The Set of Confounded Learning Beliefs is (Almost) Homeomorphic to Standard Cantor Set in Observational Learning

Xuanye Wang *

Institute for Advanced Economics Research, Dongbei University of Finance and Economics

April 30, 2021

Abstract

We provide a comprehensive characterization of the set of confounded learning beliefs in a classic observational learning model. We surprisingly find that this set could be (almost) homeomorphic to standard cantor set. This peculiar phenomenon is robust in the sense that it holds for an open set of private signals-except for a set of first category; and in the sense that it survives even if we assume all private signals have C^{∞} density functions. This peculiar phenomenon is gone if we assume all private signals have real-analytic density functions. Then the set of confounded learning beliefs must be discrete. We also prove that the set of real-analytic private signals is dense in the set of all private signals.

Keywords: Social Learning, Information Aggregation, Confounded Learning.
JEL Classification: C11, D83

1 Introduction

Smith and Sørensen (2000) shows that long run learning could be confounded in an observational learning environment. A confounded learning belief assigns positive weights to

*Email: xuanye.wang@dufe.edu.cn.

both underlying states of the world in a way so that any observable choice is equally possible across the two states. Therefore, despite all people still use their private information to make choices, late comers cannot correctly disentangle the information contained in early choices. As a result, information cannot aggregate through observing others' choices or actions, and hence the society stays "confused" in the long run. Smith and Sørensen (2000) shows that confounded learning could arise robustly if there are two groups of people with sufficiently opposite preferences.

In the past twenty years, Smith and Sørensen (2000) has become one of the standard models in the study of observational learning. But confounded learning doesn't receive enough attention in the literature. This article aims at filling this gap by providing a full characterization of the structure of the set of confounded learning beliefs. We ask the following questions: if we randomly draw an observational learning model, what will the set of confounded learning beliefs looks like? The answer is quite surprising: except for a small set in the sense of first category, for all the signals that admit confounded learning, the set of confounded learning beliefs is either a singleton or contains a (homeomorphic) cantor set. So some observational learning models are peculiar behaved in the sense that its confounded learning set looks like cantor set. Furthermore, from two perspectives we could say this peculiar phenomenon is robust: first, the set of peculiar signals is large in topological sense. Except from a set of first category, which is topologically small, the set of peculiar signals contains an open set. Second, the set of peculiar signals remains topologically large even if we impose strong smoothness assumptions. If we only consider those observational learning models whose private signals having C^∞ density functions, the peculiar C^∞ models is still topologically large among all C^∞ models.

Knowing the existence and robustness of the peculiar phenomenon, we ask the natural question-how to eliminate it? Our answer is to assume private signals have real-analytic density functions. With this real-analytic assumption, the set of confounded learning beliefs is guaranteed to be discrete. Besides, since confounded learning belief could only arise within confounding region [1] , we can further conclude that confounded learning set must be finite for bounded strength real-analytic private signals whose confounding region is bounded. Lastly, we prove that this real-analytic assumption is harmless-models with real-analytic private signals are dense among all models. Therefore, for researchers who would like to draw a generic conclusion about confounded learning beliefs, it might be good to restrict

[1]For any observational learning model, its confounding region is the set of public beliefs where both types use their private information to decide.

M	a	b
A	u	0
B	0	1

MM	a	b
A	0	v
B	1	0

Table 1: Payoff of Match Type (M) and Mismatch Type (MM)

their attention to private signals with real-analytic density functions.

This article is organized as following: in section 2 we do all the preparation works. We introduce the model, nail down the definitions and notations. We also build a mathematical framework so that we can give a slightly precise preview of our results. In section 3, we present our main results. To smooth the flow of reasoning in section 3, we put lots of important constructions into the appendix. Except subsection B.7 and section G, which shows some non-intuitive computation, readers may find interesting ideas in other parts of the appendix. We organize the appendix so that each section in the appendix fills in the details omitted in one subsection in section 3. In section 4 we conclude.

2 Set Up

In this section we first introduce the detailed model, and briefly review some work in Smith and Sørensen (2000) to nail down our notation. Then we introduce some necessary mathematical framework so that we could give a (slightly precise) preview of our results.

2.1 Model

We consider a standard observational learning model with two groups of players with sufficiently divergent preference. For readers who are familiar with the model in Smith and Sørensen (2000), the model to be presented is their model with two payoff relevant states and two groups of players, where one group has opposite preference to the other group. Below are the details.

The model is of discrete-time. In period 0, nature chooses one realization out of two potential states A and B according to a flat prior. In each period $t \geq 1$, one player arrives and needs to make a choice between actions a and b. There are two types of players-"match" type and "mismatch" type. The match type receives positive payoff iff his/her actions matches the realized underlying state. The mismatch type receives positive payoff if the chosen action is different from nature's chosen state. Table 1 gives out the payoff matrices of these two types.

Of course, player at period t doesn't directly observe nature's choice. Instead, he/she observes a noisy signal about nature's choice. This noisy signal is a state-contingent random variable $\mathcal{S}$ that takes values on an open interval $(\underline{s}, \overline{s}) \subset (0,1)$. Here state-contingence means that the signal $\mathcal{S}$'s distribution depends on the realized state. That is, $\mathcal{S}$'s distribution function is $F^A(s)$ if nature chooses state A, and is $F^B(s)$ if nature chooses state B. We could model the private signal as a direct signal, that is, $\mathcal{S} = s$ means that " the probability that nature chooses state A is s". Of course, to make things consistent, we must have

$$\frac{f^B(s)}{f^A(s)} = \frac{1-s}{s}. \tag{2.1}$$

Here $f^A(s), f^B(s)$ are the density functions of private signals under state A and B. The condition 2.1 means: when the direct signal says the probability of state being A is s, the likelihood ratio $\frac{f^B(s)}{f^A(s)}$ tells you the same thing. We assume that $\inf supp(f^A(s)) = \inf supp(f^B(s)) = \underline{s}$ and $\sup supp(f^A(s)) = \sup supp(f^B(s)) = \overline{s}$. [2] If $(\underline{s}, \overline{s}) = (0,1)$, then we say the signal is of unbounded strength since one may get a realization that is arbitrarily precise. Otherwise, we say the signal is of bounded strength. The noisy signal is private in the sense that player t's signal is only available to him/her. Besides, private signals are i.i.d. across players.

Besides a private signal realization, player t also knows the ordered choice made by early arrivers. That is, player t observes a sequence of choices $\{\alpha_1, \ldots, \alpha_{t-1}\}$, where α_k is the choice made by player in period k. For notation abbreviation, we shall refer the ordered sequence of choices up to period t as the public history h_t at period t. Player t also knows that each arriver is of match type with probability p.

2.2 Confounded Learning is Robust in the Model

Smith and Sørensen (2000) shows that confounded learning is a robust phenomenon in the above model as long as the two groups of players have sufficiently different preference. See Theorem 2(g) in Smith and Sørensen (2000). Below we briefly review how they solve the model, and why confounded learning is essentially different from other long run learning results.

Without loss of generality, assume nature chooses state A in period 0, and hence that

[2] If two density functions are only differed on a set with measure 0, they actually give the same distribution. In this case, by $\inf supp(f(s)) = \underline{s}$ we mean that: $\inf\{s | m(\{c \in (-\infty, s) | f(c) > 0\}) > 0\} = \underline{s}$. And $\sup supp(f(s)) = \overline{s}$ is similarly defined. Then it is clear that if a density function takes value 1 at $\frac{\underline{s}}{2}$ but is 0 between $\frac{\underline{s}}{2}$ and $\underline{s}$, the infimum of its support is not $\frac{\underline{s}}{2}$.

the distribution of private signals is $F^A(s)$. Let

$$\lambda_t = \frac{\Pr(B|h_t)}{\Pr(A|h_t)} \tag{2.2}$$

be the posterior likelihood ratio of the incorrect state (B) over the correct state (A), conditional on the public history h_t. Smith and Sørensen (2000) shows that $\{\lambda_t\}_{t\in\mathbb{N}}$ forms a martingale and hence converges a.s. to a limit likelihood ratio λ_∞. For any λ^* that λ_∞ takes with positive probability, the public belief λ_t converges to it with positive probability. Such a λ^* is a limit belief.

If $\lambda^* \in (u\frac{\underline{s}}{1-\underline{s}}, u\frac{\bar{s}}{1-\bar{s}})$, then even in the limit, the match type's private information still matters. For example, if a match type gets a private signal sufficiently strong for state B, then he/she shall choose action b since his/her posterior belief is close enough to $\lambda^*\frac{1-\underline{s}}{\underline{s}} > u$. Similarly, this player will choose action a upon getting a private signal sufficiently strong for state A. Thus, $(u\frac{\underline{s}}{1-\underline{s}}, u\frac{\bar{s}}{1-\bar{s}})$ is the active region for match type. Similarly, we could find $(v\frac{\underline{s}}{1-\underline{s}}, v\frac{\bar{s}}{1-\bar{s}})$ is the active region for mismatch type. The intersection $(u\frac{\underline{s}}{1-\underline{s}}, u\frac{\bar{s}}{1-\bar{s}}) \cap (v\frac{\underline{s}}{1-\underline{s}}, v\frac{\bar{s}}{1-\bar{s}})$ is called "both active region" or "confounding region" and is denoted as $\mathcal{B}$. For an interior limit belief $\lambda^* \in \mathcal{B}$, both types of players still uses his/her private information to decide. Such an interior belief could be stationary because

$$\Pr(b|\lambda^*, A) = \Pr(b|\lambda^*, B) \tag{2.3}$$

happens to hold. Given that any actions happen with equal probability under the two underlying states, observing any action doesn't provide any information about which underlying state is more plausible. As a result, the public belief cannot evolve. Such an interior limit belief $\lambda^* \in \mathcal{B}$ is called "confounded learning belief". The confounded learning set for this model is

$$\Lambda = \{\lambda^* \in \mathcal{B} | \Pr(b|\lambda^*, A) = \Pr(b|\lambda^*, B)\}. \tag{2.4}$$

If the private signal is of unbounded strength, then its confounding region $\mathcal{B} = (0, +\infty)$. However, if private signals is of bounded strength, its confounding region could be empty or degenerated. For example, we may have $v\frac{\underline{s}}{1-\underline{s}} > u\frac{\bar{s}}{1-\bar{s}}$ hold. Since confounded learning could only arise within confounding region, we only cares about those models such that associated confounding region is an non-empty open interval. Interesting readers could see appendix C.1 (Lemma C.2 to Remark C.4) for a discussion about empty and degenerate confounding region.

We close this subsection with the following observation:

Proposition 2.1 *If $u = v$, then the long run learning result in the above model cannot be confounded.*

Since in this article we focus on the structure of confounded learning set, we only consider parameters that satisfying $u \neq v$.

2.3 Mathematical Preparation and a Result Preview

In this article, we characterize an observational learning model by a five tuple $\{p, u, v, \underline{s}, \overline{s}\}$ and a set of signals

$$F_{(\underline{s},\overline{s})} = \{f(s) \in L_1(\underline{s},\overline{s}) | \int_{\underline{s}}^{\overline{s}} f(s)ds = 1; \int_{\underline{s}}^{\overline{s}} f(s)\frac{1-2s}{s}ds = 0, f(s) \geq 0\}. \tag{2.5}$$

Unlike Smith and Sørensen (2000), who model private signals as a pair of density functions $(f^A(s), f^B(s))$ that satisfies certain conditions. We make use of equation 2.1 to suppress the pair $(f^A(s), f^B(s))$ into one $f^A(s) \in F_{(\underline{s},\overline{s})}$. It is direct to verify that for any $f(s) \in F_{(\underline{s},\overline{s})}$, the associated pair $\big(f(s), f(s)\frac{1-2s}{s}\big)$ is a pair of density functions that satisfies SS's condition. The signal set $F_{(\underline{s},\overline{s})}$ is naturally a subset of $L_1(\underline{s},\overline{s})$, and inherits the subspace topology from the L_1-space. In this article all topological definitions-first category, dense .etc-are under the subspace topology of $L_1(\underline{s},\overline{s})$.

From now on, whenever we mention a signal, we mean a density function $f(s) \in F_{(\underline{s},\overline{s})}$, rather than a signal realization that's observed by some player. Sometimes, we mention $f(s) \in F_{(\underline{s},\overline{s})}$ as an "integrable signal" for the reason that the density function $f(s)$ is merely integrable. In the later part of this article, we will impose certain smoothness assumption to the signal set $F_{(\underline{s},\overline{s})}$, and study the impact of those assumptions on structure of confounded learning set. Among these "smooth signal sets", two sets are particular important. The first is the "C^k signal set"

$$F^k_{(\underline{s},\overline{s})} = F_{(\underline{s},\overline{s})} \cap C^k(\underline{s},\overline{s}), 0 \leq k \leq +\infty. \tag{2.6}$$

The second is the "real-analytic signal set"

$$F^\omega_{(\underline{s},\overline{s})} = F_{(\underline{s},\overline{s})} \cap C^\omega(\underline{s},\overline{s}). \tag{2.7}$$

Obviously, a C^k-signal is a signal whose density function is k-th continuously differentiable; a real-analytic signal is a signal whose density function is real analytic.

In this article, we will fix an arbitrary parameter tuple $(p, u, v, \underline{s}, \overline{s})$ and ask what does the confounded learning set Λ_f looks like for most signals $f(s)$ in $F_{(\underline{s},\overline{s})}$. The surprising finding-in vague language-is that: there always exists an open set of parameters U, such that for each parameter tuple in U, the confounded learning set for almost all signal $f(s)$ is either empty, or a singleton, or contains a set that is homeomorphic to a cantor set. Furthermore, if we ask how large is the set of signals that has a confounded learning set looking like a cantor set, we can say it is large in topological sense. To be specific, if we denote this set of signals as $F_{cantor} \subset F_{(\underline{s},\overline{s})}$, we find $F_{cantor} \supset \mathcal{E} - F_{fc}$, where $\mathcal{E}$ is open and F_{fc} is of first category in $F_{(\underline{s},\overline{s})}$.

We further ask whether we could eliminate those peculiar behaved signals F_{cantor} by considering "smooth signals". We find the peculiarity is quite robust to smoothness assumptions. Even if we consider the "C^∞ signal set", the same peculiarity persists. For each parameter tuple in the same open set U, there exists a topologically large set $F^\infty_{cantor} = F_{cantor} \cap C^\infty_{(\underline{s},\overline{s})}$ in $F^\infty_{(\underline{s},\overline{s})}$, such that for all $f(s) \in F^\infty_{cantor}$, associated confounded learning set Λ_f contains a set homeomorphic to a cantor set.

To eliminate the peculiar behaved signals, we must consider the "real-analytic signal set". Then we find that $F_{cantor} \cap C^\omega_{(\underline{s},\overline{s})}$ is always empty. In fact, as any real-analytic function must have a discrete zero set, we find that the confounded learning set Λ_f must be discrete for any real-analytic signal $f(s) \in F^\omega_{(\underline{s},\overline{s})}$. Furthermore, since a discrete set on a bounded region must be finite. The confounded learning set must be finite whenever the confounding region is bounded.

Lastly, we establish that $F^\omega_{(\underline{s},\overline{s})}$ is dense in $F_{(\underline{s},\overline{s})}$. In other words, for any integrable signal $f(s)$, there is a sequence of real-analytic signals $f^\omega_k(s)$ approximate it. Therefore, for any people who are interested in giving a generic result on confounded learning, it may be good to restrict themselves to real analytic signals.

3 Main Results

Arbitrarily fix a parameter tuple $\{p, u, v, \underline{s}, \overline{s}\}$, for any signal $f(s) \in F_{(\underline{s},\overline{s})}$, we introduce

$$G_f(\lambda) = \Pr(b|\lambda, B) - \Pr(b|\lambda, A) = p \int_{\underline{s}}^{\frac{\lambda}{\lambda+u}} f(s)\frac{1-2s}{s}ds + (1-p)\int_{\frac{\lambda}{\lambda+v}}^{\overline{s}} f(s)\frac{1-2s}{s}ds \quad (3.1)$$

as the characteristic function of signal $f(s)$ and denote the confounded learning set of signal $f(s)$ as Λ_f. It is direct to verify that

Lemma 3.1 *The characteristic function $G_f(\lambda)$ is absolutely continuous in λ on $\mathcal{B}$.* [3]

In this article, we mathematically view the confounded learning set Λ_f as the zero set of an absolutely continuous function on a open interval $\mathcal{B}$. That is,

$$\Lambda_f = \{\lambda \in \mathcal{B} | G_f(\lambda) = 0\}. \tag{3.2}$$

3.1 The Confounded Learning Belief Set is (Almost) Always Nowhere Dense

After establishing the mathematical characterization of confounded learning set Λ_f, we start to explore its structure. The first question we ask is: could Λ_f contains an interior point for some signal $f(s)$? Furthermore, if there are such signals, how many are them?

If we think about this question for a while, we may agree that there could be such signals, but these private signals should be rare.

Intuitively, for a private signal $f(s)$ with confounded learning belief λ^*, we know

$$G_f(\lambda^*) = \Pr(b|B, \lambda^*) - \Pr(b|A, \lambda^*) = 0. \tag{3.3}$$

If λ slightly increases above λ^*, the public belief assigns a little bit more weight to state B. So this slight increase of the public belief will change both types' decision provided that their private signal realizations are close to the margin of choosing one action over the other. However, if the private signal $f(s)$ is 0 for a small interval close to the margin, then both types choose action b with the same probability under this slightly bigger public belief λ, and hence λ is also a confounded learning belief. In short, a private signal $f(s)$ which is 0 on some interval could have a confounded learning belief set that contains an interior point. On the other hand, that confounded learning set contains an interior point implies that $G_f(\lambda)$ is locally flat. But such "local flatness" can be easily destroyed by a slight perturbation.

The above intuition is consolidated by the following two propositions, which claim that the set of private signals whose confounded learning set contains an interior point is of first

[3] Absolute continuity is in general defined on a closed bounded interval. If signal strength is unbounded, and confounding region $\mathcal{B}$ is $(0, +\infty)$, then we understand the lemma as claiming: for any ε, there is a δ, such that for any finite disjoint collection of open intervals $\{(a_i, b_i)\}_{1\leq i \leq k}$, we have $\sum_{i=1}^k |b_i - a_i| \leq \delta \Rightarrow \sum_{i=1}^k |G_f(b_i) - G_f(a_i)| \leq \varepsilon$.

category.

Proposition 3.2 *Let*

$$F_I = \{f(s) \in F_{(\underline{s},\overline{s})} | I \subset \Lambda_f\} \tag{3.4}$$

to be the set of private signals whose confounded learning belief set Λ_f *contains interval* I*. Then* F_I *is nowhere dense in* F*.*

The above proposition is proved in appendix A.

Proposition 3.3 *Let*

$$\mathring{F} = \{f(s) \in F | \Lambda_f \textit{contains an interior point}\} \tag{3.5}$$

be the set of private signals whose confounded learning belief set Λ_f *contains an interior point. Then*

$$\mathring{F} \subset \bigcup_{I \textit{ with rational endpoints}} F_I \tag{3.6}$$

and hence is of first category in F*.*

These two propositions actually say more. Since $G_f(\lambda)$ is continuous, for any $\lambda \in \mathcal{B}$, if λ is in the closure of Λ_f, λ must itself be a confounded learning belief and is in Λ_f. So the closure $\overline{\Lambda_f}$ could at most adds two points $\inf \mathcal{B}, \sup \mathcal{B}$ to Λ_f. Thus, Λ_f containing no interior point implies that $\overline{\Lambda_f}$ containing no interior point as well. As a result, we obtain our first result about the structure of confounded learning belief set for a generic model.

Theorem 3.4 *Except for a set of first category* $\mathring{F}$*, for all signals in* $F - \mathring{F}$*, the confounded learning belief set must be nowhere dense.*

3.2 There Exists a Monotonic Region on which Confounded Learning Set is either Empty or Unique

Theorem 3.4 could be re-interpreted as that the characteristic function $G_f(\lambda)$ is not locally flat for almost all signals except a set of first category $\mathring{F}$. It is natural to ask whether $G_f(\lambda)$ is monotonic somewhere for a large set of signals. The answer is yes, for all signals, there is a region on which characteristic functions $G_f(\lambda)$ are monotonic. This is established in the following lemma.

Lemma 3.5 *For any $f(s) \in F$, we have*

$$G'_f(\lambda) \le 0 \text{ on } [u,v], \text{ if } v > u;$$
$$G'_f(\lambda) \ge 0 \text{ on } [v,u], \text{ if } u > v. \tag{3.7}$$

Proof. This follows from a direct computation that

$$G'_f(\lambda) = pf(\frac{\lambda}{\lambda+u})\frac{u-\lambda}{\lambda}\frac{u}{(\lambda+u)^2} + (1-p)f(\frac{\lambda}{\lambda+v})\frac{\lambda-v}{\lambda}\frac{v}{(\lambda+v)^2}. \tag{3.8}$$

Then it is obvious that $G'_f(\lambda) \le 0$ on $\lambda \in [u,v]$, if we assume $v > u$. If we assume $u > v$, then $G'_f(\lambda) \ge 0$ on $\lambda \in [v,u]$. ■

We shall just name the region where $G_f(\lambda)$ is monotonic as the "monotonic region". If $v > u$, then the monotonic region is $[u,v]$. If $u > v$, then the monotonic region is $[v,u]$. The following proposition says that there could be at most one confounded learning belief within "monotonic region".

Proposition 3.6 *For any signal $f(s) \in F - \mathring{F}$, it could admit at most one confounded learning belief on the monotonic region.*

Proof. If we assume $v > u$, and there exists two confounded learning beliefs $\lambda_1^* < \lambda_2^*$ in $[u,v]$, we must have

$$G_f(\lambda_2^*) - G_f(\lambda_1^*) = 0 = \int_{\lambda_1^*}^{\lambda_2^*} G'_f(\lambda)d\lambda. \tag{3.9}$$

This implies that $G'_f(\lambda) = 0$ a.e. on $[\lambda_1^*, \lambda_2^*]$. Then $[\lambda_1^*, \lambda_2^*] \subset \Lambda_f$, which implies that $f(s) \in \mathring{F}$. ■

3.3 Almost No Model has an Isolated Confounded Learning Belief outside Monotonic Region

After knowing that there could be at most one confounded learning belief within the monotonic region, it is natural to ask what happens outside the region.

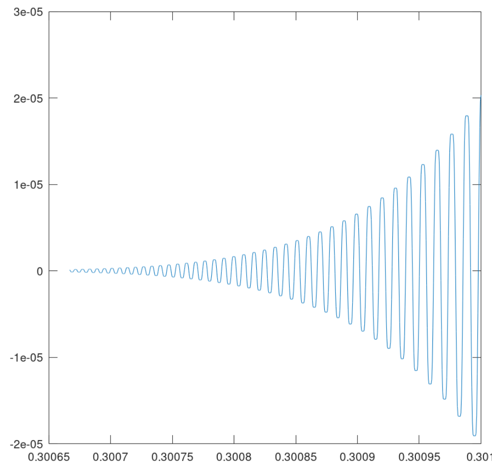


Figure 1: Oscillating Characteristic Function of Perturbed $\hat{f}(s)$

The main message, as suggested in the subsection's title, is that: the set of private signals whose confounded learning set contains an isolated point outside monotonic region, is small in the first category sense. Therefore, for almost all the signals, it either has no confounded learning belief outside of the monotonic region, or its confounded learning set must contain no isolated point. Therefore, unlike the neat structure on the monotonic region, the structure outside the monotonic region is much more complicated.

Rigorously, our result is the following theorem

Theorem 3.7 *Let*

$$F_{isolated} = \{f(s) \in F | G_f(\lambda) \textit{has an isolated zero point } \lambda^* \textit{ outside the monotonic region}\}.$$

Then $F_{isolated}$ is of first category in F.

The above theorem takes lots of effort to prove. Briefly speaking, we develop a technique that enables us to perturb any given $f(s)$ slightly [4] to a $\hat{f}(s)$ so that the characteristic function $G_{\hat{f}}(\lambda)$ oscillates around λ^*. In Figure 3.3 we graph an example with $\lambda^* = 0.3$ to illustrate what we mean by oscillating near λ^*.

In the subsection appendix B.1, we show how to break $F_{isolated}$ into countably union of sets, and how to use the existence of the above oscillating functions to show each of these

[4] Here by slightly, we actually mean that the perturbed $\hat{f}(s)$ could be made arbitrarily close to $f(s)$ under the L_1-norm.

set is nowhere dense. Interested readers could read the proof outline to get an intuition about $F_{isolated}$. Detailed proof is provided from subsections B.2 to B.7. For readers who are not interested in the proof, though not rigorous, it may be harmless to think that any private signal with isolated confounded learning belief λ^* outside monotonic region can be approximated by a sequence of such oscillating functions.

3.4 There are an Open Set of Signals Admitting Confounded Learning Outside Monotonic Region

Knowing that the confounded learning set outside the monotonic region is either empty or has a complicated structure. It is natural to ask how many signals admit confounded learning belief outside the monotonic region.

If we think about this question for a while, we realize the answer does depend on the parameter tuple $\{p, u, v, \underline{s}, \overline{s}\}$. For example, if signal strength is bounded, then the confounding region $\mathcal{B}$ could be contained in the monotonic region. Then for these parameter, the confounded learning set outside monotonic region must be empty. In lemmas C.2 and C.3 in the appendix, we characterize those parameters. We consider those parameters as uninteresting.

A much more interesting question that we try to answer is, for parameters such that part of confounding region is outside of monotonic region, how many signals admit confounded learning belief outside of monotonic region. In particular, if signal strength is unbounded, then the associated confounding region is $(0, +\infty)$. The confounding region outside monotonic region is never empty if signal is of unbounded strength.

To answer this question, our first simple observation is: if the characteristic function $G_f(\lambda)$ changes sign at λ^*, then for any signal $g(s)$ that is close enough to $f(s)$, the associated characteristic function $G_g(\lambda)$ must change sign near λ^*. This follows from the following property of $G_f(\lambda)$

Lemma 3.8 *Let $\{f_n(s)\}_{n\in\mathbb{N}}, f(s) \in L_1(\underline{s}, \overline{s})$ and $\|f_n(s) - f(s)\|_{L_1} \to 0$, then*

$$\forall \lambda \in (0, +\infty), |G_{f_n}(\lambda) - G_f(\lambda)| \to 0. \tag{3.10}$$

Therefore, a sign change of $G_f(\lambda)$ at λ^* must be shared by all signals close enough to $f(s)$. If we name a λ^* to be an essential confounded learning belief for $f(s)$ if $G_f(\lambda)$ changes sign at λ^*. Then the above simple observation can be restated as

Lemma 3.9 *The signal set $\mathcal{E} \subset F$ admitting essential confounded learning belief outside monotonic region, if non-empty, must be open in F.*

Now the question becomes whether the signal set $\mathcal{E}$ is empty.

The answer involves discussion over parameters and we delay the detailed discussion to appendix C. To briefly summarize here:

Proposition 3.10 *If we consider the set of unbounded private signals $F_{(0,1)}$, then there exists an open set $U = \{(p,u,v)|p \neq \frac{1}{2}, u \neq v\}$ in the parameter space such that for each $(p,u,v) \in U$, the signal set $\mathcal{E}$ admitting essential confounded learning belief outside monotonic region is always non-empty.*

If we consider the set of bounded signal $F_{(\underline{s},\overline{s})}$, there exists an open set $U \subset (0,1) \times \mathbb{R}_+^2 \times (0,\frac{1}{2}) \times (\frac{1}{2},1)$ such that for each $(p,u,v,\underline{s},\overline{s}) \in U$, the signal set $\mathcal{E} \subset F_{(\underline{s},\overline{s})}$ that admit essential confounded learning belief outside monotonic region is non-empty.

We prove unbounded signal case in lemma C.13 and bounded signal case in lemma C.9. The detailed characterization of set U in bounded signal case is also provided in lemma C.9. We basically discuss parameters case by case, and then for each case explicitly constructing C^∞ signals that're in $\mathcal{E}$.

3.5 Confounded Learning Set Robustly Looks like a Cantor Set

Now we know that there are enough parameters and enough signals such that confounded learning set is non-empty outside the monotonic region. From previous subsection 3.3, we also know that except for a set of first category of signals, a confounded learning set cannot contain an isolated point outside the monotonic region. From previous subsection 3.2, we know that except for a set of first category of signals, a confounded learning set must be nowhere dense. To summarize, we have shown that for each parameter tuple $\{p,u,v,\underline{s},\overline{s}\}$ in an open set U, there exists an open set $\mathcal{E}$ and two first category sets $\mathring{F}$ and $F_{isolated}$ of signals, such that for each $f(s) \in \mathcal{E} - \mathring{F} - F_{isolated}$, its confounded learning belief set outside the monotonic region must be non-empty, nowhere dense and containing no isolated point. In this subsection, we shall explore the implication for a set being non-empty, nowhere dense and containing no isolated point. We shall find that such a set must look like a cantor set.

Before we proceed, for presentation purpose, in this subsection we assume $v > u$, so that the monotonic region is $[u,v]$. Furthermore, we use $\Lambda_f^{(0,u)}$ to denote the confounded learning set's restriction on open interval $(0,u)$, and $\Lambda_f^{(v,+\infty)}$ is similarly defined.

The driven horse for our result is the following mathematical theorem.

Theorem 3.11 *If a set $S \subset \mathbb{R}$ is bounded, nowhere dense and perfect, then there exists a homeomorphism h from S to the standard cantor set $\mathcal{C}$. Besides, this homeomorphism carries the two endpoints of S to the two endpoints of $\mathcal{C}$:*

$$h(\inf S) = \inf \mathcal{C} = 0; \ \textit{and}\ h(\sup S) = \sup \mathcal{C} = 1. \tag{3.11}$$

This mathematical result is closely related to our main result in the following way. Recall that for any private signal $f(s)$ in $\mathcal{E} - \mathring{F} - F_{isolated}$, its confounded learning set on $(0,u) \cup (v,+\infty)$ is non-empty, nowhere dense, and contains no isolated point. Then the closure [5] of $\Lambda_f^{(0,u)\cup(v,+\infty)}$ is perfect [6] and nowhere dense. If $\Lambda_f^{(0,u)}$ is non-empty, then its closure is automatically bounded, nowhere dense and perfect. If $\Lambda_f^{(v,+\infty)}$ is non-empty, then the closure of its image under taking reciprocal is bounded, nowhere dense and perfect. Furthermore, because $G_f(\lambda)$ is continuous, $\overline{\Lambda_f^{(0,u)}} - \Lambda_f^{(0,u)}$ could consist of at most two endpoints $\{0,u\}$. Similarly for $\Lambda_f^{(v,+\infty)}$. Therefore, we have the following lemma:

Lemma 3.12 *For any $f(s) \in \mathcal{E} - \mathring{F} - F_{isolated}$,*

1. *if $\Lambda_f^{(0,u)} \neq \emptyset$, then $\overline{\Lambda_f^{(0,u)}}$ is bounded, perfect and nowhere dense. Furthermore, $\overline{\Lambda_f^{(0,u)}} - \Lambda_f^{(0,u)} \subset \{0,u\}$.*

2. *if $\Lambda_f^{(v,+\infty)} \neq \emptyset$, then $\overline{r(\Lambda_f^{(v,+\infty)})}$ is bounded, perfect and nowhere dense. Furthermore, $\overline{r(\Lambda_f^{(v,+\infty)})} - r(\Lambda_f^{(v,+\infty)}) \subset \{0, \frac{1}{v}\}$. Here $r(x) = \frac{1}{x}$ is the reciprocal map.*

Therefore, after knowing that a bounded, nowhere dense, perfect set is homeomorphic to a cantor set $\mathcal{C}$ with endpoints matched up, we basically know how $\Lambda_f^{(0,u)}$ and $r(\Lambda_f^{(v,+\infty)})$ looks like [7] for any $f(s) \in \mathcal{E} - \mathring{F} - F_{isolated}$.

To get a better intuition of the shape of a bounded, nowhere dense, perfect set, we draw figure 2. In the left, we show the standard construction of a cantor set: first remove the middle third open interval from $[0,1]$; then remove the two middle third open intervals from what's left; repeat this infinitely times, and then take intersection. In subsection D, we show

[5] Here to take the closure, we take the following topology on extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty,+\infty\}$. The topology is generated by all open sets in $\mathbb{R}$ together with all sets in the form of $[-\infty,a)$ and $(b,+\infty]$. With this topology, if $\Lambda_f^{(v,+\infty)}$ contains a sequence of points diverging to $+\infty$, then $+\infty$ is in the closure of $\Lambda_f^{(v,+\infty)}$.

[6] Recall that a set on $\mathbb{R}$ is call perfect if it is closed and equals to the set of its own limit points

[7] Here reciprocal map $r(x)$ is a homeomorphism from $[0,+\infty]$ to $[0,+\infty]$.

Figure 2: A Standard Cantor Set (Left) and A Bounded Nowhere Dense Perfect Set (Right)

that any bounded, perfect, nowhere dense set S is obtained by removing infinitely disjoint open intervals from $[\inf S, \sup S]$. Therefore, such a set S can be constructed in a very similar way: first remove the longest open open interval from $[\inf S, \sup S]$, then remove the next longest two open intervals from what's left, repeat this infinitely times, then take intersection. Because the infinitely many open intervals in $[\inf S, \sup S]$ is not as well organized as the middle third open intervals that are removed in the construction of cantor set $\mathcal{C}$, the set S looks like a cantor set after a stretch.

Therefore, for any $f(s)$ in $\mathcal{E} - \mathring{F} - F_{isolated}$, its confounded learning set contains a part that looks like a stretch of a standard cantor set $\mathcal{C}$ with or without two endpoints. The next theorem writes it explicitly:

Theorem 3.13 *For any* $f(s) \in \mathcal{E} - \mathring{F} - F_{isolated}$

1. *if* $\Lambda_f^{(0,u)} \neq \emptyset$ $\Lambda_f^{(0,u)}$ *looks like cantor set* $\mathcal{C}$ *in the sense that it's either homeomorphic to* $\mathcal{C}$ *or to* $\mathcal{C}$ *without one or both endpoints.*
2. *if* $\Lambda_f^{(v,+\infty)} \neq \emptyset$*, then* $\Lambda_f^{(v,+\infty)}$ *looks like cantor set* $\mathcal{C}$ *in the sense that it's either homeomorphic to* $\mathcal{C}$ *or to* $\mathcal{C}$ *without one or both endpoints.*

In fact, given any parameter $(p, u, v, \underline{s}, \overline{s}) \subset U$, we consider the set of signals $F_{confounded}$ with non-empty confounded learning set. Obviously $F_{confounded} \supset \mathcal{E}$ and hence is topologically large. After removing two sets of first category, $F_{confounded} - \mathring{F} - F_{isolated}$ is still topologically large. For any signal $f(s)$ in $F_{confounded} - \mathring{F} - F_{isolated}$, its confounded learning set must either be a singleton on the monotonic region, or contains an (almost homeomorphic) cantor set outside the monotonic region.

3.6 Same Conclusion for C^∞ Signals

In the previous section, we have completed the most surprising statement of this article. Curious reader may ask whether this statement holds because we consider all signals that are merely integrable. An integrable function could behave peculiar itself, so maybe that is the reason for us to get models whose confounded learning sets look like cantor sets. In this section we state that such "cantor-like confounded learning set phenomenon" is robust in the sense that an almost the same conclusion holds even if we assume all signals have C^k density functions for all $0 \le k \le +\infty$.

In fact, all the previous discussion for integrable signals just directly extends to C^k signals. To avoid repeating almost the same reasoning, we directly state the conclusion. Proofs in the appendices are actually done for C^k ($0 \le k \le +\infty$) signals. We choose to separately present integrable signals and C^k signals so that readers can focus on the main reasoning.

Arbitrarily choose a parameter tuple $\{p, u, v, \underline{s}, \overline{s}\}$, we consider C^k signals

$$F^k_{(\underline{s},\overline{s})} = F_{(\underline{s},\overline{s})} \cap C^k(\underline{s}, \overline{s}). \tag{3.12}$$

Similar to proposition 3.3 and theorem 3.4, we find that the set

$$\mathring{F}^k = \{f(s) \in F^k_{(\underline{s},\overline{s})} | \Lambda_f \text{ contains an interior point}\}. \tag{3.13}$$

is of first category in F^k. Thus for any $f(s) \in F^k - \mathring{F}^k$, associated confounded learning set is nowhere dense.

Proposition 3.6 directly extends to C^k signals. Monotonic region exists for C^k signals as well. For any $f(s) \in F^k - \mathring{F}^k$, there could be at most one confounded learning belief on the monotonic region.

Theorem 3.7 extends as well. Let

$$F^k_{isolated} = \{f(s) \in F^k | G_f(\lambda) \text{ has an isolated zero point } \lambda^* \text{ outside the monotonic region}\}$$

be the set of C^k signals that admit an isolated confounded learning belief outside monotonic region, then $F^k_{isolated}$ is of first category in F^k.

Proposition 3.10 holds in C^k signals. First, we could talk about $\mathcal{E}^k$ to be the set of C^k-signals that admit essential confounded learning beliefs outside monotonic region. Since we actually explicitly construct C^∞ signals in $\mathcal{E}$ in the proof of 3.10, such construction directly shows the non-emptiness of $\mathcal{E}^k$ for any parameter tuple in the same open set U.

Then we know for any $f(s)$ in $\mathcal{E}^k - \mathring{F}^k - F^k_{isolated}$, its confounded learning set outside monotonic region is nowhere dense, non-empty and contains no isolated points. The discussion of shapes of those confounded learning sets is purely mathematical, and hence holds for C^k signals as well. To summarize, we have

Theorem 3.14 *Let U be the same open set as in proposition 3.10. For each parameter tuple in U, there exists a non-empty open set $\mathcal{E}^k$ and first category sets $\mathring{F}^k$ and $F^k_{isolated}$, such that for each $f(s) \in \mathcal{E}^k - \mathring{F}^k - F^k_{isolated}$, its confounded learning set outside monotonic region is non-empty, nowhere dense and contains no isolated point, and hence looks like a cantor set in the same sense as in theorem 3.13.*

3.7 Confounded Learning Set of a Real-analytic Signal is Discrete and Real-analytic Signals are Dense in the set of Integrable Signals

Knowing that the "cantor-like confounded learning set phenomenon" survives even if we assume that all private signals have C^∞ density functions, the next natural assumption is to assume that all private signals have real-analytic density functions.

Our first observation is that real-analytic signals have real-analytic characteristic functions.

Lemma 3.15 *If $f(s) \in F^\omega(\underline{s}, \overline{s})$, then $G_f(\lambda)$ is real-analytic on confounding region $\mathcal{B}$.*

Then we immediately recall the well-known fact that the zero set of an real-analytic function must be discrete. (See Corollary 1.2.7 in Krantz and Parks (2002)). Furthermore, if the confounding region $\mathcal{B}$ is bounded, then discreteness implies finiteness. So we have

Theorem 3.16 *For any real-analytic signal $f(s) \in F^\omega_{(\underline{s},\overline{s})}$, the associated confounded learning set Λ_f must be discrete. Furthermore, for $f(s) \in F^\omega_{(\underline{s},\overline{s})}$ with $0 < \underline{s} < \overline{s} < 1$, the confounded learning set Λ_f is finite.*

Despite that we would like to summarize the second part of the above theorem as that "bounded real-analytic" signals must have finite confounded learning set, we need to carefully distinguish the "real bounded" signals and "half bounded" signals. A real bounded signal satisfies $0 < \underline{s} < \overline{s} < 1$, so a player could not obtain an arbitrarily precise signal realization for each state A and B. A half bounded signal may have $\underline{s} = 0$ or $\overline{s} = 1$. For example,

$f(s) \in F_{(0.1,1)}$ is a half bounded signal. Its confounding region is not bounded and hence we can only conclude its confounded learning set is discrete.

The sharp difference between real-analytic signal and C^∞ signal is that our key construction in section 3.3 still works for C^∞ signals: given a C^∞-signal $f(s)$ with confounded learning λ^* outside monotonic region, we could perturb it slightly to obtain a nearby C^∞-signal $\hat{f}(s)$ such that its confounded learning set has λ^* as a limit point. Such perturbation can never work for real-analytic signals for that there is no real-analytic signal could have a limit point in its zero set.

Knowing that the real-analytic assumption removes the peculiar behavior of confounded learning set, it is natural to ask how restrictive this assumption is. Actually, this assumption is almost harmless given that

Theorem 3.17 *That F^ω is dense in F, and hence is dense in F^k ($0 \leq k \leq +\infty$) as well.*

The above result relies on Whitney's approximation theorems and Lusin's theorem. The real-analytic approximation theorem implies that $C^\omega(\underline{s}, \overline{s})$ is dense in $C^\infty(\underline{s}, \overline{s})$, and the smooth approximation theorem implies that $C^\infty(\underline{s}, \overline{s})$ is dense in $C^0(\underline{s}, \overline{s})$. (See theorem 1.6.5 in Narasimhan (1968) and theorem 6.1.5 in Lee (2000)). Lusin's theorem implies that $C^0(\underline{s}, \overline{s})$ is dense in $L_1(\underline{s}, \overline{s})$. Therefore, for any signal $f(s)$, we could always find a sequence of real-analytic functions approximate it. In the appendix E, we show how to construct a sequence of real-analytic signals converging to $f(s)$ based on the existing convergent sequence of real-analytic functions.

4 Conclusion

In this paper we completely characterize the set of confounded learning beliefs in a class observational learning environment. Our primary finding is that the confounded learning set could contain a (homeomorphic) cantor set. We demonstrate that this finding is robust in the sense that models have this peculiar phenomenon is topologically large in some sense. Besides, we also find that such peculiar phenomenon survives even if we consider models whose private signals have C^∞ density functions.

We provide a solution to such peculiar phenomenon. As long as we only consider models whose private signals have real-analytic density function, confounded learning set must be discrete. For those models whose confounding region is bounded, we can further conclude their confounded learning sets are finite. We also show that it is not restrictive to only consider real-analytic signals, for the reason that the set of real-analytic signals is dense among all signals. From this perspective, we recommend researchers to restrict their attention to real-analytic private signals when they draw generic conclusion on confounded learning beliefs.

Besides, we also provide some minor results which may be of independent interest. We show that confounded learning set must be nowhere dense for almost all models. We also find the existence of monotonic region, on which a model could have at most one confounded learning belief.

A Omitted Proofs in Subsection 3.1

In this section we prove proposition 3.2. This completes the proof of theorem 3.4 that almost all private signals have nowhere dense confounded learning belief set.

Proof of proposition 3.2. We need to show that $\overline{F_I}^F$-the closure of F_I in F-contains no interior point.

First we observe that $\overline{F_I}^F = F$. This follows directly from lemma 3.8 that

$$\|f_k - f\| \to 0 \Rightarrow |G_f(\lambda) - G_{f_k}(\lambda)| \to 0,\ \forall \lambda \in (0, +\infty). \tag{A.1}$$

Now assume that F_I has an interior point $x_0 \in F_I$. Let $f^\omega(s)$ be any real-analytic private signal. We consider the sequence

$$x_k(s) = \frac{x_0 + \frac{1}{k} f^\omega}{\int_0^1 x_0(s) + \frac{1}{k} f^\omega(s) ds}. \tag{A.2}$$

Since f^ω is real-analytic, its confounded learning set must be discrete and cannot contain interval I. As a result, the confounded learning set of x_k cannot contain interval I. So $\{x_k(s)\} \subset F - F_I$, but we can easily see that $\|x_k - x_0\|_{L_1} \to 0$. This contradicts that x_0 is interior. ■

B Omitted Proofs in Subsection 3.2

In this section, we aims at showing that $F_{isolated}$ in theorem 3.7 is of first category. This is the most important and complex construction in this article. Thus we divide the proof into several subsections. We start with an outline, and then fill each part in. Besides, in this section, we don't keep the assumption that $v > u$ or private signal is of unbounded strength. The following constructions work for all these assumptions, but sometimes we need do some technique discussion to deal with different assumptions.

B.1 Outline of Proof of Theorem 3.7

To show $F_{isolated}$ is of first category, we turn to show that it is a subset of countable union of nowhere dense sets. Here and throughout this section, whenever we mention an open interval (z_1, z_2), we keep the assumption that $(z_1, z_2) \subset [0, +\infty]$ and $(z_1, z_2) \cap [u, v] = \emptyset$. It is easy to verify that

$$F_{isolated} \subset \cup_{z_1, z_2 \in \mathbb{Q}} F_{(z_1, z_2)} \tag{B.1}$$

where

$$F_{(z_1, z_2)} = \{f \in F | \exists \text{ unique } \lambda^* \in (z_1, z_2) \text{ s.t. } G_f(\lambda^*) = 0\}. \tag{B.2}$$

We turn to show

Theorem B.1 *That $F_{(z_1,z_2)}$ is nowhere dense in F. And $F_{(z_1,z_2)} \cap F^k, 0 \leq k \leq +\infty$ is nowhere dense in F^k.*

We use the following characterization for nowhere dense subset.

Definition B.2 *Let (T, τ) be a topology space. Then $S \subset T$ is nowhere dense in T if*

- *$\forall x \in S$ and $\forall$ open set O containing x, there exists a smaller open set $O' \subset O$ and $O' \cap S = \emptyset$*

This definition reflects that a nowhere dense set S is "full of holes". Any small open set O cannot be contained in S. For any point in this small open set but misses S, we could find an open hole around this point so that this hole completely misses S.

To apply definition, we start with any $f(s) \in F_{(z_1,z_2)}$ with isolated confounded learning belief $\lambda^* \in (z_1, z_2)$. We construct a sequence of private signals $\hat{f}_n \in F$ such that

- $\|f - \hat{f}_n\|_{L_1} \to 0$.
- For any n, the characteristic function $G_{\hat{f}_n}(\lambda)$ oscillates on $[\lambda^*, \lambda^* + \varepsilon_n]$ for a small ε_n. And sequence ε_n vanishes.

The second point implies that $G_{\hat{f}_n}(\lambda)$ doesn't have an unique zero on (z_1, z_2). Then for any small open ball containing $f(s)$, it contains some $\hat{f}_n \notin F_{(z_1,z_2)}$. Pick the $\hat{f}_n$ with the smallest index n, and consider its characteristic function $G_{\hat{f}_n}(\lambda)$. Since this characteristic funcion infinitely changes sign on $[\lambda^*, \lambda^*+\varepsilon_n]$, we could find a sequence $\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_k > \lambda^*$ such that

$$G_{\hat{f}_n}(\lambda_1) > 0, G_{\hat{f}_n}(\lambda_2) < 0, G_{\hat{f}_n}(\lambda_3) > 0 \dots \tag{B.3}$$

Since

$$\forall \lambda > 0, \|f - g\|_{L_1} \to 0 \Rightarrow |G_f(\lambda) - G_g(\lambda)| \to 0, \tag{B.4}$$

we could always find a sufficiently small open ball around $\hat{f}_n$ so that any g in this open ball satisfying

$$G_g(\lambda_1) > 0, G_g(\lambda_2) < 0, G_g(\lambda_3) > 0. \tag{B.5}$$

Then g has at least two essential confounded learning beliefs in (z_1, z_2). So this open ball around $\hat{f}_n$ doesn't intersect $F_{(z_1,z_2)}$. Then we prove that $F_{(z_1,z_2)}$ is nowhere dense in F.

The proof of that $F^k_{(z_1,z_2)}$ is nowhere dense in F^k follows the exactly same steps.

Now we describe the main steps in construction of $\{\hat{f}_n(s)\}_{n\in\mathbb{N}}$. We first explicitly list out conditions that $\hat{f}_n(s)$ should satisfy:

Condition B.3

1. *$\|f(s) - \hat{f}_n(s)\|_{L_1} \to 0$.*
2. *If $\lambda^* \in (z_1, z_2)$ is the unique confounded learning belief of $f(s)$ on (z_1, z_2), then $G_{\hat{f}_n}(\lambda)$ oscillate on $(\lambda^*, \lambda^* + \varepsilon_n)$, and $\varepsilon_n \searrow 0$.*
3. *(Density Property) $\int_{\underline{s}}^{\bar{s}} \hat{f}_n(s)ds = 1$; $\hat{f}_n(s) \geq 0$; (Signal Property) $\int_{\underline{s}}^{\bar{s}} \hat{f}_n(s)\frac{1-2s}{s}ds = 0$.*
4. *(Smoothness Condition) Given $f(s)$, $\hat{f}_n(s)$ must has a same or stronger smoothness assumption than $f(s)$.*

Here condition (3) guarantees that $\hat{f}_n \in F$. We refer to $\int_{\underline{s}}^{\bar{s}} \hat{f}_n(s)ds = 1, \hat{f}_n(s) \geq 0$ as that $\hat{f}_n(s)$ satisfying the density property. These are the basic conditions that a density function should satisfy. Besides, we refer to $\int_{\underline{s}}^{\bar{s}} \hat{f}_n(s)\frac{1-2s}{s}ds = 0$ as that $\hat{f}_n$ satisfies the signal property. A density function must satisfy this condition to give a pair $\left(f^A(s), f^B(s)\right)$ in an observational learning model. We give these properties names since it will be frequently mentioned in the construction.

To construct such a $\hat{f}_n(s)$, we go through the following steps:

1. We construct a function $g(\lambda)$ which could has infinitely many sign changes near the unique confounded learning belief $\lambda^* \in (z_1, z_2)$ of a given $f(s)$.

2. Vaguely speaking, we perturb the definition of $f(s)$ on a small set I_n (perturbed region) so that the perturbed characteristic function agrees with $g(\lambda)$ near λ^*.

3. If $f(s)$ is k-th continuously differentiable, after $f(s)$ has been perturbed as in step (2), the resulted function may be no longer in C_k. To solve this problem, we use the smooth gluing lemma. The smooth gluing lemma allows us to construct a C^∞ function on any closed interval with (almost) arbitrary endpoints behavior. We use it to construct a function on a small "smoothing region" which smoothly connect the $f(s)$ on perturbed and unperturbed regions.

4. After $f(s)$ has been perturbed in step (2) and smoothly glued in step (3), the resulted function is often not a private signal. We are careful in step (2) and (3) so that the non-negativity is not destroyed. But often the resulted function doesn't integrate to 1, and fails to satisfy the "signal property". The failure of signal property is more significant. To fix it, We shall find a small set-rebalance region-and add some weight on that region so that the signal property is restored. We also need to be careful about this modification so that we don't ruin the smoothness again. After the signal property is satisfied, the resulted function is a positive multiple of a private signal, we just divide a proper number to fix the "density property".

Given original $f(s)$, after all those modifications made in steps (2),(3) and (4), we obtain the desired $\hat{f}_n(s)$. By shrinking the size of the set on which the definition of $f(s)$ is changed, we could make $f(s)$ and $\hat{f}_n(s)$ arbitrarily close.

We remark that the above construction is related to Benavides (1986). However, our construction goes much deeper. First, we must perturb a density function to obtain an oscillating characteristic function. In Benavides (1986), he could directly perturb the characteristic function. Second, we must make sure the function after perturbation is still a density function. Third, we consider the perturbation up to C^∞ density functions while Benavides (1986) only considers continuous functions.

The following subsections are organized as following: from subsections B.2 to B.5 we construct the function sequence $\{\hat{f}_k(s)\}$ as in the above steps. In subsection B.6, we prove the constructed sequence converging to $f(s)$. In section B.7, we include all the details, most are non-obvious computation and constructions under different parameter assumptions.

B.2 A (Oscillation) function that infinitely changes sign near one point

For abbreviation, we refer to a function $f(s)$ which infinitely changes sign near one point λ^* as " $f(s)$ oscillates near λ^*". A function that oscillates is a oscillation function.

An immediate candidate for a oscillation function is $\sin(\frac{1}{\lambda-\lambda^*})$. But this function is not differentiable at λ^*. So we may want to consider $e^{-\frac{1}{\lambda-\lambda^*}}\sin(\frac{1}{\lambda-\lambda^*})$, which is C^∞ at λ^*. However, if we would perturb $f(s)$ to $\hat{f}(s)$ so that

$$G_{\hat{f}}(\lambda^*) = e^{\frac{1}{\lambda-\lambda^*}}\sin(\frac{1}{\lambda-\lambda^*}), \lambda \in (\lambda^*, \lambda^*+\varepsilon); \tag{B.6}$$

we shall see so-defined $\hat{f}(s)$ cannot be differentiable. The detailed reason shall be clear in the next subsection. Therefore, if we would like to consider private signals with certain smoothness assumption, we need to find an oscillation function that is *smoother* than $e^{-\frac{1}{x-\lambda^*}}\sin(\frac{1}{x-\lambda^*})$.

We consider the following function

$$g_k(\lambda) = \int_{c_k}^{\frac{1}{\lambda-\lambda^*}} \hat{h}(t)dt, \lambda \in (\lambda^*, \lambda^*+\frac{1}{c_k}), c_k > 0 \text{ to be specified} \tag{B.7}$$

where the integrand $\hat{h}(t)$ is given by

$$\hat{h}(t) = \begin{cases} e^{-(\sin t)^{-2}-t}\sin t, t \neq k\pi \text{ and } t > 0; \\ 0, t = k\pi, k \in \mathbb{N}. \end{cases} \tag{B.8}$$

[8] Let us briefly describes the motivation behind so-designed $g_k(\lambda)$. First, $\hat{h}(t)$ infinitely changes signs on $(0, +\infty)$ just like $\sin t$. If we consider the indefinite integral $\hat{g}(x) = \int_0^x \hat{h}(t)dt$, it is easy to see that $\hat{g}(x)$ goes up on regions where $\hat{h}(t) > 0$, and goes down on regions where

[8] In figure 3.3, we graph $g_k(\lambda)$ with the $\hat{h}(t) = e^{-(\sin\frac{t}{2})^{-2}-\frac{t}{100}}\sin\frac{t}{2}, t \neq k\pi$. We add some parameters so that the graph looks better. Actually, if we don't scale, the amplitude of oscillation function decreases much faster,and period of oscillation is smaller as well. We only scale $\hat{h}(t)$ in figures 3.3 and B.2.

$\hat{h}(t) < 0$. Since $\hat{h}(t)$ infinitely changes sign, $\hat{g}(x)$ goes up and down infinitely times. In other word, $\hat{g}(x)$ has a wave shape. Besides, since e^{-t} is included in $\hat{h}(t)$, $\hat{h}(t)$ vanishes smoothly at $t = +\infty$. So the amplitude of $\hat{g}(x)$ vanishes as $x \to +\infty$. The vanishing of $\hat{g}(x)$ implies that $\lim_{x\to\infty} \hat{g}(x)$ exists. Then $\overline{g}(x) = \hat{g}(x) - \lim_{x\to\infty} \hat{g}(x)$ infinitely changes sign on $(0, +\infty)$. The c_k in the definition of $g_k(\lambda)$ are those points where $\overline{g}(c_k) = 0$. Then we can check that: as x increases from c_k, $\overline{g}_k(x) \equiv \int_{c_k}^{x} \hat{h}(t)dt$ start from 0, infinitely changes sign like a wave, and amplitudes vanishes to 0. After a change of variable, we see $g_k(\lambda)$ is just $\overline{g}_k(\frac{1}{\lambda-\lambda^*})$. All the sign changes of $\overline{g}_k(x)$ on $x \in (c_k, +\infty)$ is now compressed to sign changes of $g_k(\lambda)$ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$.

To summarize, we have

Proposition B.4 *1. $g_k(\lambda)$ oscillates near λ^*.*

2. That $g_k(\lambda)$ takes value 0 at both endpoints of its domain $(\lambda^*, \lambda^*+\frac{1}{c_k})$: (1) $\lim_{\lambda\to\lambda^*} g_k(\lambda) = 0$ and (2) for each $k \in \mathbb{N}$, there exists $c_k \in ((k-1)\pi, k\pi)$ so that $g_k(\lambda^* + \frac{1}{c_k}) = 0$.

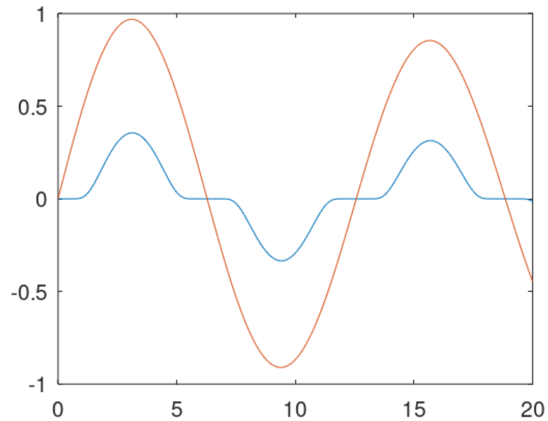
Interesting readers could read B.7.1 for the proof of proposition B.4.

Sharp reader may have realized that the above conclusion still hold if we let integrand $\hat{h}(t)$ just be $e^{-t}\sin t$. The extra $e^{-(\sin t)^{-2}}$ is there to guarantee that so-designed $\hat{h}(t)$ crosses 0 smoothly. Here by saying $e^{-t}\sin t$ doesn't cross 0 as smooth as $\hat{h}(t)$ does, we don't mean that $e^{-t}\sin t$ has a kink whenever it crosses 0. We mean $e^{-t}\sin t$ crosses 0 in a "steeper" way than $\hat{h}(t)$ does. Mathematically, the difference is that $\hat{h}_+(t)$ and $\hat{h}_-(t)$ are C^∞; but the positive part and negative part of $e^{-t}\sin t$ has kinks at all $\{k\pi\}$, which are the points where $e^{-t}\sin t$ crosses 0. In fact, at $t = k\pi$, the extra term $e^{-(\sin t)^{-2}}$ vanishes super fast as t approaches $k\pi$. The result is that $\hat{h}(t)$ is much more flatter than $e^{-t}\sin t$ near $t = k\pi$. In figure B.2, we graphed $\hat{h}(t)$ and $e^{-t}\sin(t)$. [9]

In almost the same sense, we say the so-designed $g_k(\lambda)$ is smoother than the function $e^{-\frac{1}{\lambda-\lambda^*}}\sin(\frac{1}{\lambda-\lambda^*})$, which is another candidate for the oscillation function. The next proposition states that the derivative $g_k^{(1)}(\lambda)$ of $g_k(\lambda)$ crosses 0 in a "flatter" way so that its positive and negative parts are C^∞ on its domain $(\lambda^*, \lambda^* + \frac{1}{c_k})$. The derivative of $e^{-\frac{1}{\lambda-\lambda^*}}\sin(\frac{1}{\lambda-\lambda^*})$ doesn't have that property.

Proposition B.5 *We can compute the derivative of $g_k(\lambda)$ as $g_k^{(1)}(\lambda) = -(\lambda-\lambda^*)^{-2}\hat{h}(\frac{1}{\lambda-\lambda^*})$. Its positive part $g_{k+}^{(1)}(\lambda) = \max\{g_k'(\lambda), 0\}$ and negative part $g_{k-}^{(1)}(\lambda) = -\min\{g_k'(\lambda), 0\}$ are C^∞*

[9] As in figure 3.3, we actually graphed $\hat{h}(t) = e^{-(\sin\frac{t}{2})^{-2}-\frac{t}{100}}\sin\frac{t}{2}, t \neq k\pi$ and $e^{-\frac{t}{100}}\sin(\frac{t}{2})$ to get a better figure.



in λ. [10]

Detailed proof is provided in B.7.1.

Therefore, we have constructed a sequence of oscillation functions $g_k(\lambda)$.

Recall that in the construction of $\hat{f}_k(s)$, we have selected an open interval (z_1, z_2) and a $f(s) \in F_{(z_1,z_2)}$ (or $F^k_{(z_1,z_2)}$). The unique confounded learning λ^* is determined by $f(s)$. And $g_k(\lambda)$ depends on λ^*. From now on, we automatically assume that c_k are bigger enough so that $(\lambda^*, \lambda^* + \frac{1}{c_k})$ is outside of $[u, v]$ or $[v, u]$.

B.3 Perturb to Oscillate

In this section, we would like to perturb the definition of $f(s)$ on a small set so that the characteristic function after perturbation is multiple of $g_k(\lambda)$ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$.

Recall that $f(s) \in F_{(z_1,z_2)}$ (or $F^k_{(z_1,z_2)}$) with pre-selected (z_1, z_2) that doesn't intersect $[u, v]$ or $[v, u]$. The construction is slightly different for (1) $u, v > z_1, z_2$ and (2) $u, v < z_1, z_2$. For presentation simplicity, we shall assume assumption $u, v > z_1, z_2$ until claim B.8. Interesting readers could read the appendix B.7.2 for discussions of other assumptions.

Given $h(s)$ to be a positive multiple of a private signal, we would like to have

$$G_h(\lambda) - G_h(\lambda^*) = g_k(\lambda) \text{ on } (\lambda^*, \lambda^* + \frac{1}{c_k}). \tag{B.9}$$

[10]Here we follow the standard notation in real analysis, so the "negative" part is not really negative.

By the definition of characteristic function $G_h(\lambda)$: $\forall\lambda \in (\lambda^*, \lambda^* + \frac{1}{c_k})$

$$G_h(\lambda) - G_h(\lambda^*) = p\int_{\frac{\lambda^*}{\lambda^*+u}}^{\frac{\lambda}{\lambda+u}} h(s)\frac{1-2s}{s}ds - (1-p)\int_{\frac{\lambda^*}{\lambda^*+v}}^{\frac{\lambda}{\lambda+v}} h(s)\frac{1-2s}{s}ds. \tag{B.10}$$

So if we differentiate both sides of equation B.9, we obtain

$$h(\frac{\lambda}{\lambda+u})p\overbrace{\frac{u-\lambda}{\lambda}\frac{u}{(\lambda+u)^2}}^{t_1(\lambda)} - h(\frac{\lambda}{\lambda+v})\overbrace{(1-p)\frac{v-\lambda}{\lambda}\frac{v}{(\lambda+v)^2}}^{t_2(\lambda)} = -(\lambda-\lambda^*)^{-2}\hat{h}(\frac{1}{\lambda-\lambda^*}) \tag{B.11}$$

On the left-hand side of B.11, we see that $G_h'(\lambda)$ is completely determined by $h(s)$'s values on $(\frac{\lambda^*}{\lambda^*+u}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+u}) \cup (\frac{\lambda^*}{\lambda^*+v}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+v})$, which are disjoint for small enough $\frac{1}{c_k}$. Besides, since we assume that $u, v > \lambda^* + \frac{1}{c_k}$, we have

$$h(\frac{\lambda}{\lambda+u})t_1(\lambda) \geq 0, -h(\frac{\lambda}{\lambda+v})t_2(\lambda) \leq 0 \text{ on } (\lambda^*, \lambda^* + \frac{1}{c_k}). \tag{B.12}$$

One important observation is that, we can separate $G_h'(\lambda)$ into two parts: the first part $h(\frac{\lambda}{\lambda+u})t_1(\lambda)$ is non-negative, and is completely determined by $h(s)$ on $(\frac{\lambda^*}{\lambda^*+u}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+u})$; the second part $-h(\frac{\lambda}{\lambda+v})t_2(\lambda)$ is non-positive, and is determined by $h(s)$ on $(\frac{\lambda^*}{\lambda^*+v}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+v})$. On the right-hand side, $g_k'(\lambda)$ oscillates as $g_k(\lambda)$ does

$$g_k'(\lambda)\begin{cases} \geq 0 \text{ on } \Lambda^+ = \cup_{n\geq n_0}[\lambda^* + \frac{1}{2n\pi}, \lambda^* + \frac{1}{(2n-1)\pi}]; \\ < 0 \text{ on } \Lambda^- = \cup_{n\geq n_0}(\lambda^* + \frac{1}{(2n+1)\pi}, \lambda^* + \frac{1}{2n\pi}) \cup (\lambda^* + \frac{1}{k\pi}, \lambda^* + \frac{1}{c_k}). \end{cases}$$

Here we assume that k is odd. Interesting readers can verify that $(\lambda^* + \frac{1}{k\pi}, \lambda^* + \frac{1}{c_k}) \subset \Lambda^+$ if we assumes k is even.

We shall match $G_h'(\lambda) = g_k'(\lambda)$ by "matching the positive with positive; matching the negative with the negative":

Definition B.6 *Let $l_u(\lambda) = \frac{\lambda}{\lambda+u}$. On $l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \equiv (\frac{\lambda^*}{\lambda^*+u}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+u})$,*

$$h(s)t_1(l_u^{-1}(s)) = g_{k+}'(l_u^{-1}(s)) \tag{B.13}$$

Let $l_v(\lambda) = \frac{\lambda}{\lambda+v}$*. On* $l_v(\lambda^*, \lambda^* + \frac{1}{c_k}) \equiv (\frac{\lambda^*}{\lambda^*+v}, \frac{\lambda^*+\frac{1}{c_k}}{\lambda^*+\frac{1}{c_k}+v})$*,*

$$h(s)t_2(l_v^{-1}(s)) = g'_{k-}(l_v^{-1}(s)).^{11} \tag{B.14}$$

Now it is quite clear that we must have proposition B.5. By making $g'_{k+}(\lambda)$, $g'_{k-}(\lambda)$ be C^∞, we could actually expect $h(s)$ in definition B.6 to be C^∞.

The following proposition states that $h(s)$ defined as in definition B.6 has lots of good properties, including being C^∞.

Proposition B.7 *1. If* $G_h(\lambda^*) = 0$*, then the associated characteristic function* $G_h(\lambda)$ *oscillates near* λ^**:*

$$G_h(\lambda) - G_h(\lambda^*) = g_k(\lambda) \text{ on } (\lambda^*, \lambda^* + \frac{1}{c_k}). \tag{B.15}$$

2. $h(s)$ *is* C^∞ *on the perturbed region* $l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k})$*;*

3. That $h(s)$ *is flat near* $l_u(\lambda^*)$ *and* $l_v(\lambda^*)$*:*

$$\lim_{s \to l_u(\lambda^*)^+} h^{(n)}(s) = 0 \quad ; \quad \lim_{s \to l_v(\lambda^*)^+} h^{(n)}(s) = 0; \forall n \in \mathbb{N}. \tag{B.16}$$

Here we summarize all the properties of $h(s)$ that is necessary in the entire section. The above construction of $h(s)$ should have made property (1) and (2) convincing. Interesting readers could find detailed proof in appendix B.7.2. If we think about the behavior of $h(s)$ near $l_u(\lambda^*)$, property (3) is intuitive as well. The region $\Lambda^- \subset (\lambda^*, \lambda^* + \frac{1}{c_k})$, on which $g'_{k+}(\lambda) \le 0$, is mapped by l_u to infinitely many disjoint open intervals near $l_u(\lambda^*)$. On these intervals, $h(s) = 0$ by the definition. Therefore, $\lim_{s \to l_u(\lambda^*)^+} h^{(n)}(s) = 0$ is the only possibility.

Recall that we would like to perturb $f(s)$ on the perturbed region so that the resulted characteristic function is a multiple of $g_k(\lambda)$. We choose the multiple r according to

$$\int_{l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k})} r h(s) \frac{1-2s}{s} ds = \int_{l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k})} f(s) \frac{1-2s}{s} ds. \tag{B.17}$$

This multiple r guarantees that the signal property ((3) in B.3) is not destroyed on the perturbed region. Therefore, for each k, on perturbed region $l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k})$,

[11] Despite that $h(s)t_2(l_v^{-1}(s)) \ge 0$, the negative part actually ≥ 0, as in the standard defition.

we have

$$\hat{f}_k(s) = r_k h(s), \tag{B.18}$$

where

$$r_k = \frac{\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})\cup l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds}{\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})\cup l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} h(s)\frac{1-2s}{s}ds}. \tag{B.19}$$

From the construction of $h(s)$, it is immediate that the denominator is not 0. So r_k is well-defined. It is also important to check that $r_k \neq 0$. Otherwise, $G_{\hat{f}_k}(\lambda) - G_{\hat{f}_k}(\lambda^*) = r_k[G_h(\lambda) - G_h(\lambda^*)] = 0$ doesn't provide the oscillation.

We claim this is true:

Claim B.8 *If there exists an open interval $(\lambda^*, \lambda^* + \varepsilon)$ such that $G_f(\lambda)$ is either strictly positive or strictly negative on it, then*

$$\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})\cup l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds \neq 0. \tag{B.20}$$

Therefore, since λ^* is the unique essential confounded learning belief on (z_1, z_2), r_k in B.19 is not zero.

Proof. Recall that

$$G_f(\lambda^* + \frac{1}{c_k}) - \overbrace{G_f(\lambda^*)}^{0} = p\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds - (1-p)\int_{l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds.$$

If $u, v > \lambda^*+\frac{1}{c_k}$, we have $l_u(\lambda^*, \lambda^*+\frac{1}{c_k}), l_v(\lambda^*, \lambda^*+\frac{1}{c_k}) \subset (0, \frac{1}{2})$. So the integrand $f(s)\frac{1-2s}{s} \geq 0$. Therefore, we have

$$\begin{aligned} G_f(\lambda^* + \frac{1}{c_k}) > 0 \;\; &\Rightarrow \;\; \int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds > 0; \\ G_f(\lambda^* + \frac{1}{c_k}) < 0 \;\; &\Rightarrow \;\; \int_{l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds > 0. \end{aligned} \tag{B.21}$$

Therefore,

$$\begin{aligned}
&\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})\cup l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds \\
\geq\ & \max\Big\{\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds, \int_{l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds\Big\} > 0. \qquad \text{(B.22)}
\end{aligned}$$

If $u, v < \lambda^*+\frac{1}{c_k}$, we have $l_u(\lambda^*, \lambda^*+\frac{1}{c_k}), l_v(\lambda^*, \lambda^*+\frac{1}{c_k}) \subset (\frac{1}{2}, 1)$. So the integrand $f(s)\frac{1-2s}{s} \leq 0$. Then similarly

$$\begin{aligned}
G_f(\lambda^* + \frac{1}{c_k}) > 0 \ &\Rightarrow \ \int_{l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds < 0; \\
G_f(\lambda^* + \frac{1}{c_k}) < 0 \ &\Rightarrow \ \int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds < 0. \qquad \text{(B.23)}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})\cup l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds \\
\leq\ & \min\Big\{\int_{l_u(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds, \int_{l_v(\lambda^*,\lambda^*+\frac{1}{c_k})} f(s)\frac{1-2s}{s}ds\Big\} < 0. \qquad \text{(B.24)}
\end{aligned}$$

■

Note that when the numerator of r_k is negative, the denominator of r_k is negative as well. So r_k is still positive.

To summarize, proposition B.7 (1) establishes that we could properly perturb $f(s)$ on a small region so that the perturbed characteristic function oscillates near λ^* provided that $G_{\hat{f}}(\lambda^*) = 0$. This is the main point of this subsection. The other two points in proposition B.7 will be used in the next subsection.

B.4 Smooth Gluing

Up to now, for any $f(s) \in F_{(z_1,z_2)}$ (or $F^k_{(z_1,z_2)}$), we replace $f(s)$ on two small intervals by a smooth function $rh(s), r > 0$. If we denote

$$f_{1,k}(s) = \begin{cases} rh(s), \text{ on } l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k}); \\ f(s), \text{ otherwise }. \end{cases}$$

then $G_{f_{1,k}}(\lambda) = rg_k(\lambda)$ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$. However, so-defined $f_{1,k}(s)$ needs not to be continuous at endpoints of the perturbed region. This contradicts the smoothness condition in condition B.3 if we start with a signal $f(s) \in C^k$. We shall fix this by smoothly glue the perturbed and unperturbed part.

The driving force is the following smooth gluing lemma:

Lemma B.9 (Smooth Gluing Lemma) *Given any two sequences $\{a_n\}, \{b_n\}$ such that one of the following three conditions is satisfied: (1) $a_0 \in (0, +\infty)$ and $b_0 \in (0, +\infty)$; (2) $b_0 \in (0, +\infty)$ and $a_n = 0$ for all n; (3) $a_0 \in (0, +\infty)$ and $b_n = 0$ for all n. Then for*

1. *any closed interval $[s_1, s_2]$;*
2. *any function $p(s) \in C^\infty[s_1, s_2]$ with $p(s) > 0$;*
3. *any constant $c > 0$;*

there exists a non-negative function $k(s) \in C^\infty[s_1, s_2]$ such that

1. *$k^{(n)}(s_1) = a_n$, $k^{(n)}(s_2) = b_n$;*
2. *$\int_{s_1}^{s_2} k(s)p(s)dx = c(s_2 - s_1)$.*

Roughly speaking, the smooth gluing lemma allows us to have a smooth function $k(s)$ on any closed interval $[s_1, s_2]$ with almost arbitrary endpoint behavior and almost arbitrary integral. Interesting readers could read appendix **??** for its proof.

Now we will show how to apply the smooth gluing lemma. Assuming that we start with a signal $f(s) \in C^k$, and taking $l_u(\lambda^*)$ as an example. We first exclude a special case where $f_{1,k}(s)$ is automatically of C^k at $l_u(\lambda^*)$. Recall that proposition B.7 (3) states that the perturbed $f_{1,k}(s)$ is flat to the right of $l_u(\lambda^*)$ in the sense that

$$\lim_{s \to l_u(\lambda^*)^+} f_{1,k}^{(n)}(s) = r^n \lim_{s \to l_u(\lambda^*)^+} h^{(n)}(s) = 0, \forall 0 \leq n \leq +\infty. \tag{B.25}$$

If $f(s) = 0$ on a small interval $(l_u(\lambda^*) - \varepsilon, l_u(\lambda^*))$, then

$$\lim_{s \to l_u(\lambda^*)^-} f_{1,k}^{(n)}(s) = \lim_{s \to l_u(\lambda^*)^-} f^{(n)}(s) = 0, 0 \leq n \leq k. \tag{B.26}$$

In this particular case, the perturbed $f_{1,k}(s)$ and $f(s)$ are automatically smoothly connected at $l_u(\lambda^*)$, and there is no problem need to be fixed at $l_u(\lambda^*)$.

In general, this particular case doesn't arise. If the given $f(s)$ is not flat to the left of $l_u(\lambda^*)$, then there is a sequence $\{\varepsilon_k\}$ satisfying

$$\lim_{k\to+\infty}\varepsilon_k = 0 \quad ; \quad f(l_u(\lambda^* - \varepsilon_k)) > 0, \forall k. \tag{B.27}$$

We shall define another function $t(s)$ on $l_u(\lambda^* - \varepsilon_k, \lambda^*)$ so that

$$\begin{cases} f(s), s \in (\underline{s}, l_u(\lambda^* - \varepsilon_k)); \\ t(s), s \in l_u(\lambda^* - \varepsilon_k, \lambda^*); \\ rh(s), s \in l_u(\lambda^*, \lambda^* + \frac{1}{c_k}). \end{cases} \tag{B.28}$$

is C^k and that

$$\int_{l_u(\lambda^*-\varepsilon_k,\lambda^*)} f(s)\frac{1-2s}{s}ds = \int_{l_u(\lambda^*-\varepsilon_k,\lambda^*)} t(s)\frac{1-2s}{s}ds. \tag{B.29}$$

To obtain such a $t(s)$, we apply the smooth gluing lemma. Just let $l_u(\lambda^* - \varepsilon_k)$ be s_1, $l_u(\lambda^*)$ be s_2. Let $b_n = 0$ for all $0 \le n \le +\infty$, $a_n = f^{(n)}(s_1), 0 \le n \le k$ and $a_n = 0, n > k$. Let

$$p(s) = \begin{cases} \frac{1-2s}{s}, \text{ if } [s_1, s_2] \subset (\frac{1}{2}, \overline{s}); \\ \frac{2s-1}{s}, \text{ if } [s_1, s_2] \subset (\underline{s}, \frac{1}{2}). \end{cases}$$

Then we can check that the smooth gluing lemma guarantees that a C^∞ function $t(s)$ exists so that equations B.28 and B.29 holds. We could see that such a $t(s)$ smoothly glue the perturbed region $l_u(\lambda^*, \lambda^* + \frac{1}{c_k})$ and the unperturbed region. Lastly, we remark that equation B.29 is included for the same reason as equation B.17 is used. We try not to destroy the signal property as in condition B.3.

Similarly, we could use the smooth gluing lemma to restore the smoothness around $l_v(\lambda^*), l_u(\lambda^* + \frac{1}{c_k}), l_v(\lambda^* + \frac{1}{c_k})$.

B.5 Rebalance

Let us summarize: up to now, for any $f(s) \in F^k_{(z_1,z_2)}$, we perturb it on the perturbed region and then smoothly glue the perturbed function back:

$$f_{2,k}(s) = \begin{cases} rh(s); s \in l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k}); \\ t(s); s \in l_u(\lambda^* - \varepsilon_k, \lambda^*) \cup l_u(\lambda^* + \frac{1}{c_k}, \lambda^* + \frac{1}{c_k} + \varepsilon_k) \cup l_v(\lambda^* - \varepsilon_k, \lambda^*) \cup l_v(\lambda^* + \frac{1}{c_k}, \lambda^* + \frac{1}{c_k} + \varepsilon_k) \\ f(s);\ \text{otherwise}\ . \end{cases}$$

If we consider $f(s) \in F_{(z_1,z_2)}$, then we don't need to perform the smooth gluing. Then

$$f_{2,k}(s) = \begin{cases} rh(s); s \in l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k}); \\ f(s);\ \text{otherwise}\ . \end{cases}$$

Cautious readers may have realized that we cannot directly use $f_{2,k}(s)$ as $\hat{f}(s)$. This is because: despite that

$$G_{f_{2,k}}(\lambda) - G_{f_{2,k}}(\lambda^*) = r_k g_k(\lambda), \lambda \in (\lambda^*, \lambda^* + \frac{1}{c_k}), \tag{B.30}$$

there is no guarantee that $G_{f_{2,k}}(\lambda^*) = 0$. In fact, using equation B.29, we could check that

$$G_{f_{2,k}}(\lambda^*) - G_f(\lambda^*) = (1-p) \int_{l_v(\lambda^*, \lambda^* + \frac{1}{c_k})} [r_k h(s) - f(s)] \frac{1-2s}{s} ds. \tag{B.31}$$

So in general we cannot expect $G_{f_{2,k}}(\lambda^*) = 0$. From now on, for notation abbreviation, we denote $(1-p) \int_{l_v(\lambda^*, \lambda^* + \frac{1}{c_k})} [r_k h(s) - f(s)] \frac{1-2s}{s} ds$ as e_k.

In this section, we shall add or subtract some weight from $f_{2,k}(s)$ on two small intervals near $\overline{s}$ and $\underline{s}$. We can always do this properly to balance out the extra e_k so that the resulted function's characteristic function does take value 0 at λ^*. Below we describe the outline of this rebalance.

For the chosen $f(s)$ and associated unique essential confounded learning belief $\lambda^* \in (z_1, z_2)$, let $\underline{x}$ be slightly above $\underline{s}$ so that there exists an open interval $(\underline{x}, \underline{x} + \varepsilon)$ on which $f(s) > 0$ a.e.. Similarly we select a proper $\overline{x}$ below $\overline{s}$ so that $f(s) > 0$ on some $(\overline{x} - \varepsilon, \overline{x})$.

We define two sequences $\underline{d_k}, \overline{d_k}$ such that

$$\int_{\underline{x}}^{\underline{x}+\underline{d_k}} f(s)\frac{1-2s}{s}ds = c|e_k| = \int_{\overline{x}-\overline{d_k}}^{\overline{x}} f(s)\frac{2s-1}{s}ds, \text{ where } c > \frac{1}{|2p-1|}. \tag{B.32}$$

We could do such a definition because

Lemma B.10 $\lim_{k\to+\infty} |e_k| = 0$.

The for big enough k, we could always find some small $\underline{d_k}$ and $\overline{d_k}$ so that equation B.32 holds. Lemma B.10 can be proved through a direct computation so the detail is omitted. An immediate corollary of definition B.32 is that

Corollary B.11 $\lim_{k\to+\infty} \underline{d_k} = \lim_{k\to+\infty} \overline{d_k} = 0$.

Vaguely speaking, $(\underline{x}, \underline{x} + \underline{d_k})$ and $(\overline{x} - \overline{d_k}, \overline{x})$ are the rebalance region on which we shall properly change the definition of $f_{2,k}(s)$ to obtain $f_{3,k}(s)$ so that $G_{f_{3,k}}(\lambda^*) = 0$. [12]

After introducing the rebalancing region, we see that

$$\begin{aligned} G_{f_{3,k}}(\lambda^*) - G_f(\lambda^*) &= p\int_{\underline{x}}^{\underline{x}+\underline{d_k}} [f_{3,k}(s) - f(s)]\frac{1-2s}{s}ds \\ &\quad +(1-p)\int_{\overline{x}-\overline{d_k}}^{\overline{x}} [f_{3,k}(s) - f(s)]\frac{1-2s}{s}ds + e_k \\ &= 0. \end{aligned} \tag{B.33}$$

Also, to guarantee that so-defined $\hat{f}(s)$ still satisfy the signal property ((3) in condition B.3), we must have

$$\int_{\underline{s}}^{\overline{s}} f_{3,k}(s)\frac{1-2s}{s}ds = \int_{\underline{s}}^{\overline{s}} f(s)\frac{1-2s}{s}ds. \tag{B.34}$$

Therefore, we should properly define $f_{3,k}(s)$ so that equations B.33 and B.34 hold.

After a simple transformation [13] , we could check that equations B.33 and B.34 is equiv-

[12] Here we introduce $\underline{x}, \overline{x}$ to avoid that the rebalance region takes the form $(0, \varepsilon)$. We shall see that this is important for our proof that so constructed $\overline{f}_k(s)$ converges to $f(s)$ in L_1-norm.

[13] We observe that equation B.33 is equivalent to

$$\int_{\underline{s}}^{\underline{x}+\underline{d_k}} [f_{3,k}(s) - f(s)]\frac{1-2s}{s}ds = \int_{\overline{x}-\overline{d_k}}^{\overline{s}} [f_{3,k}(s) - f(s)]\frac{2s-1}{s}ds. \tag{B.35}$$

This follows immediately from equations B.17 and B.29.

alent to

$$\int_{\underline{x}}^{\underline{x}+\underline{d_k}} f_{3,k}(s)\frac{1-2s}{s}ds = -\frac{1}{2p-1}e_k + \int_{\underline{x}}^{\underline{x}+\underline{d_k}} f(s)\frac{1-2s}{s}ds; \tag{B.36}$$

$$\int_{\overline{x}-\overline{d_k}}^{\overline{x}} f_{3,k}(s)\frac{2s-1}{s}ds = -\frac{1}{2p-1}e_k + \int_{\overline{x}-\overline{d_k}}^{\overline{x}} f(s)\frac{2s-1}{s}ds. \tag{B.37}$$

By the definition of d_k in equation B.32, we can see that the right-hand side of equations B.36 and B.37 are always positive even if $-\frac{1}{2p-1}e_k < 0$. Therefore, the right-hand side of equation B.36 is a finite multiple of $\underline{d_k}$. Similar for equation B.37. Therefore, by smooth gluing lemma, we could always define $f_{3,k}(s)$ properly on $(\underline{x}, \underline{x}+\underline{d_k})$ and $(\overline{x}-\overline{d_k}, \overline{x})$ so that equations B.36,B.37 and proper smoothness assumption hold. [14]

Now, given that so-defined $f_{3,k}(s)$ satisfies equations B.33 and B.34. We know

$$G_{f_{3,k}}(\lambda^*) = 0; \int_{\underline{s}}^{\overline{s}} f_{3,k}(s)\frac{1-2s}{s}ds = 0. \tag{B.38}$$

It is easy to verify that $f_{3,k}(s) \geq 0$ and that $f_{3,k}(s) \in C^k$ if $f(s) \in C^k$. Since in the rebalance, we don't change the definition on $l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^* + \frac{1}{c_k})$, so

$$G_{f_{3,k}}(\lambda) - G_{f_{3,k}}(\lambda^*) = r_k g_k(\lambda) \tag{B.39}$$

still holds. To complete the construction of $\hat{f}_k$, we let

$$\hat{f}_k(s) = \frac{f_{3,k}(s)}{\int_{\underline{s}}^{\overline{s}} f_{3,k}(s)ds}. \tag{B.40}$$

Then it is direct to verify that $\{f_{3,k}(s)\}_k$ satisfies conditions (2), (3), (4) in condition B.3. In the next subsection, we will show that $\hat{f}_k(s)$ converges to the original $f(s)$ in L_1-norm.

B.6 Convergence in L_1-norm

We first recall that $f_{3,k}(s) = f(s)$ except on (1) the perturbed region $l_u(\lambda^*, \lambda^*+\frac{1}{c_k}) \cup l_v(\lambda^*, \lambda^*+\frac{1}{c_k})$; (2) the smooth gluing region $l_u(\lambda^* - \varepsilon_k, \lambda^*)$, $l_u(\lambda^* + \frac{1}{c_k}, \lambda^* + \frac{1}{c_k} + \varepsilon_k)$, $l_v(\lambda^* - \varepsilon_k, \lambda^*)$,

[14] If $f(s)$ is not continuous, to apply the smooth gluing lemma, we can arbitrarily specify the endpoints behavior for $f_{3,k}(s)$ on $\underline{x}, \underline{x}+\underline{d_k}, \overline{x}, \overline{x}+\overline{d_k}$.

$l_v(\lambda^* + \frac{1}{c_k}, \lambda^* + \frac{1}{c_k} + \varepsilon_k)$; (3) the rebalance region $(\underline{x}, \underline{x} + \underline{d_k}) \cup (\overline{x} - \overline{d_k}, \overline{x})$. Furthermore,

$$\lim_{k \to +\infty} \frac{1}{c_k} = \lim_{k \to +\infty} \varepsilon_k = \lim_{k \to +\infty} \underline{d_k} = \lim_{k \to +\infty} \overline{d_k} = 0. \tag{B.41}$$

So the measure of the set, where $f_{3,k}(s) \neq f(s)$, shrinks to 0 as $k \to +\infty$. It is natural to guess that $f_{3,k}(s)$ converges to $f(s)$ in L_1-norm. The next lemma proves it.

Lemma B.12

$$\lim_{k \to +\infty} \|f_{3,k}(s) - f(s)\|_{L_1} = 0. \tag{B.42}$$

Proof. For notation abbrevity, we use P_k to denote the perturbed region, use SG_k to denote the smooth gluing region, and RB_k to denote the rebalance region of $f_{3,k}(s)$. Then

$$\|f_{3,k}(s) - f(s)\|_{L_1} = \int_{P_k \cup SG_k \cup RB_k} |f_{3,k}(s) - f(s)| ds \leq \int_{P_k \cup SG_k \cup RB_k} f_{3,k}(s) + f(s) ds. \tag{B.43}$$

Because $\lim_{k \to +\infty} m(P_k \cup SG_k \cup RB_k) = 0$, we must have

$$\lim_{k \to +\infty} \int_{P_k \cup SG_k \cup RB_k} f(s) = 0.^{15} \tag{B.44}$$

On the other hand, by the construction we have

$$\int_{P_k \cup SG_k} f(s) \frac{1 - 2s}{s} ds = \int_{P_k \cup SG_k} f_{3,k}(s) \frac{1 - 2s}{s} ds. \tag{B.45}$$

If $u, v > (z_1, z_2)$, then $P_k \cup SG_k \subset (0, \frac{1}{2})$ for large enough k. We could always find an upper bound $\overline{z} < \frac{1}{2}$ so that $\overline{z} > \sup\{P_k, SG_k\}$ for all the k in the tail. Then

$$\int_{P_k \cup SG_k} f(s) \frac{1 - 2s}{s} ds = \int_{P_k \cup SG_k} f_{3,k}(s) \frac{1 - 2s}{s} ds > \frac{1 - 2\overline{z}}{\overline{z}} \int_{P_k \cup SG_k} f_{3,k}(s) ds. \tag{B.46}$$

So $\lim_{k \to +\infty} \int_{P_k \cup SG_k} f_{3,k}(s) ds = 0$. Similarly, if $(u, v) < (z_1, z_2)$, then there exists $\underline{z} > \frac{1}{2}$ such that $\inf\{P_k \cup SG_k\} > \underline{z}$. Then following equation B.45, we have

$$\int_{P_k \cup SG_k} f(s) \frac{2s - 1}{s} ds = \int_{P_k \cup SG_k} f_{3,k}(s) \frac{2s - 1}{s} ds > \frac{2\underline{z} - 1}{\underline{z}} \int_{P_k \cup SG_k} f_{3,k}(s) ds. \tag{B.47}$$

[15] Since $f(s)$ is integrable on $(\underline{s}, \overline{s})$, for any $a \in (\underline{s}, \overline{s})$, we know the indefinite integral $\int_a^x f(s)ds$ is continuous in x and decreases to 0 as $x \downarrow 0$.

So $\lim_{k\to+\infty}\int_{P_k\cup SG_k} f_{3,k}(s)ds=0$ again.

Following equations B.36 and B.37, we have

$$\frac{1-2\overline{z}}{\overline{z}}\int_{\underline{x}}^{\underline{x}+\underline{d_k}} f_{3,k}(s)ds < \int_{\underline{x}}^{\underline{x}+\underline{d_k}} f_{3,k}(s)\frac{1-2s}{s}ds = -\frac{1}{2p-1}e_k + \int_{\underline{x}}^{\underline{x}+\underline{d_k}} f(s)\frac{1-2s}{s}ds;$$
$$\frac{2\overline{z}-1}{\overline{z}}\int_{\overline{x}-\overline{d_k}}^{\overline{x}} f_{3,k}(s)ds < \int_{\overline{x}-\overline{d_k}}^{\overline{x}} f_{3,k}(s)\frac{2s-1}{s}ds = -\frac{1}{2p-1}e_k + \int_{\overline{x}-\overline{d_k}}^{\overline{x}} f(s)\frac{2s-1}{s}ds.$$

So $\lim_{k\to+\infty}\int_{RB_k} f_{3,k}(s)ds=0$. ■

Following lemma B.12, we have

$$\lim_{k\to+\infty}\int_{\underline{s}}^{\overline{s}} f_{3,k}(s)ds = \int_{\underline{s}}^{\overline{s}} f(s)ds = 1. \tag{B.48}$$

Then we immediately have

Lemma B.13 $\lim_{k\to+\infty}\|\hat{f}_k(s)-f(s)\|_{L_1} = \lim_{k\to+\infty}\|\frac{f_{3,k}(s)}{\int_{\underline{s}}^{\overline{s}} f_{3,k}(s)ds} - f(s)\|_{L_1}=0.$

B.7 Omitted Details in the Proof

B.7.1 Omitted Details in the Construction of Oscillation Function $g_k(\lambda)$

The proof of proposition B.4 is done through the following lemma:

Lemma B.14 *Recall that*

$$\hat{g}(x) \;\; = \;\; \int_0^x \hat{h}(t)dt. \tag{B.49}$$

We have

$$\lim_{x\to+\infty}\hat{g}(x)=C. \tag{B.50}$$

Furthermore, $\hat{g}\big((2n+1)\pi\big) > C$ *and* $\hat{g}(2n\pi) < C$ *for all* $n\in\mathbb{N}$. *Thus, for each* $k\in\mathbb{N}$, *there exists an unique* $c_k\in((k-1)\pi,k\pi)$ *such that* $\hat{g}(c_k)=C$.

Intuitively, since $\hat{h}(t)$ vanishes at infinity, we could view

$$\lim_{x\to+\infty}\hat{g}(x)=\sum_{k=1}^{+\infty}\int_{(k-1)\pi}^{k\pi}\hat{h}(t)dt, \tag{B.51}$$

as a series with vanishing alternating terms. By Leibniz's rule, such a series must be convergent. Then we could easily see the existence of $c_k \in ((k-1)\pi, k\pi)$ so that

$$\lim_{x\to+\infty}\int_{c_k}^{x}\hat{h}(t)dt = 0. \tag{B.52}$$

Below are the details.

Proof. For all $n \in \mathbb{N}$, we can define

$$a_n = \int_{(n-1)\pi}^{n\pi}\hat{h}(t)dt. \tag{B.53}$$

Then $a_n > 0$ if n is odd; and $a_n < 0$ if n is even. Furthermore, $|a_n|$ decreases to 0 as following:

$$\begin{aligned}
& e^{-(n-1)\pi}e^{-1}\int_{(n-1)\pi}^{n\pi}\sin(t)dt \\
> \quad & a_n \\
= \quad & \int_{(n-1)\pi}^{n\pi}\sin(t)e^{-(\sin(t))^{-2}-t}dt \\
\overset{t'=t+\pi}{=} \quad & \int_{n\pi}^{(n+1)\pi} -\sin(t')e^{-(\sin(t'))^{2}-t'+\pi}dt' \\
= \quad & -e^{\pi}a_{n+1}.
\end{aligned} \tag{B.54}$$

Therefore by Leibniz's alternating series test, $\sum_{n=1}^{\infty} a_n$ converges to a constant C.

For any positive x that is not a multiple of π, there exists unique integer k such that $x \in (2k\pi, 2(k+1)\pi)$. Since $h(t) > 0$ on $(2k\pi,(2k+1)\pi)$ and $h(t) < 0$ on $((2k+1)\pi, 2(k+1)\pi)$, we must have

$$\hat{g}(2k\pi) < \hat{g}(x) < \hat{g}((2k+1)\pi). \tag{B.55}$$

Since $\lim_{k\to\infty}\hat{g}(2k\pi) = \lim_{k\to\infty}\hat{g}((2k+1)\pi) = C$, we must have $\lim_{x\to+\infty}\hat{g}(x) = C$. ■

The proof of proposition B.5 is primarily through computation:

Lemma B.15 *That $g_{k+}^{(1)}(\lambda)$ and $g_{k-}^{(1)}(\lambda)$ are C^∞ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$ is equivalent to that $\hat{h}_+(t) = \max\{\hat{h}(t), 0\}, \hat{h}_-(t) = -\min\{\hat{h}(t), 0\}$ are C^∞ on $(c_k, +\infty)$.*

Proof. Let $x(\lambda) = \frac{1}{\lambda-\lambda^*}$. Then it is direct to verify that

$$g_{k+}^{(1)}(\lambda) = [x(\lambda)]^2\hat{h}_-(x(\lambda)) \quad ; \quad g_{k-}^{(1)}(\lambda) = [x(\lambda)]^2\hat{h}_+(x(\lambda)). \tag{B.56}$$

Since $x(\lambda)$ is a C^∞ map from $(\lambda^*, \lambda^*+\frac{1}{c_k})$ to $(c_k, +\infty)$, that $g^{(1)}_{k+}(\lambda), g^{(1)}_{k-}(\lambda)$ is C^∞ is equivalent to that $x^2\hat{h}_+(x), x^2\hat{h}_-(x)$ is C^∞ on $(c_k, +\infty)$. Since $c_k > 0$, it is further equivalent to that $\hat{h}_+(x), \hat{h}_-(x)$ are C^∞ on $(c_k, +\infty)$. ■

The following proposition confirms that $\hat{h}_+(x), \hat{h}_-(x)$ are indeed C^∞.

Proposition B.16 *That $\hat{h}_+(x), \hat{h}_-(x)$ are C^∞ on $(0, +\infty)$.*

Proof. We first write out

$$\hat{h}_+(x) = \begin{cases} e^{-(\sin x)^{-2}-x}\sin x, x \in \cup_{n\geq 0}(2n\pi, (2n+1)\pi); \\ 0; x \in \cup_{n\geq 0}((2n+1)\pi, 2(n+1)\pi); \end{cases}$$

and that

$$\hat{h}_-(x) = \begin{cases} -e^{-(\sin x)^{-2}-x}\sin x, x \in \cup_{n\geq 0}((2n+1)\pi, 2(n+1)\pi); \\ 0, x \in \cup_{n\geq 0}(2n\pi, (2n+1)\pi). \end{cases}$$

Then we could turn to show that $e^x\hat{h}_+(x)$ and $e^x\hat{h}_-(x)$ are C^∞. For notation simplicity, let

$$E_\pi(x) = e^x\hat{h}(x) \begin{cases} = e^{-(\sin x)^{-2}}\sin x, x \neq k\pi; \\ = 0. \end{cases} \tag{B.57}$$

Then $E_{\pi+}(x) = e^x\hat{h}_+(x), E_{\pi_-}(x) = e^x\hat{h}_-(x)$.

Since being C^∞ respects multiplication and composition, it is obvious that $E_{\pi+}(x), E_{\pi-}(x)$ are C^∞ on $x \in \cup_{k\geq 1}((k-1)\pi, k\pi)$. Thus, we only need to prove the C^∞ of $E_{\pi+}(x), E_{\pi-}(x)$ at $k\pi, k \geq 0$. [16]

The following claim computes $E^{(n)}_\pi(x)$ for all $x \neq k\pi$.

Claim B.17

$$E^{(n)}_\pi(x) = e^{-(\sin x)^{-2}}(\sin x)^{-3n+1}f_n(x), x \neq k\pi; \tag{B.58}$$

where $f_n(x)$ is of C^∞ on $(0, +\infty)$ with

$$\lim_{x\to k\pi} |f^{(k)}_n(x)| < +\infty \text{ and } \lim_{x\to+\infty} |f^{(k)}_n(x)| < +\infty \tag{B.59}$$

[16] If $k = 0$, then we only need to show the C^∞ at 0^+ since we don't define $\hat{h}(t)$ for negative t.

See Appendix B.7.3 for the proof of the above claim.

As $x \to k\pi$, $\sin x \to 0$, then $E_\pi^{(n)}(x)$ contains a term $e^{-(\sin x)^{-2}}$ that vanishes exponentially and a term $(\sin x)^{-3n+1}$ that explodes polynomially. Hence $E_\pi^{(n)}(x)$ vanishes as $x \to k\pi$. To be rigorous, we prove that $E_{\pi+}(x)$ is C^∞ at $(2n+1)\pi$. The proof for $2n\pi$ and for $E_{\pi-}(x)$ are similar.

For $k\pi$ with k odd, we can compute

$$
\begin{aligned}
& \lim_{x\to k\pi^-} E_{\pi+}^{(n)}(x) = \lim_{x\to k\pi^-} E_\pi^{(n)}(x) \\
= \ & f_n(k\pi) \lim_{x\to k\pi} e^{-(\sin x)^{-2}(\sin x)^{-3n+1}} = f_n(k\pi) \lim_{t\to 0^+} e^{-t^{-2}} t^{-3n+1} = 0
\end{aligned}
$$

and

$$
\begin{aligned}
& E_{\pi+}^{(n+1)}(k\pi) = \lim_{\Delta\to 0^-} \frac{E_{\pi+}^{(n)}(k\pi+\Delta) - 0}{\Delta} \\
= \ & f_n(k\pi) \lim_{\Delta\to 0^-} \frac{e^{-(\sin(k\pi+\Delta)^{-2})}\big(\sin(k\pi+\Delta)\big)^{-3n+1}}{\sin(k\pi+\Delta)} \frac{\sin(k\pi+\Delta)}{\Delta} \\
= \ & 0
\end{aligned} \tag{B.60}
$$

Since $E_{\pi+}^{(n)}(x) = 0$ on $(k\pi, (k+1)\pi)$, we must have $\lim_{x\to k\pi^+} E_{\pi+}^{(n)}(x) = 0$. Therefore, we have shown $E_{\pi+}^{n+1}(x)$ is defined and continuous at $k\pi$. Since above two equations hold for any n, we have shown that $E_{\pi+}(x)$ is C^∞ at any $k\pi$ (k odd). ■

B.7.2 Omitted Details in "Perturbing to Oscillate"

In the main text, we have complete the construction under assumption $u, v > z_1, z_2$. In this appendix, we shall complete the construction if $u, v < z_1, z_2$.

Since λ^* is the unique confounded learning belief on (z_1, z_2). Under assumption $u, v < z_1, z_2$, we have $u, v < \lambda^*$. As a result,

$$
G_h'(\lambda) \ = \ h(\frac{\lambda}{\lambda+v})\overbrace{(1-p)\frac{\lambda-v}{\lambda}\frac{v}{(\lambda+v)^2}}^{\hat{t}_1(\lambda)} - h(\frac{\lambda}{\lambda+u})\overbrace{p\frac{\lambda-u}{\lambda}\frac{u}{(\lambda+u)^2}}^{\hat{t}_2(\lambda)}
$$

where

$$
h(l_v(\lambda))\hat{t}_1(\lambda) \geq 0 \quad ; \quad -h(l_u(\lambda))\hat{t}_2(\lambda) \leq 0 \text{ on } (\lambda^*, \lambda^* + \frac{1}{c_k}). \tag{B.61}
$$

Therefore, we need to match $h(l_v(\lambda))\hat{t}_1(\lambda)$ with $g'_{k+}(\lambda)$. While in definition B.6, we match $g'_{k+}(\lambda)$ with $h(l_u(\lambda))t_1(\lambda)$. So we have the following general definition:

Definition B.18 *If $u, v > z_1, z_2$, $h(s)$ is defined by*

$$h(s)t_1(l_u^{-1}(s)) = g'_{k+}(l_u^{-1}(s)) \text{ on } l_u(\lambda^*, \lambda^* + \frac{1}{c_k}) \tag{B.62}$$

$$h(s)t_2(l_v^{-1}(s)) = g'_{k-}(l_v^{-1}(s)) \text{ on } l_v(\lambda^*, \lambda^* + \frac{1}{c_k}); \tag{B.63}$$

where $t_1(\lambda) = p\frac{u-\lambda}{\lambda}\frac{u}{(\lambda+u)^2}$ and $t_2(\lambda) = (1-p)\frac{v-\lambda}{\lambda}\frac{v}{(\lambda+v)^2}$.
If $u, v < z_1, z_2$, $h(s)$ is defined by

$$h(s)\hat{t}_1(l_v^{-1}(s)) = g'_{k+}(l_v^{-1}(s)) \text{ on } l_v(\lambda^*, \lambda^* + \frac{1}{c_k}) \tag{B.64}$$

$$h(s)\hat{t}_2(l_u^{-1}(s)) = g'_{k-}(l_u^{-1}(s)) \text{ on } l_u(\lambda^*, \lambda^* + \frac{1}{c_k}); \tag{B.65}$$

where $\hat{t}_1(\lambda) = (1-p)\frac{\lambda-v}{\lambda}\frac{v}{(\lambda+v)^2}$ and $\hat{t}_2(\lambda) = p\frac{\lambda-u}{\lambda}\frac{u}{(\lambda+u)^2}$.

Now we turn to prove the propositios B.7 under all assumptions.

Proof of Proposition B.7, Property 1. This is done through two separate "change of variables". We have

$$G_h(\lambda) - G_h(\lambda^*) = p\int_{\frac{\lambda^*}{\lambda^*+u}}^{\frac{\lambda}{\lambda+u}} h(s)\frac{1-2s}{s}ds - (1-p)\int_{\frac{\lambda^*}{\lambda^*+v}}^{\frac{\lambda}{\lambda+v}} h(s)\frac{1-2s}{s}ds \tag{B.66}$$

By let $s = \frac{\lambda}{\lambda+u}$ for the first term, and $s = \frac{\lambda}{\lambda+v}$ for the second term, we obtain

$$\begin{aligned} & p\int_{\frac{\lambda^*}{\lambda^*+u}}^{\frac{\lambda}{\lambda+u}} h(s)\frac{1-2s}{s}ds - (1-p)\int_{\frac{\lambda^*}{\lambda^*+v}}^{\frac{\lambda}{\lambda+v}} h(s)\frac{1-2s}{s}ds \\ = & \int_{\lambda^*}^{\lambda} h(\frac{\lambda}{\lambda+u})p\frac{u-\lambda}{\lambda}\frac{u}{(\lambda+u)^2}d\lambda - \int_{\lambda^*}^{\lambda} h(\frac{\lambda}{\lambda+v})(1-p)\frac{v-\lambda}{\lambda}\frac{v}{(\lambda+v)^2}d\lambda \end{aligned} \tag{B.67}$$

We can see that under all assumptions, this becomes

$$\int_{\lambda^*}^{\lambda}[g'_{k+}(\lambda) - g'_{k-}(\lambda)]d\lambda = g_k(\lambda). \tag{B.68}$$

■

Proof of Proposition B.7, Property 2. It is easy to verify that $t_1(\lambda), t_2(\lambda), \hat{t}_1(\lambda), \hat{t}_2(\lambda)$

are C^∞ and strictly positive on $(\lambda^*, \lambda^* + \frac{1}{c_k})$. By proposition B.5, $g'_{k+}(\lambda), g'_{k-}(\lambda)$ are C^∞ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$ as well. So $\frac{g'_{k+}(\lambda)}{t_1(\lambda)}$, $\frac{g'_{k-}(\lambda)}{t_2(\lambda)}$, $\frac{g'_{k+}(\lambda)}{\hat{t}_1(\lambda)}$ and $\frac{g'_{k-}(\lambda)}{\hat{t}_2(\lambda)}$ are C^∞ since multiplication respects C^∞.

Furthermore, $l_u^{-1}(s) = \frac{us}{1-s}$ is C^∞ from $l_u(\lambda^*, \lambda^* + \frac{1}{c_k})$ to $(\lambda^*, \lambda^* + \frac{1}{c_k})$. And similar for $l_v^{-1}(s) = \frac{vs}{1-s}$. Because being C^∞ respects composition, we have $h(s)$ be C^∞ under all assumptions. ■

The proof of Property 3 is primarily through computation. Let us first recall the Faa Di Bruno's lemma which enables us to compute the high order derivatives of composition functions.

Lemma B.19 (Faa Di Bruno) *Let $f(x) \in C^\infty(I)$ for some open interval $I \subset \mathbb{R}$; and $g(x) \in C^\infty(J)$ for some open interval J with range$(f) \subset J$; then the k-th order derivative of $h(x) = g\big(f(x)\big)$ is*

$$h^{(n)}(x) = \sum \frac{n!}{k_1! \dots k_n!} g^{(k)}(f(x)) \Pi_{m=1}^{n} \Big(\frac{f^{(m)}(x)}{m!} \Big)^{k_m}; \tag{B.69}$$

where the sum is taken over all the n-tuples $\{k_1, \dots, k_n\}$ $(k_i \geq 0)$ with

$$\sum_{m=1}^{n} m k_m = n. \tag{B.70}$$

And the order k of $g^{(k)}(f(x))$ is given by

$$\sum_{m=1}^{n} k_m = k. \tag{B.71}$$

Proof of Proposition B.7, Property 3. For notation simplicity, let $\lambda_u(s) = l_u^{-1}(s)$ and $\lambda_v(s) = l_v^{-1}(s)$. Let $\phi_u(\lambda) = \frac{1}{t_1(\lambda)}, \phi_v(\lambda) = \frac{1}{t_2(\lambda)}, \hat{\phi}_u(\lambda) = \frac{1}{\hat{t}_1(\lambda)}, \hat{\phi}_v(\lambda) = \frac{1}{\hat{t}_2(\lambda)}$. With these notations: Under assumption $u, v > z_1, z_2$,

$$h(s) = \phi_u(\lambda_u(s)) g'_{k+}(\lambda_u(s)); \text{ on } l_u(\lambda^*, \lambda^* + \frac{1}{c_k}).$$
$$h(s) = \phi_v(\lambda_v(s)) g'_{k-}(\lambda_u(s)); \text{ on } l_v(\lambda^*, \lambda^* + \frac{1}{c_k}).$$

Under assumption $u, v < z_1, z_2$,

$$h(s) = \phi_v(\lambda_v(s))g'_{k+}(\lambda_v(s));\ \text{on } l_v(\lambda^*, \lambda^* + \frac{1}{c_k}).$$
$$h(s) = \phi_u(\lambda_u(s))g'_{k-}(\lambda_u(s));\ \text{on } l_u(\lambda^*, \lambda^* + \frac{1}{c_k}).$$

We first have the following claim

Claim B.20 *Under assumptions $u, v > z_1, z_2$,*

$$\lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_u(\lambda)g'_{k+}(\lambda)\big) = 0, \forall m; \Rightarrow \lim_{s \to l_u(\lambda^*)+} h^{(n)}(s) = 0, \forall n;$$
$$\lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_v(\lambda)g'_{k-}(\lambda)\big) = 0, \forall m; \Rightarrow \lim_{s \to l_v(\lambda^*)+} h^{(n)}(s) = 0, \forall n;$$

Under assumptions $u, v < z_1, z_2$,

$$\lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_v(\lambda)g'_{k+}(\lambda)\big) = 0, \forall m; \Rightarrow \lim_{s \to l_v(\lambda^*)+} h^{(n)}(s) = 0, \forall n;$$
$$\lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_u(\lambda)g'_{k-}(\lambda)\big) = 0, \forall m; \Rightarrow \lim_{s \to l_u(\lambda^*)+} h^{(n)}(s) = 0, \forall n.$$

Proof. Using Faa Di Bruno's formula, $\forall s \in l_u(\lambda^*, \lambda^* + \frac{1}{c_k})$, we have

$$h^{(n)}(s) = \sum \frac{n!}{m_1! \dots m_n!}\Big[\frac{d^m}{d\lambda^m}\big(\phi_u(\lambda_u(s))g'_{k+}(\lambda_u(s))\big)\Big]\Pi_{t=1}^{n}\Big(\frac{\lambda_u^{(t)}(s)}{t!}\Big)^{m_t}. \tag{B.72}$$

Besides, since $\lambda_u(s) = \frac{us}{1-s}$ is C^∞ on $(0, 1)$, we must have $|\lambda_u^{(t)}(l_u(\lambda^*))| < +\infty$ for all m. Then we can easily see the first statement in the claim holds. The other statements can be proved in the exactly same way. ■

The next claim gives sufficient condition for conditions in Claim B.20.

Claim B.21 *We have*

$$\lim_{\lambda \to \lambda^* +} \frac{d^n}{d\lambda^n}\big(g'_{k+}(\lambda)\big) = 0, \forall n \Rightarrow \lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_u(\lambda)g'_{k+}(\lambda)\big) = \lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_v(\lambda)g'_{k+}(\lambda)\big) = 0, \forall m;$$
$$\lim_{\lambda \to \lambda^* +} \frac{d^n}{d\lambda^n}\big(g'_{k-}(\lambda)\big) = 0, \forall n \Rightarrow \lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_v(\lambda)g'_{k-}(\lambda)\big) = \lim_{\lambda \to \lambda^* +} \frac{d^m}{d\lambda^m}\big(\phi_u(\lambda)g'_{k-}(\lambda)\big) = 0, \forall m;$$

Proof. It is obvious that

$$\frac{d^m}{d\lambda^m}\big(\phi_u(\lambda)g'_{k+}(\lambda)\big) = \sum_{n=0}^{m}\phi_u^{(m-n)}(\lambda)\frac{d^n}{d\lambda^n}\Big(g'_{k+}(\lambda)\Big). \tag{B.73}$$

Given that $\phi_u(\lambda) = \frac{(\lambda+u)^2\lambda}{up(u-\lambda)}$ can be viewed as an C^∞ function on any small open interval around λ^*, $\phi_u^{(m-n)}(\lambda^*)$ must be finite for any n, m. From similar computations, the claim follows. ■

Let $x(\lambda) = \frac{1}{\lambda-\lambda^*}$ on $(\lambda^*, \lambda^* + \frac{1}{c_k})$. Let $E_\infty(x) = -x^2e^{-x}$. Recall that we defined

$$E_\pi(x)\begin{cases} = e^{-(\sin x)^{-2}}\sin x, x \neq k\pi; \\ = 0. \end{cases} \tag{B.74}$$

It is easy to verify that

$$g'_{k+}(\lambda) = -E_{\pi-}(x(\lambda))E_\infty(x(\lambda)); g'_{k-}(\lambda) = -E_{\pi+}(x(\lambda))E_\infty(x(\lambda)). \tag{B.75}$$

Then we use Faa Di Bruno's lemma again to obtain.

$$\frac{d^n}{d\lambda^n}\big[g'_{k+}(\lambda)\big] = -\sum\frac{n!}{k_1!\dots k_n!}\Big(\sum_{t=0}^{k}E_{\pi-}^{(t)}(x(\lambda))E_\infty^{(k-t)}(x(\lambda))\Big)\Big(\Pi_{m=1}^n[(-1)^m(\lambda-\lambda^*)^{-(m+1)}]^{k_m}\Big)$$

From claim B.17, we know that $E_{\pi-}(x)$ is of C^∞ on $[c_k, +\infty)$ and $\lim_{x\to+\infty}|E_{\pi-}^{(n)}(x)| < +\infty$. As $\lambda \to \lambda^*+$, $\Pi_{m=1}^n[(-1)^m(\lambda-\lambda^*)^{-(m+1)}]^{k_m}$ explodes. Fortunately, this term explodes polynomially; we shall see that $E_\infty^{k-t}(x(\lambda))$ vanishes exponentially as λ approaches λ^* as proven in the next claim.

Claim B.22 *For all $m \geq 0$, $E_\infty^{(m)}(x) = e^{-x}p_m(x)$, where $p_m(x) = a_mx^2 + b_mx + c_m$ is a polynomial of x.*

Proof. By induction. First

$$E_\infty^{(1)}(x) = e^{-x}(x^2 - 2x). \tag{B.76}$$

Given induction hypothesis, we have

$$E_\infty^{(m+1)}(x) = e^{-x}[-a_mx^2 + (2a_m - b_m)x + b_m - c_m]. \tag{B.77}$$

■ Note that $E_{\infty}^{k-t}(x(\lambda))$ contains a $e^{-x(\lambda)}$ even if $k-t=0$. Thus, each polynomially exploding term is eliminated by the exponentially vanishing term. Hence we have the following claim

Claim B.23 $\lim_{\lambda\to\lambda^*+}\frac{d^n}{d\lambda^n}\left[g'_{k+}(\lambda)\right]=\lim_{\lambda\to\lambda^*+}\frac{d^n}{d\lambda^n}\left[g'_{k-}(\lambda)\right]=0.$

Proof. In the above statement we have proved the case for $g'_{k+}(\lambda)$. The case for $g'_{k-}(\lambda)$ is exactly the same. ■ ■

B.7.3 Other Omitted Details

Proof of Claim B.17. The proof of this lemma is done by induction. For $n=1$, we have

$$E_{\pi}^{(1)}(x)=e^{-(\sin x)^{-2}}(\sin x)^{-2}\big(\sin^2 x\cos x+\cos x\big),x\neq k\pi. \tag{B.78}$$

Using power reduction formula

$$\cos^3 x=\frac{3\cos x+\cos 3x}{4}, \tag{B.79}$$

we can write

$$f_1(x)=\frac{9}{4}\cos x-\frac{1}{4}\cos 3x. \tag{B.80}$$

So $|f_1^{(k)}(x)|\leq\frac{9}{4}+\frac{1}{4}3^k$ for all $x\in(0,+\infty)$.

Inductively,

$$E_{\pi}^{(n+1)}(x)=e^{-(\sin x)^{-2}}(\sin x)^{-3(n+1)+1}\left[(2\cos x+(1-3n)\sin^2 x\cos x)f_n(x)+\sin^3 x f_n^{(1)}(x)\right]$$

Again using the power reduction formula for trigonometric functions:

$$\cos^3 x=\frac{3\cos x+\cos 3x}{4}\qquad\sin^3 x=\frac{3\sin x-\sin 3x}{4}. \tag{B.81}$$

We can write

$$f_{n+1}(x)=C_n(x)f_n(x)+D_n(x)f_n^{(1)}(x), \tag{B.82}$$

where $C_n(x)=(\frac{9}{4}-\frac{3}{4}n)\cos x-\frac{1-3n}{4}\cos 3x$ and $D_n(x)=\frac{3}{4}\sin x-\frac{1}{4}\sin 3x$. Easy to verify

$$|C_n^{(m)}(x)|\leq|\frac{9}{4}-\frac{3n}{4}|+|\frac{1-3n}{4}|3^m\quad\text{and}\quad|D_n^{(m)}(x)|\leq\frac{3}{4}+\frac{1}{4}3^m. \tag{B.83}$$

Lastly, for any k, we have

$$f_{n+1}^{(k)}(x) = \sum_{m=0}^{k} [C_n^{(m-k)}(x) f_n^{(m)}(x) + D_n^{(m-k)}(x) f_n^{(m)}(x)]. \tag{B.84}$$

We have shown that $C_n^{(m-k)}(x)$ and $D_n^{(m-k)}(x)$ is finite. That $f_n^{(m)}(x)$ is finite follows from induction hypothesis. Therefore, we must have $f_{n+1}^{(k)}(x) < +\infty$ for any $x \in (0, +\infty)$. ■

C Omitted Proofs in Subsections 3.4

In this section we prove propositions 3.10. We start with a discussion with parameters tuple $\{p, u, v, \underline{s}, \overline{s}\}$ where $p = \frac{1}{2}$. This case is singled out for that the conclusion holds for both bounded and unbounded signals.

Lemma C.1 *For an observational learning model characterized by $\{p, u, v, \underline{s}, \overline{s}\}$ and $F_{(\underline{s}, \overline{s})}$, if $p = \frac{1}{2}$, then the model cannot admit any essential confounded learning belief outside the monotonic region.*

Proof. By direct computation, we have

$$G_f(\lambda) = \begin{cases} \frac{1}{2} \int_{\frac{\lambda}{\lambda+v}}^{\frac{\lambda}{\lambda+u}} f(s) \frac{1-2s}{s} ds; \text{ if } v > u; \\ -\frac{1}{2} \int_{\frac{\lambda}{\lambda+u}}^{\frac{\lambda}{\lambda+v}} f(s) \frac{1-2s}{s} ds; \text{ if } v < u. \end{cases}$$

Without loss of generality, we assume that $v > u$ and that $\lambda^* > v > u$ is a confounded learning belief, we shall argue that such a λ^* cannot be an essential confounded learning belief. Here

$$G_f(\lambda^*) = 0 \Rightarrow f(s) = 0 \text{ a.e. on } [\frac{\lambda^*}{\lambda^* + v}, \frac{\lambda^*}{\lambda^* + u}] \subset (\frac{1}{2}, \max\{\frac{\lambda^*}{\lambda^* + u}, \overline{s}\}). \tag{C.1}$$

Let $o \subset (\frac{1}{2}, \overline{s})$ be the largest closed interval containing $[\frac{\lambda^*}{\lambda^*+v}, \frac{\lambda^*}{\lambda^*+u}]$ such that $f(s) = 0$ almost everywhere on o. We observe that o cannot be entire $(\frac{1}{2}, \overline{s})$. Otherwise,

$$\int_{\underline{s}}^{\frac{1}{2}} f(s) \frac{1-2s}{s} ds = -\int_{\frac{1}{2}}^{\overline{s}} f(s) \frac{1-2s}{s} ds = 0 \Rightarrow f(s) = 0 \text{ a.e. on } (\underline{s}, \overline{s}). \tag{C.2}$$

This obviously contradicts that $f(s)$ is a private signal. Now let

$$\Lambda = \{\lambda \in \mathcal{B} | [\frac{\lambda}{\lambda + v}, \frac{\lambda}{\lambda + u}] \subset o\}. \tag{C.3}$$

Then we can easily see that any $\lambda \in \Lambda$ is a confounded learning belief. We can characterize the two endpoints as

$$\inf \Lambda = \inf\{\lambda \in \mathcal{B} | \frac{\lambda}{\lambda + v} \geq \inf o\} \leq \lambda^* \leq \sup \Lambda = \sup\{\lambda \in \mathcal{B} | \frac{\lambda}{\lambda + u} \leq \inf o\}. \tag{C.4}$$

If $(\inf \Lambda, \sup \Lambda) \neq \emptyset$, then any $\lambda \in \Lambda$ can not be an essential confounded learning belief. This is because $G_f = 0$ on an open interval near λ. Otherwise, if $(\inf \Lambda, \sup \Lambda) = \emptyset$, which is equivalent to that $\Lambda = \lambda^*$, then we can see that

$$\lim_{\lambda \to \lambda^* -} G_f(\lambda) < 0 \quad \text{and} \quad \lim_{\lambda \to \lambda^* +} G_f(\lambda) < 0. \tag{C.5}$$

Such a λ^* cannot be an essential confounded learning belief as well. ■

C.1 Bounded Signal Strength

Since we assume private signals are bounded, players herd on extreme public beliefs. Recall that big enough λ implies enough weight for state B, Mis-match type then herds on action a while Match type herds on action b. Similarly, for small enough λ, Mis-match type herds on action b while Match type herds on action a. If we use "confounding region" to denote the set of public beliefs where no type herds. Then we obtain

Lemma C.2 *If $v > u$, the confounding region is $(v\frac{\underline{s}}{1-\underline{s}}, u\frac{\overline{s}}{1-\overline{s}})$. If $u > v$, the confounding region $(u\frac{\underline{s}}{1-\underline{s}}, v\frac{\overline{s}}{1-\overline{s}})$.*

In fact, The "confounding region" is the only region where confounded learning could arise. After all, for those region where at most one type is active, no type is confounded with the other type at any public belief.

Therefore, for those parameters such that the confounding region is empty, or is inside $[u, v]$, then no essential confounded learning belief could exist outside monotonic region.

Lemma C.3 *If (1) $v > u$ and $v\frac{\underline{s}}{1-\underline{s}} \geq u\frac{\overline{s}}{1-\overline{s}}$ or (2) $u > v$ and $u\frac{\underline{s}}{1-\underline{s}} \geq v\frac{\overline{s}}{1-\overline{s}}$, then the confounding region is empty.*

Besides, if (3) $v \geq u\frac{\bar{s}}{1-\bar{s}} > v\frac{\underline{s}}{1-\underline{s}} \geq u$, (4) $u \geq v\frac{\bar{s}}{1-\bar{s}} > u\frac{\underline{s}}{1-\underline{s}} \geq v$, then the confounding region is contained in $[u, v]$ (or $[v, u]$ if $u > v$).

In all cases, signal set $\mathcal{E} \subset F_{(\underline{s},\bar{s})}$ is empty.

It maybe worth to point out that players could herd for moderate public beliefs for parameters satisfying (1) and (2) in lemma C.3

Remark C.4 *Consider assumption (1) in lemma C.6, for public beliefs $\lambda \in [u\frac{\bar{s}}{1-\bar{s}}, v\frac{\underline{s}}{1-\underline{s}}]$, both types actually don't use their private information. So this region is better referred as a "herding region".*

Besides, this herding region can be degenerate if $u\frac{\bar{s}}{1-\bar{s}} = v\frac{\underline{s}}{1-\underline{s}}$. In this case, $G_f(\lambda)$ does cross zero at this "degenerated herding region". We will not refer such a degenerated herding region as an essential confounded learning point, since actually no type uses private information in deciding.

Furthermore, we could verify that the only case where the characteristic function crosses zero but doesn't introduce an essential confounded learning point shows up when parameters introduce a degenerated herding region. [17] *Given that those parameters that introduces a "degenerated herding region" won't be discussed at all. It doesn't threat our definition of essential confounded learning.*

It turns out that signal $\mathcal{E}$ is not necessarily non-empty even if the confounding region is non-empty and not completely covered by $[u, v]$ (or $[v, u]$). For example, we have

Lemma C.5 *If $p > \frac{1}{2}, v > u$, and we consider bounded signal set $F_{(\underline{s},\bar{s})}$, then $G_f(\lambda) \geq 0$ on $(0, u]$. Then, even if $v \geq u\frac{\bar{s}}{1-\bar{s}} > u > v\frac{\underline{s}}{1-\underline{s}}$, so that $(v\frac{\underline{s}}{1-\underline{s}}, u)$ is in the confounding region but not in $[u, v]$, signal set $\mathcal{E}$ is still empty.*

Proof. To show that $G_f(\lambda) \geq 0$ on $(0, u]$, we simply rewrite

$$G_f(\lambda) \;=\; (2p-1)\int_{\underline{s}}^{\frac{\lambda}{\lambda+v}} f(s)\frac{1-2s}{s}ds + p\int_{\frac{\lambda}{\lambda+v}}^{\frac{\lambda}{\lambda+u}} f(s)\frac{1-2s}{s}ds \qquad \text{(C.6)}$$

[17] If the characteristic function crosses zero at a point within the confounding region, then by definition this point is an essential confounded learning. On single active region, the characteristic function takes one sign and cannot crosses zero unless at boundary points. We could further verify that if a boundary point of a single active region is also a boundary point of a confounding region, then the characteristic function must be strictly positive or negative. If a boundary point of a single point is also a boundary point of a herding region, then the characteristic function crosses zero at this boundary point only if the herding region is "degenerated".

and observe that $\frac{\lambda}{\lambda+v} < \frac{\lambda}{\lambda+u} \leq \frac{1}{2}$ for $\lambda \in (0, u]$. Then the conclusion follows immediately. ■

We also have the following three similar results:

Lemma C.6 *If $p < \frac{1}{2}, v > u$, and we consider bounded signal set $F_{(\underline{s},\overline{s})}$, then $G_f(\lambda) \leq 0$ on $[v, +\infty)$. Then, even if $u\frac{\overline{s}}{1-\overline{s}} > v > v\frac{\underline{s}}{1-\underline{s}} \geq u$, so that $(v, u\frac{\overline{s}}{1-\overline{s}})$ is in the confounding region but not in $[u, v]$, signal set $\mathcal{E}$ is still empty.*

Lemma C.7 *If $p > \frac{1}{2}, u > v$, and we consider bounded signal set $F_{(\underline{s},\overline{s})}$, then $G_f(\lambda) \geq 0$ on $[u, +\infty)$. Then, even if $v\frac{\overline{s}}{1-\overline{s}} > u > u\frac{\underline{s}}{1-\underline{s}} \geq v$, so that $(u, v\frac{\overline{s}}{1-\overline{s}})$ is in the confounding region but not in $[v, u]$, signal set $\mathcal{E}$ is still empty.*

Lemma C.8 *If $p < \frac{1}{2}, u > v$, and we consider bounded signal set $F_{(\underline{s},\overline{s})}$, then $G_f(\lambda) \leq 0$ on $(0, v]$. Then, even if $u \geq v\frac{\overline{s}}{1-\overline{s}} > v > u\frac{\underline{s}}{1-\underline{s}}$, so that $(u\frac{\underline{s}}{1-\underline{s}}, v)$ is in the confounding region but not in $[v, u]$, signal set $\mathcal{E}$ is still empty.*

Recall that a model with bounded signal strength is parameterized by $(p, u, v, \underline{s}, \overline{s})$. The entire parameter space is

$$\mathbb{B} \equiv \Big\{p \in (0,1)\Big\} \times \Big\{\mathbb{R}_+^2 - \{v \neq u\}\Big\} \times \Big\{(0, \frac{1}{2}) \cup (\frac{1}{2}, 1)\Big\}. \tag{C.7}$$

The parameters that makes the signal set $\mathcal{E}$ empty has been listed out. It turns out that for each $(p, u, v, \underline{s}, \overline{s})$ in the complement, signal set $\mathcal{E}$ is non-empty. This follow directly from the follow lemma:

Lemma C.9

1. *If $p > \frac{1}{2}, u\frac{\overline{s}}{1-\overline{s}} > v > u$, there exists a private signal $f(s)$, which is C^∞ on $(\underline{s}, \overline{s})$ such that $G_f(v) > 0$ and $G_f(u\frac{\overline{s}}{1-\overline{s}}) < 0$.*

2. *If $p < \frac{1}{2}, v\frac{\underline{s}}{1-\underline{s}} < u < v$, there exists a private signal $f(s)$, which is C^∞ on $(\underline{s}, \overline{s})$ such that $G_f(u) < 0$ and $G_f(v\frac{\underline{s}}{1-\underline{s}}) > 0$.*

3. *If $p > \frac{1}{2}, u\frac{\underline{s}}{1-\underline{s}} < v < u$, there exists a private signal $f(s)$, which is C^∞ on $(\underline{s}, \overline{s})$ such that $G_f(v) > 0$ and $G_f(u\frac{\underline{s}}{1-\underline{s}}) < 0$.*

4. *If $p < \frac{1}{2}, v\frac{\overline{s}}{1-\overline{s}} > u > v$, there exists a private signal $f(s)$, which is C^∞ on $(\underline{s}, \overline{s})$ such that $G_f(u) < 0$ and $G_f(v\frac{\overline{s}}{1-\overline{s}}) > 0$.*

Therefore, for any parameter that satisfies one of the above four conditions, signal set $O \subset F_{(\underline{s},\overline{s})}$ is non-empty.

Proof. We only prove the first case. Other cases can be similarly proved. By lemma C.5, for parameters $p > \frac{1}{2}, v > u$, no confounded learning belief exists on $(0, u]$. So we only consider $\lambda \in [v, +\infty)$. It is direct to verify that $G_f(u\frac{\overline{s}}{1-\overline{s}}) < 0$. Furthermore, we can rewrite $G_f(\lambda)$ as

$$G_f(\lambda) = (2p-1)\int_{\frac{\lambda}{\lambda+u}}^{1} f(s)\frac{2s-1}{s}ds - (1-p)\int_{\frac{\lambda}{\lambda+v}}^{\frac{\lambda}{\lambda+u}} f(s)\frac{2s-1}{s}ds \tag{C.8}$$

and check that if $f(s) \equiv 0$ on $(\frac{1}{2}, \frac{v}{u+v})$, then $G_f(v) > 0$. Therefore, we turn to show there exists a signal $f(s) \in C^\infty(\underline{s}, \overline{s})$ satisfying that $f(s) = 0$ on $(\frac{1}{2}, \frac{v}{u+v})$.

This is equivalent to show the existence of two C^∞ non-negative functions $f_1(s)$ and $f_2(s)$ such that (1) $f_1(s)$ is defined on $[\underline{s}, \frac{1}{2}]$ with k-th derivatives $f_1^{(k)}(\frac{1}{2}) = 0$ for all k, (2) $f_2(s)$ is defined on $[\frac{v}{u+v}, \overline{s}]$ with k-th derivatives $f_2^{(k)}(\frac{v}{v+u}) = 0$, and

$$\int_{\underline{s}}^{\frac{1}{2}} f_1(s)\frac{1-2s}{s}ds = c_1(\frac{1}{2} - \underline{s}) = c_2(\overline{s} - \frac{v}{v+u}) = \int_{\frac{v}{v+u}}^{\overline{s}} f_2(s)\frac{2s-1}{s}ds. \tag{C.9}$$

Then $f(s) = \frac{\bar{f}(s)}{\int_{\underline{s}}^{\overline{s}} \bar{f}(s)ds}$ is the desired C^∞ private signal.

Existence of such $f_1(s), f_2(s)$ is guaranteed by smooth gluing lemma. ■

It is quite obvious that the set of parameters described in lemma C.9 is open in $\mathbb{B}$. So the following corollary is immediate:

Corollary C.10 *Among all the parameters $\mathbb{B}$ that describe an observational learning model with bounded signal strength, there exists a set U open in $\mathbb{B}$, such that for each $(p, u, v, \underline{s}, \overline{s}) \in U$, the associated set $\mathcal{E} \subset F_{(\underline{s},\overline{s})}$ is non-empty.*

C.2 Unbounded Signal Strength

For models with unbounded signal, we have the following lemma. This is essentially the same as the conclusion in lemma C.5 to C.8.

Lemma C.11 *We have*

1. *If $p > \frac{1}{2}, v > u$, then $G_f(\lambda) > 0$ on $\lambda \in (0, u]$.*

2. *If $p < \frac{1}{2}, v > u$, then $G_f(\lambda) < 0$ on $\lambda \in [v, \infty)$.*

3. *If $p > \frac{1}{2}, u > v$, then $G_f(\lambda) > 0$ on $\lambda \in [u, \infty)$.*

4. If $p < \frac{1}{2}, u > v$, then $G_f(\lambda) < 0$ on $\lambda \in (0, v]$.

Proof. The proof is essentially the same as the proof in lemma C.5. ■

Remark C.12 *Lemma C.11 also helps in characterizing the shape of confounded learning set. It says, for each parameter triple (p, u, v), the signal set $\mathcal{E}$ that admit essential confounded learning outside monotonic region actually can only have essential confounded learning on exactly one branch-either $(0, u)$ or $(v, +\infty)$. Therefore, in theorem 3.13, $\Lambda_f^{(0,u)}$ and $\Lambda_f^{(v,+\infty)}$ cannot both be non-empty. The parameter triple (p, u, v) determines which one is non-empty.*

A similar conclusion holds for models with bounded signal strength.

The following lemma is essentially the same as lemma C.9. It proves that signal set $\mathcal{E} \subset F_{(0,1)}$ is non-empty for each parameter triple (p, u, v) where $p \neq \frac{1}{2}$.

Lemma C.13 *1. If $p > \frac{1}{2}, v > u$, then there exists a private signal $f(s)$, which is C^∞ on $(0, 1)$ satisfying that there exists a $\bar{\lambda} > v$ such that $G_f(v) > 0$ and $G_f(\bar{\lambda}) < 0$.*

2. If $p < \frac{1}{2}, v > u$, there exists a private signal $f(s)$, which is C^∞ on $(0, 1)$ satisfying that there exists a $\bar{\lambda} < u$ such that $G_f(u) < 0$ and $G_f(\bar{\lambda}) > 0$.

3. If $p > \frac{1}{2}, u > v$, there exists a private signal $f(s)$, which is C^∞ on $(0, 1)$ satisfying that there exists a $\bar{\lambda} < v$ such that $G_f(v) > 0$ and $G_f(\bar{\lambda}) < 0$.

4. If $p < \frac{1}{2}, u > v$, there exists a private signal $f(s)$, which is C^∞ on $(0, 1)$ satisfying that there exists a $\bar{\lambda} > u$ such that $G_f(u) < 0$ and $G_f(\bar{\lambda}) > 0$.

Proof. We only prove for parameters $p > \frac{1}{2}, v > u$, other cases can be similarly proved.

For these parameters, by lemma C.11, we have learned that no confounded learning belief exists on $(0, u]$. So we only consider $\lambda \in [v, +\infty)$. As in the bounded signal case, we can verify that $G_f(v) > 0$ if $f(s) = 0$ on $(\frac{1}{2}, \frac{v}{u+v})$.

If private signal is unbounded, then we don't automatically obtain an point like $u\frac{\bar{s}}{1-\bar{s}}$ so that $G_f(\lambda)$ turns negative. But it is not difficulty to construct one: for $\overline{\lambda}$ sufficiently large, we observe that

$$G_f(\bar{\lambda}) = (2p - 1) \int_{\frac{\bar{\lambda}}{\bar{\lambda}+u}}^{1} f(s) \frac{2s - 1}{s} ds - (1 - p) \int_{\frac{\bar{\lambda}}{\bar{\lambda}+v}}^{\frac{\bar{\lambda}}{\bar{\lambda}+u}} f(s) \frac{2s - 1}{s} ds. \tag{C.10}$$

To make $G_f(\bar{\lambda}) < 0$, we just need to make $f(s)$ super small on $(\frac{\bar{\lambda}}{\bar{\lambda}+u}, 1)$ and relatively big on $(\frac{\bar{\lambda}}{\bar{\lambda}+u}, \frac{\bar{\lambda}}{\bar{\lambda}+v})$. The choice is quite flexible. We can use smooth gluing lemma to find a C^∞ signal $f(s)$ satisfy those requirements. ■

D Omitted Proofs in Subsection 3.4

In this section we prove theorem 3.11. We think this result is previously known but cannot find a good reference. We reproduce the proof in details to help readers.

We shall first review the standard construction of the cantor set, and then we shall see the set S can be constructed in a similar way. The construction shall give us a direct homeomorphism.

D.1 Construction of the Standard Cantor Set

The standard cantor set is constructed in the following steps:

1. Step 1: Remove the middle third open interval $(\frac{1}{3}, \frac{2}{3})$ from closed interval $[0, 1]$. The remaining set $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ is denoted as I^1.

2. Step 2: The set I^1 consists of two closed intervals. Remove the middle third open interval for each closed interval. That is, remove $(\frac{1}{9}, \frac{2}{9})$ from $[0, \frac{1}{3}]$, and remove $(\frac{7}{9}, \frac{8}{9})$ from $[\frac{2}{3}, 1]$. Denote the union of the remaining four closed intervals $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ as I^2.

3. Repeat this, after k steps, we are left with a closed set I^k which is the union of 2^k closed intervals. In step $k+1$, we shall remove the middle third open interval for each closed interval in I^k and denote what's left as I^{k+1}.

4. The standard cantor set $\mathcal{C}$ is the intersection of all I^ks:

$$\mathcal{C} = \cap_{k \in \mathbb{N}} I^k. \tag{D.1}$$

It is well-known that the standard cantor set $\mathcal{C}$ has lots of interesting properties. What is useful to us is that we could find a dyadical expression for it.

Let a word of length k be an ordered sequence of length k consists of 0s and 1s. For example, 01001 is a word of length five; 11 is a word of length 2. In other words, w_k is a word of length k if

$$w_k \in \{0,1\}^k. \tag{D.2}$$

We allow for a word of infinite length: w is a word of infinite length if

$$w \in \{0,1\}^\infty. \tag{D.3}$$

That the standard cantor set $\mathcal{C}$ has a dyadical expression means that we could label each point in $\mathcal{C}$ by an unique word of infinite length. That is

Proposition D.1 *There exists a bijection between $\mathcal{C}$ and $\{0,1\}^\infty$.*

The dyadical expression is constructed in the following steps:

1. Step 1: Take I^1, label the closed interval $[0, \frac{1}{3}]$ to the left of the middle third open interval as I_0; label the closed interval $[\frac{2}{3}, 1]$ to the right of the middle third open interval as I_1.

2. Step 2: Take $I^2 \cap I_0 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}]$, label the left interval $[0, \frac{1}{9}]$ as I_{00}, label the right interval $[\frac{2}{9}, \frac{1}{3}]$ as I_{01}; then take $I^2 \cap I_1 = [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, label the left interval $[\frac{2}{3}, \frac{7}{9}]$ as I_{10}, label the right interval $[\frac{8}{9}, 1]$ as I_{01};

3. Repeat this. After step k, we have labeled 2^k closed intervals in I^k each with a word of length k. In step $k+1$, we intersect each labeled closed interval I_{w_k} with I^{k+1}. We could see this intersection resulted in two disjoint closed intervals. [18] Label the left interval by $k+1$ word $w_k 0$, label the right interval by $k+1$ word $w_k 1$. Here $w_k 0$ means we concatenate w_0 with a 0. For example, if $k = 3$, and $w_3 = 010$, then $w_k 0 = 0100$.

4. Then, we observe that each infinite word w gives a sequence of descending closed intervals. If we denote $(w)_{T,k}$ to be the word w truncated by its first k digits-$(0101)_{T,1} = 0, (0110)_{T,3} = 011$-then the descending sequence is given as

$$I_{(w)_{T_1}} \supset I_{(w)_{T_2}} \supset \cdots \supset I_{(w)_{T_k}} \supset \ldots \tag{D.4}$$

From the construction, the length of any closed interval in I^k is $(\frac{1}{3})^k$, then by the nested closed interval theorem we know $\cap_k I_{(w)_{T,k}}$ is a singleton. By the construction of

[18] This is because in the construction of I^{k+1}, we remove a middle third open interval from I_{w_k}.

cantor set $\mathcal{C}$, we know this singleton must be one point in $\mathcal{C}$. So we have build a map $d : \{0,1\}^\infty \to \mathcal{C}$.

5. It is easy to verify that the map d must an injection. For any two different word w^1, w^2, [19] they must be different at some digits, say the k-th digit. Then

$$d(w^1) \in I_{(w^1)_{T,k}}, d(w^2) \in I_{(w^2)_{T,k}}. \tag{D.5}$$

   But $I_{(w^1)_{T,k}} \cap I_{(w^2)_{T,k}} = \emptyset$, since they are both in I^k but are labeled by different words.

6. We can also verify d is a surjection. Actually, for any $x \in \mathcal{C}$, x must intersect an unique interval $I_{w_k} \subset I^k$. This gives out an descending closed intervals whose intersection is x, and hence x has a dyadical expression.

D.2 A Dyadical Representation for a Nowhere Dense, Perfect, Bounded Set S

In the last subsection we have reviewed the construction of a standard cantor set, and its dyadical representation. In this subsection we shall see we can build a dyadical representation for any nowhere dense, perfect, bounded set S. By mapping any point $x \in S$ to the point in $\mathcal{C}$ with the same dyadical representation, we obtain a bijection between S and $\mathcal{C}$. In the next subsection we shall show this bijection is naturally a homeomorphism.

Let $\underline{c} = \inf S, \overline{c} = \sup S$. Our first observation is that S's complement must be a disjoint union of countable open intervals:

Lemma D.2 *$S^c = (-\infty, \underline{c})\dot{\cup}(\overline{c}, +\infty)\dot{\cup}_{k\in K} O_k$, where each O_k is an open interval in $(\underline{c}, \overline{c})$, and K is an infinite countable index set.*

Proof. That S is perfect implies that it is closed. So its complement must be an open set in $\mathbb{R}$. It is well-known that all open set in $\mathbb{R}$ is disjoint union of at most countably many open intervals. It is obvious that $(-\infty, \underline{c})$ and $(\overline{c}, +\infty)$ must be contained in S^c. We are left to prove that the index K is not finite. Assume not and denote the cardinality of K as n, then we can label all open intervals in $S^c - (-\infty, \underline{c}) - (\overline{c}, +\infty)$ from left to right. If there exist a $k_0 \le n-1$ such that $\sup O_{k_0} < \inf O_{k_0+1}$, then by definition $(\sup O_{k_0}, \inf O_{k_0+1}) \subset S$.

[19] We use subscript to show the length of a word, so now we use supscript for the index of a word.

This contradicts that S is nowhere dense. If there doesn't exist such a k_0, then

$$S = \{\inf O_k\}_{k \in K} \cup \overline{c}. \tag{D.6}$$

Then S is discrete. But this contradicts that S is perfect. ■

Now we know that $S = [\underline{c}, \overline{c}] - \dot{\cup}_{k \in K} O_k$. The next lemma describes the distribution of those open intervals $\dot{\cup}_{k \in K} O_k$ in $[\underline{c}, \overline{c}]$. Roughly speaking, this set of open intervals must be everywhere in $[\underline{c}, \overline{c}]$.

Lemma D.3 *For any closed interval $[c_1, c_2] \subset [\underline{c}, \overline{c}]$ with $c_1, c_2 \in S$, there must exists infinite countable open intervals in $\{O_k\}_{k \in K}$ that are contained in $[c_1, c_2]$.*

This is an immediate corollary of the previous lemma once we realize $S \cap [c_1, c_2]$ is still bounded, nowhere dense and perfect. The last preparation for the dyadical representation for S is the following:

Lemma D.4 *For any open interval O_k, $\inf O_k, \sup O_k \in S$.*

Proof. If $\inf O_k \notin S$, then it must be covered by another open interval $O_{k'}$. But this contradicts that $O_k \cap O_{k'} = \emptyset$. Similar for $\sup O_k$. ■

We shall construct a dyadical representation for S as following:

1. Step 1: Relabel the longest [20] open interval in $\dot{\cup}_{k \in K} O_k$ as O. Then $[\underline{c}, \overline{c}] - O$ consists of two closed intervals, label the left one as J_0 and the right one as J_1. Denote $J_0 \cup J_1$ as J^1.

2. Step 2: We see that $J_0 = [\underline{c}, \inf O]$, by lemma D.4 both endpoints are in S. By lemma D.3, J_0 must contain countable many open intervals in $\{O_k\}_{k \in K}$. Among all the open intervals that's contained in J_0, relabel the longest as O_0, then $J_0 - O_0$ consists of two closed intervals, label the left one as J_{00} and the right one as J_{01}. Similarly, among the open intervals that's contained in J_1, we can relabel the longest open interval as O_1. Then $J_1 - O_1$ consists of two disjoint closed intervals. We label the left one as J_{10} and the right one as J_{11}. Denote $J_{00} \cup J_{01} \cup J_{10} \cup J_{11} = J^2$.

3. Repeat, after step k, we have selected out $2^k - 1$ open intervals out of $\{O_k\}_{k \in K}$ and relabeled them by words of different lengths [21] . After removing these open intervals

[20] Randomly pick one if there are multiple longest open intervals

[21] If we consider the word of length 0 as the empty word, then in step k, the open interval are labeled by a word of length $k - 1$. In fact, for each closed interval $J_{w_{k-1}}$ constructed in step $k - 1$, we select the longest open interval contained in $J_{w_{k-1}}$ and label it as $O_{w_{k-1}}$.

from $[\underline{c}, \overline{c}]$, we are left with a closed set J^k, consisting of 2^k disjoint closed intervals, each of which is labeled by a word of length k. The endpoints for each closed interval J_{w_k} must either be endpoints of one open interval that has been labeled in previous steps or $\{\underline{c}, \overline{c}\}$. So by lemma D.3, each closed interval I_{w_k} in J^k must contain countably many open intervals in $\{O_k\}_{k\in K}$ non-trivially. Among these intervals, choose the longest and label it as O_{w_k}. Then $J_{w_k} - O_{w_k}$ consists of two closed intervals, label the left one as $J_{w_k 0}$ and the right one as $J_{w_k 1}$.

4. We can verify that

$$S = \cap_k J^k. \tag{D.7}$$

It is direct to verify that $\cap_k J^k = [\underline{c}, \overline{c}] - \dot{\cup}_k O_{w_{k-1}}$. By definition $S = [\underline{c}, \overline{c}] - \dot{\cup}_{k\in K} O_k$ [22] . The above process could only remove less open intervals in $\{O_k\}_{k\in K}$. Thus $S \subset \cap_k J^k$. We must verify the above process removes all the intervals in $\{O_k\}_{k\in K}$. Assume not, there exists an interval O_{k_0} that is not removed in the above process. Then there must exists an descending chain of closed intervals

$$J_{w_1} \supset J_{w_2} \supset \cdots \supset J_{w_k} \supset \dots \tag{D.8}$$

such that $\cap_k J_{w_k} \supset O_{k_0}$. Besides, in the formation of this descending chain, in each step we remove the longest open interval. That O_{k_0} is not removed implies that each open interval O_{w_k} that is removed in step $k+1$ must not be shorter than O_{k_0}. This is impossible since J_{w_1} must be of finite length.

5. In the last step, the dyadical representation of S is given just as in the case of cantor set. For each infinite word $w \in \{0,1\}^\infty$, $\cap_k J_{(w)_{T,k}}$ must either be a closed interval or a singleton. Since $\cap_k J_{(w)_{T,k}} \subset S$ and S is nowhere dense, $\cap_k J_{(w)_{T,k}}$ cannot be a closed interval. Therefore, we have a map $d_S : \{0,1\}^\infty \to S$ given by $d_S(w) = \cap_k J_{(w)_{T,k}}$. Following the same argument as in the case of cantor set, it is easy to verify d is bijective.

We close this subsection with an observation that is useful in the next subsection:

[22] We would like to remind readers that $\dot{\cup}_{k\in K} O_k$ needs not to be the same as $\dot{\cup}_k O_{w_{k-1}}$. Here $\{O_k\}_{k\in K}$ represents all the open intervals as in lemma D.2. But $\{O_{w_{k-1}}\}_{k\in\mathbb{N}}$ are just all the open intervals removed in the construction of dyadical representation of S.

Lemma D.5 *Given any descending chain of closed intervals*

$$J_{(w)_{T,1}} \supset J_{(w)_{T,2}} \supset \cdots \supset J_{(w)_{T,k}} \supset \dots \tag{D.9}$$

we must have

$$\lim_{k \to +\infty} m(J_{(w)_{T,k}}) = 0, \tag{D.10}$$

where $m(J_{(w)_{T,k}})$ *is the length of* $J_{(w)_{T,k}}$*. That is, the length of closed intervals in a descending chain must shrinks to* 0*.*

The proof is easy. If the length of closed intervals in any descending chain is bounded from below, the intersection of this descending chain cannot be a single of length 0.

D.3 A Nowhere Dense, Perfect, Bounded Set is Homeomorphic to the Standard Cantor Set

In the previous two subsections, we have constructed a bijection $d : \{0,1\}^\infty \to \mathcal{C}$ and a bijection $d_S : \{0,1\}^\infty \to S$. So $d \circ d_S^{-1} : S \to \mathcal{C}$ is obviously a bijection.

In this subsection we prove that

Theorem D.6 $d \circ d_S^{-1}$ *is a homeomorphism.*

Given a sequence of infinite word $\{w^k\}_{k\in\mathbb{N}}$, we say w^k converges to an infinite word w if for any n, there exists a k_n such that

$$(w^k)_{T,n} = (w)_{T,n}, \text{ for all } k \geq k_n. \tag{D.11}$$

The next lemma claims that the convergence of dyadical representation in the above sense is the same as pointwise convergence in $\mathcal{C}$ and S:

Lemma D.7

$$\lim_{k\to+\infty} d_S(w^k) = d_S(w) \Leftrightarrow \lim_{k\to+\infty} w^k = w \Leftrightarrow \lim_{k\to+\infty} d(w_k) = d(w). \tag{D.12}$$

Proof. If $\lim_{k\to+\infty} w^k = w$, by definition, for any n, $\{d(w^k)\}_{k\geq k_n}$ and $d(w)$ are contained in the same closed interval $I_{(w)_{T,n}}$. By construction, length of $I_{(w)_{T,n}}$ shrinks to 0. Thus,

$$\lim_{k\to+\infty} w^k = w \Rightarrow \lim_{k\to+\infty} d(w^k) = d(w).$$

On the other hand, if $\lim_{k\to+\infty} d(w^k) = d(w)$, then for any n, there exists a k_n such that all $\{d(w^k)\}_{k\geq k_n}$ are contained in $I_{(w)_{T,n}}$, this by definition says that $\lim_{k\to+\infty} w^k = w$.

The proof for $\lim_{k\to+\infty} d_S(w^k) = d_S(w) \Leftrightarrow \lim_{k\to+\infty} w^k = w$ is exactly the same. ■

It is direct to see that $d \circ d_S^{-1}$ and $d_S \circ d^{-1}$ are both continuous. Then theorem 3.11 follows directly. Besides, since $\inf S$ and $\inf \mathcal{C}$ share the same dyadical representation, we know that $d \circ d_S(\inf S) = \inf \mathcal{C}$. Similarly we know $d \circ d_S(\sup S) = \sup \mathcal{C}$.

E Omitted Proof in Subsection 3.6

In this section we aims at proving the following theorem:

Theorem E.1 *Real-analytic signals is dense in the set of integrable signals under L_1-norm. To be specific, given any $f(s) \in L_1(\underline{s}, \overline{s})$ satisfying:*

$$f(s) \geq 0; \int_{\underline{s}}^{\overline{s}} f(s)ds = 1; \int_{\underline{s}}^{\overline{s}} f(s)\frac{1-2s}{s}ds = 0 \tag{E.1}$$

then there exists a sequence $g_n(s) \in C^\omega(\underline{s}, \overline{s})$ satisfying

$$g_n(s) \geq 0, \int_{\underline{s}}^{\overline{s}} g_n(s)ds = 1, \int_{\underline{s}}^{\overline{s}} g_n(s)\frac{1-2s}{s}ds = 0 \tag{E.2}$$

and

$$\lim_{n\to+\infty} \|f(s) - g_n(s)\|_{L_1(\underline{s},\overline{s})} \to 0. \tag{E.3}$$

E.1 Proof Outline

We first introduce the following two Whitney approximation theorems.

Theorem E.2 (Whitney Smooth Approximation Theorem) *For any $f(s) \in C^0(s_1, s_2)$, and any $\eta(s) \in C^0(s_1, s_2)$ with $\eta(s) > 0$ on (s_1, s_2), there exists a $g(s) \in C^\infty(s_1, s_2)$ such* *that*

$$|g(s) - f(s)| < \eta(s), \forall s \in (s_1, s_2). \tag{E.4}$$

Theorem E.3 (Whitney Analytic Approximation Theorem) *For any $f(s) \in C^\infty(s_1, s_2)$ and any $\eta(s) \in C^0(s_1, s_2)$ with $\eta(s) > 0$ on (s_1, s_2), there exists a $g(s) \in C^\omega(s_1, s_2)$ such that*

$$|f^{(n)}(s) - g^{(n)}(s)| < \eta(s), \forall n \leq \frac{1}{\eta(s)} \tag{E.5}$$

on $s \in (s_1, s_2)$.

Furthermore, Lusin's theorem directly implies that

Theorem E.4 *Given any $f(s) \in L_1(s_1, s_2)$ and any $\eta > 0$, there exists $g(s) \in C^0(s_1, s_2)$ such that*

$$\|g(s) - f(s)\|_{L_1(s_1,s_2)} < \eta. \tag{E.6}$$

Therefore, given any integral signal $f(s) \in F_{(\underline{s},\overline{s})}$, theorem E.4 says that there exists a sequence of continuous functions approximate it. Furthermore, the Whitney approximation theorems guarantee that any continuous function can be uniformly approximated by real-analytic functions. Thus, any signal $f(s) \in F_{(\underline{s},\overline{s})}$ can be approximated by a sequence of real-analytic functions in $C^\omega(\underline{s}, \overline{s})$. However, to prove theorem E.1, we need to approximate $f(s)$ by a sequence of real-analytic signals, rather than just a sequence of real-analytic functions. Our strategy is to construct a convergent sequence of real-analytic signals based on the existing convergent sequence of real-analytic function. We briefly sketch the proof strategy as following:

First, we observe that $\frac{f(s)}{s} \in L_1(\underline{s}, \overline{s})$ given that $f(s)$ is a signal and satisfies that $\int_{\underline{s}}^{\overline{s}} f(s)\frac{1-2s}{s}ds = 0$. Therefore, we could find a sequence of real-analytic functions $\{\hat{g}_n(s)\} \in C^\omega(\underline{s}, \overline{s})$ satisfying $\|\hat{g}_n(s) - \frac{f(s)}{s}\|_{L_1} \to 0$. Then obviously $\|\hat{g}_n(s)(1-2s) - \frac{f(s)}{s}(1-2s)\| \to 0$, which further implies that $\int_{\underline{s}}^{\overline{s}} \hat{g}_n(s)(1-2s)ds \equiv \varepsilon_n \to 0$. Therefore, we could slightly adjust $\hat{g}_n(s)s$ so that the "signal property" is satisfied without destroying the convergence of $\hat{g}_n(s)s$ to $f(s)$. [23]

[23] Here we start with a sequence $\hat{g}_n(s)$ approaching $\frac{f(s)}{s}$ is for the convenience of obtaining $\|\hat{g}_n(s)(1 - 2s) - f(s)\frac{1-2s}{s}\|_{L_1} \to 0$. If we starting with a sequence $\|\hat{g}_n(s) - f(s)\|_{L_1} \to 0$, then we cannot immediately conclude that $\|\hat{g}_n(s)\frac{1-2s}{s} - f(s)\frac{1-2s}{s}\|_{L_1} \to 0$ if $\underline{s} = 0$.

Let us denote the slightly perturbed sequence as $\overline{g}_n(s)$. It satisfies the signal property, and approaches $\frac{f(s)}{s}$ in L_1-norm. We could be careful in the perturbation so that $\overline{g}_n(s)$ is in $C^\omega(\underline{s}, \overline{s})$. We turn to fix the point that $\overline{g}_n(s)$ could have negative values. We could observe that $\min\{\inf_{s\in(\underline{s},\overline{s})} \overline{g}_n(s), 0\}$, which is the lowest negative value of $\overline{g}_n(s)$, must vanish. Then we could slightly lift $\overline{g}_n(s)$ by adding proper multiple of a positive real-analytic function. We must be careful so that the "slightly lift" doesn't destroy the signal property. In subsection E.4, we design a special real-analytic function to guarantee this.

Let us denote the slightly lifted sequence as $\tilde{g}_n(s)$. We shall find it is real-analytic, non-negative, satisfying "signal property" and approach $\frac{f(s)}{s}$ in L_1-norm. The only problem is that $\tilde{g}_n(s)s$ doesn't integrate to 1. To fix this, we could just let $g_n(x) = \frac{\tilde{g}_n(x)x}{\int_{\underline{x}}^{\overline{s}} \tilde{g}_n(x)xdx}$. Then $\{g_n(s)\}$ is the sequence of real-analytic signals approaching $f(s)$ in L_1-norm.

The following subsections are organized as the following: in the first subsection we show that continuous signals is dense in integrable signals; in the second subsection we show real-analytic signals is dense in continuous signals. These two subsections are the main contents. In the fourth subsection we show the construction of an real-analytic function $h(s)$ that is used for slightly lift.

E.2 Continuous Signals is Dense in Integrable Signals

Given any integrable private signal $f(s)$, we have $\frac{f(s)}{s} \in L_1(\underline{s}, \overline{s})$. Then by theorem E.4, there exists a sequence $\{\hat{g}_n(x)\} \in C^0(\underline{s}, \overline{s})$ such that

$$\|\hat{g}_n(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})} < \frac{1}{n}. \tag{E.7}$$

Since $\frac{f(x)}{x} \geq 0$ on $(\underline{s}, \overline{s})$, we must have $|\hat{g}_n^+(x) - \frac{f(x)}{x}| \leq |\hat{g}_n(x) - \frac{f(x)}{x}|$. So $\hat{g}_n^+(x)$ converges to $\frac{f(x)}{x}$ in L_1-norm as well.

Obviously we have

$$\|\hat{g}_n^+(x)(1-2x) - \frac{f(x)}{x}(1-2x)\|_{L_1(\underline{s},\overline{s})} \leq \big(\sup_{s\in(\underline{s},\overline{s})} |1-2x|\big)\|\hat{g}_n^+(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})}, \tag{E.8}$$

which further implies that

$$\int_{\underline{s}}^{\overline{s}} \hat{g}_n^+(x)(1-2x)dx \to \int_{\underline{s}}^{\overline{s}} f(x)\frac{1-2x}{x}dx = 0. \tag{E.9}$$

Let us use ε_n to denote $\int_{\underline{s}}^{\overline{s}} \hat{g}_n^+(x)(1-2x)dx$, and find a polynomial $r(x)$ such that $\int_{\underline{s}}^{\overline{s}} r(x)(1-2x) = 1$. Let

$$\overline{g}_n(x) = \hat{g}_n^+(x) - \varepsilon_n r(x). \tag{E.10}$$

We have $\int_{\underline{s}}^{\overline{s}} \overline{g}_n(x)(1-2x)dx = 0$. Let

$$\delta_n = \min\{\inf_{x\in(\underline{s},\overline{s})} \overline{g}_n(x), 0\}. \tag{E.11}$$

We can easily verify that $0 > \delta_n > -|\varepsilon_n| \sup_{x\in(\underline{s},\overline{s})} |r(x)|$. So $\delta_n \to 0$.

Let $h(x)$ be as given in lemma **??**, and $\delta = \inf_{x\in(\underline{s},\overline{s})} h(x)$. Let

$$\tilde{g}_n(x) = \overline{g}_n(x) - \frac{\delta_n}{\delta} h(x). \tag{E.12}$$

It is direct to verify that $\tilde{g}_n(x) \geq 0$ on $(\underline{s}, \overline{s})$ and $\int_{\underline{s}}^{\overline{s}} \tilde{g}_n(x)(1-2x)dx = 0$. Furthermore, it is direct to compute that

$$\|\tilde{g}_n(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})} \leq \|\hat{g}_n^+(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})} + |\varepsilon_n|\|r(x)\|_{L_1(\underline{s},\overline{s})} + |\frac{\delta_n}{\delta}|\|h(x)\|_{L_1(\underline{s},\overline{s})}. \tag{E.13}$$

So $\tilde{g}_n(x)$ converges to $\frac{f(x)}{x}$ in L_1-norm, and hence $\|\tilde{g}_n(x)x - f(x)\|_{L_1(\underline{s},\overline{s})} \to 0$.

Lastly, we can verify that $g_n(x) = \frac{\tilde{g}_n(x)x}{\int_{\underline{s}}^{\overline{s}} \tilde{g}_n(x)xdx}$ are indeed continuous signals. So for each integrable signal $f(x)$, we can approach it by a sequence of continuous signals $g_n(x)$ in L_1-norm.

E.3 Real-analytic Signals is Dense in Continuous Signals

Given any continuous private signal $f(s) \in C^0(\underline{s}, \overline{s}) \cap F_{(\underline{s},\overline{s})}$ and satisfies conditions in equation E.1. For each n, we can define

$$\hat{f}_n(x) = \begin{cases} \inf_{s\in(x,\frac{1}{n}]} \frac{f(s)}{s}, x \in (\underline{s}, \frac{1}{n}], \\ \frac{f(x)}{x}, x \in [\frac{1}{n}, \overline{s}). \end{cases}$$

It is direct to verify that $\hat{f}_n(x)$ is continuous on $(\underline{s}, \overline{s})$. By the two Whitney approximation theorems, there exists a function $\hat{g}_n(x) \in C^\omega(\underline{s}, \overline{s})$ such that $|\hat{g}_n(x) - \hat{f}_n(x)| < \frac{1}{n}$ uniformly on $(\underline{s}, \overline{s})$. Our first observation is that

Claim E.5 $\|\hat{g}_n(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})} \to 0$.

Proof. We have

$$\|\hat{g}_n(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})} \leq \|\hat{g}_n(x) - \hat{f}_n(x)\|_{L_1(\underline{s},\overline{s})} + \|\hat{f}_n(x) - \frac{f(x)}{x}\|_{L_1(\underline{s},\overline{s})}$$

$$\leq \quad \frac{\overline{s}-\underline{s}}{n} + \int_{\underline{s}}^{\frac{1}{n}} |\hat{f}_n(x) - \frac{f(x)}{x}|dx \leq \frac{\overline{s}-\underline{s}}{n} + \int_{\underline{s}}^{\frac{1}{n}} \frac{f(x)}{x}dx. \tag{E.14}$$

The last inequality holds because $\frac{f(x)}{x} \geq \hat{f}_n(x)$ on $(\underline{s}, \frac{1}{n}]$ by definition. ■

Our second observation is that

Claim E.6 $\int_{\underline{s}}^{\overline{s}} \hat{g}_n(x)(1-2x)dx \to 0$.

Proof. We first show that $\{\hat{g}_n(x)(1-2x)\}_{n\geq 3}$ is uniformly dominated by an integrable function. By the definition of $\hat{g}_n(x)$, we have

$$|\hat{g}_n(x)(1-2x) - \hat{f}_n(x)(1-2x)| < \big(\sup_{x\in(\underline{s},\overline{s})} |1-2x|\big)|\hat{g}_n(x) - \hat{f}_n(x)| < \frac{1}{n}. \tag{E.15}$$

So $0 \leq |\hat{g}_n(x)(1-2x)| < |\hat{f}_n(x)(1-2x)| + \frac{1}{n}$ on $x \in (\underline{s}, \overline{s})$. It is direct to verify that $|\hat{f}_n(x)(1-2x)| \leq |f(x)\frac{1-2x}{x}|, x \in (\underline{s}, \overline{s})$ for all $n \geq 3$. So $\{\hat{g}_n(x)(1-2x)\}_{n\geq 3}$ is uniformly dominated by $|f(x)\frac{1-2x}{x}|$ + some constant, which is integrable. Furthermore, it is direct to verify that $\hat{g}_n(x)(1-2x)$ converges to $f(x)\frac{1-2x}{x}$ pointwisely for each $x \in (\underline{s}, \overline{s})$. So the result follows. ■

The rest construction is exactly the same. Let $\varepsilon_n = \int_{\underline{s}}^{\overline{s}} \hat{g}_n(x)(1-2x)dx$, and define $\overline{g}_n(x) = \hat{g}_n(x) - \varepsilon_n r(x)$. The define δ_n as $\min\{\inf_{x\in(\underline{s},\overline{s})} \overline{g}_n(x), 0\}$. We can verify that $0 > \delta_n > \min\{\inf_{x\in(\underline{s},\overline{s})} \hat{g}_n(x), 0\} - |\varepsilon_n| \sup_{x\in(\underline{s},\overline{s})} |r(x)| > -|\varepsilon_n| \sup_{x\in(\underline{s},\overline{s})} |r(x)| - \frac{1}{n}$. So $\delta_n \to 0$. Then define $\tilde{g}_n(x) = \overline{g}_n(x) - \frac{\delta_n}{\delta}h(x)$, which is non-negative. Finally, let $g_n(x) = \frac{\tilde{g}_n(x)x}{\int_{\underline{s}}^{\overline{s}} \tilde{g}_n(x)xdx}$. We could similarly check that $g_n(x)$ converges to $f(x)$ in L_1-norm. Because we choose $r(x), h(x)$ to be real-analytic, we also have $g_n(x)$ real-analytic.

E.4 Construction of Lifting Function $h(s)$

The main lemma in this section is

Lemma E.7 *Given $(\underline{s}, \overline{s}) \subset (0,1)$, for any $a < 0, b > 1$, there exists an $\alpha \in \mathbb{R}$ such that*

$$\int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx = 0, \ \ \textit{where } h(x) = \frac{1}{2}(x-a)(b-x)e^{\alpha(x-\frac{1}{2})}. \tag{E.16}$$

Let us recall the motivation of constructing $h(x)$ is to obtain an real-analytic function which is strictly positive on $(\underline{s}, \overline{s})$ so that we can use a positive multiple of $h(x)$ to eliminate possible negative values arising in the construction of convergence signals. Obviously the above $h(x)$ meets this needs.

Now let us briefly discuss the motivation of constructing such a $h(x)$. We observe that $h(x)(1-2x) = \frac{1}{2}(x-a)(1-2x)(b-x)e^{\alpha(x-\frac{1}{2})}$, which is positive on $(0, \frac{1}{2})$ and is negative on $(\frac{1}{2}, 1)$. Besides, for positive α, the term $e^{\alpha(x-\frac{1}{2})}$ enlarges the negative values and shrinks the positive values of $h(x)(1-2x)$; for negative α, the term $e^{\alpha(x-\frac{1}{2})}$ shrinks the negative values and enlarges the positive values of $h(x)(1-2x)$.

Proof of Lemma E.7. For notation abbrevity, we use $\overline{h}(s)$ to denote $\frac{1}{2}(x-a)(b-x)(1-2x)$. If $\alpha > 0$, then we must have

$$\begin{aligned}
& \int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx \\
= \ & \int_{\underline{s}}^{\frac{1}{2}} \overline{h}(x)e^{\alpha(x-\frac{1}{2})}dx + \int_{\frac{1}{2}}^{\overline{s}} \overline{h}(x)e^{\alpha(x-\frac{1}{2})}dx \\
\leq \ & \big(\sup_{x\in(\underline{s},\frac{1}{2}]} \overline{h}(x)\big) \int_{\underline{s}}^{\frac{1}{2}} e^{\alpha(x-\frac{1}{2})}dx + \int_{\frac{1}{2}}^{\overline{s}} \overline{h}(x)dx \\
= \ & \frac{1}{\alpha}[1-e^{\alpha(\underline{s}-\frac{1}{2})}]\big(\sup_{x\in(\underline{s},\frac{1}{2}]} \overline{h}(x)\big) + \int_{\frac{1}{2}}^{\overline{s}} \overline{h}(x)dx.
\end{aligned} \tag{E.17}$$

We could easily see that the first term goes to 0 as $\alpha \to +\infty$, and that the second term is strictly negative. So for big enough positive α, $\int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx$ is bounded above by a strictly negative value. Similarly, for negative α, we have

$$
\begin{aligned}
& \int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx \\
= & \int_{\underline{s}}^{\frac{1}{2}} \overline{h}(x)e^{\alpha(x-\frac{1}{2})}dx + \int_{\frac{1}{2}}^{\overline{s}} \overline{h}(x)e^{\alpha(x-\frac{1}{2})}dx \\
\geq & \int_{\underline{s}}^{\frac{1}{2}} \overline{h}(x)dx + \big(\inf_{x\in[\frac{1}{2},\overline{s}]} \overline{h}(x)\big)\int_{\frac{1}{2}}^{\overline{s}} e^{\alpha(x-\frac{1}{2})}dx \\
= & \frac{1}{\alpha}[e^{\alpha(\overline{s}-\frac{1}{2})}-1]\big(\inf_{x\in[\frac{1}{2},\overline{s}]} \overline{h}(x)\big) + \int_{\underline{s}}^{\frac{1}{2}} \overline{h}(x)dx
\end{aligned}
\tag{E.18}
$$

We observe that the first term goes to 0 as $\alpha \to -\infty$ and the second term is strictly positive. So for α that's close enough to $-\infty$, we must have $\int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx$ is bounded from below by a strictly positive number.

To summarize, $\int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx$ is positive for large negative α; and is negative for large positive α. As $\int_{\underline{s}}^{\overline{s}} h(x)(1-2x)dx$ is continuous in α, we must have some α makes the integral be 0. ■

F Smooth Gluing Lemma

In this paper, we often need to smoothly glue two piecewisely defined function together. That is, we have $f_l(s)$ defined on $[s_1, s_2]$ and $f_r(s)$ defined on $[s_3, s_4]$, and need to define $f_m(s)$ on $[s_2, s_3]$ so that

$$
f(s) = \begin{cases} f_l(s); s \in [s_1, s_2] \\ f_m(s); s \in [s_2, s_3] \\ f_r(s); s \in [s_3, s_4] \end{cases}
$$

is of C^k on $[s_1, s_4]$. Besides, we often need to pre-specify $\int_{s_2}^{s_3} f_m(s)ds$. This job is done by the following smooth gluing lemma. It basically says that on any close interval $[s_1, s_2]$, we can construct a non-negative C^∞ function with (almost) arbitrary endpoints behavior and (almost) arbitrary integral.

Lemma F.1 (Smooth Gluing Lemma) *Given any two sequences $\{a_n\}, \{b_n\}$ such that either (1) $a_0 > 0$ and $b_0 > 0$; or (2) $a_n = 0$ for all n and $b_0 > 0$; or (3) $b_n = 0$ for all n and*

$a_0 > 0$; then for

1. *any closed interval $[s_1, s_2]$;*

2. *any function $p(s) \in C^\infty[s_1, s_2]$ with $p(s) > 0$ on $[s_1, s_2]$;*

3. *any constant $c > 0$;*

there exists a non-negative function $k(s) \in C^\infty[s_1, s_2]$ such that

1. *$k^{(n)}(s_1) = a_n$, $k^{(n)}(s_2) = b_n$;*

2. *$\int_{s_1}^{s_2} k(s)t(s)dx = c(s_2 - s_1)$.*

F.1 Proof of Smooth Gluing Lemma

There are two parts to be proved: (1) matching endpoints behavior; (2) matching integral. We first discuss how to match endpoint behavior. In other words, given two sequences of numbers $\{a_n, b_n\}$ as in smooth gluing lemma F.1, how to construct a C^∞ function $k(x)$ on $[s_1, s_2]$ such that $k^{(n)}(s_1) = a_n, k^{(n)}(s_2) = b_n$.

The trivial case is $a_n = b_n = 0$ for all $0 \le n \le +\infty$. In this case, just let $k(x) = 0$ on $[s_1, s_2]$. A much complicated and general case is that $a_0 > 0$ and $b_0 > 0$. In this case, we turn to the following Besicovitch's theorem for help. (See Theorem 3.2.2 in Krantz and Parks (2002))

Theorem F.2 (Besicovitch) *Given two sequences $\{a_n\}_{n \in \mathbb{N}}$ and $\{b_n\}_{\mathbb{N}}$, there exists a smooth function $f(x)$ on $[0, 1]$ such that*

1. *$f(x)$ is real analytic on $(0, 1)$;*

2. *$f^{(n)}(0) = n!a_n, f^{(n)}(1) = n!b_n$.*

Besicovitch's theorem basically says that we could construct a function, being C^∞ on $[0, 1]$ and being real-analytic on $(0, 1)$, with arbitrary pre-specified endpoints behavior. With a simple scaling, Besicovitch's theorem seems to have done the job for us. However, in this paper we we want to keep the functions non-negative so that they could be positive multiple of density functions. Besicovitch's $f(x)$ is not guaranteed to be non-negative.

To remove possible negative values of $f(x)$, we multiply $f(x)$ with some function that we call smooth dent function [24] . A typical smooth dent function $d(x)$ is C^∞ and satisfy:

$$d(x)\begin{cases} = 1;\ \text{if } x \in [s_1, s_1+\varepsilon) \cup (s_2-\varepsilon, s_2] \\ = 0;\ \text{if } x \in (s_1+\varepsilon+\varepsilon', s_2-\varepsilon-\varepsilon') \end{cases}$$

for some small $\varepsilon, \varepsilon' > 0$. Given that $d(x) = 1$ around s_1, s_2, $d(x)f(x)$ has the same behavior as $f(x)$ has at s_1, s_2. Besides, as long as the possible negative values of $f(x)$ is within $(s_1+\varepsilon+\varepsilon', s_2-\varepsilon-\varepsilon')$ for proper parameters, $d(x)f(x)$ is guaranteed to be non-negative. As long as $f(s_1) > 0, f(s_2) > 0$, due to $f(x)$'s continuity, we must be able to find small enough ε and ε' so that $f(x) > 0$ on $s \in (s_1, s_1+\varepsilon+\varepsilon') \cup (s_2-\varepsilon-\varepsilon', s_2)$.

The third case is one of a_0, b_0 is 0. The smooth gluing lemma only deals with two special cases: $a_0 > 0$ and $b_n = 0, 0 \le n \le +\infty$; or $a_n = 0, 0 \le n \le +\infty$ and $b_0 > 0$. [25] Taking $a_n = 0$ for all n, $b_n > 0$ for example. In this case, there is no need to keep $d(x) = 1$ around s_1. We could multiply a smooth transition function

$$d(x)\begin{cases} = 1;\ \text{if } x \in (s_2-\varepsilon, s_2] \\ = 0;\ \text{if } x \in [s_1, s_2-\varepsilon-\varepsilon') \end{cases}$$

Here by letting $d(x) = 0$ around s_1 doesn't make endpoint behavior of $d(x)f(x)$ and $f(x)$ different, for the reason that $f(x)$ just need to be flat at s_1

Now we turn to discuss "matching integral". Given any $p(s) > 0$ on $[s_1, s_2]$ and $c > 0$, we must further modify the above $d(x)f(x)$ so that its integral equals to $c(s_2 - s_1)$. Take $a_0 > 0, b_0 > 0$ as an example, First, recall $d(x)$'s definition depends on $\varepsilon, \varepsilon'$ and $d(x)f(x)p(x) = 0$ except on at most two intervals with length $\varepsilon + \varepsilon'$. So by choosing $\varepsilon + \varepsilon'$ small enough, we can always obtain that

$$\int_{s_1}^{s_2} d(x)f(x)p(x)dx < c(s_2 - s_1). \tag{F.1}$$

Now we add a smooth bump function to $d(x)f(x)p(x)$. A typical smooth bump function is

[24] A smooth dent function can easily be obtained by transition of a smooth bump function. See chapter 2, section 4 in Lee (2000) for construction of a smooth bump function.

[25] Here we actually assume that if we want $k(x) = 0$ on one endpoints, then we also want $k^{(n)}(x) = 0$ for all n. This solves all the problems whenever we need smooth gluing lemma. Thanks to proposition B.7 (3). In fact, a problem could arise if we allow $b_0 = 0$ and $b_1 > 0$. In this case, any $k(x)$ such that $k(s_2) = 0$ and $k^{(1)}(s_2) > 0$ must satisfy that $k(x) < 0$ on $(s_2 - \varepsilon, s_2)$. Then by multiplying a smooth dent function we cannot remove all the negative values of $k(x)$.

a C^∞ function $b(x)$:

$$b(x) = \begin{cases} 0, x \in [s_1, s_1+\varepsilon) \cup (s_2-\varepsilon, s_2]; \\ h, x \in (s_1+\varepsilon+\varepsilon', s_2-\varepsilon-\varepsilon'). \end{cases}$$

By adjusting the height h, we could always have

$$\int_{s_1}^{s_2} [d(x)f(x)p(x)+b(x)]dx = c(s_2-s_1) = \int_{s_1}^{s_2} [d(x)f(x)+\frac{b(x)}{p(x)}]p(x)dx. \tag{F.2}$$

Since $b(x)$ is identically 0 around s_1, s_2, $d(x)f(x)+\frac{b(x)}{p(x)}$ has the same behavior as $d(x)f(x)$ at s_1, s_2. So $d(x)f(x)+\frac{b(x)}{p(x)}$ is the $k(x)$ that we want to construct.

G Other Omitted Proofs

Proof of Proposition 2.1. This follows from a direct computation. If $u=v$ and λ^* is a confounded learning belief, then

$$G_f(\lambda^*) = \Pr(b|\lambda^*, B) - \Pr(b|\lambda^*, A) = (2p-1)\int_{\underline{s}}^{\frac{\lambda^*}{\lambda^*+u}} f(s)\frac{1-2s}{s}ds = 0. \tag{G.1}$$

This certainly cannot hold if $p \neq \frac{1}{2}$. If $p=\frac{1}{2}$, then $\Pr(b|\lambda, B) = \Pr(b|\lambda, A) = \frac{1}{2}$ for all λ. In this case, any public history contains no information. So the long run belief stays at the prior. We prefer not to call such a long run belief the confounded learning belief, for the reason that in this case no learning happens at all. ■

Proof of Lemma 3.1. If we denote $G_1(x) = \int_{\underline{s}}^{x} f(s)\frac{1-2s}{s}ds$ and $G_2(x) = \int_{x}^{\overline{s}} f(s)\frac{1-2s}{s}ds$ on $x \in (\underline{s}, \overline{s})$, then

$$G_f(\lambda) = G_1(\frac{\lambda}{\lambda+u}) + G_2(\frac{\lambda}{\lambda+v}), \forall \lambda \in \mathcal{B}. \tag{G.2}$$

Since $G_1(x), G_2(x)$ are indefinite integral, they must be absolutely continuous. Furthermore, we could verify that $l_u(\lambda) = \frac{\lambda}{\lambda+u}, l_v(\lambda) = \frac{\lambda}{\lambda+v}$ are Lipschitz. Given that the composition of Lipschitz function with absolutely continuous function is again absolutely continuous, the conclusion follows. ■

Proof of Lemma 3.8. For any $\lambda > 0$, we have

$$\begin{aligned}
|G_{f_n}(\lambda) - G_f(\lambda)| &= |p \int_{\underline{s}}^{\frac{\lambda}{\lambda+u}} [f_n(s) - f(s)] \frac{1-2s}{s} ds + (1-p) \int_{\frac{\lambda}{\lambda+v}}^{\overline{s}} [f_n(s) - f(s)] \frac{1-2s}{s} ds| \\
&= |1-2p| |\int_{\frac{\lambda}{\lambda+v}}^{\overline{s}} [f_n(s) - f(s)] \frac{1-2s}{s} ds| \\
&\leq |1-2p| \|f_n - f\|_{L_1(\underline{s},\overline{s})} \max_{s \in [\frac{\lambda}{\lambda+v}, \overline{s}]} |\frac{1-2s}{s}|. \qquad \text{(G.3)}
\end{aligned}$$

Then the result follows directly. ■

Proof of Lemma 3.15. Let

$$G_1(x) = p \int_{\underline{s}}^{x} f(s) \frac{1-2s}{s} ds \quad ; \quad G_2(x) = (1-p) \int_{x}^{\overline{s}} f(s) \frac{1-2s}{s} ds. \qquad \text{(G.4)}$$

Then

$$G_f(\lambda) = G_1(\frac{\lambda}{\lambda+u}) + G_2(\frac{\lambda}{\lambda+v}). \qquad \text{(G.5)}$$

If $f(s)$ is analytic on $(\underline{s}, \overline{s})$, then as multiplication of two real-analytic functions, $f(s)\frac{1-2s}{s}$ is also analytic on $(\underline{s}, \overline{s})$. So for any $x^* \in (\underline{s}, \overline{s})$,

$$f(s) \frac{1-2s}{s} = \sum_{n=0}^{+\infty} a_n (s - x^*)^n \text{ on } [x^* - \rho, x^* + \rho]. \qquad \text{(G.6)}$$

Here we assume $[x^* - \rho, x^* + \rho]$ to be some compact interval within the interval of convergence, so that $\sum_{n=0}^{+\infty} a_n (s - x^*)^n$ uniformly converges to $f(s)\frac{1-2s}{s}$. Then for each $x \in [x^* - \frac{\rho}{2}, x^* + \frac{\rho}{2}]$, we have

$$\begin{aligned}
G_1(x) &= G_1(x^* - \frac{\rho}{2}) + \int_{x^* - \frac{\rho}{2}}^{x} \sum_{n=0}^{+\infty} a_n (s - x^*)^n dx \\
&= G_1(x^* - \frac{\rho}{2}) + \sum_{n=0}^{+\infty} \frac{a_n}{n+1} (x - x^*)^{n+1} - \sum_{n=0}^{+\infty} \frac{a_n}{n+1} (-\frac{\rho}{2})^{n+1}. \qquad \text{(G.7)}
\end{aligned}$$

Here we can integrate term by term since $\sum_{n=0}^{+\infty} a_n (x - x^*)^n$ is uniformly convergent on $[x^* - \rho, x^* + \rho]$. Besides, the radius of convergence of $\sum_{n=0}^{+\infty} \frac{a_n}{n+1} (x - x^*)^{n+1}$ is the same as $\sum_{n=0}^{+\infty} a_n (s - x^*)^n$. Therefore, formula $G.7$ expands $G_1(x)$ at x^* as an uniformly convergent series on $(x^* - \frac{\varrho}{2}, x^* + \frac{\varrho}{2})$. Hence $G_1(x)$ is real analytic on $(\underline{s}, \overline{s})$. Similarly we can prove that $G_2(x)$ is real analytic on $(\underline{s}, \overline{s})$ as well.

Recall that $l_u(\lambda) = \frac{\lambda}{\lambda+u}, l_v(\lambda) = \frac{\lambda}{\lambda+v}$. We can verify that $l_u^{-1}(\underline{s}, \overline{s}) \cap l_v^{-1}(\underline{s}, \overline{s})$ is the confounding region $\mathcal{B}$. Then for any $\lambda \in \mathcal{B}$, $G_1(\frac{\lambda}{\lambda+u})$ is the composition of two real-analytic functions and is hence real-analytic on $\mathcal{B}$. Similarly, $G_2(\frac{\lambda}{\lambda+v})$ is real-analytic on $\mathcal{B}$. So the conclusion follows. ■

References

Benavides, T. D. (1986): "How Many Zeros Does a Continuous Function Have?" *The American Mathematical Monthly*, 93, 464–466.

Krantz, S. G. and H. R. Parks (2002): *A Primer of Real Analytic Functions*, 2 ed.

Lee, J. M. (2000): *Introduction to Smooth Manifolds.*

Narasimhan, R. (1968): *Analysis on Real and Complex Manifolds.*

Smith, L. and P. Sørensen (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68, 371–398.