NBER WORKING PAPER SERIES

PERSUADING WITH ANECDOTES

Nika Haghtalab
Nicole Immorlica
Brendan Lucier
Markus Mobius
Divyarthi Mohan

ABSTRACT

We study a model of social learning and communication using hard anecdotal evidence. There are two Bayesian agents (a sender and a receiver) who wish to communicate. The receiver must take an action whose payoff depends on their personal preferences and an unknown state of the world. The sender has access to a collection of n samples correlated with the state of the world, which we think of as specific anecdotes or pieces of evidence, and can send exactly one of these samples to the receiver in order to influence her choice of action. Importantly, the sender's personal preferences may differ from the receiver's, which affects the seller's strategic choice of which anecdote to send.

We show that if the sender's communication scheme is observable to the receiver (that is, the choice of which anecdote to send given the set they receive), then they will choose an unbiased and maximally informative communication scheme, no matter the difference in preferences. Without observability, however, even a small difference in preferences can lead to a significant bias in the choice of anecdote, which the receiver must then account for. This can significantly reduce the informativeness of the signal, leading to substantial utility loss for both sides. One implication is informational homophily: a receiver can rationally prefer to obtain information from a poorly-informed sender with aligned preferences, rather than a knowledgeable expert whose preferences may differ from her own.

Nika Haghtalab
University of California, Berkeley
nika@berkeley.edu

Nicole Immorlica
Microsoft Research
nicimm@microsoft.edu

Brendan Lucier
Microsoft Research
brlucier@microsoft.com

Markus Mobius
School of Information
University of Michigan
4322 North Quad 105 S. State St.
Ann Arbor, MI 48109-1285
and NBER
markusmobius@gmail.com

Divyarthi Mohan
Princeton University
dm23@cs.princeton.edu

Persuading with Anecdotes[*]

Nika Haghtalab
UC Berkeley [†]

Nicole Immorlica
Microsoft Research [‡]

Brendan Lucier
Microsoft Research [§]

Markus Mobius
Microsoft Research [¶]

Divyarthi Mohan
Princeton University [‖]

Anecdotal thinking comes naturally; science requires training.

- Michael Shermer

Abstract

We study a model of social learning and communication using hard anecdotal evidence. There are two Bayesian agents (a sender and a receiver) who wish to communicate. The receiver must take an action whose payoff depends on their personal preferences and an unknown state of the world. The sender has access to a collection of n samples correlated with the state of the world, which we think of as specific anecdotes or pieces of evidence, and can send exactly one of these samples to the receiver in order to influence her choice of action. Importantly, the sender's personal preferences may differ from the receiver's, which affects the seller's strategic choice of which anecdote to send.

We show that if the sender's communication scheme is observable to the receiver (that is, the choice of which anecdote to send given the set they receive), then they will choose an unbiased and maximally informative communication scheme, no matter the difference in preferences. Without observability, however, even a small difference in preferences can lead to a significant bias in the choice of anecdote, which the receiver must then account for. This can significantly reduce the informativeness of the signal, leading to substantial utility loss for both sides. One implication is informational homophily: a receiver can rationally prefer to obtain information from a poorly-informed sender with aligned preferences, rather than a knowledgeable expert whose preferences may differ from her own.

1 Introduction

Economists usually assume that people learn about the world by updating the parameters of some underlying model as new evidence arrives. Such models can be efficiently communicated to others: for example, abstracts of academic papers might summarize main results in the form of model parameters such as the "elasticity of demand" in a certain industry or an overall toxicity score of a new radiation treatment. This type of communication is very natural for modeling learning amongst experts who have already agreed on a common set of models that provides a "language" for their field of study.

However, communication between experts can be incomprehensible to non-experts who have no understanding of such models. Such agents instead often rely on *anecdotal evidence*. Consider, for example, a

consumer/investor who is trying to decide how much money to invest in a given mutual fund. An optimal investment decision depends on the state of the future economy. If the economy is growing quickly, the investor would like to invest a lot of money; if it is growing slowly or shrinking, she may want to invest less. A typical investor lacks expertise in economic analysis and so relies on information from other actors, such as politicians, newspapers or financial analysts. This information is often provided in the form of anecdotes, say the percent increase in the number of jobs in a given sector this quarter. In fact, newspaper articles often simply report a selection of related facts on a topic.

Some sources of information are transparent about how they select the anecdotes they communicate. For example, a newspaper might credibly announce to report in an unbiased way because readers might cancel their subscriptions if they find out the newspaper does not stick to that strategy. On the other hand, a politician who seeks re-election might claim to be unbiased but then systematically only report anecdotes that support the view that the economy is booming.

We model these situations as a communication game between a sender and a receiver. Both players have to take an action and care about the state of the world (for example, the true state of the economy) and their *moral stance* which we think of their "ideal" action. The optimal actions of sender and receiver with different moral stances will generally differ even if they have access to the same information. In our model, only the sender observes informative signals about the state of the world which we refer to as anecdotes. She can select one these signals to send to the receiver. Importantly, we assume throughout that anecdotes are always *truthful*: the sender cannot make up "fake news", for example.

The sender in our model cares about the action of the receiver: she would like him to take an action close to hers. This creates a *persuasion temptation* for the sender: for example, if the politician would like to persuade the sender that the economy is booming so that he invests his money in the stock market, the politician might select a more positive anecdote. We assume that both sender and the receiver in our model are Bayesian. Hence, a Bayesian receiver will anticipate this bias and properly account for it. This might lead to unraveling where the politician chooses successively more biased anecdotes which the receiver adjusts for more and more. We show that even for small differences in moral stances between sender and receiver the anecdotes that are communicated can be highly biased when the sender communication strategy is not observable to the receiver.

In equilibrium, the politician will not be able change the mean posterior beliefs of her voters and the receiver will *always* learn *something* because anecdotes are always truthful. However, this does not imply that the sender's attempts to persuade the receiver are without costs. We show that when anecdotes follow a single-peaked distribution such as the normal or Laplace distribution then the most informative anecdotes are those close to the peak (which the sender approximates through her posterior mean). The intuition is that anecdotes are dense around the peak and are therefore very informative to the receiver. Senders whose communication strategy is transparent or who have the same moral stance as the receiver will select such anecdotes.

On the other hand, when moral stances differ and the sender cannot commit to an observable communication scheme then the sender will tend to choose anecdotes from the tail of the anecdote distribution. While this does not succeed in persuading the receiver to take a biased action it does destroy precision: anecdotes in the tail are more thinly distributed and hence reveal information about the true state of the world to the receiver with more noise.

This gives to "informational homophily" where receivers prefer to communicate with senders with similar moral stances because this eliminates the persuasion temptation of the sender. In fact, we show that a receiver might sometimes prefer to talk to a sender with access to just a few signals (or even just one) compared to a well-informed expert who has access to a vast number of anecdotes but has a different moral stance from the receiver. This will be the case for fat-tailed distributions (such as the Laplace distribution) because the sender can more likely access extreme anecdotes which are less informative to the receiver. This insight can explain why receivers might not seek experts with different background.

There is a rich literature in economics on sender-receiver games where the sender can send costless signals to the receiver. On a high-level, we can distinguish two themes in the literature, namely *framing* versus *selection*. Papers on framing give the sender a rich message space and the sender generally suppresses

some of the accuracy of her own signals when transmitting messages to the receiver if there is a conflict of interest between sender and receiver. In contrast, papers on selection restrict the message space of the sender to a subset of the set of signals that she received (which might include revealing everything or staying silent). Our paper belongs to the selection literature. However, our model also exhibits the tradeoff between persuasion and message accuracy that comes out of the framing literature - however, in our model this tradeoff arises not because the sender "fudges" a signal (as in the framing literature) but because the sender is tempted to select more extreme (and hence less informative) anecdotes.

Our paper is organized as follows. Section 2 discusses related work. Section 3 introduces our communication game and some basic analysis of our model that we will rely on later. Section 4 analyzes the case where the communication strategy of the sender is transparent to the receiver. Section 5 analyzes the case where the sender's strategy is not observable to the receiver and will only be guesses correctly in equilibrium. Section 6 concludes.
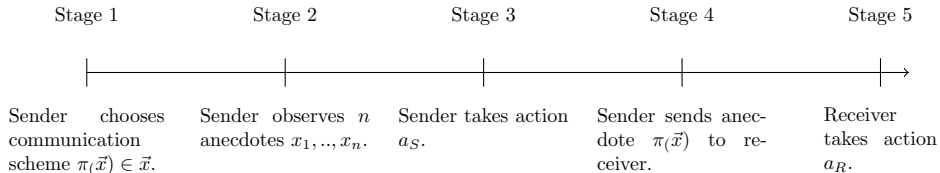
2 Related Work

We model is related to two strands of literature, namely papers on *framing* and papers on *selection*. The framing literature allows the sender to send any type of signal: the Bayesian persuasion literature assumes that the sender can commit to particular signaling scheme while the "cheap-talk" literature assumes no such commitment. Kamenica and Gentzkow (2011) introduce the Bayesian persuasion model where the sender commits to sending a signal that is consistent with her information in a Bayesian sense. In that model, as in ours, there is a state of the world distributed according to a common prior. The sender commits to a signaling scheme, mapping observations about the state of the world to an arbitrary signal. She then observes the state and transmits the corresponding signal to the receiver. The receiver then picks an action. The sender's payoff is a function of the receiver's action, and so the sender wishes to "persuade" the receiver to take particular actions. Kamenica and Gentzkow (2011) provide a characterization of the optimal signaling scheme. Our work approaches a similar question for a constrained signaling problem, where the sender's signaling scheme is restricted to take the form of sending one of a collection of anecdotes in a translation-invariant manner. This constraint imposes friction that limits the sender's ability to persuade, and indeed we find that the optimal choice of the sender under these restrictions will be to communicate as informatively as possible about the state of the world. Crawford and Sobel (1982) introduced the "cheap-talk" model where the sender has no commitment power. The sender can again choose an arbitrary signaling scheme, but is not able to commit to the signaling strategy in advance. Crawford and Sobel (1982) show that despite the lack of commitment, a non-trivial amount of information can be communicated at equilibrium, and moreover such equilibria take the form of sending a coarsening of the signal available to the sender.

Given our restriction that the sender must choose from a subset of available anecdotes, our work fits into the literature on selection. The literature on voluntary disclosure was introduced by Grossman and Hart (1980), Grossman (1981), and Milgrom (1981). These papers consider the setting of a seller who can choose whether to disclose information about a product to a buyer, and can make this choice based on the information itself. Similar to our model of communication via anecdotes, the seller in these papers cannot arbitrarily distort information about the product, but rather simply choose whether or not to reveal it. Importantly, the seller cannot necessarily commit to their revelation strategy in advance. The main result is that in every sequential equilibrium, the seller fully discloses her information. This so-called unraveling is driven by the fact that the seller can not commit to a signaling scheme.

Milgrom (1981) further analyze a setting where the seller is constrained in the amount of information she can reveal. Namely, she has access to a set of data points about her product (akin to our anecdotes) and can only reveal a fixed number of them (e.g., just one as in our setting). He then shows that the seller always reveals the most favorable information. We see a similar unraveling in our setting with an unobserved signaling scheme – the sender ends up sending an extreme signal, but not necessarily the most extreme signal (due to the structure of payoffs which differs from that of Milgrom (1981)). This contrasts to the setting of an observed signaling scheme, which can be interpreted as having no unraveling (the sender sends the most informative signal).

Figure 1: Timing of sender-receiver game

Stage 1	Stage 2	Stage 3	Stage 4	Stage 5

Sender chooses communication scheme $\pi_{(\vec{x})} \in \vec{x}$.

Sender observes n anecdotes $x_1, .., x_n$.

Sender takes action a_S.

Sender sends anecdote $\pi(\vec{x})$ to receiver.

Receiver takes action a_R.

There are other settings with partial unraveling. Dye (1985) and Jung and Kwon (1988) show unraveling breaks down when the receiver is uncertain whether the sender is informed. Martini (2018) shows unraveling breaks down when information is multi-dimensional. Ali et al. (2020) describe a disclosure game in which senders (buyers in their setting) gain by coarsening their information. Fishman and Hagerty (1990) consider a setting, similar to ours, in which the sender is restricted in the number of signals she can disclose and study the optimal amount discretion a designer should permit the sender.

In our paper, as in much of the literature, the unraveling, or partial unraveling, is driven by a lack of the power to commit on the part of the sender prior to observing the signals. A recent work also studies the power of the receiver to commit to a mapping from received information to sender payoff as in mechanism design (the difference being that the sender is restricted to voluntary disclosure strategies). Hart et al. (2017) define a disclosure game in which the equilibrium outcome without commitment coincides with the optimal outcome with commitment.

3 Model

Let us start with an overview of our model and agent interactions. We consider two agents, a sender and a receiver who play a communication game. Sender and receiver share a common prior μ about the state of the world θ. Both the sender and the receiver take actions $a_S \in \mathbb{R}$ and $a_R \in \mathbb{R}$ during this game that are determined both by their respective *moral stances*, m_S, m_R, and a state of the world θ that is unknown to both of them. The sender observes n i.i.d *anecdotes* (or signals) $x_1, .., x_n$ that are informative about the state of the world before taking her action a_S. Throughout, we denote $\vec{x} = \{x_1, \ldots, x_n\}$ to be the set of anecdotes observed by the sender. The sender sends one of her signals to the receiver according to *communication scheme* $\pi : \mathbb{R}^n \to \mathbb{R}$ such that $\pi(\vec{x}) \in \vec{x}$. The receiver then forms her own belief about the state of the world based on this anecdote $\pi(\vec{x})$ and her moral stance m_R and selects action a_R. When choosing a communication scheme the sender receives higher utility when the receiver's actions a_R is closer to the sender's action a_S. While the sender is restricted to sending an observed anecdote $\pi(\vec{x}) \in \vec{x}$, the choice of π can influence the receiver towards taking actions that have more utility to the sender.

For example, this model can capture communication between a politician (sender) who wants the economy (the state of the world) to do well to get re-elected. She communicates to a voter (receiver) whose moral stance might be neutral ($m_R = 0$) so that she is neither inclined to take very optimistic or pessimistic actions. The politician would like the voter to take actions that would suggest that the economy is doing better than it actually does and therefore might be tempted to reveal a very positive anecdote about the economy in her speeches that address the voter.

We next describe in detail the information sets and preferences of both agents in this model. For this purpose, we summarize the precise timing of the our game in Figure 1. In stage 1, the sender chooses the *communication scheme* π that determines what type of information the sender will reveal to the receiver. In stage 2, the sender observes n anecdotes $\vec{x}$. The sender then takes an action a_S in stage 3 that is not observed by the receiver. In stage 4 the sender selects one of her anecdotes according to her communication scheme and sends this anecdote $\pi(\vec{x})$ to the receiver. The receiver uses that anecdote to inform her own action a_R in stage 5.

4

Stage 1. The sender selects a *communication scheme* $\pi : \mathbb{R}^k \to \mathbb{R}$ such that $\pi(\vec{x}) \in \vec{x}$. This describes what anecdote the sender will communicate to the receiver for any vector of anecdotes $\vec{x} = (x_1, .., x_n)$. Our main results focus on a natural class of communication schemes that we call *translation invariant schemes*.

Definition 1. *For a common prior μ over θ, we call a communication scheme π_μ translation invariant if for all $\vec{x} \in \mathbb{R}^k$ and finite $d \in \mathbb{R}$ we have $\pi_{\mu \oplus d}(\vec{x} + d) = \pi_\mu(\vec{x}) + d$, where $\vec{x} + d := \{x_1 + d, \ldots, x_n + d\}$ and $\mu \oplus d$ is the translation of prior μ by d, i.e., $(\mu \oplus d)(\theta + d) = \mu(\theta)$. Note that when μ is a diffuse prior, $\mu \oplus d = \mu$ and we drop subscript of μ for brevity.*

In other words, translation-invariance implies that the choice of anecdote is independent of where where 0 is on the real line. Examples of translation invariant communication schemes include, sending the anecdote closest to the posterior mean of μ given $\vec{x}$. Similarly, always sending the minimum or maximum signal are translation invariant schemes. On the other hand, sending the minimum signal when all anecdote realizations are negative and the maximum signal otherwise is not a translation invariant scheme.

The sender cares about the receiver's action in the sense that she wants a_R to be as close as possible to her own action. Formally, she chooses her communication scheme to minimize the following *disagreement loss* function:

$$\mathbb{E}_{\vec{x}}(a_S(\vec{x}) - a_R(\pi(\vec{x}), \pi_R))^2 \tag{1}$$

Note, that the sender takes the expectation over all possible realizations of anecdotes. Her own action a_S will depend on all the anecdotes she observes. In contrast, the receiver's action a_R depends on the anecdote $\pi(\vec{x})$ that the sender decides to transmit as well as π_R, which is the communication scheme that the receiver expects the sender to use.

We consider two different environments. In the *observable regime* the receiver observes the sender's communication scheme and we simply have $\pi_R = \pi$. In contrast, in the *non-observable regime* the receiver does not observe π – instead we consider the choices of π_R and π in equilibrium.

Stage 2. Sender and receiver share a common *diffuse prior* about the state of the world $\theta \sim N(0, \infty)$. The sender observes a vector of conditionally i.i.d. *anecdotes* $\vec{x} = (x_1, \ldots, x_n)$ where each anecdote x is generated by adding noise ϵ drawn from some finite-variance distribution F to the true state of the world θ such that $x = \theta + \epsilon$.

Three examples of anecdote distributions F we will refer to later in the paper are the uniform distribution over $[-1, 1]$ (a distribution with finite domain), the normal distribution $N(0, 1)$ (a distribution with infinite domain and "thin tails") and the Laplace distribution with density $f(\epsilon) = \frac{1}{2} \exp(-|x|)$ (a distribution with infinite domain and "fatter tails").

Stage 3. The sender cares both about the state of the world θ as well as her *moral stance* $m_S \in \mathbb{R}$ when choosing an action. Formally, she will perform the following quadratic loss minimization:

$$a_S \in \mathrm{argmin}_a \, \mathbb{E}_{\theta|\vec{x}} \left[\alpha(a - \theta)^2 + (1 - \alpha)(a - m_S)^2 \right] \tag{2}$$

Here, $\alpha \in (0, 1)$ captures how the sender trades off the importance of the state of the world and her moral stance. Note that we assume that the sender chooses her action without taking into account how it affects the disagreement loss from stage 1. We make this assumption to simplify our analysis but it can be justified, for example, by assuming that the sender's total utility is a weighted average of her instrumental utility (captured by the loss function in equation 2) and the disagreement loss and then placing most of the weight on the instrumental utility part.

We note that the solution to the minimization problem in Equation 2 has a closed form. In particular, let $D_S(\vec{x})$ denote the sender's posterior belief about the state of the world θ based on her observed set of anecdotes $\vec{x}$. Let the sender's posterior mean of this distribution be $\theta_S(\vec{x})$. The sender's optimal action can then be derived as:

$$a_S(\vec{x}) = \alpha \theta_S(\vec{x}) + (1 - \alpha)m_S \tag{3}$$

Stage 4. The sender now sends anecdote $\pi(\vec{x})$ to the receiver.

Stage 5. The receiver forms a posterior belief after observing the signal $x = \pi(\vec{x})$ received from the sender. While the signal x is observed by the sender, the communication scheme π may or may not be directly observable. We say that the receiver believes that the sender uses communication scheme π_R – if the sender's scheme is observable then $\pi_R = \pi$; otherwise this equality pins down the equilibrium communication scheme of the sender. We denote the receiver's posterior belief with $D_R(x, \pi_R)$. The receiver forms this posterior by Bayesian updating, given the diffuse prior over θ and the distribution over $\pi_R(\vec{x})$ given that $\vec{x}$ is drawn according to conditionally i.i.d. draws from F. We use $\theta_R(x, \pi_R)$ to describe the posterior mean of the receiver.

The receiver has her own moral stance m_R which can differ from the receiver's moral stance. She faces the same type of loss function as the sender when choosing an action and therefore her optimal action will be:

$$a_R(x, \pi_R) = \alpha \theta_R(x, \pi_R) + (1 - \alpha) m_R \tag{4}$$

3.1 Preliminary Analysis

We now derive a few insights for our model that hold for both the observable and non-observable regimes.

First of all, we define the *bias* $\delta(\pi)$ of a communication scheme π to be the mean difference between the posterior mean $\theta_S(\vec{x})$ and the reported anecdote :

$$\delta(\pi) = \mathop{\mathbb{E}}_{\vec{x} \sim F^n} \left[\pi(\vec{x}) - \theta_S(\vec{x}) \right],$$

where $F^n = \prod_{i=1}^n F$ is the product distribution.

Our first observation is that the posterior mean of the receiver, given an observed signal x and receiver's presumed communication scheme π_R, is simply a debiasing shift of signal x.

Lemma 1. *For any translation invariant communication scheme π_R and signal x in the range of π_R, $\theta_R(x, \pi_R) = x - \delta(\pi_R)$.*

Proof: Fix translation invariant scheme π, and write y for the random variable $\pi(\vec{x}) - \theta$, the difference between the anecdote returned by π and the state θ. Since π is translation invariant, the distribution of y is independent of θ. Write f_π for its density function. Then by construction we have $\delta(\pi) = \int_y y f_\pi(y) dy$.

Now fix anecdote x and suppose $\pi(\vec{x}) = x$. Given this realization of x, the conditional distribution over the state of the world $\theta = x - y$ has density

$$\frac{f_\pi(y) g(x - y)}{\int_y f_\pi(y) g(x - y) dy}$$

where g is the unconditional density of θ. As θ is drawn from a diffuse prior, we can treat g as a constant function. The conditional distribution over $\theta = x - y$ therefore has density $f_\pi(y)$. The expected state of the world is therefore $x - \int_y y f_\pi(y) dy$ which is simply $x - \delta(\pi)$.

This helps us to decompose the sender's disagreement loss in stage 1 when choosing the optimal communication scheme π_S:

$$
\begin{aligned}
\mathop{\mathbb{E}}_{\vec{x}}(a_S(\vec{x}) - a_R(\pi(\vec{x}), \pi_R))^2 &= \mathop{\mathbb{E}}_{\vec{x}} \left[(\alpha \theta_S(\vec{x}) + (1 - \alpha) m_S) - (\alpha \theta_R(\pi(x), \pi_R) + (1 - \alpha) m_R) \right]^2 \\
&= \alpha^2 \mathop{\mathbb{E}}_{\vec{x}} (\theta_S(\vec{x}) - \theta_R(\pi(\vec{x}), \pi_R))^2 + \\
&+ 2\alpha(1 - \alpha) \mathop{\mathbb{E}}_{\vec{x}} (\theta_S(\vec{x}) - \theta_R(\pi(\vec{x}), \pi_R))(m_S - m_R) + \\
&+ (1 - \alpha)^2 (m_S - m_R)^2
\end{aligned}
\tag{5}
$$

6

We know from lemma 1 that $\theta_R(\pi(\vec{x}), \pi_R) = \pi(\vec{x}) - \delta(\pi_R)$. Hence we have:

$$
\begin{aligned}
\underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \theta_R(\pi(\vec{x}), \pi_R)) &= \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta(\pi_R)) \\
&= \delta(\pi_R) - \delta(\pi)
\end{aligned}
$$

We denote the difference between the sender's and receiver's moral stance with $\Delta = m_S - m_R$ which allows us to simplify expression 5:

$$
\begin{aligned}
\underset{\vec{x}}{\mathbb{E}}(a_S(\vec{x}) - a_R(\pi(\vec{x}), \pi_R))^2 &= \alpha^2 \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \theta_R(\pi(\vec{x}), \pi_R))^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta + (1-\alpha)^2\Delta^2 \\
&= \alpha^2 \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta(\pi_R))^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta + (1-\alpha)^2\Delta^2 \\
&= \underbrace{\alpha^2 \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta(\pi))^2}_{\text{signaling loss}} + \\
&+ \underbrace{\alpha^2\left(\delta(\pi_R) - \delta(\pi)\right)^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta}_{\text{persuasion temptation}} + \\
&+ \underbrace{(1-\alpha)^2\Delta^2}_{\text{Fundamental loss}} \quad\quad\quad\quad\quad (6)
\end{aligned}
$$

We can now see that the disagreement loss has 3 components. The last component is the *fundamental loss* that arises because the sender's and receiver's moral stances differ: it is constant and cannot be affected by the sender. We can therefore ignore this term in our analysis.

The second component is the *persuasion temptation*. It is always equal to 0 in the observable regime because $\delta(\pi_R) = \delta(\pi)$. However, in the non-observable regime the sender can always reduce her disagreement loss slightly off-equilibrium by either increasing the bias $\delta(\pi)$ of her communication scheme (when $\Delta > 0$) or decreasing the bias (when $\Delta < 0$): effectively this will move the receiver's action closer to hers.[1]

The first term is the signaling loss that arises from the fact that the sender can only transmit a single coarse anecdote but not her posterior mean: this expression is the square distance between the sender's posterior mean and the bias-corrected posterior of a receiver who uses the correct communication scheme π. While we know that the expected distance $\mathbb{E}[\theta_S(\vec{x} - \pi(\vec{x}) - \delta(\pi))]$ is 0 in expectation the variance is generally bounded away from 0.

Proposition 1. *In the observable regime the sender's optimal communication scheme minimizes her signaling loss. Therefore, her optimal communication scheme does not depend on the moral stances of the sender or receiver.*

In the unobservable regime, on the other hand, the receiver trades off her persuasion temptation against increases in the signaling loss: the sender is tempted to always slightly increase the bias of her communication scheme in order to move the receiver's action close to hers. This results in the sender selecting a tail signal which has higher variance for distributions such as the normal or Laplace distribution: intuitively, signals in the tail are more thinly distributed and therefore the signaling "error" tends to be larger. This increases the sender's signaling loss. This unraveling will stop when the marginal increase in the signaling loss exceeds the marginal temptation benefit.

Proposition 2. *In the unobservable regime the sender's optimal communication scheme is equal to the optimal scheme in the observable regime when sender and receiver have the same moral stances. Otherwise, the sender's signaling loss is weakly worse off in the non-observable regime compared to the observable regime.*

When sender and receiver share the same moral stance the temptation to persuade disappears and the sender simply minimizes the signaling loss just as in the observable regime. Otherwise, in equilibrium the

---

[1]To see this, assume that $\delta(\pi_R) = \delta(\pi)$ and then either increase or decrease the bias $\delta(\pi)$. The first term of the persuasion temptation is quadratic and hence second-order while the second term is linear in the bias and hence first-order.

temptation loss is 0 and the sender's utility is solely determined by the signaling loss. However, the signaling loss is minimized in the observable regime and therefore the sender cannot be better off in the non-observable regime.

3.2 Characterization of Optimal Communication Scheme

We now characterize the optimal communication scheme precisely. To that end, we introduce the notion of the *center* of a communication scheme $\pi(\vec{x})$.

Definition 2. *The center $r(\delta)$ of a communication scheme $\pi(\vec{x})$ with bias δ is defined as the offset from the posterior mean $\theta_S(\vec{x})$ such that when one selects the anecdote that is closest to $\theta_S(\vec{x}) + r(\delta)$ the mean value is $\theta_S(\vec{x}) + \delta$.*

One might naively expect that the center is exactly δ: however, this is generally not true. For example consider a scheme with a negative bias $\delta < 0$ and symmetric and single-peaked signal distribution F. In that case, the density of the anecdotes is increasing at δ and therefore the expected value of the closest anecdote will be greater than δ. In this case, we will generally observe $r(\delta) < \delta$ for $\delta < 0$. One important exception is the case of symmetric, single-peaked distributions with zero bias $\delta = 0$. In that case, it is easy to see that $r(\delta) = 0$.

We can now characterize the optimal translation-invariant communication scheme.

Theorem 1. *The optimal translation invariant communication scheme π with bias δ selects the closest signal to $\theta_S(\vec{x}) + r(\delta)$.*

Proof: We decompose the signaling loss from equation 5:

$$
\begin{aligned}
\alpha^2 \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta(\pi))^2 &= \alpha^2 \underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta(\pi) + r(\delta) - r(\delta))^2 \\
&= \alpha^2 [\underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta))^2 + (\delta(\pi) - r(\delta))^2 + \\
&\quad + 2 \underbrace{\underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta))(\delta(\pi) - r(\delta))]}_{=r(\delta) - \delta(\pi)} \\
&= \alpha^2 \left[\underset{\vec{x}}{\mathbb{E}}(\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta))^2 - (\delta(\pi) - r(\delta))^2 \right] \quad (7)
\end{aligned}
$$

Let's fix the bias δ and find the optimal communication scheme that minimizes the sender's objective Equation 5. It is easy to see that this scheme has to minimize $\mathbb{E}_{\vec{x}}(\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta))^2$ which occurs when we choose whatever anecdote is closest to the center $r(\delta)$. This scheme has exactly the right bias δ by construction. This shows that it is the optimal scheme with bias δ.

Note, that not all biases δ can also be implemented. For example, when F is symmetric and senders have only access to a single anecdote ($n = 1$) the only implementable scheme is the unbiased scheme. More generally, it is easy to see that the largest implementable negative bias is the *minimum scheme* (or the maximum scheme for positive biases) where the sender always selects the smallest anecdote (largest anecdote). We denote the bias of the minimum and maximum schemes with δ_{min} and δ_{max}, respectively.

Lemma 2. *The range of implementable biases is $R = [\delta_{min}, \delta_{max}]$. Any bias $\delta \in (\delta_{min}, \delta_{max})$ can be implemented by a communication scheme with some center $r(\delta)$. For $r \to -\infty$ $(r \to \infty)$ the bias of the communication scheme converges to δ_{min} (δ_{max}). For $n \geq 2$ the minimum (maximum) bias can be implemented by a center-based scheme if and only if the distribution F is bounded on the left (bounded on the right).*

The proof is immediate: if the distribution is bounded on the left by some number a, for example, then any $r \leq a$ implements the minimum scheme. On the other hand, if F is unbounded on the left and there are at least two anecdotes then there are always realizations where the center-based scheme selects the larger of the two signals.

3.3 Receiver's Loss

There is a close relationship between the signaling loss of the sender and the loss of the receiver. To see this consider the definition of the receiver's loss for taking an action a:

$$\text{receiver's loss} \quad = \quad \mathbb{E}_{\theta | \pi_R(\vec{x}), \pi_R} \left[\alpha(a - \theta)^2 + (1 - \alpha)(a - m_R)^2 \right]$$

Fix some communication scheme π of the sender, and for notational convenience write θ_R for the random variable $\theta_R(\pi(\vec{x}), \pi)$ where the randomness is over the realization of $\vec{x}$. We can then plug in the receiver's optimal choice of action $a = \alpha\theta_R + (1 - \alpha)m_R$ and take expectations *with respect to the sender's information set at stage 1 before she observes the anecdotes.* This expectation is over the realization of θ and the observed anecdotes $\vec{x}$.

$$
\begin{aligned}
\text{receiver's loss} \quad = \quad & \mathbb{E}\left[\alpha\left(\alpha\theta_R + (1-\alpha)m_R - \theta\right)^2 + (1-\alpha)\left(\alpha\theta_R + (1-\alpha)m_R - m_R\right)^2 \right] \\
= \quad & \mathbb{E}\left[\alpha\left(\alpha\theta_R + (1-\alpha)(m_R - \theta + \theta) - \theta\right)^2 + (1-\alpha)\alpha^2(m_R - \theta_R)^2 \right] \\
= \quad & \mathbb{E}\left[\alpha\left((1-\alpha)(m_R - \theta) + \alpha(\theta_R - \theta)\right)^2 + (1-\alpha)\alpha^2(m_R - \theta_R)^2 \right] \\
= \quad & \alpha(1-\alpha)^2\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^3\,\mathbb{E}\left[(\theta_R - \theta)^2\right] + 2\alpha^2(1-\alpha)E_\theta\left[(m_R - \theta)\underbrace{\mathbb{E}_{\vec{x}|\theta}(\theta_R - \theta)}_{=0}\right] + \\
+ \quad & (1-\alpha)\alpha^2\,\mathbb{E}\left[(m_R - \theta_R)^2\right] \\
= \quad & \alpha(1-\alpha)^2\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^3\,\mathbb{E}\left[(\theta_R - \theta)^2\right] + (1-\alpha)\alpha^2\,\mathbb{E}\left[(m_R - \theta + \theta - \theta_R)^2\right] \\
= \quad & \alpha(1-\alpha)^2\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^3\,\mathbb{E}\left[(\theta_R - \theta)^2\right] + (1-\alpha)\alpha^2\,\mathbb{E}\left[(m_R - \theta)^2\right] \\
+ \quad & (1-\alpha)\alpha^2\,\mathbb{E}\left[(\theta - \theta_R)^2\right] + 2(1-\alpha)\alpha^2\,\mathbb{E}[(m_R - \theta)\underbrace{\mathbb{E}_{\vec{x}|\theta}(\theta - \theta_R)}_{=0}] \\
= \quad & \alpha(1-\alpha)\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^2\,\mathbb{E}\left[(\theta - \theta_R)^2\right] \\
= \quad & \alpha(1-\alpha)\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^2\,\mathbb{E}\left[(\theta - \theta_S(\vec{x}) + \theta_S(\vec{x}) - \theta_R)^2\right] \\
= \quad & \alpha(1-\alpha)\,\mathbb{E}\left[(m_R - \theta)^2\right] + \alpha^2\,\mathbb{E}(\theta - \theta_S(\vec{x}))^2 + \alpha^2\,\mathbb{E}\left[(\underbrace{\theta_S(\vec{x}) - \theta_R}_{=\theta_S(\vec{x}) - \pi(\vec{x}) + \delta_S})^2\right] + \\
+ \quad & 2\alpha^2 E_{\vec{x}}[(\theta_S(\vec{x}) - \theta_R)\underbrace{\mathbb{E}_{\theta|\vec{x}}(\theta - \theta_S(\vec{x}))}_{=0}] \\
= \quad & \underbrace{\alpha(1-\alpha)\,\mathbb{E}\left[(m_R - \theta)^2\right]}_{\text{fundamental loss}} + \underbrace{\alpha^2\,\mathbb{E}\left[(\theta - \theta_S(\vec{x}))^2\right]}_{\text{sender uncertainty}} + \underbrace{\alpha^2\,\mathbb{E}\left[(\theta_S(\vec{x}) - \pi(\vec{x}) + \delta_R)^2\right]}_{\text{sender's signaling loss }(\delta_R = \delta_S)} \quad (8)
\end{aligned}
$$

The expected receiver's loss (evaluated at stage 1) has three components. The *fundamental loss* is the unavoidable loss because the receiver trades off her moral stance against the desire to take an action close to the state of the world (and she can satisfy neither perfectly). The *sender uncertainty* captures the loss from the sender not knowing the true state but only the posterior mean $\theta_S(\vec{x})$. Neither the fundamental loss nor the sender uncertainty are influenced by the sender's communication scheme.

The third term is the loss that arises because the receiver only learns about the sender's posterior mean $\theta_S(\vec{x})$ from a single anecdote and is exactly equal to the sender's signaling loss. We have therefore established the following relation between the receiver's expected loss and the sender's signaling loss.

Lemma 3. *The receiver's expected loss is equal to the sender's signaling loss plus a constant.*

In other words, any communication scheme that increases the sender's loss in equilibrium is also worse for the receiver.

3.4 Discussion of Modeling Assumptions

We now discuss some of the assumptions we made in this Section.

Communication. Unlike in standard models of communication and learning our agents can only communicate a raw signal but not a posterior belief. We believe that this type of communication is common among "non-experts" or in short format social media. There are two interrelated reasons for this. One is that the communication medium may lack the depth or breadth of communication needed to synthesize pieces of observations into a coherent new view that can then be shared with others. For example, a recent academic paper on COVID-19 vaccines synthesized prior studies and facts into one package that includes summary statistics, such as "More than 3% of recipients of Moderna COVID-19 vaccine develop severe temporary side effects including fatigue (9.7%) , myalgia (8.9%), arthralgia (5.2%), headache (4.5%), pain (4.1%)." Hussey (2020). In contrast, a tweet is often retelling a personal or familiar anecdote such as "Got my 1st Moderna shot: I felt slightly flu-y last night. A little achy, slight headache, and tired." Another reason for this is that the sender and receiver may not have access to a common model (or "language" in the form of posterior beliefs) that would allow them to update some parameters and then efficiently communicate the mean and standard deviations of those parameters. For example, newspapers typically inform their readers about a new event mainly by reporting a selection of "stories", e.g., with headlines such as "A bribery scheme involving Politician X is discovered" rather than summary statistics that gives full account of the all politicians, their financial ties, and whether occurrence of corruption within politicians has increased in the past decade. In this case, even expert journalists are advised to use "examples" (or anecdotes) when communicating with non-expert reader who might otherwise be overwhelmed by expert jargon.

This restriction to communicating anecdotes drives our main results through a particular mechanic: more extreme anecdotes are noisier. While we consider the model of anecdotes to be a contribution of this paper, we note that any communication scheme which retains this noise structure would result in conceptually similar results. For example, one might argue that facts can be strategically manipulated up to a point, perhaps through the use of framing. In such a world, the sender can communicate signals that are within a fixed interval of an anecdote. This preserves the mechanic that more extreme anecdotes are noisier, and so individuals with more aligned utilities will have more accurate communication.

Timing. In our model, the sender chooses a communication scheme when she only knows her own moral stance and the moral stance of the receiver but has not yet seen any anecdotes. We believe that this assumption is most natural in an environment where senders adopt a number of "communication postures" when dealing with different types of receivers and then stick to this postures without trying to adapt their communication to specific signal realizations. For example, a sender might adopt the posture "always tell people who know me and people who agree with me the anecdote closest to the mean anecdote". Alternatively, she might adopt the posture "tell people who don't know me and are left of me the most right-wing anecdote I know". The former posture might be applied to friends while the latter posture might be adopted by a politician who is sampling from biased anecdotes when communicating in an attempt to gather more support.

Translation invariance. This assumption captures the idea that agents should not favor any arbitrary ranges of signals given that they have a diffuse prior and essentially know nothing about the true state. Translation invariance greatly simplifies our analysis because it allows the receiver to correct for any bias in the sender's signaling scheme by adding a single number to the received anecdote. However, translation invariance is not without loss of generality: there are pathological examples were the optimal signaling scheme is non-translation invariant (see Appendix B.3).

Utility function. Our work focuses on a particular functional form of utilities. We propose that both the sender and receiver are choosing actions based on squared loss functions, and that the sender further selects a communication scheme to minimize squared loss between her action and that of the receiver. The former of these assumptions is simply a microfounding of our model. More generally, the sender and receiver can each choose any point (fixed relative to the state of the world), and the alignment of the pair becomes the difference between these ideal points. The receiver chooses an action minimizing distance to her posterior about the location of her ideal point, and the sender chooses a communication scheme trying to minimize the squared loss between the receiver's action and her own ideal point. Seemingling more restrictive is the latter choice – that the sender cares about the squared distance of the receiver's action to her ideal point. We believe our general intuitions and results follow in any model where the sender's loss is a convex (equivalently, utility is concave) function of the receiver's action, with the minimum (equivalently, maxmimum) obtained at the sender's ideal point (e.g., her action in our model). With both these generalizations, it will still be the case that more aligned individuals have more accurate communication.

Memory. We present our model as a communication game with a sender and receiver. However, there is an alternative interpretation in the case where the sender is "communicating" with her future self by deciding what anecdote to store in memory for retrieval by the future self. In this setting, it makes sense to assume that there is no difference in the moral stances of the current and future self. Therefore, the current self will store the anecdote that minimizes her signaling loss (e.g. the observable regime).

4 Communication under Observability

In this section we analyze the case where the sender's communication is observable, meaning that the sender can commit to a communication scheme. We have already shown in Proposition 1 that the sender's optimal communication scheme will minimize signaling loss, and in particular it does not depend on any differences in moral stances. In other words, observability guarantees that the sender will seek to minimize the receiver's loss subject to the constraint that she can only send a single signal.

What communication scheme minimizes signaling loss? Intuitively, for symmetric single-peaked signal distributions it should be approximately optimal to select an signal close to the sender's posterior mean. After all, this is where we expect the signals to be most densely distributed, so communicating such an signal should minimize noise and give the receiver the best possible information about the sender's posterior belief.

It turns out that this intuition can be derived most crisply for the case of many signals, which we analyze next. Later in Section 4.3 we discuss some special considerations that arise when analyzing the case of few signals.

4.1 Observability: Asymptotic Analysis

For this part, we focus on the following class of *well-behaved* signal distributions.

Definition 3. *We say that the signal distribution F is well-behaved if the following holds. (1) The distribution is single-peaked and symmetric with variance 1. (2) The Fisher information $c_{\mathcal{I}}$ is not zero.[2] (3) Let $g(x) = f'(x)/f(x)$. That is, $g(x) = \frac{d \log f(x)}{dx}$. We assume that $|g'(x)| \leq c_1$ for all x, and some constant $c_1 > 0$. That is, $|g(x)| \leq c_1|x| + c_2$ [3] (4) F has exponential tail. That is, there is a constant $Q > 0$, such that for $x > Q$, we have $1 - F(x) \leq c_3 \exp(-|x|)$ for a constant $c_3 > 0$, and $x < -Q$ we have $F(x) \leq c_3 \exp(-|x|)$.*

For example, the normal distribution $F \sim N(0, 1)$ and the Laplace distribution with density $f(\epsilon) = \frac{1}{2} \exp(-|x|)$ are well-behaved. Recall that $f(\epsilon)$ is the density of the signal distribution at $\theta + \epsilon$, an offset of ϵ from the true state of the world.

We can now state our main result which asymptotically calculates the signaling loss of the unbiased communication scheme for well-behaved distributions.

---

[2]Note that Fisher information is 0 if and only when f is (piece-wise) constant.
[3]Note that, for $x < 1$ we bound $|g(x)| \leq c_2$ and otherwise we can bound $|g(x)| \leq c_1$.

Figure 2: Intuition for proof of theorem 2
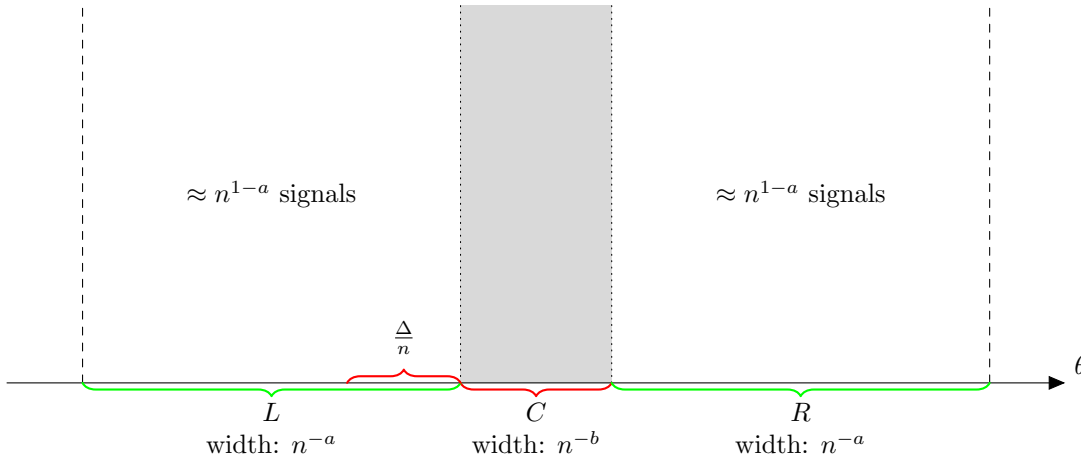
We assume $a < 1$ and $1 < b < 2a - \frac{1}{2}$ which ensures that the intervals L and R contain many signals but the collective influence of these signals on the posterior mean is $O(n^{\frac{1}{2} - 2a})$ and hence smaller than $\frac{1}{n}$. It also implies that if we consider a posterior that is contained in C then a rearrangement of signals in L or R will keep the posterior mean within C with high likelihood *and* the probability of signals drawn from the interval C goes to 0. For example, $a = \frac{4}{5}$ and $b = \frac{12}{11}$ satisfy these conditions.

Theorem 2. *Suppose that the signal distribution is well behaved. Then the unbiased communication scheme with $\delta = r(\delta) = 0$ which selects the closest signal to the sender's posterior mean $\theta_S(\vec{x})$ has loss:*

$$\frac{\alpha^2}{2n^2 f(0)^2} + o\left(\frac{1}{n^2}\right) \tag{9}$$

We can derive a similar theorem for biased communication schemes with $\delta \neq 0$.

Proposition 3. *Suppose that the signal distribution is well behaved. Then the biased communication scheme with bias δ has signaling loss:*

$$\frac{\alpha^2}{2n^2 f(\delta(\pi))^2} + o\left(\frac{1}{n^2}\right) \tag{10}$$

These two results together imply that the unbiased communication scheme is asymptotically optimal, and the optimal communication scheme is asymptotically unbiased.

Corollary 1. *For any well-behaved signal distribution, the unbiased communication scheme strictly dominates any biased signaling scheme for sufficiently large n and is optimal among all unbiased signaling schemes up to $o(1/n^2)$.*

This result confirms our intuition that the optimal communication scheme sends the signal closest to the posterior mean. Intuitively, the sender would like to send *precisely* the posterior mean to the receiver. However, since she can only send a signal she has to do with the second-best which is to send the signal closest to the posterior mean. When we interpret our model as a model of memory where the current self communicates with her future self by storing a single anecdote in memory we can think of the anecdote closest to the posterior mean as the "most representative anecdote".

4.2   Overview of Theorem 2 and Proposition 3

In this section we see an overview of the proofs of Theorem 2 and Proposition 3. The full proofs are relegated to Appendix A.

We begin by relating the sender's loss to the expected squared distance between the signal sent and the sender's posterior mean. For any communication π that is unbiased (meaning that $\delta(\pi) = 0$) the signaling loss is given by $\alpha^2 \mathbb{E}_{\vec{x} \sim F}[(\theta_S(\vec{x} - \pi(x))^2]$. Let π_0 denote the unbiased communication scheme that simply returns the signal closest to $\theta_S(\vec{x})$, the sender's posterior mean. Let $X_0 = \min |\theta_S(\vec{x}) - x_i|$ be the distance between the this closest signal and the posterior mean. Then the signaling loss of π_0 is precisely equal to $\alpha^2 \mathbb{E}_{\vec{x} \sim F}[X_0^2]$ (and moreover π_0 is the optimal unbiased communication scheme).

We can similarly describe the signaling loss of biased communication schemes. For any $\delta \geq 0$, let π_δ denote the communication scheme that returns the signal closest to $\theta_S(\vec{x}) + \delta$. That is, $\pi_\delta(\vec{x}) = \text{argmin}_{x_i \in \vec{x}} |\theta_S(\vec{x}) + \delta - x_i|$. Similar to the unbiased case, we can let $X_\delta = \min |\theta_S(\vec{x}) - x_i + \delta|$ be the distance between $\theta_S(\vec{x}) + \delta$ and $\pi_\delta(\vec{x})$ the signal closest to $\theta_S(\vec{x}) + \delta$.

We note that, for $\delta \neq 0$, the communication scheme π_δ may not have bias δ. That is, it might be that $\delta(\pi_\delta) \neq \delta$. This is because signals may not be symmetrically distributed around the point $\theta_S(\vec{x}) + \delta$. However, the distance ($\mathbb{E}[X_\delta^2]$) between $\pi_\delta(\vec{x})$ and $\theta_S(\vec{x}) + \delta$, is at most the distance between $\pi(\vec{x})$ and $\theta_S(\vec{x}) + \delta$ for any communication scheme π. Moreover, when π has bias $\delta(\pi) = \delta$ this is precisely $1/\alpha^2$ times the signaling loss of π. Thus, the signaling loss of any communication scheme π with bias $\delta \neq 0$ is at most $\alpha^2 \mathbb{E}[X_\delta^2]$.

We can summarize the discussion above about biased and unbiased signaling loss as follows.

Observation 1. *To prove Theorem 2 it is enough to show that $\mathbb{E}[X_0^2] \leq \frac{1}{2n^2 f(0)^2} + o(\frac{1}{n^2})$.*

Observation 2. *To prove Proposition 3 it is enough to show that $\mathbb{E}[X_\delta^2] \geq \frac{1}{2n^2 f(\delta)^2} + o(\frac{1}{n^2})$.*

Given these observations, it remains to bound $\mathbb{E}[X_0^2]$ from below and $\mathbb{E}[X_\delta^2]$ from above for $\delta > 0$. To gain some intuition, imagine first what would happen if the sender had *foresight* and knew the value of θ in advance, so that $\theta_S(\vec{x}) = \theta$. In this case, X_δ is simply the minimum distance between $\theta + \delta$ and any of n independent random draws from the signal distribution. One can then derive an expression for the expected minimum distance as a function of $f(\delta)$, the density of the signal distribution at $\theta + \delta$, and indeed one can check that these bounds would satisfy the requirements for Observations 1 and 2.

When the sender does not have foresight, the posterior mean $\theta_S(\vec{x})$ depends on the realized signals $\vec{x}$, and this introduces correlation between the signal realizations and the value of $\theta_S(\vec{x}) + \delta$. We therefore cannot model X_δ using independent draws from the signal distribution. Indeed, as we will see in Section 4.3, these correlations can significantly impact $\mathbb{E}[X_\delta^2]$ when the number of signals is small.

Our approach is to argue that as n grows large, the impact of these correlations grows small. Small enough, in fact, that the correlation between $\theta_S(\vec{x}) + \delta$ and the signal closest to that point becomes small enough that it is dominated by the statistical noise that would anyway be present if signals were drawn independently of $\theta_S(\vec{x})$. We argue this in three steps.

Step 1: We argue that it suffices to focus on cases where $\theta_S(\vec{x})$ falls within a narrow interval. Let $I = [-n^{-\frac{1}{2}+\epsilon}, n^{-\frac{1}{2}+\epsilon}]$ for some $\varepsilon > 0$. Using the law of large numbers, we argue that $\theta_S(\vec{x}) \in I$ with all but exponentially small probability (in n). The contribution to $\mathbb{E}[X_\delta^2]$ from events where $\theta_S(\vec{x}) \notin I$ is therefore negligible and can be safely ignored. This allows us to assume that $\theta_S(\vec{x}) \in I$.

Step 2: To reduce the impact of correlation we won't focus on the exact value of $\theta_S(\vec{x})$, but rather an interval in which it falls. To this end we partition I into subintervals of width n^{-b}, where b is chosen so that any given interval is unlikely to contain a signal. One such subinterval contains the posterior mean $\theta_S(\vec{x})$; call that subinterval C. We then consider longer subintervals L and R to the left and right of C, respectively, of width n^{-a} chosen large enough that we expect many signals to appear in each. See Figure 2.

We bound the impact of correlation by showing that if we condition on the number of signals that appear in L and R, then the actual arrangement of signals within those subintervals (keeping all other signals fixed) has only negligible effect on the posterior mean. Specifically, given any arrangement of the signals within L and R, the probability that the posterior mean falls within C remains large. (See Corollary 8 for more details.)

13

This implies that there is negligible correlation between the joint density function of a fixed number of k signals in $L \cup R$ and the event that $\theta_S(\vec{x}) \in C$. For a general δ, we consider the interval $C_\delta = C + \delta$ containing $\theta_S(\vec{x}) + \delta$ and the intervals L_δ and R_δ to the left and right of C_δ. We similarly, bound the correlation between the joint density function of a fixed number of k signals in $L_\delta \cup R_\delta$ and the event that $\theta_S(\vec{x}) + \delta \in C_\delta$. We formally show this in Lemmas 19 and 20 [4].

Step 3: The analysis in Step 2 conditioned on the number of signals that fall in $L \cup R$. We now consider the distribution of this number of signals. We show a concentration result: with high probability, the number of signals that lie in $L \cup R$ is close to what would be expected if the signals were distributed independently and not correlated with $\theta_S(\vec{x})$ (through the identity of interval C). See Appendix A for the proof, and Lemma 15 for the proof that it suffices to consider only this high-probability event.

Given this concentration result, we can focus on bounding the expected value of X_δ^2 given the numbers of signals in L and R. From the analysis in Step 2, we can view these signals as (approximately) independently distributed within L and R. We can therefore bound the expected squared distance between interval C and the closest signal to interval C by performing an explicit calculation for independent signals. We still do not know the value of $\theta_S(\vec{x})$ within interval C (and we have not bounded the impact of correlation on that value), but C is sufficiently narrow that this uncertainty has limited impact on $\mathbb{E}[X_\delta^2]$. We conclude that the impact of correlation on $\mathbb{E}[X_\delta^2]$ is absorbed in lower-order terms, which yields the bounds required by Observations 1 and 2. This gives us the required results of Theorem 2 and Proposition 3.

4.3   Observability: Few Signals

Corollary 1 shows that for well-behaved distribution, biased communication schemes are asymptotically suboptimal. One might wonder whether this result extends to all finite sample cases. It turns out that this is not true. In this section we present examples showing that the optimal communication scheme can include a significant amount of bias when n is small.

4.3.1   Two Signals

We begin by considering the case $n = 2$, where there are just two samples and assume the signal distribution F is symmetric around 0. In this case, the strictly optimal communication scheme π is the one that always returns the minimum signal (or, by symmetry, always returns the maximum signal).

Proposition 4. *When $n = 2$ and signals are drawn from a distribution symmetric around the state of the world θ, the communication scheme that always sends the min signal (or always sends the max signal) minimizes the cost of the sender amongst all translation invariant schemes under observability.*

The reasoning behind Proposition 4 is that, after receiving two signals $\vec{x} = (x_1, x_2)$, the sender's posterior mean $\theta_S(\vec{x})$ will always equal the empirical mean $\bar{x} = \frac{1}{2}(x_1 + x_2)$. In particular, this means that each signal has the same distance to the posterior mean. Upon receiving a signal, the receiver has two sources of uncertainty: the distance of the signal to the sender's posterior mean, and whether the posterior mean is larger or smaller than the received signal. The former quantity is the same for all possible communication schemes, so the only lever by which to reduce variance is to reduce uncertainty about the latter. But this uncertainty is minimized by removing it entirely, which is done by always returning the smaller signal (or, symmetrically, by always returning the larger signal).

4.3.2   Three Signals

The result for two signals is very crisp, but somewhat unsatisfying. It relies crucially on the mechanical fact that when there are $n = 2$ signals, every communication scheme has the same distribution over distance

---

[4]For these lemmas we assume that the density function θ_S is sufficiently "nice" in C. Refer to Section A.2.2 for details about this assumption, and why we can make this assumption without loss of generality.

between signal and sender posterior mean. We next show that this observation is not specific to the case of two signals. When there are $n = 3$ signals, the optimal signaling scheme may still be biased.

Proposition 5. *When $n = 3$ and signals are drawn from a uniform distribution around θ, the optimal translation-invariant signaling scheme under observability has non-zero bias.*

The idea behind Proposition 5 is that, conditional on the value of the sender's posterior mean, the conditional density over signal realizations is not necessarily single-peaked. For the setting of Proposition 5, the optimal unbiased scheme is precisely the one that always returns the middle signal. However, for uniform distributions, the correlation between the posterior mean and the minimum and maximum signals is stronger than the correlation between the posterior mean and the middle signal. One can therefore communicate more information about the posterior mean through a biased communication scheme that sometimes returns the minimum signal (or, by symmetry, sometimes returns the maximum signal).

Discussion   Why does bias help? Recall that there is intrinsic error in the sender's posterior mean. This variance is unavoidable. But it can introduce correlation with particular samples. This correlation can be used to help minimize variance between the posterior mean and the signal passed to the receiver. This is why, for the uniform case, it is helpful to bias toward more extreme signals: even though they are not more informative than the moderate signals when it comes to the true state of the world, they are more informative with respect to the sender's posterior mean. The interplay between these two sources of errors therefore introduces incentive for the sender to systematically bias their communication.

4.3.3   Relaxing Translation Invariance

The next thing we note is that the restriction to translation invariance is necessary: there are examples with $n = 2$ signals and where the distribution F is symmetric around 0, but yet the optimal (possibly not translation invariant) communication scheme for the sender is not translation invariant.

Proposition 6. *When $n = 2$ and signals are drawn from a distribution symmetric around θ, the optimal signaling scheme under observability for the sender may not be translation-invariant.*

We know from Proposition 4 that the communication scheme that always returns the minimum signal is optimal among translation invariant communication schemes. So to prove Proposition 6 it suffices to find any communication scheme that improves upon this minimal-signal scheme. The idea behind the construction is to set up distribution F to be bimodal, so that each pair of signals (x_1, x_2) will either come from the same mode (and have small difference) or from different modes (and have large difference). This splits signal realizations into two cases, but a translation-invariant scheme cannot distinguish these two cases for the receiver. However, one can construct a non-translation-invariant scheme that correlates the choice of signal with which case has occurred. If the two modes have sufficiently low variance, the extra value of providing a hint about same-mode versus different-mode dominates the loss due to being ambiguous about whether the posterior mean is higher or lower than the provided signal.

5   Communication under Non-observability

In this Section we analyze the case where the sender's communication scheme is unobservable. We already know from Proposition 2 that non-observability matters when the moral stances of sender and receiver differ and non-observability increases the sender's loss (as well as the receiver's). We start by first characterizing the equilibrium communication scheme under non-observability and then properties of these equilibria. In particular, we show that due to unraveling even small differences in the moral stances can cause large bias in the sender's communication scheme.

5.1 Characterizing the Sender's Optimal Communication Scheme

We start by recalling the decomposition of the sender's disagreement loss from equation 5 into signaling loss, persuasion temptation and fundamental loss. Since the latter is constant only the first two components matter for the sender's optimization problem. Hence the sender is minimizing the following expression:

$$\underbrace{\alpha^2 \left[\mathbb{E}_{\vec{x}} (\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta))^2 - (\delta(\pi) - r(\delta))^2 \right]}_{\text{signaling loss}} +$$

$$+ \underbrace{\alpha^2 \left(\delta(\pi_R) - \delta(\pi) \right)^2 + 2\alpha(1 - \alpha) \left(\delta(\pi_R) - \delta(\pi) \right) \Delta}_{\text{persuasion temptation}} \tag{11}$$

In contrast to the observable environment we analyzed in the previous section, we can no longer ignore persuasion temptation. This term is equal to zero in equilibrium as the recipient will correctly anticipate the sender's bias such that $\delta_R = \delta$. However, the sender takes the receiver's anticipated bias as fixed when choosing her communication scheme and hence the persuasion term will matter for off-equilibrium deviations. In particular, we note that the sender can always lower her loss by moving her communication scheme to the left when $\Delta < 0$ (or moving to the right in the opposite case).[5] Intuitively, the sender is tempted to increase her bias in order to move the receiver's action in her preferred direction.

Next, we note that Theorem 1 characterizes the optimal anecdote $\pi(\vec{x})$ as the closest anecdote to $\theta_S(\vec{x}) + r(\delta)$. One difficulty here is that the posterior mean $\theta_S(\vec{x})$ is often difficult to calculate for signal distributions other than the normal distribution. Moreover, it is a function of the entire anecdote vector $\vec{x}$ and therefore the conditional distribution of the closest anecdote to $\theta_S(\vec{x}) + r(\delta)$ depends on $\theta_S(\vec{x})$. This was a major technical difficulty we had to deal with in Section 4.1.

In this Section, we will circumvent this technical problem by assuming that the sender has **foresight** and *knows* the precise state of the world θ already in Stage 1. She therefore does not have to estimate the state of the world and $\theta_S(\vec{x}) = \theta$. We expect that the difference between the optimal communication with and without foresight is small especially for large n because we know that the posterior mean converges to the true state at rate $\sqrt{n}$. Using the foresight assumption we can derive (a) a precise condition for finding the center $r(\delta)$ of a communication scheme with bias δ and (b) solve the sender's first-order condition in Equation 11.

Let's start with finding a condition that pins down the center $r(\delta)$. Denote by y the distance between the center r and the closest anecdote (out of n total anecdotes) where y is positive if the closest anecdote is to the right and negative if the closest anecdote is to the left. We can derive the density of y as follows:

$$g(y) = nf(r + y) \left[1 - (F(r + |y|) - F(r - |y|)) \right]^{n-1}. \tag{12}$$

To see this, note that there are n anecdotes that can take the value $r + y$. At the same time all the other $n - 1$ anecdotes cannot fall into the interval $[r - |y|, r + |y|]$. We define $H(r, y) = 1 - (F(r + y)) - F(r - y))$ which allows us to write the density as:

$$g(y) = nf(r + y)H(r, |y|)^{n-1}. \tag{13}$$

We can now calculate the expected bias of communication scheme with center $r(\delta)$, which has to be δ:

$$\delta = r + \int_{-\infty}^{\infty} ynf(r + y)H(r, |y|)^{n-1}dy \tag{14}$$

---

[5]The off-equilibrium temptation incentive is apparent when starting from an equilibrium where $\delta_R = \delta$. By slightly lowering δ (for $\Delta < 0$) the change in the quadratic term is second-order but the linear term is first-order.

This equation implicitly defines $r(\delta)$. We can simplify it further:

$$
\begin{aligned}
\delta &= r + \int_0^\infty ynf(r+y)H(r,|y|)^{n-1}dy + \int_{-\infty}^0 ynf(r+y)H(r,|y|)^{n-1}dy \\
&= r + \int_0^\infty ynf(r+y)H(r,y)^{n-1}dy - \int_\infty^0 (-y)nf(r-y)H(r,y)^{n-1}dy \\
&= r + \int_0^\infty yn\left[f(r+y) - f(r-y)\right]H(r,y)^{n-1}dy \\
&= r - \int_0^\infty yn\frac{\partial H(r,y)}{\partial r}H(r,y)^{n-1}dy
\end{aligned}
\tag{15}
$$

This proves the next lemma.

Lemma 4. *Assume the sender has foresight. Then the center $r(\delta)$ of her optimal communication scheme with bias δ satisfies:*

$$
r(\delta) = \delta + \int_0^\infty yn\frac{\partial H(r(\delta),y)}{\partial r}H(r(\delta),y)^{n-1}dy
\tag{16}
$$

We turn next to the sender's loss minimization problem (11). Note that when $n = 1$ the sender's problem is trivial (as there is only one signal to send), so suppose $n \geq 2$. Recalling that random variable y corresponds to $\theta_S(\vec{x}) - \pi(\vec{x}) + r(\delta)$, we can rewrite (11) using the density $g(y)$ as follows:

$$
\underbrace{\alpha^2\left[\int_{-\infty}^\infty y^2nf(r(\delta)+y)H(r,|y|)^{n-1}dy - (\delta(\pi) - r(\delta))^2\right]}_{\text{signaling loss}} +
$$

$$
+ \underbrace{\alpha^2\left(\delta(\pi_R) - \delta(\pi)\right)^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta}_{\text{persuasion temptation}}
$$

$$
\begin{aligned}
&= \alpha^2\left[\int_0^\infty y^2n\left(f(r(\delta)+y) + f(r(\delta)-y)\right)H(r,y)^{n-1}dy - (\delta(\pi) - r(\delta))^2\right] \\
&+ \alpha^2\left(\delta(\pi_R) - \delta(\pi)\right)^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta \\
&= \alpha^2\left[-\int_0^\infty y^2n\frac{\partial H(r(\delta),y)}{\partial y}H(r,y)^{n-1}dy - (\delta(\pi) - r(\delta))^2\right] + \\
&+ \alpha^2\left(\delta(\pi_R) - \delta(\pi)\right)^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta \\
&= \alpha^2\left[\int_0^\infty 2yH(r(\delta),y)^ndy - (\delta(\pi) - r(\delta))^2\right] + \alpha^2\left(\delta(\pi_R) - \delta(\pi)\right)^2 + 2\alpha(1-\alpha)\left(\delta(\pi_R) - \delta(\pi)\right)\Delta \\
&+ \alpha^2 y^2 H(r(\delta),y)^n\big|_{y=0}^\infty
\end{aligned}
\tag{17}
$$

Consider the final term $\alpha^2 y^2 H(r(\delta),y)^n\big|_{y=0}^\infty$. This expression is equal to 0 when evaluated at $y = 0$. We also claim that $\lim_{y\to\infty} \alpha^2 y^2 H(r(\delta),y)^n \to 0$. To see why, write μ for the mean of the noise distribution F and σ^2 for its (finite) variance. For any $y > 2|\mu - r(\delta)|$, observe that if $|x - \mu| \leq y/2$ then $|x - r(\delta)| \leq y$. By Chebyshev inequality we have

$$
H(r(\delta),y) \leq \Pr_{x\sim F}[|x - \mu| \geq y/2] \leq 4\sigma^2/y^2
$$

and hence for $n \geq 2$ we have $\lim_{y\to\infty} y^2 H(r(\delta),y)^n = 0$. We therefore conclude $\alpha^2 y^2 H(r(\delta),y)^n\big|_{y=0}^\infty = 0$.

The sender minimizes (17) with respect to $\delta(\pi)$. Hence we can write her first-order condition as

$$
\alpha^2 \left[\underbrace{\int_0^\infty 2ynr'(\delta)\frac{\partial H(r(\delta),y)}{\partial r}H(r(\delta),y)^{n-1}dy}_{=2r'(\delta)(r-\delta) \quad \text{by Lemma 4}} -2(\delta(\pi)-r(\delta))(1-r'(\delta)) \right]
$$
$$
+2\alpha^2(\delta(\pi)-\delta(\pi_R)) - 2\alpha(1-\alpha)\Delta = 0
$$

Which implies that

$$
-2\alpha^2(\delta(\pi_R)-r(\delta)) - 2\alpha(1-\alpha)\Delta = 0. \tag{18}
$$

We summarize our finding in the following lemma.

Lemma 5. *Fixing the belief π_R of the receiver, a sender with foresight under non-observability will best-respond by choosing a communication scheme with bias δ and center $r(\delta)$ that satisfies*

$$
\delta(\pi_R) - r(\delta) = -\frac{1-\alpha}{\alpha}\Delta. \tag{19}
$$

At equilibrium, we must have $\delta(\pi_R) = \delta(\pi)$. Lemma 5 therefore directly implies the following condition on the sender's equilibrium strategy.

Lemma 6. *At equilibrium, a sender with foresight under non-observability chooses a communication scheme with bias δ and center $r(\delta)$ that satisfies*

$$
\delta - r(\delta) = -\frac{1-\alpha}{\alpha}\Delta. \tag{20}
$$

We combine Lemmas 4 and 6 and characterize the sender's equilibrium communication scheme.

Proposition 7. *Assume the sender has foresight. At equilibrium under non-observability, the sender's communication scheme has center r^* which satisfies:*

$$
\int_0^\infty yn\frac{\partial H(r^*,y)}{\partial r}H(r^*,y)^{n-1}dy = \frac{1-\alpha}{\alpha}\Delta \tag{21}
$$

Moreover, this scheme has bias $\delta = r^* - \frac{1-\alpha}{\alpha}\Delta$.

5.2   Receiver's Loss

We now calculate the loss of a receiver when the sender uses her optimal communication scheme at equilibrium. This allows us to analyze whether a receiver prefers listening to a more informed expert whose moral stance differs from hers versus listening to a less informed sender whose moral stance is closer to hers.

From Lemma 3 we know that the expected receiver loss is equal to the sender's signaling loss plus a constant. Therefore we can deduce from equation 17:

$$
\begin{aligned}
\text{receiver's loss} &= \text{constant} + \alpha^2 \left[\int_0^\infty 2yH(r(\Delta),y)^n dy - (\delta(\Delta)-r(\Delta))^2 \right] \\
&= \text{constant} + \alpha^2 \left[\int_0^\infty 2yH(r(\Delta),y)^n dy - \frac{(1-\alpha)^2}{\alpha^2}\Delta^2 \right]
\end{aligned} \tag{22}
$$

Note, that we write the equilibrium bias δ and center r as a function of the difference in the moral stance Δ (we obtain these functions from Proposition 7). We can derive both sides with respect to Δ:

$$
\begin{aligned}
\frac{d(\text{receiver's loss})}{d\Delta} &= \alpha^2 \left[\underbrace{\int_0^\infty 2yn \frac{dr(\Delta)}{d\Delta} \frac{\partial H(r(\Delta), y)}{\partial r} H(r(\Delta), y)^{n-1} dy}_{=2\frac{dr(\Delta)}{d\Delta}\frac{1-\alpha}{\alpha}\Delta \quad \text{by Proposition 7}} - 2\frac{(1-\alpha)^2}{\alpha^2}\Delta \right] \\
&= 2\alpha(1-\alpha)\Delta \underbrace{\left[\frac{dr(\Delta)}{d\Delta} - \frac{1-\alpha}{\alpha} \right]}_{=\frac{d\delta(\Delta)}{d\Delta} \quad \text{by Lemma 6}} = 2\alpha(1-\alpha)\Delta \frac{d\delta(\Delta)}{d\Delta} \quad (23)
\end{aligned}
$$

Lemma 7. *The receiver's loss function $L_R(\Delta)$ satisfies:*

$$
\frac{dL_R(\Delta)}{d\Delta} = 2\alpha(1-\alpha)\Delta \frac{d\delta(\Delta)}{d\Delta} \quad (24)
$$

This finally allows us to compare the receiver's utility when listening to senders with different moral stances.

Corollary 2. *Assume that the anecdote distribution F is single-peaked and symmetric. Then the receiver's loss is minimized when he talks to like-minded senders ($\Delta = 0$) and increases on both sides otherwise. Moreover, $\frac{d\delta(\Delta)}{d\Delta} > 0$, thus $sgn\left(\frac{dL_R(\Delta)}{d\Delta} \right) = sgn(\Delta)$.*

Proof: Assume that $\delta(\Delta)$ is not always increasing over $[-\infty, 0]$. We know that for single-peaked and symmetric F we have $\delta = 0$ at $\Delta = 0$ and for $\Delta < 0$ we have $\delta < 0$ because $\frac{\partial H(r,y)}{\partial r} < 0$ and $r < \delta < 0$ by Proposition 7. Then there has to be some $\Delta < \Delta' < 0$ such that $\delta = \delta'$. But this implies that $r = r'$. Therefore $\delta - r = \delta' - r'$. However, this is not possible by Proposition 7.

The above proposition tells us that the equilibrium bias of the sender's optimal scheme increases with Δ. We can in fact show that for single-peaked symmetric distributions we converge to the minimum or maximum schemes.

Corollary 3. *Assume that the anecdote distribution F is single-peaked and symmetric. Then the sender's optimal communication scheme under non-observability converges to the minimum scheme as $\Delta \to -\infty$ (in the sense that $r(\delta) \to -\infty$) and converges to the maximum scheme as $\Delta \to \infty$ (in the sense that $r(\delta) \to \infty$).*

Proof: Consider $\Delta < 0$ (the proof for positive Δ is analogous). For single-peaked symmetric distributions we have $r < \delta < 0$ and $\delta - r = -\frac{1-\alpha}{\alpha}\Delta$. This implies $r < \frac{1-\alpha}{\alpha}\Delta$. Hence, $r(\Delta) \to -\infty$ as $\Delta \to -\infty$.

This result shows convergence to the minimum (maximum) scheme when fixing n and $\Delta \to -\infty$. The reverse holds as well.

Corollary 4. *Assume that the anecdote distribution F is single-peaked and symmetric. Consider a sender and receiver with fixed moral stance difference Δ. Then the sender's optimal communication scheme under non-observability converges to the minimum scheme as $n \to \infty$ if $\Delta < 0$ and converges to the maximum scheme as $n \to \infty$ if $\Delta > 0$.*

Proof: Denote the center of the optimal communication scheme with n signals with $r(n)$. Assume that $\Delta < 0$ and that the optimal communication scheme does not converge to the minimum scheme. Then there is some $\tilde{r}$ such that for any N there is some $n > N$ such that $r(n) > \tilde{r}$. However, as we increase n the density of anecdotes in any ϵ interval around $\tilde{r}$ increases and therefore it has to be case that for some sufficiently large n we have $\delta(n) - r(n) < \epsilon$. By choosing ϵ small enough we can eventually find an n where $\delta(n) - r(n) < -\frac{1-\alpha}{\alpha}\Delta$ which contradicts Proposition 7.

19

5.3 Informational Homophily

We can build on Corollary 3 and characterize the receiver's loss in the limit as $\Delta \to -\infty$ and $\Delta \to \infty$.

Corollary 5. *Assume that the anecdote distribution F is single-peaked and symmetric. Then the receiver's loss $L(\Delta)$ can be written as $L(\Delta) = C + \alpha^2 l(\Delta)$ where C is a positive constant and the function $l(\Delta)$ satisfies the following properties:*

$$\begin{aligned}
l(0) &= 0 \\
l'(\Delta) &< 0 \quad for \; \Delta < 0 \\
l'(\Delta) &> 0 \quad for \; \Delta > 0 \\
\lim_{\Delta \to -\infty} l(\Delta) &= \alpha^2 V_{min} \\
\lim_{\Delta \to \infty} l(\Delta) &= \alpha^2 V_{max}
\end{aligned} \tag{25}$$

where V_{min} and V_{max} are the variance of the minimum and maximum anecdote, respectively.

We can now answer the following question: will a receiver prefer to a well-informed but biased expert rather than a less-informed sender whose moral stance is closer to the receiver's. It turns out that the answer depends on the distribution.

Theorem 3. *Consider a receiver who can choose between two senders: (a) well-informed expert with access to n signals and difference in moral stance (compared to receiver) equal to $\Delta \neq 0$; (b) a sender with access to a single signal who shares the same moral stance as the receiver. If the expert's anecdotes are drawn from a normal distribution the receiver will always prefer to talk to the biased expert. In contrast, if the expert's anecdotes are drawn from a Laplace distribution then the receiver might prefer the sender with a single signal to the expert.*

Proof: According to the Fisher-Tippett-Gnedenko theorem the variance of the extremal anecdotes converge to 0 with n for the normal distribution but are bounded away from 0 for the Laplace distribution Fisher and Tippett (1928); Gnedenko (1943). Hence, by corollary 4 the result follows.

6 Conclusion

We analyze communication between a sender a receiver where the sender sends one of her signals to the receiver. Signals are always correct (no "fake news") but they are not necessarily representative of the information set of the sender. In particular, when the sender has a motive to persuade the sender she will tend to send signals from the tail of her distribution. The receiver will undo this bias but there is a cost: tail signals are less informative than more representative signals. Hence, the sender's attempts to persuade the receiver lowers the precision of the transmitted information.

We show that receivers do not necessarily want to talk to the best informed senders. If the signal distribution has "fat tails" the receiver might prefer to talk to a less informed sender whose preferences are more aligned with her own. Such experts have access to more anecdotes from the tail which limits the ability of the receiver to learn even if the expert is very well informed. In contrast, if signals are distributed according to a "thin-tail" distribution such as the normal distribution receivers will generally prefer talking to the biased expert. This gives rise to a type of informational homophily where receivers seek out like-minded senders because these senders tell more representative anecdotes from which they can learn more efficiently.

References

S Nageeb Ali, Greg Lewis, and Shoshana Vasserman. 2020. Voluntary disclosure and personalized pricing. In *Proceedings of the 21st ACM Conference on Economics and Computation.* 537–538.

Vincent P Crawford and Joel Sobel. 1982. Strategic information transmission. *Econometrica: Journal of the Econometric Society* (1982), 1431–1451.

Ronald A Dye. 1985. Disclosure of nonproprietary information. *Journal of accounting research* (1985), 123–145.

R. A. Fisher and L. H. C. Tippett. 1928. Limiting forms of the Frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society* 24, 2 (1928), 180–190. `https://doi.org/10.1017/S0305004100015681`

Michael J Fishman and Kathleen M Hagerty. 1990. The optimal amount of discretion to allow in disclosure. *The Quarterly Journal of Economics* 105, 2 (1990), 427–444.

B. Gnedenko. 1943. Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *Annals of Mathematics* 44, 3 (1943), pp. 423–453. `https://doi.org/10.2307/1968974`

Sanford J Grossman. 1981. The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics* 24, 3 (1981), 461–483.

Sanford J Grossman and Oliver D Hart. 1980. Disclosure laws and takeover bids. *The Journal of Finance* 35, 2 (1980), 323–334.

Sergiu Hart, Ilan Kremer, and Motty Perry. 2017. Evidence games: Truth and commitment. *American Economic Review* 107, 3 (2017), 690–713.

J. A. Hartigan. 1983. *Asymptotic Normality of Posterior Distributions.* Springer New York, 107–118. `https://doi.org/10.1007/978-1-4613-8242-3_11`

Colleen Hussey. 2020. *Moderna's COVID-19 Vaccine Candidate Meets its Primary Efficacy Endpoint in the First Interim Analysis of the Phase 3 COVE Study.* `https://investors.modernatx.com/news-releases/news-release-details/modernas-covid-19-vaccine-candidate-meets-its-primary-efficacy`

Woon-Oh Jung and Young K Kwon. 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting research* (1988), 146–153.

Emir Kamenica and Matthew Gentzkow. 2011. Bayesian persuasion. *American Economic Review* 101, 6 (2011), 2590–2615.

Giorgio Martini. 2018. Multidimensional Disclosure. (2018).

Paul R Milgrom. 1981. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* (1981), 380–391.

A   Proofs from Section 4.1

A.1   More Detailed Overview of Theorem 2 and Proposition 3

We begin by revisiting the three-step overview of Theorem 2 and Proposition 3 from Section 4.2, but described in more formal detail.

We note that, when $|\delta| > 2(\log n)^2$, $\mathbb{E}[X_\delta^2] \geq \left(2(\log n)^2 - (\log n)^2\right)^2 \left(1 - ne^{-(\log n)^2}\right) \geq O((\log n)^4)$. Since, for a *well-behaved* distribution, the probability of there exists a signal $> (\log n)^2$ is at most $ne^{-(\log n)^2}$. Hence we will focus only on $|\delta| \leq 2(\log n)^2$.

Step 1: We note the the posterior distribution of θ converges to a normal distribution with expectation $\theta^* = 0$ and variance $\mathcal{I}^{-1}(\theta)/n$, where $I(\theta^*)$ is the Fisher information of θ^*Hartigan (1983). Let $I = [-n^{-\frac{1}{2}+\epsilon}, n^{-\frac{1}{2}+\epsilon}]$ for some $\varepsilon > 0$. Therefore, the probability that $\theta_S \notin I$ is at most $\exp(-n^{2\epsilon}A)$ for some $A > 0$.

Note that, for any δ we have, $\mathbb{E}[X_\delta^2] \geq \Pr[\theta_S \in I] \cdot \mathbb{E}[X_\delta^2|\mathbb{1}\{\theta_S \in I\}]$. Further, in Section A.2.2, we bound the contribution to the cost when $\theta_S \notin I$ to be $O(\exp(-n^{2\epsilon}A))$.

$$\mathbb{E}[X_0^2] = \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in I\}] + \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}]$$
$$\leq \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in I\}] + O((\exp(-n^{2\epsilon}A))$$

Therefore, we will focus on $\theta_S \in I$.

Step 2: We partition I into intervals of length $1/n^b$, where $b = 12/11 (> 1)$. Let $\mathcal{P}$ denote this partition of I. Let C be one such interval in $\mathcal{P}$. Our goal is to bound the distance of the closest signal to $\theta_S + \delta$ conditioned on $\theta_S \in C$ (and equivalently $\theta_S + \delta \in C + \delta$).

We consider the intervals L_δ and R_δ of length $1/2n^a$, where $a = 4/5 (< 1)$, to the left and right interval $C + \delta$ respectively [6]. For $\delta = 0$ we usually drop the subscripts and denote the intervals as L and R. Our choice of a allows L_δ (and R_δ) to be wide enough to contain many signals, while small enough that they cover only a fraction of $I + \delta$ and hence contain at most $O(n^{1-a})$. Figure 2 shows one such interval.

A key lemma here shows that, given a well-behaved distribution, any re-arrangement of signals within $L_\delta \cup R_\delta$ has only a negligible effect on the posterior mean. In other words, given $\theta_S \in C$, no matter how we arrange $k = O(f(\delta)n^{1-a})$ many signals in $L_\delta \cup R_\delta$ the posterior mean remains in C almost all the time (see Corollary 8).

This implies that there is negligible correlation between the joint density function of k signals in $\vec{x}_{[k]} \in L_\delta \cup R_\delta$ and the event $\theta_S \in C$, informally stated in Lemma 8. We formally show this in Lemmas 19 and 20 [7].

Lemma 8 (Informal). *Let $\vec{x}_{[k]}$ be a subset of k (at most n^{1-a}) signals in $L_\delta \cup R_\delta$, and $K(k)$ be the event that there are exactly k anecdotes $L_\delta \cup R_\delta$. Then $f(\vec{x}_{[k]}|\theta_S \in C \wedge K(k)) = f(\vec{x}_{[k]})(1 + O(n^{-1/2}))$.*

Moreover, the density function f of a signal within $I + \delta$ is reasonably "flat" and is approximately close to $f(\delta)$ (see Claims 1 and 4). Hence, even conditioned on $\theta_S \in C$, the density of the anecdotes in $L_\delta \cup R_\delta$ is essentially uniformly distributed.

Step 3: Let K be the event that there are at least $k = (1 - \varepsilon)f(0)n^{1-a}$ in $L \cup R$ (recall that L and R are the correspoding intervals around C). We show that with high probability the number of signals in $L \cup R$, of length n^{-a}, is at least $(1 - \varepsilon)f(0)n^{1-a}$, see Appendix A for the proof. In Lemma 15 (in Appendix A.2.2), we show that the contribution of the low probability event $\overline{K}$, is bounded by $o(1/n^2)$.

---

[6] We set b and a to these values to ensure that intervals have desirable properties. Refer to Appendix A for more details

[7] For these Lemmas we assume that the density function θ_S is sufficiently "nice" in all $C \in \mathcal{P}$. Refer to Section A.2.2 for details about this assumption, and why we can make this assumption wlog.

$$\mathbb{E}[X_0^2 \mathbb{1}(\theta_S \in I)] \leq \sum_{C \in \mathcal{P}} \mathbb{E}[X_0^2 \mathbb{1}(\theta_S \in C \wedge K)] + \sum_{C \in \mathcal{P}} \mathbb{E}[X_0^2 \mathbb{1}(\theta_S \in C \wedge \overline{K})]$$

$$\leq \sum_{C \in \mathcal{P}} \mathbb{E}[X_0^2 \mathbb{1}(\theta_S \in C \wedge K)] + o(1/n^2)$$

Hence for each $C \in \mathcal{P}$ we just focus on bounding $\mathbb{E}[X_0^2 \cdot \mathbb{1}(K)|\theta_S \in C]$, towards this we prove the following Lemma.

Lemma 9. *The distribution of the random variable $X_0 \cdot \mathbb{1}(K)$ conditioned on $\theta_S \in C$ is stochastically dominated by the exponential distribution with $\lambda = 2nf(0)$. That is, $\Pr[X_{(0)} \cdot \mathbb{1}(K) > d|\theta_S \in C] < \exp[-\lambda d](1 - O(n^{-1/2}))$.*

Similarly for any $|\delta| \leq 2(\log n)^2$, we define K_δ to be the event that there are at most $(1 + \varepsilon_1)f(\delta)n^{1-a}$ in $L_\delta \cup R_\delta$ (where L_δ and R_δ are the corresponding intervals around $C + \delta$). We show that with high probability the number of signals in $L_\delta \cup R_\delta$), of length n^{-a}, is in $(1 \pm \varepsilon_1)f(\delta)n^{1-a}$, see Corollary 7 in Appendix A for the proof. Analogous to Lemma 9, we get the following lemma lower bounding the CDF of $X_\delta^2 \cdot \mathbb{1}(K_\delta)$.

Lemma 10. *The distribution of the random variable $X_\delta \cdot \mathbb{1}(K_\delta)$ conditioned on $\theta_S \in C$ stochastically dominates the exponential distribution with $\lambda = 2nf(\delta)(1 + O(n^{-1/10}))$. That is, for $d < n^{-9/10}$, $\Pr[X_{(\delta)} \cdot \mathbb{1}(K) > d|\theta_S \in C] \geq \exp[-\lambda d](1 - o(1))$.*

Let $Z(\lambda)$ be the random variable with exponential distribution. We note that, $\mathbb{E}[Z(\lambda)^2] = \frac{2}{\lambda^2}$. Therefore, by Lemmas 9 and 10 we get that $\mathbb{E}[X_0^2] \leq \frac{1}{2n^2 f(0)^2} + o(\frac{1}{n^2})$ and $\mathbb{E}[X_\delta^2] \geq \frac{1}{2n^2 f(\delta)^2} - o(\frac{1}{n^2})$. This gives us the required results of Theorem 2 and Proposition 3.

A.2 Proof of Theorem 2

In this section we prove Theorem 2, that $\mathbb{E}[X_0^2] \leq \frac{1}{2n^2 f(0)} + o(1/n^2)$. In Section A.2.1, we consider the high probability event with the following *desirable* properties:

- $\theta_S \in C$, where $C \in \mathcal{P}$, that is, $\theta_S \in I$.

- $\theta_S \in C$, where C is not *weak* (see Definition 4).

- $\theta_S \in C$, where there are *sufficiently* many signals in $N(C) = L \cup R$.

In Lemma 12 we bound the loss contributed by this high probability event. Further, in Section A.2.2 we bound the loss from the "rare events" that some desirable property does not hold. We bound the loss from the event when $\theta_S \notin I$ (in Lemma 13), when C is a weak interval (in Lemma 14), and when there are very few signals in $N(C)$ (in Lemma 15).

With this we prove the theorem,

Proof of Theorem 2. Recall that the signaling loss of the unbiased scheme π_0 is $L(\pi_0, 0) = \alpha^2 \mathbb{E}[X_0^2]$. Given $\theta_S \in C$, let K_0 be the event that there are *sufficiently* many signals in $N(C)$. We see that,

$$\mathbb{E}[X_0^2] = \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}] + \sum_{C \in W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C\}] + \sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C\}]$$

$$= \underbrace{\mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}] + \sum_{C \in W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C\}] + \sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C, \overline{K_0}\}]}_{\text{rare events}} \quad (26)$$

$$+ \sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C, K_0\}] \quad (27)$$

$$\leq \frac{1}{2n^2 f(0)^2} + o(1/n^2)$$

This is because by Lemma 12 the term in Eq. (27) is $\sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \in C, K_0\}] \leq \frac{1}{2n^2 f(0)^2} + o(1/n^2)$, and from Lemmas 13, 14, and 15 we see that all terms in Eq. (26) contribute at most $o(1/n^2)$.

$\square$

A.2.1 Contribution of High Probability Event

In this section we explain what the desirable properties are and why they are useful. We start by defining what is a *weak interval*.

Definition 4. *Let $\tau(\cdot)$ be the pdf of the posterior mean $\theta_S(\vec{x})$. We say that an interval $C \in \mathcal{P}$ is weak if $\tau(\overline{\theta}) \leq c_2 n^{1+1/22} e^{-n^{1/22\alpha}}$ for all $\overline{\theta} \in C$. Let $W_I \subset \mathcal{C}_I$ be the set of all such intervals C.*

By Claim 7 we see that $\theta_S \in C$ for some $C \in W_I$ is $O(n^{-4\log n + 1})$. Observe that, if $\tau(\overline{\theta}) \geq c' n^{1/22} n^{-4\log n + 1}$ then we can get $\tau(\overline{\theta})(1 + O(n^{-\frac{1}{22}}))$, by Claim 6.

Next we show that there are *sufficiently* many signals in L and R. We start by proving the following claim that $f(x) = f(0)(1 + O(1/\sqrt{n}))$ for all $x \in I = [-n^{-\frac{1}{2}+\epsilon}, n^{-\frac{1}{2}+\epsilon}]$.

Claim 1. *Given any well-behaved distribution with pdf f, for all $x \in I$, we have $f(x) = f(0)(1 + O(1/\sqrt{n}))$.*

Proof. Without loss of generality we assume that $x > 0$, since $f(-x) = f(x)$. By mean value theorem we see that $f(x) = f(0) + xf'(\tilde{x})$ for some $\tilde{x} \in [0, x]$. By our assumption on g' we get $|g(x)| \leq cx^m$ for some constants $c > 0$ and $m \geq 0$. This implies $|f'(x)| \leq cx^m f(x)$ for all $x > 0$. Since f is non-increasing in $(0, \infty)$ we see that $f(x) \leq f(\tilde{x}) \leq f(0)$. Moreover, $f'(\tilde{x}) \leq 0$, so we have $f'(\tilde{x}) \geq -cx^m f(x)$. By mean value theorem we have,

$$\begin{aligned}
f(x) &= f(0) + xf'(\tilde{x}) \\
&\geq f(0) - xc\tilde{x}^m f(\tilde{x}) \quad &&(\text{Since } f'(\tilde{x})/f(\tilde{x}) \geq -c\tilde{x}^m) \\
&\geq f(0)(1 - xc\tilde{x}^m) \quad &&(\text{Since } f(0) \geq f(\tilde{x})) \\
&\geq f(0)(1 - cx^{m+1}) \quad &&(\text{Since } \tilde{x} \leq x)
\end{aligned}$$

Therefore for all $x \in I$ we have $f(x) \geq f(0)(1 - c(n^{(-\frac{1}{2}+\epsilon)(m+1)}))$. Note that $m \geq 0$, hence we get $f(x) = f(0)(1 - O(n^{-1/2+\epsilon}))$.

$\square$

Using the above claim that $f(x)$ is approximately $f(0)$ for $x \in I$, we bound the number of signals in a subset $A \subset I$.

Claim 2. *Given any interval $A \subset I$ of length ℓ, the expected number signals in A is $n\ell f(0)(1 - O(n^{-1}))$. Let $Y(A)$ be the number of signals in A. For any $0 < \varepsilon < 1$, we have*

$$\Pr[Y(A) \leq (1 - \varepsilon) \mathbb{E}[Y(A)]] \leq \exp\left(-\frac{\varepsilon^2 f(0)n\ell}{2}\right).$$

Proof. Let $Y_i = 1$ if $x_i \in A$ and 0 otherwise. So we have, $\sum_{i=1}^{n} Y_i = Y(A)$. By Claim 1 we have $f(x) = f(0)(1 + O(1/\sqrt{n}))$ for all $x \in I$. Therefore we have $\Pr[x_i \in A] = \int_A f(x)dx = f(0)(1 + O(1/\sqrt{n})) \int_A dx = f(0)(1 - O(1/\sqrt{n}))\ell$. Note that, Y_i are i.i.d. random variables, and $\mathbb{E}[Y(A)] = f(0)(1 - O(1/\sqrt{n}))n\ell$. By using Chernoff bound we get

$$\Pr[Y(A) \leq (1 - \varepsilon) \mathbb{E}[Y(A)]] \leq \exp\left(-\frac{\varepsilon^2 f(0)n\ell(1 - O(n^{-1/2}))}{2}\right).$$

$\square$

We partition the interval I into intervals J of length n^{-a}/M for $M = n^{1/22}$. Let $\mathcal{J}$ denote the partition. Note that, the size of $\mathcal{J}$ is $n^{-1/2+\epsilon+a}M$.

Lemma 11. *Let $k_m = f(0)n^{1-a}/M$ For each $J \in \mathcal{J}$. Let $Y(J)$ be the number signals in interval J (of length n^{-a}/M). $\Pr[|Y(J) - k_m| \geq \varepsilon_1 k_m] \leq \exp\left(-\frac{\varepsilon_1^2 f(0)n^{1-a}}{3M}\right)$.*

Moreover, the probability that there is a $J \in \mathcal{J}$ with $|Y(J) - k_m| \geq n^{-1/20}k_m$ is at most $O(\exp^{-n^{1/22}})$

Proof. By directly invoking Claim 2 on J of length n^{-a}/M we get $\Pr[|Y(J) - k_m| \geq \varepsilon_1 k_m] \leq \exp\left(-\frac{\varepsilon_1^2 f(0)n^{1-a}}{3M}\right)$. Note that, the size of $\mathcal{J}$ is $n^{-1/2+\epsilon+a}M$. Therefore by union bound, we get $\Pr[\exists J \in \mathcal{J} : |Y(J) - k_m| \geq \varepsilon_1 k_m] \leq (n^{-1/2+\epsilon+a}M) \cdot \exp\left(-\frac{\varepsilon_1^2 f(0)n^{1-a}}{3M}\right)$ which is at most $O(\exp^{-n^{1/20}})$, for $\varepsilon_1 = O(n^{-1/22})$ and $M = n^{1/100}$. $\square$

We will now only focus on the case where *all $J \in \mathcal{J}$* has sufficiently many signals, which immediately implies the following Corollary 6. In Lemma 15 we bound the loss of the rare event that it is not the case.

Corollary 6. *Let $k_0 = (1 - \varepsilon)f(0)n^{1-a}(1 - 2/M)$. For all $C \in \mathcal{P}$, let $N(C) = L \cup R$ (of size n^{-a}). Let $Y(C)$ be the number signals in interval $N(C)$. If all $J \in \mathcal{J}$ have at least k_m signals, then $Y(C)$ is at least k_0.*

That is, $\Pr[Y(C) \leq k_0] \leq O(\exp^{-n^{1/20}})$.

Proof. Note that any $N(C) \subset I$ of length n^{-a} contains at least $M - 2$ many intervals $J \in \mathcal{J}$. By Lemma 11, we have that all $J \in \mathcal{J}$ has at least $(1 - \varepsilon)f(0)n^{1-a}$ many signals with high probability. Therefore $N(C)$ contain at least k_0 many signals with probability $1 - O(\exp^{-n^{1/20}})$. Note this is regardless of which $C \in \mathcal{P}$ we are considering. $\square$

We will now only consider the event where all the desirable properties hold. For each $C \in \mathcal{P} \setminus W_I$, let K_0 denote the event that there are at least k_0 signals in $N(C)$.

Lemma 9. *Fix any $C \in \mathcal{P} \setminus W_I$. The distribution of the random variable $X_0 \cdot \mathbb{1}(K_0)$ conditioned on $\theta_S \in C$ is stochastically dominated by the exponential distribution with $\lambda = 2nf(0)$. That is, $\Pr[X_{(0)} \cdot \mathbb{1}(K_0) > d | \theta_S \in C] < \exp[-\lambda d]$.*

Proof. Let $k_0 = (1 - \varepsilon)f(0)n^{1-a}(1 - 2/M)$. For all $d > n^{-a} + n^{-b}$ we have $\Pr[X_{(0)} \cdot \mathbb{1}(K_0) > d | \theta_S \in C] = 0$. For all $d < n^{-a}/2 + n^{-b}$, let E_d denote the event that there is no signal in B_d, where B_d is the interval of length $2d$ centered around θ_S. For all $k_0 \leq k \leq n$, let $K_0(k)$ denote the event that there are exactly k signals in $N(C)$.

25

$$\Pr[X_{(0)} \cdot \mathbb{1}(K_0) > d | \theta_S \in C]$$

$$= \Pr[\forall x_i \notin B_d, \exists x_1, x_2, \ldots, x_{k_0} \in N(C) \text{ and } K_0 | \theta_S \in C]$$

$$\leq \Pr[\exists x_1, x_2, \ldots, x_{k_0} \in N(C) \setminus B_d \text{ and } K_0 | \theta_S \in C]$$

$$= \Pr[K_0 | \theta_S \in C] \cdot \Pr[\exists x_1, x_2, \ldots, x_{k_0} \in N(C) \setminus B_d \ | K_0, \theta_S \in C]$$

$$= \Pr[K_0 | \theta_S \in C] \int_{\vec{x}_{[k_0]} \in N(C) \setminus B_d} f(\vec{x}_{[k_0]} \ | K_0, \theta_S \in C) \int_{\tilde{x} \in A(\vec{x}_{[k_0]})} f(\tilde{x} \ | \vec{x}_{[k_0]}, \theta_S \in C, K_0) d\vec{x}$$

$$= \Pr[K_0 | \theta_S \in C] \int_{\vec{x}_{[k_0]} \in N(C) \setminus B_d} f(\vec{x}_{[k_0]} \ | K_0, \theta_S \in C) d\vec{x}_{[k_0]}$$

$$= \Pr[K_0 | \theta_S \in C] \int_{\vec{x}_{[k_0]} \in N(C) \setminus B_d} \left(f(\vec{x}_{[k_0]} | K_0) \cdot \frac{\Pr[\theta_S \in C \ | \vec{x}_{[k_0]}, K_0]}{\Pr[\theta_S \in C \ | K_0(k)]} \right) d\vec{x}_{[k]} \qquad \text{(Bayes rule)}$$

We will use Lemma 19 and the fact that $\tau(\bar{\theta})$ is approximately constant in $C \pm n^{-2a+\epsilon}$.

$$\leq \Pr[K_0 | \theta_S \in C] \cdot \frac{\Pr[\theta_S \in C \pm k_0 n^{-1-a} \ | K_0]}{\Pr[\theta_S \in C | K_0]} \int_{\vec{x}_{[k_0]} \in L \setminus B_d} f(\vec{x}_{[k_0]} | K_0) d\vec{x}_{[k_0]} \qquad \text{(By Lemma 19)}$$

$$= \frac{\Pr[\theta_S \in C \pm k_0 n^{-1-a} \text{ and } K_0]}{\Pr[\theta_S \in C]} \Pr[\vec{x}_{[k_0]} \in N(C) \setminus B_d | K_0]$$

$$= \frac{\Pr[\theta_S \in C \pm k_0 n^{-1-a} \ | K_0]}{\Pr[\theta_S \in C]} \Pr[\vec{x}_{[k_0]} \in N(C) \setminus B_d] \qquad \text{(Bayes rule)}$$

$$\leq \frac{\Pr[\theta_S \in C \pm k_0 n^{-1-a} \ | K_0]}{\Pr[\theta_S \in C]} \cdot \left(1 - \frac{2d}{n^{-a}}(1 - O(1/n^{1/4})) \right)^{k_0} \qquad \text{(By Claim 1)}$$

$$= \frac{\Pr[\theta_S \in C \pm k_0 n^{-1-a}]}{\Pr[\theta_S \in C]} \cdot \frac{\Pr[K_0 | \theta_S \in C \pm k_0 n^{-1-a}]}{\Pr[K_0]} \cdot \left(1 - \frac{2d}{n^{-a}}(1 - O(1/n^{1/4})) \right)^{k_0}$$

$$\leq \left(1 + O(n^{-1/2}) \right) \cdot \frac{\Pr[K_0 | \theta_S \in C \pm k_0 n^{-1-a}]}{\Pr[K_0]} \cdot \left(1 - \frac{2d}{n^{-a}}(1 - O(1/n^{1/4})) \right)^{k_0} \qquad \text{(Since } C \notin W_I\text{)}$$

$$\leq \left(1 + O(n^{-1/2}) \right) \cdot \frac{1}{(1 - O(\exp(-n^{1/20}))} \cdot \left(1 - \frac{2d}{n^{-a}}(1 - O(1/n^{1/4})) \right)^{k_0} \qquad \text{(By Corollary 6)}$$

$$\leq \left(1 + O(n^{-1/2}) \right) \cdot (1 + O(\exp(-n^{1/20})) \cdot \left(1 - \frac{2d}{n^{-a}}(1 - O(1/n^{1/4})) \right)^{k_0}$$

Recall that $k_0 = (1 - \varepsilon)f(0)n^{1-a}(1 - 2/M)$. Let $\varepsilon = O(n^{-1/22})$ and $M = n^{1/22}$. Since we can bound $1 - x \leq e^{-x}$ we get,

$$\Pr[X_{(0)} \cdot \mathbb{1}(K_0) > d | \theta_S \in C] \leq \left(1 + O(n^{-1/2}) \right) \cdot \exp\{-2nf(0)(1 - \varepsilon)(1 - 2/M)(1 - O(n^{-1/4})) \cdot d\}$$

$$\leq \left(1 + O(n^{-1/2}) \right) \cdot \exp\{-2nf(0)(1 - O(n^{-1/22}))(1 - O(n^{-1/4})) \cdot d\}$$

$\square$

We finally bound the cost of the event with all the desirable properties.

Lemma 12. *Fix any $C \in \mathcal{P} \setminus W_I$. Then we have, $\mathbb{E}[X_0^2 \cdot \mathbb{1}(K_0) | \theta_S \in C] \leq \frac{1}{2n^2 f(0)^2} + o(1/n^2)$.*

Proof. Let $Z(\lambda)$ be the random variable with exponential distribution. We observe that $\mathbb{E}[X_0^2 \cdot \mathbb{1}(K_0) | \theta_S \in C] \leq \mathbb{E}[Z(\lambda)^2]$ for $\lambda = 2nf(0)(1 - O(n^{-1/22}))(1 - O(n^{-1/4}))$, because of the stochastic dominance proved

above in Lemma 9. Moreover, $\mathbb{E}[Z(\lambda)^2] = \frac{2}{\lambda^2}$. Hence we get $\mathbb{E}[X_0^2 \cdot \mathbb{1}(K_0)|\theta_S \in C] \leq \frac{1}{2n^2 f(0)^2}(1 - O(n^{-1/22}))(1 - O(n^{-1/4})) \leq \frac{1}{2n^2 f(0)^2} + o(1/n^2)$.

$\square$

In Section A.2.2 we bound the loss due to the rare events of $\theta_S \notin I$, $C \in W_I$, and $\overline{K_0}$, that is, the number of signals in $N(C)$ at most $(1 - \varepsilon_1)n^{1-a}f(0)(M-2)/M$. We show that these contribute up to $o(1/n^2)$ loss. With this we complete the proofs of Theorem 2.

A.2.2 Contribution of Rare Events

In this section we bound the loss from the rare events from Eq. (26).

Lemma 13. $\mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}] \leq O\left(\exp(-\frac{n^{2\epsilon}A}{2})\right)$ for some constant $A > 0$.

Proof. Recall that $\theta_S(\vec{x})$ is the MMSE estimator and $\theta_S - \theta^* \to^d \mathcal{N}(0, C_{\mathcal{I}}/n)$. Hence the probability that $\theta_S \notin I$ is at most $\exp(-n^{2\epsilon}A)$ for some $A > 0$. Let $P = \exp(-n^{2\epsilon}A)$. To bound $\mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}]$ we see that,

$$\mathbb{E}[X_0^2 \cdot \mathbb{1}\{\theta_S \notin I\}] = \int_0^\infty \Pr[X_0^2 > y \wedge \theta_S \notin I]dy$$

$$= \int_0^{1/P^{1/2}} \Pr[\theta_S \notin I]\Pr[X_0^2 > y|\theta_S \notin I]dy + \int_{1/P^{1/2}}^\infty \Pr[X_0^2 > y \wedge \theta_S \notin I]dy$$

$$\leq P^{1/2} + 2\int_{1/P^{1/2}}^\infty \Pr[X_i > \theta_S + \sqrt{y} \wedge \theta_S > 0]dy \qquad \text{(For any arbitrary choice of } i)$$

$$\leq P^{1/2} + 2\int_{1/P^{1/2}}^\infty \Pr[X_i > \sqrt{y}]dy$$

$$\leq P^{1/2} + 2\int_{1/P^{1/2}}^\infty (e^{-\sqrt{y}})dy \qquad \left(\text{Since } \int_x^\infty f(z)dz < e^{-x} \text{ for all } x > Q\right)$$

$$\leq P^{1/2} + 4(1/P^{1/4} + 1)\exp\left(-1/P^{1/4}\right)$$

$$= O\left(\exp(-\frac{n^{2\epsilon}A}{2})\right)$$

Recall that, $P = \exp(-n^{2\epsilon}A$. So we have $P^{1/2} = O\left(\exp(-\frac{n^{2\epsilon}A}{2})\right)$. Since xe^{-x} is $O(e^{-x})$ for x sufficiently large, the term $4(1/P^{1/4} + 1)\exp\left(-1/P^{1/4}\right) = O\left(\exp\left(-\exp(\frac{n^{2\epsilon}A}{4})\right)\right)$.

$\square$

For a well-behaved distribution we have $1 - F(x) \leq c_3 e^{-x}$ for all $x > Q$.

Claim 3. *Let T_Q be the event that all $|x_i| > Q$. Then $\mathbb{E}[X_0^2\mathbb{1}(T_Q \wedge \theta_S \in I)] \leq o(1/n^2)$.*

Proof. Since $\theta_S \in I = [-n^{-1/2+\epsilon}, n^{-1/2+\epsilon}]$, and all signals $|x_i| > Q$ are outside I, we have that $X_0 = \min_i |x_i - \theta_S| \leq |x_i| + n^{-1/2+\epsilon}$ for all x_i. Let $t(x) = (|x| + n^{-1/2+\epsilon})^2$. Hence we have,

$$\mathbb{E}[X_0^2 \mathbb{1}(T_Q \wedge \theta_S \in I)] \leq \mathbb{E}[(|x_1| + n^{-1/2+\epsilon})^2 \mathbb{1}(T_Q \wedge \theta_S \in I)] \qquad \text{(For an arbitrary choice of } i = 1)$$

$$= \mathbb{E}[t(x_1)\mathbb{1}(T_Q)|\theta_S \in I]$$

$$= \int_{x_1:|x_1|>Q} \cdots \int_{\vec{x}_{-1}} t(x_1)f(x_1) \prod_{i \neq 1} f(x_i)\mathbb{1}(|x_i| > Q) \cdot \mathbb{1}(\theta_S(\vec{x}) \in I)d\vec{x}$$

$$\leq \int_{x_1:|x_1|>Q} \cdots \int_{\vec{x}_{-1}} t(x_1)f(x_1) \prod_{i \neq 1} f(x_i)\mathbb{1}(|x_i| > Q)d\vec{x}$$

$$\leq \int_{x_1:|x_1|>Q} t(x_1)f(x_1)dx_1 \left(2\exp(-Q)\right)^{n-1} \qquad (\text{ By tail bound Assumption of } f)$$

$$\leq \exp\left(-\frac{Q(n-1)}{2}\right) 2 \int_Q^\infty (x + n^{-1/2+\epsilon})^2 f(x)dx$$

$$\leq \exp\left(-\frac{Q(n-1)}{2}\right) \cdot Q \cdot O(1) \qquad (\text{ By tail bound Assumption of } f)$$

$\square$

Recall that $W_I \subset \mathcal{P}$ is the set of all intervals C such that $\tau(\overline{\theta}) < c'n^{1/22}n^{-4\log n+1}$ for all $\overline{\theta} \in C$.

Lemma 14. *Then* $\sum_{C \in W_I} \mathbb{E}[X_0^2 \mathbb{1}\{\theta_S \in C\}] \leq O\left(Q^2 n^{-4\log n+1}\right) + o(1/n^2)$.

Proof. Let $\overline{T}_Q$ be the event that there is some $|x_i| \leq Q$. Since $\theta_S \in I$ and there is some $|x_i| \leq Q$, we have that $X_0 \leq Q + n^{-\frac{1}{2}+\epsilon}$. Thus,

$$\sum_{C \in W_I} \mathbb{E}[X_0^2 \mathbb{1}\{\theta_S \in C\} \cdot \mathbb{1}\{\overline{T}_Q\}] \leq (Q + n^{-\frac{1}{2}+\epsilon})^2 \Pr[\theta_S \in W_I] \leq (Q + n^{-\frac{1}{2}+\epsilon})^2 O\left(n^{-4\log n+1}\right)$$

where the last inequality follows from Claim 7(proved in section A.4)that $\Pr[\theta_S \in W_I] \leq O\left(n^{-4\log n+1}\right)$.

Moreover, by Claim 3 proved above, we have $\sum_{C \in W_I} \mathbb{E}[X_0^2 \mathbb{1}\{\theta_S \in C\} \cdot \mathbb{1}\{T_Q\}] \leq \mathbb{E}[X_0^2 \mathbb{1}(T_Q \wedge \theta_S \in I)] \leq o(1/n^2)$. Thus proving the lemma.

$\square$

Lemma 15. *Let* $k_0 = (1 - O(n^{-1/22}))f(0)n^{1-a}$. *Let* $\overline{K_0}$ *be the event such that* $Y(C) \leq k_0$. *Then,* $\sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_0^2 \cdot \mathbb{1}\{\overline{K_0} \text{ and } \theta_S \in C\}] \leq (Q + n^{-1/2})^2 O\left(\exp\left(-n^{1/20}f(0)\right)\right) + O(\exp^{-Qn}) \leq o(1/n^2)$.

Proof. Consider the case where $N(C)$ doesn't have k_0 signals. Let A_0 denote that event. By Claim 3, when all $|x_i| > Q$ and $\theta_S \in I$, we bound the expected X_0^2 by $O(\exp^{-Qn})$. If there is even a single $|x_i| \leq Q$ (denoted by the event $\overline{T}_Q$), then we can bound X_0^2 by $(Q + n^{-1/2+\epsilon})^2$ because $\theta_S \in I = [-n^{-1/2+\epsilon}, n^{-1/2+\epsilon}]$. By corollary 6, we get that $\Pr[A_0 \cap \overline{T}_Q \cap \theta_S \in C] \leq \Pr[A_0] \leq \exp\left(-\frac{n^{-1/11}f(0)n^{1-a}}{3M}\right)$.

$$\mathbb{E}[X_0^2 \cdot \mathbb{1}\{\overline{K_0} \text{ and } \theta_S \in C\}] \leq \mathbb{E}[X_0^2 \cdot \mathbb{1}\{A_0 \text{ and } \theta_S \in C\} \cdot \mathbb{1}\{T_Q\}] + \mathbb{E}[X_0^2 \cdot \mathbb{1}\{A_0 \text{ and } \theta_S \in C\} \cdot \mathbb{1}\{\overline{T}_Q\}]$$

$$\leq O(\exp^{-Qn}) + (Q + n^{-1/2+\epsilon})^2 \exp\left(-\frac{n^{-1/11}f(0)n^{1-a}}{3M}\right)$$

$$\leq o(1/n^2)$$

$\square$

A.3   Proof of Proposition 3

In this section we prove Proposition 3 that, for all sufficiently large n and all δ, $\mathbb{E}[X_\delta^2] \geq \frac{1}{2n^2 f(\delta)^2} - o(1/n^2)$. We focus only on $|\delta| \leq 2(\log n)^2$ and δ such that $f(\delta) \geq n^{-1/100}$ [8]. Similar to the proof of Theorem 2 we consider the high probability event where all the following desirable properties hold:

- $\theta_S \in C$, where $C \in \mathcal{P}$, that is, $\theta_S \in I$.

- $\theta_S \in C$, where C is not *weak* (see Definition 4).

- $\theta_S \in C$, where there are *sufficiently few* signals in $N(C_\delta) = L_\delta \cup R_\delta$, an interval of length n^{-a}, and there are no signals in C_δ.

Since these hold with high probability we prove Proposition 3 as follows.

Proof Proposition 3. Recall that the signaling loss of a biased scheme π with $\delta(\pi) = \delta$ is $L(\pi, \delta(\pi)) \geq L(\pi_\delta, \delta) = \alpha^2 \mathbb{E}[X_\delta^2]$. Given $\theta_S \in C$, let K_δ be the event that there are *sufficiently few* signals in $N(C_\delta)$ and there are no signals in C_δ. We see that,

$$
\begin{aligned}
\mathbb{E}[X_\delta^2] &\geq \sum_{C \in \mathcal{P} \setminus W_I} \mathbb{E}[X_\delta^2 \cdot \mathbb{1}\{\theta_S \in C, K_\delta\}] \\
&\geq \sum_{C \in \mathcal{P} \setminus W_I} \Pr[\{\theta_S \in C, K_\delta\}] \cdot \underbrace{\mathbb{E}[X_\delta^2 \mid \theta_S \in C, K_\delta]}_{\text{conditional expectation}} \qquad (28) \\
&\geq \frac{1}{2n^2 f(\delta)^2} - o(1/n^2)
\end{aligned}
$$

This is because by Lemma 17 the conditional expectation term in Eq. (28) is $\mathbb{E}[X_\delta^2 \mid \theta_S \in C, K] \geq \frac{1}{2n^2 f(\delta)^2}(1 - o(1))$. We note that $\{\vec{x} : \forall J_\delta \in \mathcal{J}_\delta \text{ with suffiencetly few signals}\} \subset K_\delta$ for all $C \in \mathcal{P}$, and by Lemma 16 we see that $\Pr[\overline{K_\delta}] \leq \Pr[\exists J_\delta \in \mathcal{J}_\delta \text{ with too many signals}] \leq O(\exp\{f(\delta)n^{1/20}\})$. By Lemma 14 we bound the probability of C is weak. Thus we have,

$$
\sum_{C \in \mathcal{P} \setminus W_I} \Pr[\{\theta_S \in C, K_\delta\}] \geq \Pr[\forall J_\delta \in \mathcal{J}_\delta \text{ with sufficiently few signals, and } \theta_S \notin W_I]
$$

$$
\geq 1 - O(\exp\{f(\delta)n^{1/20}\}) - O(\exp\{f(\delta)n^{1/55}\})O(n^{4\log n - 1}).
$$

$\square$

We start by showing a more general version of Claim 1.

Claim 4. *Given any well-behaved distribution with pdf f, for all $x \in I + \delta$, we have $f(x) = f(\delta)(1 + O(1/n))$.*

Proof. By mean value theorem we see that $f(x) = f(\delta) + (x - \delta)f'(\tilde{x})$ for some $\tilde{x} \in [\delta, x]$. By our assumption on g' we get $|g(x)| \leq cx^m$ for some constant $c > 0$. This implies $|f'(x)| \leq c|x|^m f(x)$ for all x. Since $|\delta| < 2(\log n)^2$ and $x \in I + \delta$, we have $|x| \leq 2(\log n)^2 + n^{-1/2+\epsilon}$. By mean value theorem we have,

$$
\begin{aligned}
|f(x) - f(\delta)| &= |(x - \delta)f'(\tilde{x})| \\
&\leq |x - \delta|c|\tilde{x}|^m f(\tilde{x}) \qquad (\text{Since } |f'(\tilde{x})| \leq c|\tilde{x}|^m f(\tilde{x})) \\
&\leq c(n^{-1/2+\epsilon})(2(\log n)^2 + n^{-1/2+\epsilon})^m f(\tilde{x}) \\
&\leq c'(n^{-1/2+\epsilon})n^{1/4}f(\tilde{x}) \qquad (\text{Since } (\log n)^2 m = o(n^{1/4})) \\
&\leq f(\delta) + c'n^{-1/4+\epsilon}f(\tilde{x})
\end{aligned}
$$

---

[8] When $f(\delta) \leq n^{-1/100}$ we see that $\mathbb{E}[X_\delta^2] \geq \Omega(\frac{n^{100}}{n^2}) >> \mathbb{E}[X_0^2]$.

Without loss of generality we assume that $x, \delta > 0$, since f is symmetric. Thus we get $f(\tilde{x}_i)$ is between $f(x)$ and $f(\delta)$, because f is single-peaked.

Suppose $f(x) \geq f(\tilde{x}) \geq f(\delta)$, then $f(x) \leq f(\delta) + c'n^{-1/4+\epsilon}f(\tilde{x}) \leq f(\delta) + c'n^{-1/4+\epsilon}f(x)$. So we get $f(x)(1 - c'n^{-1/4+\epsilon}) \leq f(\delta)$. Thus, $f(x) \leq f(\delta)\frac{1}{1-c'n^{-1/4+\epsilon}} \leq f(\delta)(1 + c''n^{-1/4+\epsilon})$ for some constant $c'' > 0$.

Similarly if $f(x) \leq f(\tilde{x}) \leq f(\delta)$, then $f(x) \geq f(\delta) - c'n^{-1/4+\epsilon}f(\tilde{x}) \geq f(\delta) - c'n^{-1/4+\epsilon}f(\delta)$. Thus we get, $f(x) \geq f(\delta)(1 - O(n^{-1/4+\epsilon}))$.

Therefore for $\delta < 2(\log n)^2$ and all $x \in I + \delta$ we have $f(x) = f(\delta)(1 - O(n^{-\frac{1}{4}+\epsilon}))$. $\qquad\square$

Claim 5. *Given any interval $A \subset I + \delta$ of length ℓ, the expected number signals in A is $n\ell f(\delta)(1 - O(n^{-1/4}))$. Let $Y(A)$ be the number of signals in A. For any $0 < \varepsilon_1 < 1$, we have*

$$\Pr[|Y(A) - \mathbb{E}[Y(A)]| \geq \varepsilon_1)\,\mathbb{E}[Y(A)]] \leq \exp\left(-\frac{\varepsilon_1^2 f(\delta)n\ell}{3}\right).$$

Proof. Let $Y_i = 1$ if $x_i \in A$ and 0 otherwise. So we have, $\sum_{i=1}^{n} Y_i = Y(A)$. By Claim 4 we have $f(x) = f(\delta)(1 + O(n^{-1/4+\epsilon}))$ for all $x \in I + \delta$. Therefore we have $\Pr[x_i \in A] = \int_A f(x)dx = f(\delta)(1 + O(n^{-1/4+\epsilon}))\int_A dx = f(\delta)(1 + O(n^{-1/4+\epsilon}))\ell$. Note that, Y_i are i.i.d. random variables, and $\mathbb{E}[Y(A)] = f(\delta)(1 + O(n^{-1/4+\epsilon}))n\ell$. By using Chernoff bound we get

$$\Pr[Y(A) \geq (1 + \varepsilon)\,\mathbb{E}[Y(A)]] \leq \exp\left(-\frac{\varepsilon_1^2 f(\delta)n\ell(1 + O(n^{-1/4+\epsilon}))}{3}\right).$$

$\qquad\square$

We again partition $I + \delta$ into intervals J_δ of size n^{-a}/M. Exactly following Lemma 11 we see that *all* $J_\delta \subset I + \delta$ have $(1 \pm \varepsilon)f(\delta)n^{1-a}(1 + O(n^{-1/4+\epsilon})$ many signals in L_δ and R_δ.

Lemma 16. *Let $k_m = f(\delta)n^{1-a}/M$. The probability that there is a $J_\delta \in \mathcal{J}_\delta$ with more than $(1 + \varepsilon)k_m$ signals (or less than $(1 - \varepsilon)k_m$ is $O(\exp\{\varepsilon^2 f(\delta)n^{1-a}/3M\})$.*

Corollary 7. *Let $k_\delta = f(\delta)n^{1-a}$, let $k^* = (1 + \varepsilon_1)k_\delta$, and $k' = (1 - \varepsilon_1)k_\delta(1 - 2/M)$. Let $Y(N_\delta)$ be the number signals in interval $N(C_\delta)$ (of length n^{-a}). $\Pr[(Y(N_\delta) \notin [k', k^*]] \leq \exp\left(-\frac{\varepsilon_1^2 f(\delta)n^{1-a}}{3M}\right).$*

Let K_δ denote the event that there are at most $(1 + \varepsilon_1)k_\delta$ and at least $(1 - \varepsilon_1)k_\delta(1 - 2/M)$ many signals in $N(C_\delta)$, and there are no signals in C_δ.

Lemma 10. *The distribution of the random variable X_δ conditioned on $K_\delta, \theta_S \in C$ stochastically dominates (up to a factor of $(1 - o(1))$) the exponential distribution with $\lambda = 2nf(\delta)(1 + O(n^{-1/10}))$. That is, for $d < n^{-9/10}$, $\Pr[X_{(\delta)} > d \,|K_\delta, \theta_S \in C] \geq \exp[-\lambda d]\frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]}.$*

Proof. Let $k^* = (1 + \varepsilon_1)k_\delta$, and $k' = (1 - \varepsilon_1)k_\delta(1 - 2/M)$. For all $k' \leq k \leq k^*$, define $K_\delta(k)$ to be event where $\#x_i \in N(C_\delta) = k$, and there are no signals in C_δ .

For all $d < n^{-a} + n^{-b}$, let E_d denote the event that $\forall x \in L, x \notin B_d$, where $B(d)$ is the interval of length $2d$ centered around $\theta_S + \delta$.

$$\Pr[X_{(\delta)} \cdot \mathbb{1}(K_\delta) > d|\theta_S \in C]$$
$$= \Pr[\forall x_i \notin B_d, \exists x_1, x_2, \ldots, x_{k'} \in L \text{ and } K_\delta|\theta_S \in C]$$
$$= \sum_{k=k'}^{k^*} \Pr[\forall x_i \notin B_d, \exists x_1, x_2, \ldots, x_k \in L \text{ and } K_\delta(k)|\theta_S \in C]$$
$$= \sum_{k=k'}^{k^*} \Pr[\exists x_1, x_2, \ldots, x_k \in L \setminus B_d \text{ and } K_\delta(k)|\theta_S \in C]$$
$$= \sum_{k=k'}^{k^*} \Pr[K_\delta(k)|\theta_S \in C] \cdot \Pr[\exists x_1, x_2, \ldots, x_k \in L \setminus B_d \ |K_\delta(k), \theta_S \in C]$$
$$= \sum_{k=k'}^{k^*} \Pr[K_\delta(k)|\theta_S \in C] \int_{\vec{x}_{[k]} \in L \setminus B_d} f(\vec{x}_{[k]} \ |K_\delta(k), \theta_S \in C) \int_{\tilde{x} \in A(\vec{x}_{[k]}) \setminus (L \cup C)} f(\tilde{x} \ |\vec{x}_{[k]}, \theta_S \in C, K_\delta(k))d\vec{x}$$
$$= \sum_{k=k'}^{k^*} \Pr[K_\delta(k)|\theta_S \in C] \int_{\vec{x}_{[k]} \in L \setminus B_d} f(\vec{x}_{[k]} \ |K_\delta(k), \theta_S \in C)d\vec{x}_{[k]}$$
$$= \sum_{k=k'}^{k^*} \Pr[K_\delta(k)|\theta_S \in C] \int_{\vec{x}_{[k]} \in L \setminus B_d} \left(f(\vec{x}_{[k]}|K_\delta(k)) \cdot \frac{\Pr[\theta_S \in C \ |\vec{x}_{[k]}, K_\delta(k)]}{\Pr[\theta_S \in C \ |K_\delta(k)]} \right) d\vec{x}_{[k]} \qquad (\text{ Bayes rule})$$

We will use Lemma 20 and the fact that $\tau(\bar{\theta})$ is approximately constant in $C \notin W_I$.

$$\geq \sum_{k=k'}^{k^*} \Pr[K_\delta(k)|\theta_S \in C] \cdot \frac{\Pr[\theta_S \in C \setminus E \ |K_\delta(k)]}{\Pr[\theta_S \in C|K_\delta(k)]} \int_{\vec{x}_{[k]} \in L \setminus B_d} f(\vec{x}_{[k]}|K_\delta(k))d\vec{x}_{[k]} \qquad (\text{By Lemma 20})$$
$$= \sum_{k=k'}^{k^*} \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta(k)]}{\Pr[\theta_S \in C]} \Pr[\vec{x}_{[k]} \in L \setminus B_d|K_\delta(k)]$$
$$\geq \sum_{k=k'}^{k^*} \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta(k)]}{\Pr[\theta_S \in C]} \cdot \left(1 - \frac{2d}{n^{-a}}(1 + O(1/n^{1/4})) \right)^k \qquad (\text{By Claim 4})$$
$$\geq \left(1 - \frac{2d}{n^{-a}}(1 + O(1/n^{1/4})) \right)^{k^*} \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]}$$

We bound $\left(1 - \frac{2d}{n^{-a}}\right)^{f(\delta)(1+\varepsilon_1)n^{1-a}}$ by observing that $(1 - x) \geq e^{-x-x^2}$ for $x < 1/2$. We will choose of $\varepsilon_1 = O(n^{-1/20})$ and consider $d \leq n^{-a-1/20} = n^{-17/20}$, this gives us $\varepsilon_3 = O(n^{-1/20})$.

$$\Pr[X_{(\delta)} \cdot \mathbb{1}(K_\delta) > d|\theta_S \in C] \geq \left(1 - \frac{2d}{n^{-a}}(1 + O(1/n^{1/4})) \right)^{k^*} \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]}$$
$$\geq \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]} \exp(-2dn(1 + \varepsilon_2)f(\delta)) \exp(-(2d)^2 n^{1+a} f(\delta)(1 + \varepsilon_2))$$
$$\geq \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]} \exp(-2dn(1 + \varepsilon_3)f(\delta)) \qquad (\text{For } 2d < n^{-a}\varepsilon_3)$$

$\square$

Finally, for each $C \in \mathcal{P} \setminus W_I$, we bound the loss $\mathbb{E}[X_\delta^2 \cdot \mathbb{1}\{K_\delta, \theta_S \in C\}]$.

Lemma 17. *Let K_δ be the event such that $(1 - \varepsilon_1)(1 - 2/M)k_\delta \leq Y(L_\delta) \leq (1 + \varepsilon_1)k_\delta$, and no signals in C_δ. Then for all $C \in \mathcal{P} \setminus W_I$,*

1. for $f(\delta) > n^{-1/100}$, we have $\mathbb{E}[X_\delta^2 \cdot \mathbb{1}\{K_\delta, \theta_S \in C\}] \geq \frac{1}{2n^2 f(\delta)^2}(1 - o(1))\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]$,

2. for $f(\delta) \leq n^{-1/100}$, we have $\mathbb{E}[X_\delta^2] \geq c_5 \frac{n^{100}}{n^2}$.

Proof. Let $Z(\lambda)$ be the random variable with exponential distribution. We observe that $\mathbb{E}[X_\delta^2 \cdot \mathbb{1}(K_\delta)|\theta_S \in C] \geq \mathbb{E}[Z(\lambda)^2 \cdot \mathbb{1}\{d < (n^{-9/10})^2\}] \cdot \frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]}$ for $\lambda = 2nf(\delta)(1 + \varepsilon_3)$, because of the stochastic dominance proved above in Lemma 10. Moreover, $\mathbb{E}[Z(\lambda)^2 \cdot \mathbb{1}\{d < (n^{-17/20})^2\}] = \frac{2}{\lambda^2}(1 - O(\lambda n^{-17/20} \exp^{-\lambda n^{-17/20}}))$. Hence we get $\mathbb{E}[X_\delta^2 \cdot \mathbb{1}(K_\delta)|\theta_S \in C] \geq \frac{1}{2n^2 f(\delta)^2}(1 - o(1))\frac{\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]}{\Pr[\theta_S \in C]}$, for $f(\delta) \geq c_4 n^{-1/100}$ for a constant $c_4 > 0$.

We finish by noting that, for sufficiently large n, when $f(\delta) = O(\frac{1}{n^{1/100}})$, we get $\mathbb{E}[X_\delta^2] \geq c_5(\frac{n^{1/100}}{n^2}) >> \frac{1}{2n^2 f(0)^2} \geq \mathbb{E}[X_0^2]$ for a constant $c_5 > 0$.

$\square$

Proof of Proposition 3. By lemma 17, for a sufficiently large n and any $\delta < 2(\log n)^2$ and $f(\delta) > n^{-1/100}$, we have $\mathbb{E}[X_\delta^2] \geq \sum_{C \in \mathcal{P} \setminus W_I} \Pr[\theta_S \in C] \mathbb{E}[X_\delta^2|\theta_S \in C] \geq \sum_{C \in \mathcal{P} \setminus W_I} \frac{1}{2n^2 f(\delta)^2}(1 - o(1))\Pr[\theta_S \in C \setminus E \text{ and } K_\delta]$.

Given $E \subset C$ as the union of first and last $kn^{-1-a+\epsilon}$ length interval, we define $2E \subset C$ to be the union of the first and last $2kn^{-1-a+\epsilon}$ intervals. Since K_δ is the event that there are $(1 \pm \varepsilon)k_\delta$ many signals in $N(C_\delta)$ and there are no signals in C_δ, that $\Pr[\theta_S \in C \setminus E] \geq \Pr[\theta_S \in C \setminus 2E|Y(N(C_\delta) \cup C_\delta) \in (1 \pm \varepsilon)k_\delta]$, this is because if $\theta_S \in C \setminus 2E$ and $\vec{x}_{[k]} \in N(C_\delta) \cup C_\delta$ then rearranging the signals $\vec{x}_{[k]}$ by moving the signals in C_δ into $N(C_\delta)$ changes θ_S by at most $kn^{-1-a}f(\delta)$. Moreover, we have $\Pr[K_\delta] = \Pr[Y(N(C_\delta) \cup C_\delta) \in (1 \pm \varepsilon)k_\delta] \cdot \Pr[Y(C_\delta) = 0 \,|Y(N(C_\delta) \cup C_\delta) \in (1 \pm \varepsilon)k_\delta]$. The probability that there are no signals in C_δ (of length $n^{-12/11}$) is at least $(1 - O(n^{-1/11}))$. Thus we get,

$$\mathbb{E}[X_\delta^2] \geq \sum_{C \in \mathcal{P} \setminus W_I} \frac{1}{2n^2 f(\delta)^2}(1 - o(1))\Pr[\theta_S \in C \setminus 2E \text{ and } Y(N(C_\delta) \cup C_\delta) \in (1 \pm \varepsilon)k_\delta](1 - O(n^{-1/11}))$$

Let $\mathcal{G}_\delta$ denote the event that all $J_\delta \in \mathcal{J}_\delta$ has $(1 \pm \varepsilon)k_\delta/M$ signals. Thus we get,

$$\mathbb{E}[X_\delta^2] \geq \sum_{C \in \mathcal{P} \setminus W_I} \frac{1}{2n^2 f(\delta)^2}(1 - o(1))\Pr[\theta_S \in C \setminus 2E \text{ and } \mathcal{G}_\delta](1 - O(n^{-1/11}))$$

$$= (1 - O(n^{-1/11}))\Pr[\mathcal{G}_\delta \text{ and } \exists C \in \mathcal{P} \setminus W_I \text{ s.t. } \theta_S \in C \setminus 2E]$$

$$\geq (1 - O(n^{-1/11}))\left(1 - \Pr[\overline{\mathcal{G}_\delta}] - \Pr[\overline{\exists C \in \mathcal{P} \setminus W_I \text{ s.t. } \theta_S \in C \setminus 2E}]\right)$$

$$\geq (1 - O(n^{-1/11}))\left(-\Pr[\overline{\mathcal{G}_\delta}] + \Pr[\exists C \in \mathcal{P} \setminus W_I \text{ s.t. } \theta_S \in C \setminus 2E]\right)$$

Recall that, $\Pr[\theta_S \in I] \geq (1 - \exp(-n^{2\epsilon}A)$, and by Claim 7 (proved in section A.4) we have $\Pr[\theta_S \in W_I] \leq O(n^{-4\log n + 1})$. By Lemma 16 we have $\Pr[\overline{\mathcal{G}_\delta}] \leq O(\exp\{\varepsilon^2 f(\delta) n^{1-a}/3M\})$. Hence we get $\mathbb{E}[X_\delta^2] \geq \frac{1}{2n^2 f(\delta)^2}(1 - o(1))$.

$\square$

A.4 Proofs of Common Lemmas

In this section we will prove all the lemmas that are common to both Theorem 2 and Proposition 3.

In Lemma 18, we characterize the effect of a single signal x on the posterior mean $\theta_S(\vec{x})$. This lemma directly implies Corollary 8, where we show that if for any signals $\vec{x}$ rearranging at most k signals in L (and R) to get $\vec{y}$ guarantees that the new posterior mean is $\theta_S(\vec{y}) \in \theta_S(\vec{x}) \pm O(kn^{-1-a})$.

Lemma 18. *For any signals $\vec{x}$ observed by the sender we have,*

$$\left|\frac{\partial \theta_S(\vec{x})}{\partial x_i}\right| \leq c_1 Var_{\theta \sim \mathcal{D}_S(\vec{x})}[\theta] + 2\theta_S(\vec{x})^2,$$

where $Var_{\theta \sim \mathcal{D}_S(\vec{x})}[\theta]$ is the variance of the senders posterior distribution.

Proof. Let $h_S(\theta|\vec{x})$ denote the pdf of the sender's posterior distribution, $h(\theta)$ be the (constant) pdf of the diffuse prior. Note that the pdf of a signal x given that the state of the world is θ (denoted by $\hat{f}(x|\theta)$) equals $f(x-\theta)$, where f is the pdf of $\mathcal{D}(0)$. Recall that when the sender observes $\vec{x}$ they update their posterior in a Bayesian way. Hence we have,

$$h_S(\theta|\vec{x}) = \frac{\prod_i \hat{f}(x_i|\theta)h(\theta)}{\int_{\hat{\theta}} \prod_i \hat{f}(x_i|\hat{\theta})h(\hat{\theta})d\hat{\theta}}$$

$$= \frac{\prod_i f(x_i-\theta)}{\int_{\hat{\theta}} \prod_i f(x_i-\hat{\theta})d\hat{\theta}}$$

and the sender's posterior mean is

$$\theta_S(\vec{x}) = \frac{\int_{\theta} \theta \prod_i f(x_i-\theta)d\theta}{\int_{\theta} \prod_i f(x_i-\theta)d\theta} \tag{29}$$

We want to understand $\frac{\partial \theta_S(\vec{x})}{\partial x_i}$ which is the effect of a single signal x_i on the posterior mean. We can deduce

$$
\begin{aligned}
\frac{\partial \theta_S(\vec{x})}{\partial x_i} &= \frac{\int \theta f'(x_i-\theta)\prod_{j\neq i} f(x_j-\theta)d\theta \int \prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&\quad - \frac{\int \theta \prod_j f(x_j-\theta)d\theta \int f'(x_i-\theta)\prod_{j\neq i} f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&= \frac{\int \theta g(x_i-\theta)\prod_j f(x_j-\theta)d\theta \int \prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&\quad - \frac{\int \theta \prod_j f(x_j-\theta)d\theta \int g(x_i-\theta)\prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2}
\end{aligned} \tag{30}
$$

where $g(y) = \frac{f'(y)}{f(y)}$.

Using the mean value theorem we can write $g(x_i-\theta) = g(x_i) - g'(\tilde{x}_i)\theta$ for $\tilde{x}_i \in [x_i-\theta, x_i]$. We then obtain:

$$
\begin{aligned}
\frac{\partial \theta_S(\vec{x})}{\partial x_i} &= \frac{\int \theta \left(g(x_i) - g'(\tilde{x}_i)\theta\right)\prod_j f(x_j-\theta)d\theta \int \prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&\quad - \frac{\int \theta \prod_j f(x_j-\theta)d\theta \int \left(g(x_i) - g'(\tilde{x}_i)\theta\right)\prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&= -\frac{\int \theta g'(\tilde{x}_i)\theta \prod_j f(x_j-\theta)d\theta \int \prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2} \\
&\quad + \frac{\int \theta \prod_j f(x_j-\theta)d\theta \int g'(\tilde{x}_i)\theta \prod_j f(x_j-\theta)d\theta}{\left(\int \prod_j f(x_j-\theta)d\theta\right)^2}
\end{aligned} \tag{31}
$$

Next, we note that for a well behaved distribution, $|g'(\tilde{x}| \leq c_1$. Hence we can bound,

33

We can then simplify:

$$\left| \frac{\partial \theta_S(x)}{\partial x_i} \right| \leq [c_1] \left(\frac{\int_\theta \theta^2 \prod_j f(x_j - \theta) d\theta}{\int_\theta \prod_j f(x_j - \theta) d\theta} + \left(\frac{\int_\theta \theta \prod_j f(x_j - \theta) d\theta}{\int_\theta \prod_j f(x_j - \theta) d\theta} \right)^2 \right) \tag{32}$$

$$= [c_1] \left(E_{\theta \sim D_S(\vec{x})}(\theta^2 | x) + E_{\theta \sim D_S(\vec{x})}(\theta | x)^2 \right)$$

$$= c_1 \left(Var_{\theta \sim D_S(\vec{x})}[\theta] + 2 E_{\theta \sim D_S(\vec{x})}(\theta | x)^2 \right) \tag{33}$$

$\square$

Next we bound the shift in θ_S when rearranging k signals in L (or R) using Lemma 18. We note that $Var_{\theta \sim D_S(\vec{x})}[\theta] = O(1/n)$ and $\theta_S \in I$.

Corollary 8. *Assume that the posterior mean $\theta_S(\vec{x})$ lies within interval $C \subset I$. Consider a subset of k signals in a subset A of length ℓ. Any rearrangement of these signals within A changes the posterior mean by $O(k \ell n^{-1+2\epsilon})$.*

Proof. We prove this by using mean value theorem on the function $\theta_S : \mathbb{R}^n \to \mathbb{R}$. Given $\vec{x}$, consider a subset of signals $\vec{x}_{[k]} \in A$. Let $\vec{y}$ be any vector such that $y_i \in A$ for all $i \in [k]$ and $y_i = x_i$ for the rest. By mean value theorem we get for some $\vec{z}$ such that:

$$\theta_S(\vec{y}) = \theta_S(\vec{x}) + \nabla \theta_S(\vec{z}) \cdot (\vec{y} - \vec{x})$$

Note that $|y_i - x_i| \leq \ell$ for all $i \in [k]$ and $y_i - x_i = 0$ otherwise. That is, at most k terms with $|y_i - x_i| \neq 0$. Hence we get,

$$|\theta_S(\vec{y}) - \theta_S(\vec{x})| = |\nabla \theta_S(\vec{z}) \cdot (\vec{y} - \vec{x})|$$

$$\leq \sum_{i=1}^{k} c_1 \left(Var_{\theta \sim D_S(\vec{x})}[\theta] + 2\theta_S(\vec{x})^2 \right) |y_i - x_i| \quad \text{(By Lemma 18)}$$

$$= c_1 O(n^{-1+2\epsilon}) \sum_{i=1}^{k} \ell$$

Since $x_i, y_i \in A$ we bound $|x_i - y_i| \leq \ell$. Further, by bounding $Var_{\theta \sim D_S(\vec{x})}[\theta]$ by $O(1/n)$, and $\theta_S(\vec{x})$ by $n^{-1/2+\epsilon}$, we get $\theta_S(\vec{y}) = \theta_S(\vec{x}) \pm k\ell O(n^{-1+2\epsilon})$ when rearranging at most k signals in each A.

$\square$

Next we show that the density of the posterior mean is nice in the interval C.

Claim 6. *Assume that the density function f has exponential (or thinner) tails. Let $\tau(\cdot)$ be the density function of the posterior mean. Then for all $\bar{\theta} \in C$ we have $\tau(\bar{\theta} + \epsilon') = \tau(\bar{\theta}) \left(1 + O(n^{-1/22}) \right) + O(e^{-n^{1/22\alpha}})$ for all $0 < \epsilon' \leq 1/n^b$.*

Proof. Fix a posterior mean $\bar{\theta} \in [-n^{-\frac{1}{2}+\epsilon}, n^{-\frac{1}{2}+\epsilon}]$ and consider all the signal draws $X(\bar{\theta}) = \{\vec{x} | \bar{\theta}(\vec{x}) = \bar{\theta}\}$ that generate this posterior mean. We know that $\tau(\bar{\theta}) = \int_{\vec{x} \in X(\bar{\theta})} \prod_{x_i \in \vec{x}} f(x_i) dx$. Now consider $\bar{\theta} + \epsilon'$. We can couple all the signal realizations in $X(\bar{\theta} + \epsilon')$ and $X(\bar{\theta})$ by considering uniform shifts of the corresponding $\vec{x}$ by ϵ'. This follows from the assumption of a diffuse prior (if we shift all signals by ϵ' we also shift the posterior mean by ϵ'). That is, $\tau(\bar{\theta} + \epsilon') = \int_{\vec{x} \in X(\bar{\theta})} \prod_{x_i \in \vec{x}} f(x_i + \epsilon') dx$.

Next, consider the probability of observing x versus the coupled signal realizations $x + \epsilon'$.

$$\prod_{x_i \in \vec{x}} f(x_i + \epsilon') = \prod_{x_i \in \vec{x}} [f(x_i) + f'(\tilde{x}_i)\epsilon'] \tag{34}$$

Recall that, by our assumption on g' we have $|f'(x)| \leq c|x|^m f(x)$.

- Note that, for all $|x_i| < 4(\log n)^2$, we have

$$
\begin{aligned}
|f(x_i + \epsilon') - f(x_i)| &= |\epsilon' f'(\tilde{x}_i)| \\
&\leq |\epsilon' c(|\tilde{x}|^m)| f(\tilde{x}) \\
&\leq |\epsilon' c(|4\log n|^{2m})| f(\tilde{x}) \\
&\leq |c'(n^{-b+1/22})| f(\tilde{x}) \qquad (|\epsilon'| \leq n^{-12/11} \text{ and } (\log n)^{2m} = o(n^{1/22}))
\end{aligned}
$$

  Note that, wlog we can assume that $\text{sign}(x_i + \epsilon) = \text{sign}(x_i)$ because f is symmetric. Thus we get $f(\tilde{x}_i)$ is between $f(x_i)$ and $f(x_i + \epsilon')$, because f is single-peaked. If $f(x_i) \leq f(\tilde{x}_i) \leq f(x_i + \epsilon')$ we get $0 \leq f(x_i + \epsilon') - f(x_i) \leq c'(n^{-b+1/22})|f(\tilde{x}_i) \leq c'(n^{-b+\epsilon})|f(x_i + \epsilon')$. Thus, $f(x_i) \leq f(x_i + \epsilon) \leq f(x_i)\left(\frac{1}{1-c'(n^{-b+\epsilon}}\right) \leq f(x_i)(1 + c''n^{-b+\epsilon})$.

  Similarly if $f(x_i) \geq f(\tilde{x}_i) \geq f(x_i + \epsilon)$, then we get $f(x_i) \geq f(x_i + \epsilon) \geq f(x_i)\left(\frac{1}{1+c'(n^{-b+\epsilon}}\right) \geq f(x_i)(1 - c'n^{-b+\epsilon})$.

- Further, by our assumption that f has exponential tails we have $\Pr_{\vec{x}}[\exists x_i \text{ s.t. } |x_i| > 4(\log n)^2] \leq ne^{-4(\log n)^2} = (n^{-4\log n + 1})$.

If $x_i \in [-4(\log n)^2, 4(\log n)^2]$ for all i, then we bound $\prod_{x_i \in \vec{x}} f(x_i + \epsilon') = \prod_i f(x_i)(1 + O(n^{-b+\frac{1}{22}})) = (\prod_i f(x_i))(1 + O(n^{-b+\frac{1}{22}}))^n = (\prod_i f(x_i))(1 + O(n^{1-b+\frac{1}{22}})) = (\prod_i f(x_i))(1 + O(n^{-\frac{1}{22}}))$ for $b = 12/11$. Hence we get

$$
\tau(\overline{\theta} + \epsilon') = \int_{-n^{1/22\alpha}}^{n^{1/22\alpha}} \left(\prod f(x_i)\right)(1 + O(n^{-\frac{1}{22}}))\mathbb{1}\{\vec{x} \in X(\hat{\theta})\}d\vec{x} \qquad + O(n^{-4\log n + 1})
$$

Therefore, $\tau(\overline{\theta} + \epsilon') = \tau(\overline{\theta})(1 + O(n^{-\frac{1}{22}})) + O(n^{-4\log n+1})$ for all $0 < \epsilon' < 1/n^b$ and $\overline{\theta} \in C$. □

Recall that $W_I \subset \mathcal{P}$ is the set of all intervals C such that $\tau(\overline{\theta}) < c'n^{1/22}n^{-4\log n+1}$ for all $\overline{\theta} \in C$ and some constant $c' > 0$. We show that the total probability mass of these intervals is $O(n^{-4\log n+1})$.

Claim 7. $\Pr[\theta_S \in W_I] \leq O(n^{-4\log n+1})$.

Proof. This is simply because there are at most $2n^{b-\frac{1}{2}+\epsilon}$ many intervals in $\mathcal{C}$ (since each interval is of size n^{-b}). Therefore,

$$
\Pr[\theta_S \in W_I] \leq \sum_{C \in W_I} \int_C \tau(\overline{\theta})d\overline{\theta} \leq 2n^{b-\frac{1}{2}+\epsilon}\left(n^{-b} \cdot c'n^{1+1/22-4\log n}\right) \leq O(n^{-4\log n+6/11+\epsilon}).
$$

□

Using Corollary 8 we bound the correlation between the events $\theta_S \in C$ and any realization of k anecdotes in L_δ (and R_δ). Observe that, if $\tau(\overline{\theta}) \geq c'n^{1/22}n^{-4\log n+1}$ then we can get $\tau(\overline{\theta})(1 + O(n^{-\frac{1}{22}}))$.

Lemma 19. *Fix any $C \in \mathcal{P} \setminus W_I$. Let $K_\delta(k)$ denote the event that there are (at most) k signals in each L_δ (and R_δ). Let $\vec{x}_L$ be a subset of k (at most $c'n^{1-a}$) signals in L_δ. Then $\Pr[\theta_S \in C \mid K_\delta(k), x_L] \leq \Pr[\theta_S \in C \pm kn^{-1-a}|K_\delta(k)]$.*

Proof. Let $A(x_L) = \{\tilde{x} : \theta(x_L \cup \tilde{x}) \in C\}$ for any subset of k signals $x_L \in L_\delta$. By Corollary 8 we know that by changing x_L to any y_L (in L_δ) for each $\tilde{x} \in A(x_L)$ we get $\tilde{y} = \tilde{x}$ such that $\theta_S(\vec{y}) = \theta_S(\vec{x}) + O(kn^{-1-a+\epsilon})$. If $\theta_S(\vec{x}) \in C$ then $\theta_S(\vec{y}) \in C \pm kn^{-1-a+\epsilon}$.

$$
\Pr[\theta_S \in C|x_L, K_\delta(k)] \leq \Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}|y_L, K_\delta(k)] \qquad \text{For all } y_L \text{ of size } k
$$

So if for some c we have,

$$\Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}|y_L, K_\delta(k)] \geq c$$
$$\Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}, K_\delta(k) \text{ and } y_L] \geq c \Pr[K_\delta(k), y_L]$$
$$\int_{y_L} \Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}, K_\delta(k) \text{ and } y_L] d\vec{y}_{[k]} \geq \Pr[K_\delta(k), \exists y_1, \ldots, y_k \in L]$$
$$\Pr[\theta_S \in C \pm kn^{-1-a+\epsilon} \text{ and } K_\delta(k)] \geq c \Pr[K_\delta(k)]$$
$$\Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}|K_\delta(k)] \geq c$$

Therefore, $\Pr[\theta_S \in C \pm kn^{-1-a+\epsilon}|K_\delta(k)] \geq \Pr[\theta_S \in C|x_L]$.

$\square$

Similarly, we have a lower bound on $\Pr[\theta_S \in C|K_\delta(k), x_L]$.

Lemma 20. *Fix any $C \in \mathcal{P} \setminus W_I$. Let $K_\delta(k)$ denote the event that there are k signals in each L_δ (and R_δ), and no signals in C_δ. Let $\vec{x}_L$ be a subset of k (at most $c'n^{1-a}$) signals in L_δ. Then $\Pr[\theta_S \in C|K_\delta(k), x_L] \geq \Pr[\theta_S \in C \setminus E|K_\delta(k)]$.*

Proof. Let $A(x_L) = \{\tilde{x} : \theta(x_L \cup \tilde{x}) \in C\}$ for any subset of k signals $x_L \in L_\delta$. By Corollary 8 we know that by changing x_L to any y_L (in L_δ) for each $\tilde{x} \in A(x_L) \setminus (L_\delta \cup C_\delta)$ we get $\tilde{y} = \tilde{x}$ such that $\theta_S(\vec{y}) = \theta_S(\vec{x}) + O(kn^{-1-a+\epsilon})$. Define $E \subset C$ as the union of the first and last $kn^{-1-a+\epsilon}$ length sub-interval of C. If $\theta_S(\vec{y}) \in C \setminus E$ then $\theta_S(\vec{x}) \in C$.

$$\Pr[\theta_S \in C|K_\delta(k), x_L] \geq \Pr[\theta_S \in C \setminus E|K_\delta(k), y_L] \quad \text{For all } y_L \text{ of size } k$$

So if for some c we have,

$$\Pr[\theta_S \in C \setminus E|K_\delta(k), y_L] \leq c$$
$$\Pr[\theta_S \in C \setminus E, K_\delta(k) \text{ and } y_L] \leq c \Pr[K_\delta(k), y_L]$$
$$\Pr[\theta_S \in C \setminus E, K_\delta(k)] \leq c \Pr[K_\delta(k)]$$
$$\Pr[\theta_S \in C \setminus E|K_\delta(k)] \leq c$$

Therefore we have, $\Pr[\theta_S \in C \setminus E|K_\delta(k)] \leq \Pr[\theta_S \in C|K_\delta(k), x_L]$.

$\square$

B Proofs from Section 4.3

B.1 Two Signals

In this section we prove Proposition 4, that the optimal translation-invariant signaling scheme may be biased, even when there are two signals drawn from a symmetric distribution and when the moral stances of the sender and receiver coincide.

We first note that the sender's posterior mean will always be equal to the average of the two signals received.

Lemma 21. *For any symmetric distribution F and $n = 2$, we have $\theta_S(x_1, x_2) = \frac{x_1+x_2}{2}$.*

Proof. Fix some realization of the observed signals x_1 and x_2, and view θ and $\epsilon_1, \epsilon_2 \sim F$ as random variables. Choose any $\theta' \in \mathbb{R}$, and set $\epsilon_1' = x_1 - \theta'$ and $\epsilon_2' = x_2 - \theta'$. Then realization $(\theta^*, \epsilon_1, \epsilon_2) = (\theta', \epsilon_1', \epsilon_2')$ is consistent with the observed signals x_1 and x_2.

Now define $\theta'' = x_1 + x_2 - \theta'$, $\epsilon_1'' = -\epsilon_2'$, and $\epsilon_2'' = -\epsilon_1'$. Note then that realization $(\theta^*, \epsilon_1, \epsilon_2) = (\theta'', \epsilon_1'', \epsilon_2'')$ is also consistent with the observed signals. Also, since F has mean 0 and is symmetric about its mean, and since θ^* is drawn from a diffuse prior, these two realizations have equal probability density, which is to say that $D_S(x_1, x_2)$ assigns equal probability density to θ' and θ''.

We conclude that given x_1 and x_2, the posterior distribution $D_S(x_1, x_2)$ places equal likelihood on θ' and $x_1 + x_2 - \theta'$, for all $\theta' \in \mathbb{R}$. As the average of each such pair is $(x_1 + x_2)/2$, we conclude that this must be the mean of the posterior as claimed. $\qquad\square$

Fix any arbitrary signaling scheme π. Define random variable d by $d = x' - \pi(\vec{x})$ where x' denotes whichever of the two signals is not sent by π. Then $|d| = |x_1 - x_2|$ which is independent of π; the signaling scheme π determines only the distribution over the sign of d. Since $\theta_S(\vec{x}) = \frac{x_1 + x_2}{2}$ we have that $\theta_S(\vec{x}) = \pi(\vec{x}) + d/2$. Moreover, the bias of π is precisely $\delta(\pi) = -\mathbb{E}[d/2]$, and hence we have that $\theta_R(\pi(\vec{x}), \pi) = \pi(\vec{x}) + \mathbb{E}[d/2]$. Let $h(\cdot)$ denote the pdf of the random variable d. Note that $h(\cdot)$ doesn't depend on the state of the world.

Lemma 22. *The sender's optimal signaling scheme π maximizes $\mathbb{E}\left[d^2\right] - \mathbb{E}\left[d\right]^2$.*

Proof. Since moral stances are assumed to be aligned, (5) implies that the sender's loss is proportional to

$$
\begin{aligned}
\mathbb{E}_{\vec{x}}(\theta_S(\vec{x}) - \theta_R(\pi(\vec{x}), \pi))^2 &= \mathbb{E}_{\vec{x}}(\|\pi(\vec{x}) + d/2 - \pi(\vec{x}) - \mathbb{E}(d/2)\|^2) \\
&= \mathbb{E}_{\vec{x}}(d/2 - \mathbb{E}(d/2))^2 \\
&= \mathbb{E}(d^2)/4 - \mathbb{E}(\mathbb{E}(d)^2)/4
\end{aligned}
$$

which is proportional to $\mathbb{E}(d^2) - \mathbb{E}(\mathbb{E}(d)^2)$ as required. $\qquad\square$

Note that the first term $\mathbb{E}(d^2)$ is only depends on the distribution on the signals and is independent of the signaling scheme. Hence we want to find a scheme π such that $\mathbb{E}(d)^2$ is maximized.

As π is translation invariant, the probability with which it returns the minimum signal is a function of $|d|$. So write $p(|d|)$ for the probability that $\pi(\vec{x})$ returns the lesser of the two signals when $|d| = |x_1 - x_2|$. We can describe π as sending the minimum signal if some event A happens, otherwise send the maximum signal, where event A only depends on $|d|$ and randomness. For example, the *always send min* signaling has $p(|d|) = 1$ for all distances d. For the *send a random signal* scheme, $p(|d|) = 1/2$ for all d.

We can write $h(d) = h(d \wedge A) + h(d \wedge \bar{A})$. Note that when $d < 0$, we have $h(d \wedge \bar{A}) = 0$. Similarly when $d > 0$, we have $h(d \wedge A) = 0$, since $d > 0$ if and only if the lesser signal was sent.

Recall that the sender wants to maximize $\mathbb{E}_{\vec{x}}(\mathbb{E}(d|\pi(\vec{x}))^2)$ by lemma 22. Fix any arbitrary translation invariant signaling scheme π. Let $A = A(\pi)$ be the event that describe when to send the min signal in π. Let $d > 0$. We have $h(d) = h(d \wedge A)$, since $h(d \wedge \bar{A}) = 0$ for $d > 0$.

$$
\begin{aligned}
h(d \wedge A) &= 2 \int_{\epsilon=-\infty}^{\infty} f\left((\epsilon_1 = \epsilon, \epsilon_2 = \epsilon + d) \wedge A\right) d\epsilon \\
&= 2 \int_{\epsilon=-\infty}^{\infty} f\left(A | \epsilon_1 = \epsilon, \epsilon_2 = \epsilon + d\right) f(\epsilon_1 = \epsilon, \epsilon_2 = \epsilon + d) d\epsilon \\
&= 2 \int_{\epsilon=-\infty}^{\infty} p(d) f(\epsilon) f(\epsilon + d) d\epsilon \\
&= p(|d|) \frac{1}{\sqrt{\pi}} \exp\left[\frac{-d^2}{4}\right]
\end{aligned}
$$

Similarly for $d < 0$, we have $h(d) = h(d \wedge \bar{A})$, and we get $h(d \wedge \bar{A}) = (1 - p(|d|))\sqrt{\pi} \exp\left[\frac{-d^2}{4}\right]$.

$$\mathbb{E}[d] = \int_{d=-\infty}^{\infty} d \cdot h(d) dd$$

$$= \int_{d=0}^{\infty} d \cdot h(d \wedge A) dd + \int_{d=-\infty}^{0} d \cdot h(d \wedge \bar{A}) dd$$

$$= \int_{d=0}^{\infty} d \cdot p(|d|) \frac{1}{\sqrt{\pi}} \exp\left[\frac{-d^2}{4}\right] dd + \int_{d=0}^{\infty} (-d) \cdot (1 - p(|d|)) \frac{1}{\sqrt{\pi}} \exp\left[\frac{-d^2}{4}\right] dd$$

$$= 2 \int_{d=0}^{\infty} d \cdot p(|d|) \frac{1}{\sqrt{\pi}} \exp\left[\frac{-d^2}{4}\right] dd - \frac{2}{\sqrt{\pi}}$$

Recall that we want to maximize $|\mathbb{E}[d]|$, which amounts to maximizing (or minimizing) $2 \int_{d=0}^{\infty} d \cdot p(|d|) \frac{1}{\sqrt{\pi}} \exp\left[\frac{-d^2}{4}\right] dd$, which happens when $p(|d|) = 1$ for all d (or $p(|d|) = 0$ for all d). This exactly corresponds to the *always send min* signaling scheme (or *always send max* respectively).

B.2 Three Signals

In this section we prove Proposition 5, which is that the optimal signaling scheme may be biased even for $n = 3$ signals drawn from a uniform distribution.

Suppose that F is the uniform distribution on $[-1/2, 1/2]$. For notational convenience, we will additively shift all values so that samples are drawn from $[0, 1]$, which is to say that we equivalently shift values such that $\theta = 1/2$.

Given 3 signals $x_1 \le x_2 \le x_3$, the sender's posterior is a uniform distribution on $[x_3 - 1/2, x_1 + 1/2]$. The posterior mean is therefore $\theta_S(\vec{x}) = \frac{x_1 + x_3}{2}$.

Optimal Unbiased Scheme Consider the optimal unbiased scheme, call it π_0. As noted in Section 4.2, this scheme sends the closest signal to $\theta_S(\vec{x})$. Since θ_S is the midpoint of the interval $[x_1, x_3]$, and since x_2 falls in that interval, the optimal unbiased scheme always sends signal x_2.

Let's calculate the mean squared error of signal x_2 relative to θ. The CDF of x_2 is given by

$$F(y) = \Pr[x_2 < y] = y^3 + 3y^2(1 - y)$$

for $y \in [0, 1]$, since the first term is the probability all three samples are less than y, and the second term is the probability that two of the three samples are less than y. Now write $d = |\theta - x_2| = |1/2 - x_2|$. Then 1 minus the CDF of d is given by

$$G(z) = \Pr[d > z] = \Pr[y < (1/2 - z)] + \Pr[y > (1/2 + z)] = 2\Pr[y < (1/2 - z)] = 2F(1/2 - z)$$

for $z < 1/2$, and $G(z) = 0$ for $z \ge 1/2$. Here we used that $\Pr[y > (1/2 + z)] = \Pr[y < (1/2 - z)]$ by symmetry.

The total loss of signaling scheme π_0 is therefore

$$\mathbb{E}[d^2] = \int_0^\infty \Pr[d^2 > z] dz$$

$$\int_0^\infty \Pr[d > \sqrt{z}] dz$$

$$\int_0^\infty 2F(1/2 - \sqrt{z}) dz$$

$$= 1/20$$

where the final equality is via numerical calculation.

A Better Biased Scheme We'll now build a scheme with strictly less loss than π_0. Write π_δ for the scheme that returns whichever of the three points is closest to $\theta_S(\vec{x}) + \delta$, for some $\delta \geq 0$. We will eventually choose $\delta = 1/5$, but for now we'll proceed with general δ.

Which point does π_δ return? Write x^* for the random variable representing the point that π_δ returns. Recall that $\theta_S(\vec{x}) = (x_1 + x_3)/2$, so $\theta_S(\vec{x}) + \delta$ is always closer to x_3 than x_1. The distance to point x_3 is $|\theta_S(\vec{x}) + \delta - x_3| = (x_3 - x_1)/2 + \delta$, and the distance to point x_2 is $|\theta_S(\vec{x}) + \delta - x_2| = (x_3 + x_1)/2 + \delta - x_2$. So the point x_2 will be closest precisely if $x_2 > x_1 + 2\delta$. To summarize: $x^* = x_2$ if $x_2 > x_1 + 2\delta$, otherwise $x^* = x_3$.

As before, let's work out the CDF for x^*. What is the probability that $x^* < y$? If all three points are less than y (which happens with probability y^3) then x^* certainly is. On the other hand, if $x_2 > y$, then certainly $x^* > y$ as well. If $x_2 < y$ and $x_3 > y$ (which happens with probability $3y^2(1 - y)$), then $x^* < y$ only if $x^* = x_2$, which occurs if and only if $x_2 > x_1 + 2\delta$. The conditional probability of that last event is equivalent to the probability that two random variables, each drawn uniformly from $[0, y]$, are at least distance 2δ apart from each other. So we can write the CDF as

$$
\begin{aligned}
F[y] &= \Pr[x^* < y] \\
&= y^3 + 3y^2(1 - y)\Pr[|x_1 - x_2| > 2\delta \mid x_2 < y] \\
&= y^3 + 3y^2(1 - y) \cdot 2 \cdot \int_0^{y-2\delta} \frac{1}{y} \cdot \frac{y - (x + 2\delta)}{y} dx.
\end{aligned}
$$

To justify the last equality, consider drawing one point uniformly from $[0, y]$, so with uniform density $\frac{1}{y}$. What is the probability that a second drawn point is at least 2δ larger? If the first point (call it x) is greater than $y - 2\delta$ the probability is 0. Otherwise it is $\frac{y-(x+2\delta)}{y}$. Integrating over x gives the probability of this event. We then double that probability to account for the possibility that the first point drawn is the larger one.

Now write $d = |x^* - \delta - \theta| = |x^* - (1/2 + \delta)|$. This will be the distance between the receiver's action and θ, if the receiver shifts the received signal x^* by δ. Note that this may not be the optimal action of the receiver, but the optimal action performs at least as well as $\mathbb{E}[d^2]$.

Now 1 minus the CDF of d is given by

$$
G(z) = \Pr[d > z] = \begin{cases} F(1/2 + \delta - z) + 1 - F(1/2 + \delta + z) & \text{if } 0 < z < \frac{1}{2} - \delta, \\ F(1/2 + \delta - z) & \text{if } \frac{1}{2} - \delta < z < \frac{1}{2} + \delta, \\ 0 & \text{if } z > \frac{1}{2} + \delta. \end{cases}
$$

Note that unlike the case of π_0, the fact that $\delta > 0$ breaks symmetry in the calculation of G. But the reasoning is the same: $d > z$ precisely if either x^* is greater than $1/2 + \delta + z$ or x^* is less than $1/2 + \delta - z$.

Finally, as before, the total loss of the scheme π_δ is

$$
\begin{aligned}
\mathbb{E}[d^2] &= \int_0^\infty \Pr[d^2 > z]dz \\
&\quad \int_0^\infty \Pr[d > \sqrt{z}]dz \\
&\quad \int_0^\infty G(\sqrt{z})dz
\end{aligned}
$$

For $\delta = 1/5$, this integral evaluates to approximately 0.036, which is less than 1/20.

Intuition and Discussion. Why is π_δ better than π_0? In this case, $\theta_S(\vec{x}) = (x_1 + x_3)/2$, so $\theta_S(\vec{x})$ is highly correlated with x_1 and x_3 and much less correlated with x_2. This is very specific to the uniform distribution. By selecting the point closest to $\theta_S(\vec{x}) + 1/5$, we are trading off probability of returning x_2 with probability of returning x_3. Because of the improved correlation with x_3, the location of x_3 is more highly concentrated, given $\theta_S(\vec{x})$, than the location of x_2. So by targeting an "expected" location of x_3 relative to $\theta_S(\vec{x})$ (in this case, $\theta_S(\vec{x}) + 1/5$), we can reduce the variance of the distance to the closest point.

B.3 Importance of Translation Invariance

In this section we prove Proposition 6, that the optimal signaling scheme may not be translation invariance, even when there are two signals drawn from a symmetric distribution and when the moral stances of the sender and receiver coincide.

Suppose that $n = 2$ and F is the uniform distribution over $\{-11, -10, 10, 11\}$. Fix a signaling scheme π and write define random variables $\hat{x}$ and x' such that $\hat{x} = \pi(x_1, x_2)$ and x' is the other signal. Write $d = \hat{x} - x'$. Then as argued in Appendix B.1, the sender's goal is to maximize $\mathbb{E}_{\hat{x}}[\mathbb{E}[d|\hat{x}]^2]$. We had actually argued further in Appendix B.1 that for translation invariant signals this expression is equal to $\mathbb{E}[d]^2$, but since we now consider non-translation-invariant signaling schemes we will maintain the more general form of this expression.

By Proposition 4, the optimal translation-invariant signaling scheme is to send the minimum signal. Under this scheme the value d is always non-negative, and the distribution of d is independent of $\hat{x}$. By enumerating the possibilities for $\epsilon_1 \sim F$ and $\epsilon_2 \sim F$ (recalling that $x_i = \theta + \epsilon_i$), we have that d is distributed as follows:

$$d = \begin{cases} 0 & \text{w.p. } 4/16 \\ 1 & \text{w.p. } 4/16 \\ 20 & \text{w.p. } 2/16 \\ 21 & \text{w.p. } 4/16 \\ 22 & \text{w.p. } 2/16. \end{cases}$$

We therefore have that $\mathbb{E}_{\hat{x}}[\mathbb{E}[d|\hat{x}]^2] = \mathbb{E}[d]^2 = (10.75)^2 = 115.5625$, for the scheme that always sends the minimum signal.

We now construct a non-translation-invariant signaling scheme that achieves a higher value of $\mathbb{E}_{\hat{x}}[\mathbb{E}[d|\hat{x}]^2]$. Our signaling scheme will behave precisely as the one that sends the minimum signal, except that when $|x_1 - x_2| = 1$ it returns whichever of x_1 or x_2 has an odd integer part. Note that this always uniquely specifies either x_1 or x_2. Moreover, conditional on the event that $|x_1 - x_2| = 1$ (but not on the exact values of x_1 and x_2), it is equally likely for the signal returned by this scheme to be the larger or smaller of the two.

Conditional on $\hat{x}$ having an odd integer part, the value of d is distributed as follows:

$$d = \begin{cases} -1 & \text{w.p. } 4/20 \\ 0 & \text{w.p. } 4/20 \\ 1 & \text{w.p. } 4/20 \\ 20 & \text{w.p. } 2/20 \\ 21 & \text{w.p. } 4/20 \\ 22 & \text{w.p. } 2/20. \end{cases}$$

Thus, for any such $\hat{x}$, we have $\mathbb{E}_{\hat{x}}[d|\hat{x}]^2 = (8.4)^2 = 70.56$. Conditional on $\hat{x}$ having an even integer part, the value of d is distributed as follows:

$$d = \begin{cases} 0 & \text{w.p. } 4/12 \\ 20 & \text{w.p. } 2/12 \\ 21 & \text{w.p. } 4/12 \\ 22 & \text{w.p. } 2/12. \end{cases}$$

Thus, for any such $\hat{x}$, we have $\mathbb{E}_{x_1, x_2}[d|\hat{x}]^2 = (14)^2 = 196$. As each case is equally likely, we conclude that $\mathbb{E}_{\hat{x}}[\mathbb{E}_{x_1, x_2}[d|\hat{x}]^2] = (70.56 + 196)/2 = 133.28$. This is larger than 115.5625, as required.