θ_g) or bad

[6] This is adapted from a different example proposed by Shengwu Li in private correspondence.

(θ_b)—with equal probability. At the beginning of the game, nature will draw the true state. Only the broadcaster observes the state, not the other two players. Therefore, the true state is the broadcaster's private information.

In this game, all players except for the broadcaster will take turns to choose either a safe option (s) or a risky option (r). Without loss of generality, we let player i be the i-th mover for any $i \in \{1, 2\}$. Each of these two players only takes one action and all actions of all players are public. Moreover, before each player $i \in \{1, 2\}$ makes a move, the broadcaster makes a public announcement about the state being good (g) or bad (b).

This is a three-player four-stage multi-stage game with observed actions where player B moves at odd stages ($t = 1, 3$) with the action set $A_B(h^t) = \{g^{h^t}, b^{h^t}\}$ and each player $i \in \{1, 2\}$ moves at stage $2i$ with the action set $A_i(h^{2i}) = \{s^{h^{2i}}, r^{h^{2i}}\}$. Formally, the coarsest valid partition of this game satisfies PHC since the available actions at different public histories are labeled differently.[7] We drop the superscripts h^t and h^{2i} when there is no risk of confusion.

The broadcaster will get one unit payoff for each truthful announcement and 0 otherwise. Therefore, it is strictly optimal for the broadcaster to always report the true state. For other players, they will get 0 for sure if they choose the safe option s. If they choose the risky option r, they will get $\alpha \in (0, 1)$ if the state is θ_g and -1 if the stage is θ_b.

In summary, the broadcaster has private information about the true state, and is incentivized to always truthfully report the state. All other players do not know the true state, but they are incentivized to choose r in their turn if they are sufficiently confident about the state being θ_g. The standard equilibrium theory predicts the broadcaster will always announce the true state and every player $i \in \{1, 2\}$ will choose s if the announcement is θ_b and choose r otherwise.

To highlight how the belief updating processes differ under SCE and CSE, we will focus on the case in which the state is good (θ_g) and characterize player i's best response to the broadcaster's announcement(s) for $i \in \{1, 2\}$. Because the solutions have the structure that if it is optimal for some player i to choose r, then it is also optimal for $i + 1$ to choose r, in the following, we characterize the solutions by finding the first player to choose r. Notice that there are three possible cases, $\{1, 2, \bar{s}\}$, which corresponds to the cutoff player or the case where both players choose s. We use i^* and i^{**} to denote the cutoff players predicted by SCE and CSE, respectively. Claim 3 shows that, in every (χ_S, ψ_S)-SCE, the two players other than the broadcaster would take identical actions—either $i^* = 1$ (both choose r) or $i^* = \bar{s}$ (both choose s).

[7] See Section 3.5 and Cohen and Li (2023) pp. 32-33 for discussion about labeling effects in SCE.

Claim 3. *In a (χ_S, ψ_S)-SCE, when the broadcaster announces g, the cutoff player i^* is characterized as the following:*

1. *$i^* = 1$ if and only if $\chi_S(1-\psi_S) \leq \frac{2\alpha}{(1+\alpha)}$, and*
2. *$i^* = \bar{s}$ if and only if $\chi_S(1-\psi_S) \geq \frac{2\alpha}{(1+\alpha)}$.*

Proof. Consider the case where the broadcaster announces g and fix any player $i \in \{1,2\}$. By Claim 2, we can obtain that player i's belief about θ_g is

$$\begin{aligned}\mu_i^{(\chi_S,\psi_S)}(\theta_g \mid \text{observing } g \text{ for } i \text{ times}) &= 0.5\chi_S(1-\psi_S) + [1-\chi_S(1-\psi_S)] \\ &= 1 - 0.5\chi_S(1-\psi_S).\end{aligned}$$

In other words, every player i will hold the same belief. Therefore, it is optimal for player i to choose r if and only if

$$\alpha\,[1-0.5\chi_S(1-\psi_S)] - 0.5\chi_S(1-\psi_S) \geq 0 \iff \chi_S(1-\psi_S) \leq \frac{2\alpha}{(1+\alpha)}.$$

Because every player has the same belief, they will take the same action. Coupled with the calculation above, we can conclude that

1. $i^* = 1$ if and only if $\chi_S(1-\psi_S) \leq \frac{2\alpha}{(1+\alpha)}$, and
2. $i^* = \bar{s}$ if and only if $\chi_S(1-\psi_S) \geq \frac{2\alpha}{(1+\alpha)}$,

which completes the proof. □

On the contrary, in the framework of χ-CSE, each player i will gradually update their beliefs as they gather more observations of g announcements, so player 2 will be better informed than player 1. Claim 4 shows that, for intermediate values of χ, the first mover would choose the safe option but the second mover would be confident enough about state being good and choose the risky option in a χ-CSE.

Claim 4. *In a χ-CSE, when the broadcaster announces g, the cutoff player i^{**} is characterized as the following:*

1. *$i^{**} = 1$ if and only if $\chi \leq \frac{2\alpha}{1+\alpha}$,*
2. *$i^{**} = 2$ if and only if $\chi \in \left[\left(\frac{2\alpha}{1+\alpha}\right), \left(\frac{2\alpha}{1+\alpha}\right)^{\frac{1}{2}}\right]$, and*

3. $i^{**} = \bar{s}$ *if and only if* $\chi \geq \left(\frac{2\alpha}{1+\alpha}\right)^{\frac{1}{2}}$.

Proof. Consider the case where the broadcaster announces g and fix any player $i \in \{1, 2\}$. By Claim 1, we can obtain that player i's belief about θ_g is

$$\mu_i^\chi(\theta_g \mid \text{observing } g \text{ for } i \text{ times}) = \chi \mu_i^\chi(\theta_g \mid \text{observing } g \text{ for } i-1 \text{ times}) + (1-\chi)$$

As we iteratively apply Claim 1, we can obtain that for any player i, the belief about θ_g is

$$\mu_i^\chi(\theta_g \mid \text{observing } g \text{ for } i \text{ times}) = 1 - 0.5\chi^i,$$

suggesting that players are more certain about the state being good as they observe more g announcements. As a result, it is optimal for player i to choose r if and only if

$$\alpha\left[1 - 0.5\chi^i\right] - 0.5\chi^i \geq 0 \iff \chi \leq \left(\frac{2\alpha}{1+\alpha}\right)^{\frac{1}{i}}.$$

This completes the proof since the RHS is strictly increasing in i. □

It is noteworthy that the above example as well as Claim 3 and 4 can be easily extended to the case with n players (other than the broadcaster) for any $n > 2$. In a (χ_S, ψ_S)-SCE, the n players would take the same action, while in a χ-CSE, the early movers would choose s but the late movers would choose r for some intermediate value of χ. Furthermore, for any value of $\chi < 1$, there is a critical value n_χ such that for all $n > n_\chi$, there will be a switch point at which later movers will all choose r.

The (χ_S, ψ_S)-SCE predicts that in this example, even though later players ($i \geq 2$) repeatedly observe the broadcaster's behavior, they only believe the first announcement contains useful information. Under the SCE framework, a cursed player i acts as if she fully understands the correlation between the broadcaster's announcements, so collecting more of the broadcaster's announcements does not provide new information for updating her belief. In contrast, the χ-CSE predicts that a (partially) cursed player i will eventually learn the true state if she can observe the broadcaster's announcements infinitely many times. Under the CSE framework, although a cursed player i may be unsure about the broadcaster's private information at the beginning, she will gain confidence from observing that the broadcaster makes the same announcement over time. In this sense, player i gradually learns how another player's strategy depends on his type as she observes the other's actions over time, which is fundamentally different from the learning process characterized in SCE, where beliefs can get stuck after the first stage of the game.

3.2 Difference in Treating Public Histories

In addition to the difference in the belief updating process, another key difference between the CSE and SCE is about the *publicness* of public histories. In the framework of CSE, players have correct understanding about how choosing different actions would result in different histories and they know all other players know this. In other words, from the perspective of CSE, *public histories are essentially public.*

However, public histories are not necessarily public in the framework of SCE. Players in SCE are allowed to be cursed about endogenous information—namely, players may incorrectly believe other players will not respond to changes of actions. Technically speaking, when the coarsest valid partition is not consistent with public histories, i.e., PHC is violated, there is some information set where the player would believe regardless of what he chooses, others will have the same perception about the course of game play.

To demonstrate how this difference leads to different predictions in a specific game, we next consider a simple signaling game and compare χ-CSE and (χ_S, ψ_S)-SCE. This comparison illustrates when PHC is violated, the (χ_S, ψ_S)-SCE can look dramatically different from χ-CSE.

Illustrative Example (Fong et al. (2023))

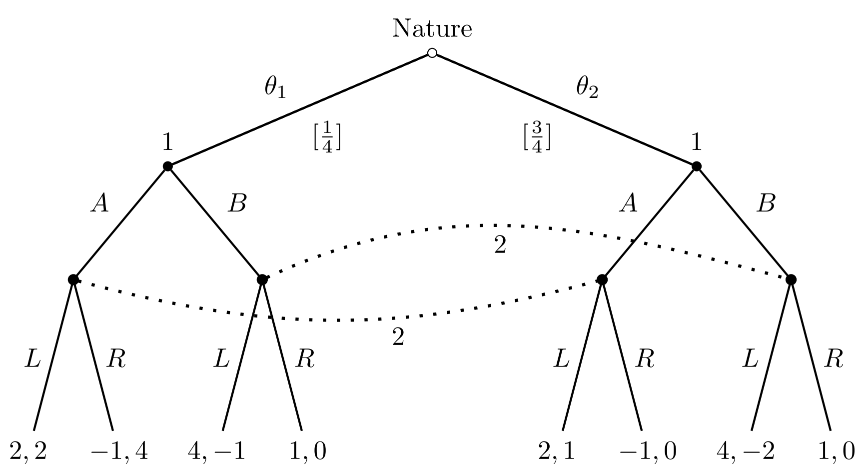

Figure 1: Game Tree for the Illustrative Example

Consider the signaling game depicted in Figure 1 (Example 1 in Section 4.1 of Fong et al., 2023): The sender has two possible types drawn from the set $\Theta = \{\theta_1, \theta_2\}$ with $\Pr(\theta_1) = 1/4$. The receiver does not have any private information. After the sender's type is drawn, the sender observes his type and decides to send a message $m \in \{A, B\}$ (or any mixture between the two). After that, the receiver decides between action $a \in \{L, R\}$ (or any mixture between the two), and the game ends. In the following, we will focus on equilibrium in *pure* behavioral strategies and use a four-tuple $[(m(\theta_1), m(\theta_2)); (a(A), a(B))]$ to denote a behavioral strategy profile.

In Claim 5 and Claim 6, we summarize the solutions of χ-CSE and (χ_S, ψ_S)-SCE, respectively.

Claim 5. *There are two pure pooling χ-CSE, which are:*

1. $[(A, A); (L, R)]$ *is a pooling χ-CSE for any $\chi \in [0, 1]$.*
2. $[(B, B); (R, R)]$ *is a pooling χ-CSE if and only if $\chi \leq 8/9$.*

Proof. First observe that after player 1 chooses B, it is strictly optimal for player 2 to choose R for all beliefs $\mu_2(\theta_1|B)$, and after player 1 chooses A, it is optimal for player 2 to choose L if and only if

$$2\mu_2(\theta_1|A) + [1 - \mu_2(\theta_1|A)] \geq 4\mu_2(\theta_1|A) \iff \mu_2(\theta_1|A) \leq 1/3.$$

Equilibrium 1. If both types of player 1 choose A, then $\mu_2(\theta_1|A) = 1/4$, so it is optimal for player 2 to choose L. Given $a(A) = L$ and $a(B) = R$, it is optimal for both types of player 1 to pick A as $2 > 1$. Hence, $[(A, A); (L, R)]$ is a pooling χ-CSE for any $\chi \in [0, 1]$.

Equilibrium 2. In order to support $m(\theta_1) = m(\theta_2) = B$ to be an equilibrium, player 2 has to choose R at the off-path information set A, which is optimal if and only if $\mu_2(\theta_1|A) \geq 1/3$. In addition, by the χ-dampened updating property (see Proposition 3 of Fong et al., 2023), we know in a χ-CSE, the belief system satisfies

$$\mu_2(\theta_2|A) \geq \frac{3}{4}\chi \iff \mu_2(\theta_1|A) \leq 1 - \frac{3}{4}\chi.$$

Therefore, the belief system has to satisfy that $\mu_2(\theta_1|A) \in \left[\frac{1}{3}, 1 - \frac{3}{4}\chi\right]$, requiring $\chi \leq 8/9$. This completes the proof as it is straightforward to verify that for any $\mu \in \left[\frac{1}{3}, 1 - \frac{3}{4}\chi\right]$, $\mu_2(\theta_1|A) = \mu$ satisfies χ-consistency. We skip the details and refer readers to the Appendix B of Fong et al. (2023). □

Claim 6. *There are four pure (χ_S, ψ_S)-SCE, which are:*

1. $[(A, A); (L, R)]$ *is a (χ_S, ψ_S)-SCE if and only if $\chi_S \leq 1/3$.*

2. $[(B,B);(L,R)]$ *is a* (χ_S,ψ_S)*-SCE if and only if* $\chi_S \geq 1/3$.

3. $[(B,A);(L,R)]$ *is a* (χ_S,ψ_S)*-SCE if and only if* $\chi_S = 1/3$.

4. $[(B,B);(R,R)]$ *is a* (χ_S,ψ_S)*-SCE if and only if* $\chi_S(1-\psi_S) \leq 8/9$.

Proof. First, we can observe that the presence or absence of a PHC restriction on the coarsest partition does not affect player 2's best response to player 1's strategy. Since player 2 moves at the last stage after receiving player 1's message, player 2's conjecture about the other player's (and nature's) strategy is independent of the condition on the publicness of public history. Therefore, we can still apply Claim 2 to pin down player 2's beliefs and the corresponding best responses under player 1's different strategies. For player 2, it is strictly optimal to choose R after seeing $m = B$ for any belief, and it is optimal to choose L if and only if $\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) \leq 1/3$. By Claim 2, we can find that:

1. $m(\theta_1) = m(\theta_2) = A$: $\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) = 1/4$, so it is optimal for player 2 to choose L when $m = A$, implying that $[(A,A);(R,R)]$ is not a (χ_S,ψ_S)-SCE.

2. $m(\theta_1) = m(\theta_2) = B$: $\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) \in [\frac{1}{4}\chi_S(1-\psi_S), 1-\frac{3}{4}\chi_S(1-\psi_S)]$, so $m(A) = R$ can be supported as a best response if and only if $\chi_S(1-\psi_S) \leq 8/9$; alternatively, $m(A) = L$ can be supported as a best response for all $(\chi_S,\psi_S) \in [0,1]^2$.

3. $m(\theta_1) = A$ and $m(\theta_2) = B$: $\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) = 1-\frac{3}{4}\chi_S(1-\psi_S)$, so it is optimal for player 2 to choose L when receiving A if and only if $\chi_S(1-\psi_S) \geq 8/9$.

4. $m(\theta_1) = B$ and $m(\theta_2) = A$: $\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) = \frac{1}{4}\chi_S(1-\psi_S) \leq 1/3$, so it is optimal for player 2 to choose L when receiving A.

Player 1's conjecture about player 2's strategy, however, depends on whether player 1 realizes that the message m will be public. In the framework of SCE, the element of the coarsest valid partition at the message m is

$$P(m) = \{(\theta, m') : \theta \in \{\theta_1,\theta_2\}, m' \in \{A,B\}\},$$

which is the same for any $m \in \{A,B\}$. This implies player 1 would (incorrectly) believe player 2 will behave the same regardless of which message is sent. Given a strategy profile, $\sigma = [(m(\theta_1), m(\theta_2)); (a(A), a(B))]$, a type θ' player 1's (sequentially and typically) cursed conjecture about player 2's strategy when seeing the message $\tilde{m} \in \{A, B\}$ becomes

$$\begin{aligned}\bar{\sigma}_2^{(\theta')}(a'|\tilde{m}) &= \mu[\sigma](\{((\theta, m), a') : (\theta, m) \in P(\tilde{m})\} |\ \theta' \cap P(\tilde{m})) \\ &= \frac{\sigma_\lambda(\theta') \sum_{m \in \{A,B\}} \sigma_1(m|\theta')\sigma_2(a'|m)}{\sigma_\lambda(\theta')} \\ &= 1 \cdot \sigma_2(a'|m = m(\theta')) + 0 \cdot \sigma_2(a'|m \neq m(\theta')) \\ &= \sigma_2(a'|m(\theta')).\end{aligned}$$

In other words, with probability χ_S, player 1 conjectures that player 2's off-path strategy would be the same as her on-path strategy. Notice that both types of player 1 have the same payoff function. Hence:

1. $a(A) = a(B) = R$: It is optimal for player 1 to send the message B as $1 > -1$, implying $[(B, B); (R, R)]$ is a (χ_S, ψ_S)-SCE when $\chi_S(1 - \psi_S) \leq 8/9$.

2. $a(A) = L$ and $a(B) = R$: It is optimal for player 1 to send A if and only if

$$2 \geq 4\chi_S + (1 - \chi_S) \iff \chi_S \leq 1/3.$$

   Alternatively, it is optimal for player 1 to send B if and only if

$$1 \geq -\chi_S + 2(1 - \chi_S) \iff \chi_S \geq 1/3.$$

   Therefore, $[(A, A); (L, R)]$ is a (χ_S, ψ_S)-SCE when $\chi_S \leq 1/3$, and $[(B, B); (L, R)]$ is a (χ_S, ψ_S)-SCE when $\chi_S \geq 1/3$. Moreover, $[(B, A); (L, R)]$ is also a (χ_S, ψ_S)-SCE when $\chi_S = 1/3$. Note that $[(A, B); (L, R)]$ cannot be supported as a (χ_S, ψ_S)-SCE since it requires two contradicting conditions to hold ($\chi_S = 1/3$ and $\chi_S(1 - \psi_S) \geq 8/9$).

This completes the characterization of (χ_S, ψ_S)-SCE. □

To visualize the effect of PHC, we plot the solutions characterized in Claim 6 in Figure 2. In a (χ_S, ψ_S)-SCE, a cursed player 1 would incorrectly believe player 2 would behave the same regardless of which m is chosen. As a result, (χ_S, ψ_S)-SCE differs dramatically from χ-CSE when PHC is violated. The pooling equilibrium $[(A, A); (L, R)]$ is a (χ_S, ψ_S)-SCE if and only if players are not too cursed about endogenous information, i.e., $\chi_S \leq 1/3$, while this is a χ-CSE for any $\chi \in [0, 1]$. Moreover, when players are sufficiently cursed about endogenous information ($\chi_S \geq 1/3$), there is an additional (χ_S, ψ_S)-SCE $[(B, B); (L, R)]$ which is neither a PBE, a CE nor a χ-CSE. Lastly, it is worth mentioning that there exists an knife-edge case that $[(B, A); (L, R)]$ is a separating (χ_S, ψ_S)-SCE if and only if $\chi_S = 1/3$.

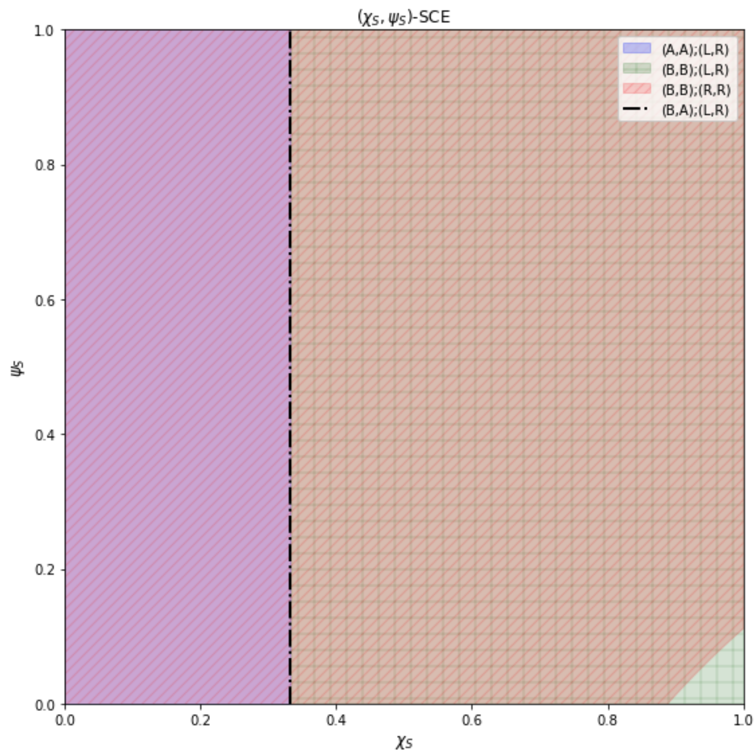

Figure 2: The solution set of (χ_S, ψ_S)-SCE

Since the incentives of both types of player 1 are perfectly aligned, the existence of such a separating (χ_S, ψ_S)-SCE is unexpected.

In summary, a key difference between CSE and SCE is about the way of treating public histories. From the illustrative example, we can find that when the coarsest valid partition is not consistent with the public histories, (χ_S, ψ_S)-SCE and χ-CSE make extremely different predictions. In the following sections, we will explore the implications of the difference in the publicness of public histories. Surprisingly, the impact of the publicness of public histories not only arises in multistage games with incomplete information but also appears in games of perfect information, which will be formally discussed in the next section.

3.3 Difference in Games of Complete Information

A significant difference between CSE and SCE is about the predictions of the games with complete information, i.e., $|\Theta| = 1$. As shown in Claim 7, in games with complete information, CSE coincides with sequential equilibrium. As the type space is singleton, there is no possibility for players to make mistaken inferences about the types of other players, so neglect of the correlation between types and actions is a moot issue, as is also the case in CE.

Claim 7. *If $|\Theta| = 1$, χ-CSE is equivalent to the sequential equilibrium for any $\chi \in [0, 1]$.*

Proof. Let $\Theta = \{(\bar{\theta}_1, \ldots, \bar{\theta}_n)\}$ and consider any behavioral strategy profile. Because $|\Theta| = 1$, for any player i and public history h^{t-1}, we can obtain that $\mu_i(\bar{\theta}_{-i}|\bar{\theta}_i, h^t) = 1$. Consequently, the average behavioral strategy profile of $-i$ is simply:

$$\bar{\sigma}_{-i}(a^t_{-i}|\bar{\theta}_i, h^{t-1}) = \mu_i(\bar{\theta}_{-i}|\bar{\theta}_i, h^t)\sigma_{-i}(a^t_{-i}|\bar{\theta}_i, h^{t-1}) = \sigma_{-i}(a^t_{-i}|\bar{\theta}_i, h^{t-1})$$

for any $a^t_{-i} \in A_{-i}(h^{t-1})$, suggesting $\sigma^{\chi}_{-i}(a^t_{-i}|\bar{\theta}, h^{t-1}) = \sigma_{-i}(a^t_{-i}|\bar{\theta}_{-i}, h^{t-1})$ for any $\chi \in [0, 1]$. Since the perception about others' strategy profile always aligns with the true strategy profile, χ-CSE is equivalent to the sequential equilibrium for any $\chi \in [0, 1]$. □

In contrast, SCE does not coincide with the sequential equilibrium in games with complete information. In fact, this is even true for games of *perfect information*, i.e., nature is not a player and every information set is a singleton. This is illustrated with the following simple game of perfect information. The intuition of this phenomenon ties in with the discussion earlier in the note, since in this example PHC is violated. The coarsest partition is not consistent with the public history—even though the type space and information sets are singleton sets. As a result, when the coarsest partition bundles multiple public histories, players neglect how their current action affects another player's future action.

One of the motivations for SCE is to extend the notion of cursedness to *endogenous* information, i.e., observed actions, which are endogenous to the game. However, it is useful to distinguish between private endogenous information and public endogenous information. In CSE, cursedness arises because players neglect the jointness of other players' actions and their *private information*, but understand the link between other players' actions and *public information.* In multistage games with public histories, there is no "private endogenous" information. We conjecture that if SCE were modified so that the coarsest valid partition is required to be measurable with respect to public histories, then cursedness with respect to *private endogenous* information could still arise, but the phenomenon of cursedness with respect to *public endogenous* information would not arise.

Illustrative Example[8]

Consider the two-player game of perfect information in Figure 3. In the first stage, player 1 makes a choice from $A_1 = \{B, R\}$. After observing player 1's decision, player 2 then makes a choice from $A_2 \in \{b, r\}$. The payoffs are shown in the game tree where $x \in \mathbb{R}$ and

[8]This example is adapted from the game depicted in Figure 5 of Cohen and Li (2023).

$y < 1$. In the following, we will denote the players' (pure) behavioral strategy profile by $[a_1; (a_2(B), a_2(R))]$.

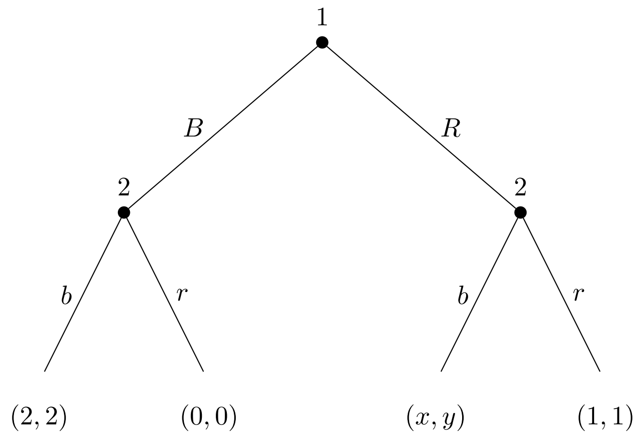

Figure 3: A Two-Player Game with Complete Information

In this game, the only subgame perfect Nash equilibrium is $[B; (b, r)]$. However, Claim 8 shows that $[R; (b, r)]$ can also be supported as a (χ_S, ψ_S)-SCE if χ_S is sufficiently large.

Claim 8. *For any $x \in \mathbb{R}$ and $y < 1$, $[R; (b, r)]$ is a (χ_S, ψ_S)-SCE if and only if $\chi_S \geq 1/2$.*

Proof. First, player 2's best responses to B and R are b and r, respectively. Note that player 1's information set is singleton, so coarsening it has no effect on player 2's conjecture about, and thus best response to, player 1's action. Second, under the coarsest partition, $P(h^1 = R)$ is $\{a_1 : a_1 \in \{B, R\}\}$. Given $[R; (b, r)]$, player 1 conjectures that player 2 when observing B will choose r with probability χ_S and choose b with probability $1 - \chi_S$. Therefore, player 1 will not have an incentive to deviate to B if and only if $1 \geq 2(1 - \chi_S) \iff \chi_S \geq 1/2$. □

It is noteworthy that, *given any x and $y < 1$*, choosing R is a (χ_S, ψ_S)-SCE strategy for player 1 if $\chi_S > 1/2$ and this threshold is independent of x and y. In other words, under the SCE framework, a cursed player 1 could choose R in equilibrium even when such choice is extremely risky for player 1 due to a huge potential loss (e.g., $x = -1000000$) or when player 2 is extremely unlikely to go for b conditional on observing R due to a large negative payoff (e.g., $y = -1000000$), and such possibility does not change in the values of x and y.

3.4 Difference in consistency with Subgame Perfection and Sequential Rationality

The example above shows that SCE can yield predictions that violate subgame perfection, which cannot happen in CSE. A similar issue arises in the context of the signaling game analyzed in section 3.2. In the CSE analysis of the signaling game, players correctly understand that choosing different actions would bring the game to different public histories although they update their beliefs about the types via the χ-cursed Bayes' rule. They further believe that future players will choose optimally (perhaps with non-Bayesian beliefs) at future public histories. Thus, CSE has built into it the property of sequential rationality, analogous to subgame perfection in games of perfect information.[9]

As we observed in the signaling game, when $\chi_S \geq 1/3$, there exists a (χ_S, ψ_S)-SCE where both types of player 1 pool at B and player 2 will choose L and R in public histories A and B, respectively. This solution violates sequential rationality in the sense that, in this SCE, player 1 incorrectly believes player 2 would behave the same at both public histories. That is, player 1 believes player 2 will irrationally choose R in response to A. But if player 1 correctly believes player 2 will rationally choose L at public history A, then it would be profitable for both types of player 1 to choose A.

In summary, the violation of sequential rationality of SCE in these examples is indeed a consequence of the coarsest valid partition being incompatible with public histories. In the next section, we show how the coarseness of the partition can also be affected by how the actions are labelled, which in CSE is an inessential technical detail of the formal game representation. For some additional discussion about the labeling issue see Cohen and Li (2023).

3.5 Difference in the Effect of Re-labeling Actions

From the discussion in the previous section, we can find that the violation of subgame perfection of SCE is a consequence of coarsening public histories into an information set, which in fact, is sensitive to the labels of actions available at these public histories. The requirement of an information set is that the set of actions for every history in this information set is exactly the same. Therefore, if the actions at every public histories are all labelled differently, then the coarsest valid partition is consistent with the public histories. This observation is stated in Claim 9.

Definition 6. *A multi-stage game with observed actions is* ***scrambled*** *if for any $i \in N$, any $t < T$, and any $h, h' \in \mathcal{H}^t$ such that $h \neq h'$, then*

$$\forall\ s \in A_i(h) \ \ \textit{and} \ \ s' \in A_i(h') \implies s \neq s'.$$

[9] Here we use the term sequential rationality instead of subgame perfection, because in the signaling game there are no proper subgames.

Claim 9. *A scrambled multi-stage game with observed actions satisfies PHC.*

Proof. If not, there is some $P(\theta_i, h^t)$ which contains two public histories $h, h' \in \mathcal{H}^t$ where $h \neq h'$. However, because the game is scrambled, $A_i(h) \neq A_i(h')$. Therefore, h and h' cannot belong to the same cell of a partition under any partition, which yields a contradiction. □

This observation provides an alternative interpretation of PHC—we can view this as an additional requirement of SCE such that the solution concept is immune to the effect of re-labeling actions.[10] On the other hand, because the average behavioral strategy of CSE is defined at every public history, the immunity of CSE is built in the model setup.

The illustrative example in Section 3.2 can also demonstrate the effect of re-labeling. Let $A_2(A)$ and $A_2(B)$ be the action sets at the public history A and B, respectively. If $A_2(A) = \{L, R\}$ and $A_2(B) = \{L', R'\}$, then this is a scrambled game and hence satisfies PHC. As characterized in Claim 10, we can find that CSE and SCE are equivalent if $\chi = \chi_S(1 - \psi_S)$. However, if $A_2(A) = A_2(B) = \{L, R\}$, then two public histories A and B belong to the same information set under the coarsest valid partition. As shown in Claim 6, the (χ_S, ψ_S)-SCE solution looks dramatically different.

Claim 10. *If the signaling game depicted in Figure 1 is scrambled, then there are two pure pooling (χ_S, ψ_S)-SCE, which are:*

1. $[(A, A); (L, R')]$ *is a pooling* (χ_S, ψ_S)*-SCE for any* $(\chi_S, \psi_S) \in [0, 1]^2$.
2. $[(B, B); (R, R')]$ *is a pooling* (χ_S, ψ_S)*-SCE if and only if* $\chi_S(1 - \psi_S) \leq 8/9$.

Proof. If the signaling game is scrambled, then by PHC, we can find that the partition illustrated in Figure 1 is the coarsest valid partition. In addition, we can observe that at the public history $m = B$, it is strictly optimal for player 2 to choose R' for any belief. On the other hand, at the public history $m = A$, it is optimal for player 2 to choose L if and only if $\mu_2^{(\chi_S, \psi_S)}(\theta_1|A) \leq 1/3$, as shown in the proof of Claim 5.

Equilibrium 1. If both types of player 1 choose A, then by Claim 2, we can find that $\mu_2^{(\chi_S, \psi_S)}(\theta_1|A) = 1/4$, so it's optimal for player 2 to choose L at the public history $m = A$. Because of PHC, player 1 knows player 2 is choosing a contingent strategy for different public histories. As a result, given $a(A) = L$ and $a(B) = R'$, it is optimal for both types of player 1 to choose A. This shows under PHC, $[(A, A); (L, R')]$ is a pooling (χ_S, ψ_S)-SCE for any $(\chi_S, \psi_S) \in [0, 1]^2$.

[10] To deal with this issue, Cohen and Li propose the concept of *casual SCE*. See Appendix D of Cohen and Li (2023) for details.

Equilibrium 2. To support $m(\theta_1) = m(\theta_2) = B$ to be an equilibrium, because of PHC, player 2 has to choose R at the public history $m = A$. By Claim 2 and the consistency requirement, the belief system has to satisfy

$$\mu_2^{(\chi_S,\psi_S)}(\theta_1|A) \in \left[\frac{1}{3},\ 1 - \frac{3}{4}\chi_S(1-\psi_S)\right],$$

which is valid if and only if $\chi_S(1-\psi_S) \leq 8/9$. Finally, because of PHC, both players know that player 2's strategy is a contingent strategy which is responsive to different public histories. Consequently, we can use the standard equilibrium argument to show that there is no separating pure strategy equilibrium. This completes the characterization of SCE under PHC. □

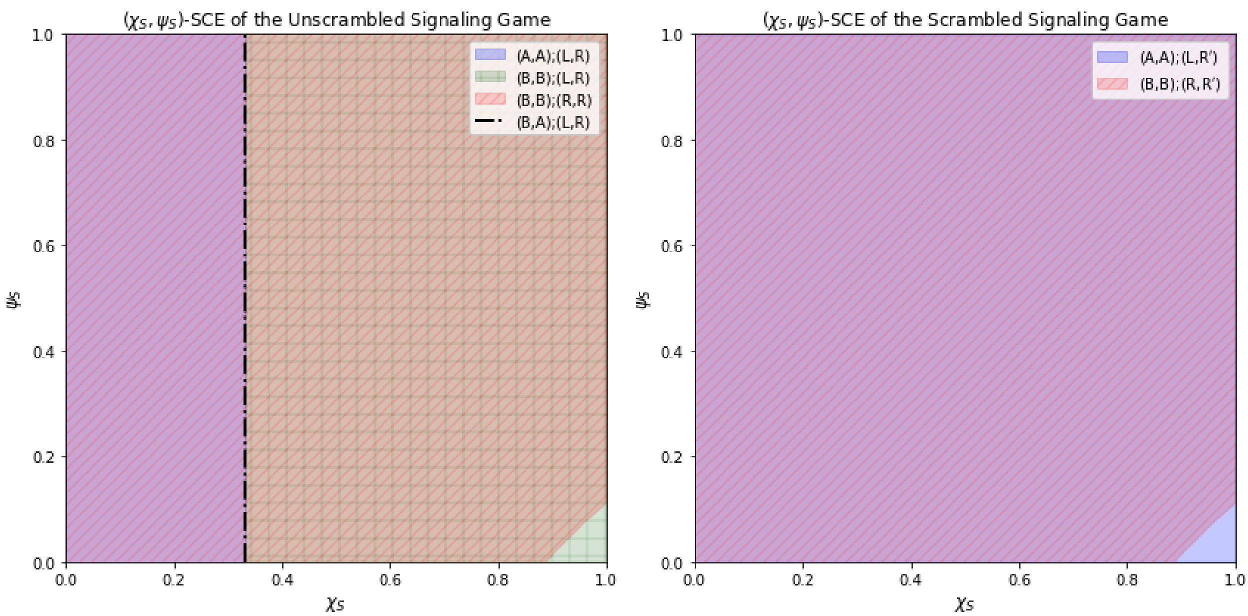

Figure 4: (χ_S, ψ_S)-SCE of an unscrambled and scrambled signaling game

Figure 4 depicts the SCE of the signaling game when it is unscrambled (i.e., PHC is violated) and scrambled (i.e., PHC is satisfied). With PHC, all players understand that every player's strategy is a contingent strategy conditional on the public histories. Therefore, if this two-stage signaling game is scrambled, (χ_S, ψ_S)-SCE coincides with χ-CSE when $\chi = \chi_S(1-\psi_S)$. Moreover, the strategy profiles that could only be supported as a (χ_S, ψ_S)-SCE in the unscrambled signaling game by a relatively large χ_S (i.e., $[(B, B); (L, R')]$ and $[(B, A); (L, R')]$) would no longer be a (χ_S, ψ_S)-SCE when PHC is required.

3.6 Difference in One-Stage Simultaneous-Move Games

The last important difference between CSE and SCE is in their relations with the standard cursed equilibrium in one-stage games. As shown by Fong et al. (2023), CSE coincides with the standard CE for any one-stage game. Yet, Cohen and Li (2023) show that SCE in one-stage games is equivalent to *independently cursed equilibrium* (ICE), under which players are cursed about not only the dependence of opponents' actions on private information but also the correlation between opponents' actions. In the following, we will first summarize the definitions of CE and ICE, and then illustrate how substantial their predictions can differ in an example of a three-player game.

Definition 7 describes the standard CE in an one-stage game. Under CE, a player fails to account for how the other players' action profile may depend on their types, and best responds to the *average strategy profile of the other players.*

Definition 7 (Cursed Equilibrium, Eyster and Rabin, 2005)**.** *A strategy profile σ is a cursed equilibrium (CE) if for each player i, type $\theta_i \in \Theta_i$ and each $a_i^1 \in A_i(h_\emptyset)$ such that $\sigma_i(a_i^1|\theta_i, h_\emptyset) > 0$,*

$$a_i^1 \in \underset{a_i^{1'} \in A_i(h_\emptyset)}{\text{argmax}} \sum_{\theta_{-i}\in\Theta_{-i}} \mathcal{F}(\theta_{-i}|\theta_i) \times \sum_{a_{-i}^1 \in A_{-i}(h_\emptyset)} \left[\sum_{\theta_{-i}\in\Theta_{-i}} \mathcal{F}(\theta_{-i}|\theta_i)\sigma_{-i}(a_{-i}^1|\theta_{-i}, h_\emptyset) \right] u_i(\theta_i, \theta_{-i}, a_i^{1'}, a_{-i}^1),$$

Definition 8 provides the definition of ICE in an one-stage game. Under ICE, a player fails to account for how each player's action may depend on her own type, and how it may correlate with another player's action (via the correlation in type distribution). Therefore, a player would best respond as if the *average strategies across the other players are independent.*

Definition 8 (Independently Cursed Equilibrium)**.** *A strategy profile σ is an independently cursed equilibrium (CE) if for each player i, type $\theta_i \in \Theta_i$ and each $a_i^1 \in A_i(h_\emptyset)$ such that $\sigma_i(a_i^1|\theta_i, h_\emptyset) > 0$,*

$$a_i^1 \in \underset{a_i^{1'} \in A_i(h_\emptyset)}{\text{argmax}} \sum_{\theta_{-i}\in\Theta_{-i}} \mathcal{F}(\theta_{-i}|\theta_i) \times \sum_{a_{-i}^1 \in A_{-i}(h_\emptyset)} \prod_{j\in N\backslash\{i\}} \left[\sum_{\theta_j\in\Theta_j} \mathcal{F}(\theta_j|\theta_i)\sigma_j(a_j^1|\theta_j, h_\emptyset) \right] u_i(\theta_i, \theta_{-i}, a_i^{1'}, a_{-i}^1),$$

Claim 11 and 12 summarize the relations between CSE, SCE and CE in one-stage games.

The proofs can be found in Fong et al. (2023) and Cohen and Li (2023), respectively.

Claim 11 (Proposition 4 in Fong et al., 2023)**.** *For any one-stage game and for any* $\chi \in [0,1]$, χ*-CSE and* χ*-CE are equivalent.*

Claim 12 (Theorem 2.7 in Cohen and Li, 2023)**.** *For any finite one-stage game, ICE and SCE are equivalent.*

Although CE and ICE are equivalent in an one-stage two-person game, we show in the following example that the sets of CE and ICE may be non-overlapping in a three-person game. This finding suggests that CSE and SCE are generally different when there are more than two players.

Illustrative Example

Consider the following three-player one-stage game. Player 1 and 2 have two possible types drawn from the set $\Theta = \{b, r\}$ with the joint distribution $\mathcal{F}(\theta_1 = \theta_2 = b) = \mathcal{F}(\theta_1 = \theta_2 = r) = 0.5 - \epsilon$ and $\mathcal{F}(\theta_1 = b, \theta_2 = r) = \mathcal{F}(\theta_1 = r, \theta_2 = b) = \epsilon$ where $\epsilon \in (0, 0.5)$. Player 3 has no private information. Each player makes a choice from the set $A = \{b, r, m\}$. Player 1 and 2 will get one unit of payoff if his choice matches his type (and 0 otherwise). Player 3 will get one unit of payoff if his choice matches player 1's and 2's choices when $a_1 = a_2$, or if he chooses m when $a_1 \neq a_2$ (and 0 otherwise).

To summarize, player 1 and 2 have private information, and their payoffs will be maximized if their actions match their types. Player 1's (and 2's) type is b or r with equal probabilities. However, their types can be the same with probability $1 - 2\epsilon$. Player 3 has no private information, and his goal is to guess his opponents' actions by following them when they act the same and choosing m when they act differently.

Claim 13. *When* $\epsilon < 1/6$*, player 3 will choose* b *or* r *in CE but* m *in ICE. That is, CE and ICE do not overlap.*

Proof. In both CE and ICE, $a_i^*(\theta_i) = \theta_i$ for $i \in \{1, 2\}$, which is a strictly dominant strategy for Player 1 and 2. Player 3's expected payoff of choosing a_3 under CE is thus

$$\mathbb{E}u_3(a_3, a_{-3}|\sigma_{-3}^*) = \begin{cases} \mathcal{F}(\theta_1 = \theta_2 = a_3) & \text{if } a_3 = b \text{ or } r \\ \mathcal{F}(\theta_1 \neq \theta_2) & \text{if } a_3 = m. \end{cases}$$

Therefore, it is optimal for player 3 to choose b or r in CE if

$$0.5 - \epsilon > 2\epsilon \iff \epsilon < 1/6.$$

Alternatively, player 3's expected payoff of choosing a_3 under ICE is

$$\mathbb{E}u_3(a_3, a_{-3}|\sigma^*_{-3}) = \begin{cases} \mathcal{F}(\theta_1 = a_3)\mathcal{F}(\theta_2 = a_3) & \text{if } a_3 = b \text{ or } r \\ \mathcal{F}(\theta_1 = b)\mathcal{F}(\theta_2 = r) + \mathcal{F}(\theta_1 = r)\mathcal{F}(\theta_2 = b) & \text{if } a_3 = m \end{cases}$$

$$\Longrightarrow \mathbb{E}u_3(a_3, a_{-3}|\sigma^*_{-3}) = \begin{cases} 0.25 & \text{if } a_3 = b \text{ or } r \\ 0.5 & \text{if } a_3 = m. \end{cases}$$

Therefore, it is optimal for player 3 to choose m for all $\epsilon \in [0, 0.5]$. □

Players' private information may serve as a coordinating device for the players' actions if their types are correlated. As a result, when a player neglects the possible correlation in the other players' strategies under ICE, he would respond as if he neglects the dependence of types across players in the prior distribution.

References

Cohen, S. and S. Li (2023): "Sequential Cursed Equilibrium," *arXiv preprint arXiv:2212.06025v3 First version posted December 13, 2022.*

Eyster, E. and M. Rabin (2005): "Cursed equilibrium," *Econometrica*, 73, 1623–1672.

Fong, M.-J., P.-H. Lin, and T. R. Palfrey (2023): "Cursed Sequential Equilibrium," *arXiv preprint arXiv:2301.11971.*

Fudenberg, D. and J. Tirole (1991): "Perfect Bayesian equilibrium and sequential equilibrium," *Journal of Economic Theory*, 53, 236–260.

Jehiel, P. (2005): "Analogy-based expectation equilibrium," *Journal of Economic Theory*, 123, 81–104.

Jehiel, P. and F. Koessler (2008): "Revisiting games of incomplete information with analogy-based expectations," *Games and Economic Behavior*, 62, 533–557.

Kreps, D. M. and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50, 863–894.