Robustly Obedient Data Driven Persuasion under State-dependent Sender Payoffs

Shabab Neel
neelshababhasin@gmail.com
Institute of Health Economics, University of Dhaka
Bangladesh

Abstract

We study a data-driven persuasion framework with state-dependent Sender preferences. Under unrestricted signaling, state-dependent payoffs can destroy the global bilinear structure required for standard minimax arguments. By restricting attention to robustly obedient direct signaling, where recommendations remain optimal for the Receiver across all observationally consistent priors, we restore bilinearity. We show that the resulting robustly obedient game has a pure-strategy saddle point whenever the robustly obedient policy set is nonempty. We characterize when this restriction is without loss relative to unrestricted persuasion, and derive both a local exactness result and a quantitative bound on the loss from robust obedience under Receiver action stability. Turning to computational complexity, we prove that deciding whether a target robust payoff is achievable is coNP-complete in succinct multi-variable environments, even with two recommended actions. Finally, we establish that this computational bottleneck can be resolved via dynamic programming that yields an XP algorithm parameterized by the width of the interaction hypergraph and the number of actions.

ACM Reference Format:
Shabab Neel. 2026. Robustly Obedient Data Driven Persuasion under State-dependent Sender Payoffs. In . ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Keywords: Bayesian Persuasion, Information Design, Robust Obediency, Computational Complexity, Parameterized Algorithms

1 Introduction

Bayesian persuasion typically assumes that the Sender knows the Receiver's prior over states. In many applications, however, the Sender may have to act without observing this prior while having information about the Receiver's behavior. [Ros26] develops a data-driven persuasion framework in which the Sender observes an empirical aggregate distribution of actions and uses it to construct the set of priors that rationalize the observed behavior. The Sender then chooses a signal to maximize her worst-case payoff over this identified set. His analysis shows that the resulting problem can have a saddle-point characterization under state-independent restrictions, but also that such a characterization can fail in more general environments.

This paper studies a natural extension in which the Sender's payoff depends on both the Receiver's action and the state, $v(a, \omega)$. This extension matters because state dependence can make the Sender's payoff respond directly to uncertainty about the underlying prior, even when the Receiver's preferences remain unchanged. More importantly, under unrestricted signaling, a fixed message can induce different Receiver actions at different priors in the identified set. The Sender's payoff is then generally only piecewise linear in the prior, so the linear structure underlying standard minimax arguments is lost.

In this paper, we study this problem in a robustly obedient direct signaling setting: every recommendation that is sent with positive probability must be optimal for the Receiver at every prior in the data-driven identified set. This is a restriction on implementation rather than an assumption that the Receiver mechanically follows recommendations independently of his prior. Within this class, the Receiver's action following each recommendation is invariant across the identified set, which restores bilinearity of the Sender's payoff in the signaling policy and prior. Beyond establishing economic existence and structural robustness, this framework also allows us to study the exact algorithmic complexity of robust persuasion under succinct environments that can help us uncover both sharp hardness limits and structured tractability via hypergraph decompositions.

The main result, Theorem 3.1, shows that whenever the robustly obedient policy set is nonempty, the resulting game admits a pure-strategy saddle point between the Sender and the adversarial prior[1][2]. It also shows that, for a maxmin-optimal policy, a worst-case prior can be selected from an extreme point of the data-driven identified set. Observation 1 complements this result by establishing, for every fixed robustly obedient policy, the linearity and Fenchel biconjugacy of the Sender's payoff as a function of the prior, together with fixed-policy vertex attainment. Then Claim 1 establishes that restricting the Sender to robustly obedient policies entails no loss in optimal payoff if and only if the optimal policy for the unrestricted problem naturally satisfies obedience across every prior in the data-driven uncertainty set. Theorem 4.3 shows that when an unrestricted

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference'17, Washington, DC, USA
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

[1] The result is deliberately restricted to the robustly obedient subclass $\Pi_{RO}(\alpha)$. It does not assert saddle-point existence for the unrestricted signaling problem. The latter remains potentially nonlinear because Receiver actions may vary with the prior, as discussed in Section 2.

[2] The "adversarial prior" is a representation of the Sender's maxmin evaluation over the data-consistent set $\mathcal{P}(\alpha)$, rather than an additional economic player choosing the prior strategically. The saddle-point formulation is therefore a minimax representation of the Sender's robust objective.

optimal policy has a strictly positive obedience margin at a baseline prior, it remains robustly obedient and therefore globally optimal whenever the uncertainty set around that prior is sufficiently small relative to the margin. Finally, Theorem 4.4 proves that any nominal optimal policy can be perturbed into a strictly robustly obedient policy via convex combination with an interior policy. This is the economic theory part of the paper.

Now turning to computational aspects of the paper, Claims 2 and 3 along with Proposition 5.1 characterize structural properties of the posteriors, while Theorems 5.2 and 5.3 establish coNP-completeness for the decision problem of robust persuasion under succinct representations. Finally, Theorem 5.4 demonstrates that this computational bottleneck can be bypassed via dynamic programming, which yields an XP algorithm parameterized by tree width and the number of actions. Thus, the results identify a tractable robustly implementable subclass of data-driven persuasion in which state-dependent payoffs coexist with a clean convex-analytic solution and explicit computational guarantees.

1.1 Related Works

The paper builds on the landmark Bayesian persuasion framework of [KG11], which already permits state-dependent Sender and Receiver preferences. our contribution is therefore not state dependence itself, but its interaction with data-driven prior uncertainty and [Ros26] provides the closest benchmark of it. Observed action frequencies generate a set of rationalizing priors, and the Sender maximizes her worst-case payoff over that set. Rosenthal's results also show why prior variation cannot generally be ignored. Outside environments satisfying additional structural conditions, saddle-point existence may fail. We retain his identification framework but impose robust obedience to recover a bilinear game when Sender payoffs are state-dependent.

The broader literature studies persuasion under uncertainty about beliefs, but the source and timing of that uncertainty differ. [Kos22] considers a Sender who is ignorant of the Receiver's prior and chooses an information structure against the worst prior. [LR26] study public persuasion with receivers holding heterogeneous priors. In contrast, our ambiguity set is not exogenously specified, but rather endogenously generated by observed behavioral data through Rosenthal's rationalization criterion. [DP22] likewise develop a robust persuasion criterion, but their robustness concerns uncertainty about information available to Receivers and strategy selection, not a data-identified prior set. [PS26] study persuasion without priors using minimax regret, which differs from our maxmin objective and identified-set construction. And [HF26] explicitly study robust obedience when beliefs vary over a neighborhood of a nominal prior. Their analysis reinforces the importance of requiring obedience uniformly across possible beliefs, while this paper's robust-obedience condition is defined instead over the economically generated set of priors consistent with observed actions. This distinction links our restriction to the emerging robust-information-design literature while preserving the distinctive data-driven structure of [Ros26].

Bayesian persuasion has already been studied computationally, pioneered by [DX16] who proved polynomial-time approximability for i.i.d. actions but #P-hardness for independent actions. Later work extended this setting to limited signal spaces, established hardness under communication bounds, and to combinatorial receiver actions, showed NP-hardness across matroid and path constraints [GHHS22, FS22]. These works analyze standard persuasion under explicit priors, whereas our framework adopts state-dependent sender payoffs and data-driven prior ambiguity, while using robust obedience to restore bilinearity.

Furthermore, succinct representations in game theory, such as Boolean circuit games, frequently elevate equilibrium decisions to promise-S_2^P, Σ_2^P, or NEXP, while robust optimization with polyhedral uncertainty yields coNP-hardness [FIKU08, SV12, CG19]. We prove that evaluating robust persuasion in succinct multi-variable environments mirrors this intractability that yield coNP-completeness. To bypass worst-case hardness, CS literature leverages graph-theoretic structures. [Fre90] showed tree-structured constraint satisfaction problems are linearly solvable, and [PCdG+19] demonstrated that treewidth dictates exact inference in graphical models. Importing these structural decomposition techniques, we develop an XP dynamic programming algorithm parameterized by hypergraph treewidth and action cardinality.

2 Setup

We study a strategic communication game between a Sender (she) and a Receiver (he) in a finite-state, finite-action environment with state-dependent preferences. Let Ω be a finite set of states of the world and A a finite set of Receiver actions. A representative state is denoted by $\omega \in \Omega$ and a representative action by $a \in A$.

The Receiver's preferences are represented by the state-dependent utility function

$$u : \Omega \times A \to \mathbb{R},$$

and the Sender's preferences are represented by

$$v : \Omega \times A \to \mathbb{R}.$$

The state of the world is distributed according to an underlying prior

$$\mu_0 \in \operatorname{int} \Delta(\Omega),$$

where $\Delta(\Omega)$ denotes the simplex of probability distributions over Ω. The Receiver observes the true prior μ_0, whereas the Sender does not. Instead, the Sender observes aggregate behavioral data that restrict the set of priors that may be consistent with the observed behavior.

$$\frac{\pi(s \mid \omega)\mu(\omega)}{\sum_{\omega' \in \Omega} \pi(s \mid \omega')\mu(\omega')}.$$

The Receiver then chooses an optimal action

$$a^*(s, \mu) \in \arg\max_{a \in A} \sum_{\omega \in \Omega} \mu_s(\omega) u(a, \omega).$$

$$V(\pi \mid \mu) = \sum_{\omega \in \Omega} \sum_{s \in S} \pi(s \mid \omega)\mu(\omega) v(a^*(s, \mu), \omega)$$

The Sender does not know μ_0, and therefore evaluates a signaling policy against the set of priors that are consistent with the observed behavioral data.

Although the true prior satisfies $\mu_0 \in \operatorname{int} \Delta(\Omega)$, I define the identified set $\mathcal{P}(\alpha)$ as the set of all priors in the closed simplex $\Delta(\Omega)$ that rationalize the observed aggregate behavior. Thus, $\mathcal{P}(\alpha)$ may contain boundary priors even though the true prior is interior.

This distinction is deliberate: the interiority assumption is imposed on the data-generating prior, whereas $\mathcal{P}(\alpha)$ represents the set of priors that the observed data do not rule out. Using the closed rationalizing set preserves compactness, which is required for the existence and minimax results below. The true prior μ_0 is therefore guaranteed to belong to $\mathcal{P}(\alpha)$, while boundary points are included as observationally admissible limiting priors.

Throughout, α is treated as the observed population action distribution used for rationalization. No sampling uncertainty or statistical estimation of α is modeled. The identified set $\mathcal{P}(\alpha)$ is therefore conditional on the observed distribution. We impose the following maintained condition throughout the paper: no action in $\mathrm{supp}(\alpha)$ is statewise strictly dominated under the Receiver's preferences. [Ros26] shows that this condition is equivalent to the existence of a posterior under which each such action is optimal. Hence the identified set is non-empty.

Lemma 2.1. *For any observed action distribution $\alpha \in \Delta(A)$ where no action $a \in \mathrm{supp}(\alpha)$ is statewise strictly dominated under u, the data-driven identified set $\mathcal{P}(\alpha)$ is a non-empty, compact, polyhedral set in $\mathbb{R}^{|\Omega|}$. Consequently, $\mathcal{P}(\alpha)$ satisfies the Farkas–Minkowski (FM) property.*

2.1 Robustly Obedient Direct Signaling

The polyhedral structure of $\mathcal{P}(\alpha)$ does not, by itself, imply that the Sender's payoff is globally linear in the prior. The reason is that the Receiver knows the true prior and therefore may choose different actions following the same message under different priors.

This point can be made formally.

Remark 1. *For a fixed signaling policy π, the Receiver's best responses are determined by finitely many linear inequalities in the prior μ. Consequently, if different Receiver action profiles occur on the identified set $\mathcal{P}(\alpha)$, the Sender's payoff is described by different linear functionals on the corresponding regions and therefore need not be globally represented by a single linear functional.*

Remark 1 identifies the source of the difficulty: the Receiver's action following a message may vary with the unknown prior. We therefore restrict attention to a class of mechanisms for which recommendations remain optimal throughout the entire data-driven identified set.

A direct signaling policy takes the action set itself as the message space, $S = A$, so that message $a \in A$ is interpreted as a recommendation to choose action a.

Definition 2.2 (Robust Obedience). A direct signaling policy $\pi : \Omega \to \Delta(A)$ is robustly obedient with respect to $\mathcal{P}(\alpha)$ if, for every prior $\mu \in \mathcal{P}(\alpha)$, every recommendation $a \in A$, and every alternative action $a' \in A$, $\sum_{\omega\in\Omega} \pi(a \mid \omega)\mu(\omega)\big[u(a,\omega) - u(a',\omega)\big] \geq 0$. Equivalently, whenever recommendation a is sent with positive probability under a prior $\mu \in \mathcal{P}(\alpha)$, the Receiver weakly prefers following a to every alternative action. Thus, the recommendation is incentive compatible simultaneously for every prior in the identified set.

We interpret a robustly obedient recommendation as an implemented Receiver action. Equivalently, within the restricted problem, the Receiver follows the recommendation whenever it is a best response. Thus weak robust obedience is sufficient for implementation, including at ties[3]. The restriction is motivated by the Sender's prior uncertainty. Since the Sender does not know which $\mu \in \mathcal{P}(\alpha)$ is the true prior, robust obedience guarantees that the recommended action is optimal for the Receiver throughout the entire set of priors that the observed data fail to rule out. It is therefore an implementation requirement under prior ambiguity, rather than an assumption that the Receiver mechanically follows recommendations independently of his posterior.

$$V(\pi \mid \mu) = \langle c(\pi), \mu\rangle \qquad \forall \mu \in \mathcal{P}(\alpha).$$

Thus, robust obedience restores the global linear structure that is not guaranteed under unrestricted signaling. For every fixed robustly obedient policy, the Sender's worst-case prior problem is therefore

$$\min_{\mu\in\mathcal{P}(\alpha)} \langle c(\pi), \mu\rangle,$$

i.e., the minimization of a continuous linear functional over the compact polyhedral identified set. This structure permits the convex-analytic and duality arguments developed in the remainder of the paper.

3 Robustly Obedient Policy

Let $B := \mathrm{supp}(\alpha)$[4]. For each action $a \in A$, define the Receiver's obedience region by

$$M_a := \left\{ \nu \in \Delta(\Omega) : \sum_{\omega\in\Omega} \nu(\omega)\big[u(a,\omega) - u(a',\omega)\big] \geq 0 \quad \forall a' \in A \right\}.$$

Thus, M_a is the set of posterior beliefs at which a is a best response for the Receiver. By Lemma 2.1, under the maintained assumption that no action in B is statewise strictly dominated, the identified set

$$\mathcal{P}(\alpha) = \sum_{a\in B} \alpha(a) M_a$$

is nonempty, compact, convex, and polyhedral[5].

We restrict attention to direct signaling policies that recommend only actions in B:

$$\Pi_B := \{\pi \in \Pi(A \mid \Omega) : \pi(a \mid \omega) = 0 \quad \forall a \notin B,\ \forall \omega \in \Omega\}.$$

For $\pi \in \Pi_B$, $\mu \in \mathcal{P}(\alpha)$, and $a \in B$, let

$$q_a^\pi(\mu) := \sum_{\omega\in\Omega} \pi(a \mid \omega)\mu(\omega)$$

denote the ex ante probability of recommendation a. Whenever $q_a^\pi(\mu) > 0$, the posterior induced by recommendation a is

$$\mu_a^\pi(\omega;\mu) := \frac{\pi(a \mid \omega)\mu(\omega)}{\sum_{\tilde\omega\in\Omega} \pi(a \mid \tilde\omega)\mu(\tilde\omega)}.$$

[3] This implementation convention is specific to the restricted robust-obedience problem. [Ros26] instead specifies a Sender-preferred selection from the Receiver's best-response set. The present restriction treats the recommended action itself as the implemented best response. Imposing strict robust obedience would alternatively eliminate tie cases, but is unnecessarily restrictive for the present result.

[4] The restriction to recommendations in $B = \mathrm{supp}(\alpha)$ aligns the signaling policy with the action support used in the data-rationalization criterion. Allowing recommendations outside B would require specifying how such recommendations are reconciled with the observed aggregate distribution and its rationalizing representation. This paper does not consider that extension.

[5] Here $\sum_{a\in B} \alpha(a) M_a$ denotes the weighted Minkowski sum $\{\sum_{a\in B} \alpha(a)\mu_a : \mu_a \in M_a\ \forall a \in B\}$. Equivalently, $\mathcal{P}(\alpha)$ is the linear image of the product set $\prod_{a\in B} M_a$ under the map $(\mu_a)_{a\in B} \mapsto \sum_{a\in B} \alpha(a)\mu_a$.

Define the robustly obedient policy set as

$$\Pi_{RO}(\alpha) := \left\{ \pi \in \Pi_B : \sum_{\omega \in \Omega} \pi(a \mid \omega)\mu(\omega)\big[u(a,\omega) - u(a',\omega)\big] \geq 0 \quad \forall a \in B, \forall a' \in A, \forall \mu \in \mathcal{P}(\alpha) \right\}.$$

This restriction makes the recommendation a optimal for the Receiver for every $\mu \in \mathcal{P}(\alpha)$. Weak set inclusion ($\Pi_{\mathrm{RO}}(\alpha) \subseteq \Pi_{\mathrm{unrestricted}}$) yields $W_{\mathrm{RO}}(\alpha) \leq W(\alpha)$.

Fact 1. *Let Ω and A be finite, and let $\alpha \in \Delta(A)$ satisfy the conditions of Lemma 2.1.*

(1) *The set $\Pi_{RO}(\alpha)$ is nonempty if and only if there exists a policy $\pi \in \Pi_B$ such that, for every $a \in B$ and every $\mu \in \mathcal{P}(\alpha)$, either $q_a^\pi(\mu) = 0$, or $\mu_a^\pi(\cdot;\mu) \in M_a$.*
(2) *A sufficient condition for $\Pi_{RO}(\alpha) \neq \varnothing$ is the existence of a collection $\{\Omega_a\}_{a \in B}$ of pairwise disjoint subsets of Ω whose union is Ω, such that $\delta_\omega \in M_a$ for every $a \in B$ and every $\omega \in \Omega_a$. Under this condition, the deterministic direct policy $\pi(a \mid \omega) = \begin{cases} 1, & \omega \in \Omega_a, \\ 0, & \omega \notin \Omega_a, \end{cases}$ belongs to $\Pi_{RO}(\alpha)$.*

Thus, the obedience requirement factors through posterior beliefs rather than the prior directly. Fact 1 provides a simple sufficient condition: if the state space can be partitioned into regions Ω_a, one per recommended action, such that each degenerate belief δ_ω already makes a optimal, then the deterministic policy that recommends a on Ω_a is robustly obedient. Convexity of each M_a ensures that the mixed posteriors generated on Ω_a remain in M_a. We can now state the main result of this section.

Theorem 3.1. *Let Ω and A be finite. Let $\alpha \in \Delta(A)$ satisfy the conditions of Lemma 2.1, and suppose that $\Pi_{RO}(\alpha) \neq \varnothing$. Then $\max_{\pi \in \Pi_{RO}(\alpha)} \min_{\mu \in \mathcal{P}(\alpha)} V(\pi \mid \mu) = \min_{\mu \in \mathcal{P}(\alpha)} \max_{\pi \in \Pi_{RO}(\alpha)} V(\pi \mid \mu)$. Moreover, there exist $\pi^s \in \Pi_{RO}(\alpha)$ and $\mu^s \in \mathcal{P}(\alpha)$ such that $V(\pi \mid \mu^s) \leq V(\pi^s \mid \mu^s) \leq V(\pi^s \mid \mu) \quad \forall \pi \in \Pi_{RO}(\alpha), \forall \mu \in \mathcal{P}(\alpha)$. Finally, for every fixed $\pi \in \Pi_{RO}(\alpha)$, the inner minimization $\min_{\mu \in \mathcal{P}(\alpha)} V(\pi \mid \mu)$ is attained at an extreme point of $\mathcal{P}(\alpha)$. In particular, every maxmin-optimal policy π^s admits a worst-case prior $\widehat{\mu} \in \mathrm{ext}(\mathcal{P}(\alpha))$.*

The proof follows a standard minimax argument. The policy set $\Pi_{RO}(\alpha)$ is shown to be nonempty, compact, and convex and it is the intersection of the simplex constraints defining Π_B with finitely many closed half-spaces, one per robust-obedience inequality. The identified set $\mathcal{P}(\alpha)$ is compact and convex by Lemma 2.1. Under robust obedience, the Sender's payoff $V(\pi \mid \mu)$ is jointly continuous and bilinear in (π, μ): affine in π for each fixed μ, and affine in μ for each fixed π. Sion's minimax theorem therefore applies, which yields the stated equality of maxmin and minimax values and the existence of a saddle point (π^s, μ^s). Vertex attainment of the inner minimization follows because, for fixed π, $V(\pi \mid \cdot)$ is a continuous linear functional on the polytope $\mathcal{P}(\alpha)$, and a linear functional attains its minimum over a nonempty polytope at an extreme point. "Pure strategy" here means neither player needs to randomize over policies or priors. A signaling policy π, as usual, may randomize its recommendation conditional on the state[6].

[6]The maintained assumptions guaranteeing $\mathcal{P}(\alpha) \neq \varnothing$ do not by themselves guarantee the existence of a policy that is obedient for every prior in $\mathcal{P}(\alpha)$. Fact 1 therefore provides a sufficient condition under which the restricted policy class is nonempty.

We record one further consequence of robust obedience, namely that the fixed-policy problem admits an explicit convex-analytic representation.

Observation 1. *Fix $\pi \in \Pi_{RO}(\alpha)$, and define $c_\omega(\pi) := \sum_{a \in B} \pi(a \mid \omega)v(a,\omega) \quad$ for each $\omega \in \Omega$. Then $g_\pi(\mu) := V(\pi \mid \mu) = \sum_{\omega \in \Omega} c_\omega(\pi)\mu(\omega) = \langle c(\pi), \mu \rangle$ is a finite-valued continuous linear functional on $\mathbb{R}^{|\Omega|}$. In particular, g_π is proper, convex, and lower semicontinuous, and satisfies $g_\pi = g_\pi^{**}$. Its Fenchel conjugate is $g_\pi^*(z) = \begin{cases} 0, & z = c(\pi), \\ +\infty, & z \neq c(\pi). \end{cases}$*

The biconjugacy statement is included to record explicitly the convex-analytic structure of the fixed-policy problem. It is immediate here because g_π is an everywhere-finite continuous linear functional. However, Theorem 3.1 does not generally guarantee that the saddle-point prior μ^s itself is a vertex.

4 When is Robust Obedience Optimal?

Previous section showed us that saddle point can be recovered when restricted to Robust Obediency. In this section, we analyze when a robustly obedient policy is optimal. We begin with two examples to set up the economic intuition behind our main results.

Example 4.1. Consider a startup (Sender) communicating with an investor (Receiver) deciding between a_1 (Invest) and a_0 (Decline) and $\Omega = \{\omega_L, \omega_H\}$. Receiver utility: $u(a_1, \omega_H) = 1, u(a_0, \omega_H) = 0$; $u(a_0, \omega_L) = 1, u(a_1, \omega_L) = 0$. Receiver threshold is $c = 0.5$, giving $M_{a_1} = \{p(H) \geq 0.5\}$ and $M_{a_0} = \{p(H) \leq 0.5\}$. Sender utility is state-dependent: $v(a_1, \omega_H) = 2, v(a_1, \omega_L) = 1$, and $v(a_0, \cdot) = 0$. $P(\alpha) = [p_{\min}, p_{\max}] = [0.2, 0.4] \subset M_{a_0}$.

Let $\pi(a_1 \mid \omega_H) = x$ and $\pi(a_1 \mid \omega_L) = y$.

(1) Recommendation a_1 obedience: $D_{a_1}(p) = px(1) + (1-p)y(-1) = px - (1-p)y$. Since $D'_{a_1}(p) = x + y \geq 0$, obedience holds across $[0.2, 0.4]$ iff it holds at $p_{\min} = 0.2$: $0.2x - 0.8y \geq 0 \implies y \leq \frac{1}{4}x$.
(2) Recommendation a_0 obedience: $D_{a_0}(p) = (1-p)(1-y)(1) + p(1-x)(-1) = (1-p)(1-y) - p(1-x)$. Since $D'_{a_0}(p) = -(2-x-y) \leq 0$, obedience holds across $[0.2, 0.4]$ iff it holds at $p_{\max} = 0.4$: $0.6(1-y) - 0.4(1-x) \geq 0 \implies y \leq \frac{2}{3}x + \frac{1}{3}$. Choosing $x^* = 1, y^* = 0.25$ satisfies both constraints. Sender payoff under π^* at prior p is $V(p) = 2p(1) + 1(1-p)(0.25) = \frac{1}{4} + \frac{7}{4}p$. The worst case over $[0.2, 0.4]$ occurs at $p = 0.2$: $W_{\mathrm{RO}}(\alpha) \geq V(0.2) = 0.4 + 0.8(0.25) = 0.6$.

At $p = 0.2$, any message m inducing a_1 satisfies $\pi(m \mid \omega_L) \leq \frac{1}{4}\pi(m \mid \omega_H)$. Summing probabilities across all a_1-inducing messages yields $Y \leq \frac{1}{4}X$. Sender payoff at $p = 0.2$ under any unrestricted policy π is:

$$V(\pi \mid 0.2) = 0.2(2X) + 0.8(1Y) = 0.4X + 0.8Y \leq 0.4X + 0.8\left(\frac{1}{4}X\right) = 0.6X \leq 0.6.$$

Thus $W(\alpha) \leq 0.6$. Combining $W_{\mathrm{RO}}(\alpha) \geq 0.6$ and $W(\alpha) \leq 0.6$ proves $W_{\mathrm{RO}}(\alpha) = W(\alpha) = 0.6$.

Now, we see another example where there is a gap between the unrestricted signaling and robust obediency.

Example 4.2. Consider a two-state, two-action environment $\Omega = \{\omega_1, \omega_2\}$ with $p = p(\omega_2)$, and $A = \{a_0, a_1\}$. Receiver's utility:

$u(a_1, \omega_1) = -1, u(a_1, \omega_2) = 1$, and $u(a_0, \omega_1) = u(a_0, \omega_2) = 0$. The receiver prefers a_1 iff $-(1-p) + p \geq 0 \iff 2p - 1 \geq 0 \iff p \geq \frac{1}{2}$, giving $M_{a_1} = [\frac{1}{2}, 1]$ and $M_{a_0} = [0, \frac{1}{2}]$. Sender's utility: $v(a_1, \omega_1) = 1$, $v(a_1, \omega_2) = 0$, and $v(a_0, \omega_1) = v(a_0, \omega_2) = 0$. Observed action distribution: $\alpha(a_1) = \frac{3}{5}$ and $\alpha(a_0) = \frac{2}{5}$.

The identified set is:

$$P(\alpha) = \alpha(a_1)M_{a_1} + \alpha(a_0)M_{a_0} = \tfrac{3}{5}[\tfrac{1}{2}, 1] + \tfrac{2}{5}[0, \tfrac{1}{2}] = \left[\tfrac{3}{10}, \tfrac{4}{5}\right].$$

Note that $P(\alpha)$ straddles the receiver's threshold $\frac{1}{2}$.

For a direct policy $\pi = (x_1, x_2)$, obedience for a_1 requires $-(1-p)x_1 + px_2 \geq 0$ for all $p \in [\frac{3}{10}, \frac{4}{5}]$, or equivalently $x_1 \leq \frac{p}{1-p}x_2$. The tightest constraint occurs at $p = \frac{3}{10}$, giving $x_1 \leq \frac{3}{7}x_2$. Since $x_2 \leq 1$, the optimal direct policy has $x_2^* = 1$ and $x_1^* = \frac{3}{7}$. Checking obedience for a_0 at $p = \frac{4}{5}$ yields $\frac{1}{5}(1 - \frac{3}{7}) \geq 0$, which holds. The robustly obedient benchmark value is:

$$W_{\mathrm{RO}}(\alpha) = \left(1 - \tfrac{4}{5}\right) \cdot \tfrac{3}{7} = \tfrac{1}{5} \cdot \tfrac{3}{7} = \tfrac{3}{35}.$$

Now construct an unrestricted signaling policy with three messages $M = \{m_1, m_2, m_0\}$ and conditional probabilities:

	m_1	m_2	m_0
ω_1	$\frac{3}{11}$	$\frac{4}{11}$	$\frac{4}{11}$
ω_2	$\frac{7}{11}$	$\frac{4}{11}$	0

Each row sums to 1. The posteriors induced are: * m_1 induces $p_1 = \frac{7p}{3+4p} \in [\frac{1}{2}, \frac{28}{31}] \subset M_{a_1}$, so the receiver chooses a_1. * m_2 induces $p_2 = p$, straddling $\frac{1}{2}$. * m_0 induces $p_0 = 0$, so the receiver chooses a_0.

Evaluating the sender's worst-case payoff over $p \in [\frac{3}{10}, \frac{4}{5}]$ yields $W(\alpha) \geq \frac{7}{55}$. Comparing the two benchmarks:

$$W_{\mathrm{RO}}(\alpha) = \tfrac{3}{35} = \tfrac{33}{385}, \quad W(\alpha) \geq \tfrac{7}{55} = \tfrac{49}{385}.$$

Thus,

$$W(\alpha) - W_{\mathrm{RO}}(\alpha) \geq \tfrac{49}{385} - \tfrac{33}{385} = \tfrac{16}{385} > 0,$$

yielding:

$$\boxed{W(\alpha) > W_{\mathrm{RO}}(\alpha)}$$

and demonstrating a strict gap between unrestricted persuasion and robustly obedient persuasion.

The two examples illustrate the two possible regimes. In the first, robust obedience is without loss as the identified set lies entirely within a single obedience region, so any recommendation that is optimal at the worst-case prior remains optimal throughout $P(\alpha)$, and the unrestricted and robust values coincide. In the second, the identified set straddles the Receiver's indifference threshold. An unrestricted policy then exploits the Receiver's prior-dependent best response and achieves a strictly higher worst-case payoff than any robustly obedient policy. The gap arises precisely because robust obedience forbids the Sender from conditioning the implemented action on the unknown prior. The natural question is therefore when the two benchmarks coincide.

Claim 1. *Assume the unrestricted problem admits a maxmin-optimal finite-message direct policy $\pi^* \in \Pi_B$. Then $W_{\mathrm{RO}}(\alpha) = W(\alpha)$ if and only if each recommended action $a \in B$ sent under π^* is Receiver-optimal for every prior $p \in P(\alpha)$.*

We next give a sufficient condition that is verifiable from primitives, by requiring the optimal policy to maintain a strict incentive margin at a nominal prior $\bar{p}$ and the identified set to be sufficiently small in ℓ_1-radius around $\bar{p}$.

Theorem 4.3. *Let $\bar{p} \in P(\alpha)$, and let $\pi^* \in \Pi_B$ be an unrestricted maxmin-optimal direct policy. Define the obedience margin for recommendation $a \in B$ against $a' \in A \setminus \{a\}$ at prior p by: $D^{\pi^*}_{a,a'}(p) := \sum_{\omega \in \Omega} \pi^*(a \mid \omega)p(\omega)[u(a, \omega) - u(a', \omega)]$. Suppose π^* satisfies a strict Receiver incentive margin $\gamma > 0$ at $\bar{p}$: $\gamma := \min_{a \in B, a' \in A \setminus \{a\}} D^{\pi^*}_{a,a'}(\bar{p}) > 0$. Let $L_R := \max_{a \in B, a' \in A \setminus \{a\}, \omega \in \Omega} |u(a, \omega) - u(a', \omega)| > 0$. Define the prior uncertainty radius $r_1(P(\alpha), \bar{p}) := \sup_{p \in P(\alpha)} \|p - \bar{p}\|_1$. If $r_1(P(\alpha), \bar{p}) < \frac{\gamma}{L_R}$, then $\pi^* \in \Pi_{\mathrm{RO}}(\alpha)$ and $W_{\mathrm{RO}}(\alpha) = W(\alpha)$.*

Theorem 4.3 says that if the optimal unrestricted policy has a strict incentive margin at a nominal prior and the identified set is small enough relative to that margin, robust obedience is free, meaning the same policy is robustly obedient and the gap closes. The threshold γ / L_R is interpretable as the amount of prior uncertainty the Receiver's incentives can absorb before some recommendation loses its optimality somewhere in $P(\alpha)$. This is a local, non-asymptotic condition. It requires neither a continuum of states nor any statistical approximation, but it is also a knife-edge. Once the radius exceeds the threshold, the argument gives no guarantee, and the example above shows the gap can indeed be positive. The next result addresses the intermediate regime: when no perfectly robust policy is available but a uniformly obedient policy exists elsewhere in the simplex, we can construct a robustly obedient policy by mixing and bound the resulting welfare loss.

Theorem 4.4. *Let $\bar{p} \in P(\alpha)$ and $r_1 := r_1(P(\alpha), \bar{p}) = \sup_{p \in P(\alpha)} \|p - \bar{p}\|_1$. Let $\pi^* \in \Pi_B$ be an unrestricted maxmin-optimal direct policy that is obedient at the nominal prior $\bar{p}$ (i.e., $D^{\pi^*}_{a,a'}(\bar{p}) \geq 0$ for all $a \in B, a' \neq a$). Suppose there exists a policy $\pi_{\mathrm{int}} \in \Pi_B$ satisfying $D^{\pi_{\mathrm{int}}}_{a,a'}(p) \geq \eta > 0$ for all $a \in B, a' \neq a, p \in P(\alpha)$.*

(i) ***[Policy Robustification]:*** *For $\delta = \frac{L_R r_1}{\eta + L_R r_1}$, the perturbed policy $\pi_\delta := (1 - \delta)\pi^* + \delta\pi_{\mathrm{int}}$ belongs to $\Pi_{\mathrm{RO}}(\alpha)$.*

(ii) ***[Payoff Bound under Receiver Action Stability]:*** *If π^* satisfies Receiver Action Stability under π_δ across $P(\alpha)$ (that is, $a^*(a, p; \pi_\delta) = a^*(a, p; \pi^*) = a$ for all $a \in B, p \in P(\alpha)$), then: $0 \leq W(\alpha) - W_{\mathrm{RO}}(\alpha) \leq C \cdot r_1(P(\alpha), \bar{p})$, where $C = \frac{2M_v L_R}{\eta}$ with $M_v = \max_{a \in B, \omega \in \Omega} |v(a, \omega)|$.*

The proof of Theorem 4.4 proceeds in two steps. For part (i), the Lipschitz bound on $D^{\pi^*}_{a,a'}$ in p — inherited from linearity and the utility range L_R — implies that π^* loses at most $L_R r_1$ of obedience margin as p ranges over $P(\alpha)$. Mixing in a policy π_{int} that is uniformly obedient with slack η restores the margin. Choosing $\delta = L_R r_1 / (\eta + L_R r_1)$ exactly equates the worst-case loss of the original policy with the guaranteed gain from the interior policy, so the mixture π_δ is robustly obedient. For part (ii), Receiver Action Stability ensures that the implemented action profile is unchanged by the mixture, so the Sender's payoff under π_δ differs from that under π^* only through the mixing weight. Since π_{int} and π^* are both probability distributions over B at each state, the ℓ_1-distance between them is at most 2, yielding the $O(r_1)$ bound on the welfare gap with constant $2M_v L_R / \eta$.

5 Computational Complexity of Robust Persuasion

In direct information design with prior uncertainty, evaluating whether a persuasion strategy achieves a targeted payoff threshold requires reasoning over the sender's worst-case payoff across all priors in the identified set $P(\alpha)$. When the state space is explicitly listed, linear programming methods make this optimization computationally straightforward. However, in economic applications involving complex multi-dimensional state spaces—such as combinatorial auctions, contests, multi-asset financial networks, or decentralized voting—the state space $\Omega = \{0,1\}^d$ is succinctly represented via d Boolean attributes, causing $|\Omega| = 2^d$ to grow exponentially in the description length of the game. This section establishes the computational complexity behind the robust persuasion decision problem. We examine how prior uncertainty and action space cardinality patterns govern the frontier between exponential intractability and algorithmic tractability.

We begin by examining whether robust persuasion is computationally tractable when the sender faces prior uncertainty that collapses to a known, single prior. A natural hypothesis is that if the identified set $P(\alpha)$ is a singleton $\{\bar{\mu}\}$, computing worst-case payoffs should simplify to standard expected utility maximization. However, when the action set is succinctly specified via combinatorial structures, deciding robust persuasion remains intractable even under a singleton uniform prior.

Claim 2. *Consider the succinct robust-persuasion decision problem with Boolean-hypercube states, polynomial-time utility circuits, rational payoffs, and a succinctly represented combinatorial action set. Then deciding whether $W_{\mathrm{RO}}(\alpha) \geq \tau$ is NP-hard under polynomial-time many-one reductions, even when: $B = \{a_0, a_1\}$ and $\alpha(a_0) = \alpha(a_1) = 1/2$; the identified set is generated by the original rationalization formula; $P(\alpha)$ is a singleton consisting of a known uniform product prior; all utilities are polynomial-time computable and have polynomial-length encodings.*

The simplex $\Delta(\Omega)$ has affine dimension $N-1$. The $K-1$ comparisons against actions other than a supply at most $K-1$ additional linear inequalities. At a point in M_a, at most $K-1$ independent active inequality normals can reduce the local dimension. Thus a nonempty M_a has dimension at least

$$(N-1) - (K-1) = N - K.$$

For the Minkowski sum, choose any $a \in B$. Since $\alpha(a) > 0$ and all other summands are nonempty, variation in M_a induces variation in $P(\alpha)$; hence a non-singleton M_a makes $P(\alpha)$ non-singleton. In particular, with $N = 2^{d+1}$ states and $K = \mathrm{poly}(d)$ actions, the identified set cannot be forced to equal the uniform prior by this rationalization formula. But it doesn't work for arbitrary singleton priors, because even one inequality can isolate a simplex vertex. The valid obstruction is specifically about pinning down an interior prior such as the uniform distribution.

Proposition 5.1. *Suppose $\mu^\circ \in \operatorname{int}\Delta(\Omega)$ and $M_a = \{\mu^\circ\}$. Then*

$$|A| - 1 \geq |\Omega|.$$

Now suppose $P(\alpha) = \{\mu^\circ\}$. A Minkowski sum of nonempty sets with strictly positive weights can be a singleton only if every summand is a singleton. Hence $M_a = \{\mu^\circ\}$ for every $a \in \operatorname{supp}(\alpha)$, so the bound applies. For $\Omega = \{0,1\}^{d+1}$, this requires at least $2^{d+1}+1$ actions. Thus the singleton-uniform-prior route cannot yield the polynomial-action result.

The proposition can be generalized to cover the boundary case.

Claim 3. *If $M_a = \{\mu^a\}$, then*

$$|\operatorname{supp}(\mu^a)| \leq |A| - 1.$$

Thus, in any realistic economic model with a polynomial action space, the identified set $P(\alpha)$ must be robustly full-dimensional and contain a continuum of plausible priors.

5.1 Intractability under Polynomial Action Sets and Full-Dimensional Uncertainty

Having shown that realistic action spaces necessarily lead to full-dimensional uncertainty sets $P(\alpha)$, we analyze the computational complexity of robust persuasion under polynomial action sets $|A| = \mathrm{poly}(d)$. In this environment, the sender must ensure robust obedience against an entire polyhedral set of priors $P(\alpha)$ spanning the full dimension of $\Delta(\Omega)$.

Theorem 5.2. *Consider the decision problem that, given a succinctly represented finite-state robust-obedience instance and a rational threshold τ, asks whether*

$$W_{\mathrm{RO}}(\alpha) \geq \tau.$$

This problem is coNP-hard, even when the observed action distribution has support of size two, the number of actions is polynomial in the input size, and the rationalization set $P(\alpha)$ is full-dimensional in the probability simplex.

The proof sketch relies on a polynomial-time reduction from 3-UNSAT. Given a 3-CNF formula $\phi(x_1, \ldots, x_d)$ with m clauses, we construct an instance with state space $\Omega = \{0,1\}^{d+1}$ and only $m+4$ receiver actions. Rather than isolating a single prior, we construct auxiliary actions d_0, e_0 and clause actions $\{c_1, \ldots, c_m\}$ that carve out a full-dimensional polytope $P(\alpha) \subset \Delta(\Omega)$. The clause actions c_j penalize probability allocations on assignments that falsify clause C_j.

If ϕ is unsatisfiable (a 3-UNSAT instance), every state falsifies at least one clause, forcing any prior $\mu \in P(\alpha)$ to assign at least half its probability weight to the $z = 0$ states ($\mu_0 \geq 1/2$). The sender can then achieve a robust payoff of at least $1/2 \geq \tau = 2^{-(d+1)}$ using the constant policy. Conversely, if ϕ is satisfiable under assignment x^*, we can construct a specific worst-case prior $\mu^* \in P(\alpha)$ supported on $(0, x^*)$ and $(1, 0^d)$ under which the sender receives expected utility zero for all supported actions, driving $W_{\mathrm{RO}}(\alpha) = 0 < \tau$. This establishes a sharp reduction: $\phi \in$ 3-UNSAT $\iff W_{\mathrm{RO}}(\alpha) \geq \tau$.

Now a critical question remains: could robust persuasion be even harder, given that checking robust obedience requires evaluating a universal quantifier over infinitely many priors $\mu \in P(\alpha)$ combined with an existential choice of statewise policies? We resolve this question in the next result by establishing that the problem belongs to coNP.

Theorem 5.3. *Consider the original robust-persuasion setup with succinct state space $\Omega = \{0,1\}^d$, explicitly represented action set A, and $B = \operatorname{supp}(\alpha) = \{a_0, a_1\}$. Assume:*

(1) *The utilities $u(a, \omega)$ and $v(a, \omega)$, the distribution α, and the threshold τ have polynomial-bit rational encodings, and utility values can be evaluated in polynomial time at a specified state.*
(2) *Each rationalization set M_b, for $b \in B$, is nonempty. Thus $P(\alpha)$ is nonempty, compact, and polyhedral, as in Lemma 2.1.*
(3) *The input length includes the d-bit state description, and invalid inputs are handled separately in polynomial time.*

Then the decision problem

$$\text{DECISION-}W_{\text{RO}} = \{\mathcal{I} : W_{\text{RO}}(\alpha) \geq \tau\}$$

belongs to coNP.

The mathematical mechanism establishing this result transforms a continuous optimization problem over an exponential state space into a concise verification certificate using dual linear programming and Helly's Theorem.

5.2 Parameterized XP Algorithms

To bypass the worst-case intractability in practical applications, we identify structural restrictions on state interactions that admit efficient deterministic computation. In many economic settings, state variables exhibit localized interaction topologies. We formalize this structure via the treewidth w of the interaction hypergraph and parameterize algorithm performance.

Theorem 5.4. *Consider the original robust-persuasion instance*

$$\mathcal{I} = (d, A, \alpha, u, v), \qquad \Omega = \{0, 1\}^d,$$

with $k = |A|$ and $B = \text{supp}(\alpha)$. Retain the standing assumptions that the rationalization sets M_b, $b \in B$, are nonempty and that $P(\alpha)$ is the compact polyhedral set of Lemma 2.1. Suppose, in addition, that:

(1) *The input includes a tree decomposition of the interaction hypergraph of width w. Every local utility component depends only on variables contained together in some bag of this decomposition.*
(2) *Each utility $u(a, \cdot)$, $a \in A$, and each relevant sender utility $v(a, \cdot)$, $a \in B$, is a sum of at most d local components.*
(3) *Every local component is integer-valued and has absolute value at most R, where R is unary-encoded or bounded by a polynomial in the input length L.*
(4) *The probabilities $\alpha(b)$ and threshold τ are rational with polynomial bit length.*

Then deciding whether

$$W_{\text{RO}}(\alpha) \geq \tau$$

can be done deterministically in time

$$\boxed{2^{O(w)}(dR)^{O(k)}\,\text{poly}(L).}$$

Consequently, the decision problem belongs to XP parameterized by (w, k).

The execution strategy follows in two dynamic steps: state space compression via dynamic programming and then by solving a reduced linear program.

First, despite the state space containing 2^d individual states, the receiver and sender utilities at state ω depend only on a vector of utility evaluations called the state's utility type:

$$V(\omega) = ((u(a, \omega))_{a \in A}, (v(a, \omega))_{a \in B}) \in \mathbb{Z}^r, \quad r \leq 2k.$$

Because each local utility component is bounded by R and at most d components sum together, each coordinate of $V(\omega)$ lies in $\{-dR, \ldots, dR\}$. The total number of attainable utility types is bounded by $|\mathcal{V}| \leq (2dR + 1)^{2k}$, which is polynomial in d for fixed $k = |A|$. Using a dynamic programming algorithm over the supplied tree decomposition of width w, we construct table entries $D_t(\sigma, s)$ at each bag t counting valid partial state assignments that yield partial utility sum s. This allows us to compute the complete set of attainable utility types $\mathcal{V}$ deterministically in time $2^{O(w)}(dR)^{O(k)}\,\text{poly}(L)$.

Second, we can show that whenever an optimal robustly obedient policy exists, there exists an optimal policy that is type-constant, recommending identical action probabilities $\pi(\cdot \mid \omega) = p_v(\cdot)$ for all states ω sharing the same utility type $V(\omega) = v$. This type-constancy follows from the existence of shared global dual witnesses across states. Finally, a reduced LP can be formulated whose policy variables $p_v(a)$ are defined only over the polynomial set of types $v \in \mathcal{V}$, rather than the 2^d individual states. The reduced LP contains $(dR)^{O(k)}$ variables and constraints with polynomial-bit rational coefficients. Solving this LP deterministically yields an overall runtime of $2^{O(w)}(dR)^{O(k)}\,\text{poly}(L)$.

As a result, when the interaction hypergraph has bounded treewidth $w = O(1)$ and the number of available actions $k = |A| = O(1)$ is fixed, robust persuasion can be solved in deterministic polynomial time $O(\text{poly}(L))$. This places DECISION-W_{RO} in the parameterized complexity class XP that provides an efficient, exact computational pipeline for structured, high-dimensional robust persuasion problems.

6 Concluding Remarks

Studying state-dependent data-driven persuasion opens two distinct methodological and algorithmic paths to resolve the prior uncertainty inherent in the framework. The first path, which this paper pursues, is to restrict signaling to robustly obedient direct mechanisms. By requiring recommendations to remain optimal for the Receiver across the entire data-identified set $\mathcal{P}(\alpha)$, we stabilize Receiver actions and restore bilinearity, which yields a pure-strategy saddle point while keeping decision-making within coNP. From a computational perspective, while exact evaluation is coNP-complete, our parameterized tractability result opens some compelling algorithmic directions like developing efficient polynomial-time approximation schemes for the worst-case robust value, characterizing fixed-parameter tractable regimes under alternative bounded-width decompositions, exploring the interactive query complexity of checking robust obedience when utilities are accessed via black-box sampling rather than explicit circuit representations, and so on. The second path is to allow fully unrestricted signaling, thereby retaining the changing, prior-dependent Receiver actions of the original [Ros26] model. Under state dependence, this path destroys bilinearity and introduces severe non-convexities and non-smoothness. Fully characterizing this unrestricted setting is a promising direction for future research. Doing so will require invoking nonsmooth analysis and variational geometry to analyze the boundaries where optimal Receiver actions jump across the polyhedral identified set.

A Proof of Lemma 2.1

By the primitives of the model, Ω and A are finite sets. For any action $a \in \text{supp}(\alpha)$, the assumption that a is not statewise strictly dominated implies that there exists at least one posterior belief $\mu_a \in \Delta(\Omega)$ such that a is an optimal response for the Receiver. Rosenthal's rationalization criterion establishes that a prior μ belongs to the identified set $\mathcal{P}(\alpha)$ if and only if there exists a family of posteriors $\{\mu_a\}_{a\in\text{supp}(\alpha)}$ such that for each a, $a \in a^*(\mu_a)$ and the prior is a Bayes-plausible convex combination of these posteriors: $\mu = \sum_{a\in\text{supp}(\alpha)} \alpha(a)\mu_a$. Since $\alpha \in \Delta(A)$ is a probability distribution and each μ_a exists, the weighted sum μ necessarily exists. Therefore, $\mathcal{P}(\alpha) \neq \emptyset$.

For any action a, the set of rationalizing posteriors $M_a = \{\mu \in \Delta(\Omega) : \sum_{\omega\in\Omega} \mu(\omega)[u(a,\omega) - u(a',\omega)] \geq 0 \ \forall a' \in A\}$ is defined by a finite system of linear inequalities, as A and Ω are finite. Such a set is by definition a polyhedron. The identified set $\mathcal{P}(\alpha)$ is the Minkowski-like weighted sum of these polyhedral sets: $\mathcal{P}(\alpha) = \sum_{a\in\text{supp}(\alpha)} \alpha(a)M_a$. In finite dimensions, the image of a polyhedral set under a linear transformation—specifically the mapping that defines this weighted sum—is itself a polyhedral set.

The simplex $\Delta(\Omega)$ is a bounded and closed set in $\mathbb{R}^{|\Omega|}$, and is therefore compact. Since every $\mu_a \in \Delta(\Omega)$, their convex combination μ must also lie within $\Delta(\Omega)$. Because $\mathcal{P}(\alpha)$ is a polyhedral set, it is closed in finite-dimensional Euclidean space. Consequently, as a closed subset of the compact simplex $\Delta(\Omega)$, the set $\mathcal{P}(\alpha)$ is compact.

Now we represent the polyhedral set $\mathcal{P}(\alpha)$ as a finite constraint system $S = \{\langle a_t^*, \mu\rangle \leq b_t, t \in T\}$ where $|T| < \infty$. The characteristic cone of this system is defined as $K = \text{cone co}\{\bigcup_{t\in T} \text{epi } f_t^* \cup \text{epi } \sigma_C\}$. Because the system S is finite and linear, K is a polyhedral cone. In finite dimensions, polyhedral cones are always closed in the w^*-topology. A constraint system satisfies the Farkas–Minkowski property if its characteristic cone is w^*-closed. Therefore, $\mathcal{P}(\alpha)$ satisfies the FM property. ■

B Proof of Remark 1

Since the denominator is independent of the Receiver's action, this is equivalent to

$$a^*(s,\mu) \in \arg\max_{a\in A} \sum_{\omega\in\Omega} \pi(s \mid \omega)u(a,\omega)\mu(\omega).$$

The Receiver's objective from action a is therefore

$$\langle \mathbf{U}_s(a), \mu\rangle,$$

which is linear in μ.

$$\{\mu \in \mathcal{P}(\alpha) : \langle \mathbf{U}_s(a) - \mathbf{U}_s(a'), \mu\rangle \geq 0 \quad \forall a' \in A\}.$$

Because A is finite, $\mathcal{P}_s(a)$ is defined by finitely many linear inequalities in μ. By Lemma 2.1, $\mathcal{P}(\alpha)$ is polyhedral; therefore each $\mathcal{P}_s(a)$ is a polyhedral subset of $\mathcal{P}(\alpha)$. Intersecting these regions across the finitely many messages produces a finite collection of polyhedral regions on which the Receiver's action profile is fixed.

$$c_\omega^R(\pi) := \sum_{s\in S} \pi(s \mid \omega)v(a_{s,R},\omega).$$

$$V(\pi \mid \mu) = \langle c^R(\pi), \mu\rangle, \qquad \mu \in R.$$

Thus $V(\pi \mid \cdot)$ is linear on each region R. Therefore, unless the induced action profiles generate the same Sender payoff coefficient vector, $V(\pi \mid \cdot)$ need not be representable by a single linear functional on all of $\mathcal{P}(\alpha)$. ■

C Proof of Fact 1

For Part 1, fix $\pi \in \Pi_B$, $a \in B$, and $\mu \in \mathcal{P}(\alpha)$. If $q_a^\pi(\mu) > 0$, then

$$\sum_{\omega\in\Omega} \pi(a \mid \omega)\mu(\omega)\big[u(a,\omega) - u(a',\omega)\big] = q_a^\pi(\mu) \sum_{\omega\in\Omega} \mu_a^\pi(\omega;\mu)\big[u(a,\omega) - u(a',\omega)\big].$$

Since $q_a^\pi(\mu) > 0$, the robust-obedience inequality holds for every $a' \in A$ if and only if

$$\sum_{\omega\in\Omega} \mu_a^\pi(\omega;\mu)\big[u(a,\omega) - u(a',\omega)\big] \geq 0 \quad \forall a' \in A,$$

which is precisely the condition $\mu_a^\pi(\cdot;\mu) \in M_a$.

If $q_a^\pi(\mu) = 0$, then $\pi(a \mid \omega)\mu(\omega) = 0$ for every ω, because all of these terms are nonnegative. Thus every robust-obedience inequality associated with recommendation a holds with equality. This proves Part 1.

For Part 2, consider the deterministic policy in the statement. Fix $a \in B$ and $\mu \in \mathcal{P}(\alpha)$. If

$$\mu(\Omega_a) := \sum_{\omega\in\Omega_a} \mu(\omega) = 0,$$

then $q_a^\pi(\mu) = 0$, so robust obedience for recommendation a is automatic.

Suppose instead that $\mu(\Omega_a) > 0$. The posterior induced by recommendation a is

$$\mu_a^\pi(\omega;\mu) = \begin{cases} \dfrac{\mu(\omega)}{\mu(\Omega_a)}, & \omega \in \Omega_a, \\ 0, & \omega \notin \Omega_a. \end{cases}$$

Therefore,

$$\mu_a^\pi(\cdot;\mu) = \sum_{\omega\in\Omega_a} \frac{\mu(\omega)}{\mu(\Omega_a)}\delta_\omega.$$

By assumption, $\delta_\omega \in M_a$ for every $\omega \in \Omega_a$. Since M_a is convex, it follows that

$$\mu_a^\pi(\cdot;\mu) \in M_a.$$

Part 1 then implies $\pi \in \Pi_{RO}(\alpha)$. ■

D Proof of Theorem 3.1

The set Π_B is nonempty, compact, and convex in $\mathbb{R}^{|A||\Omega|}$. It is defined by the finite linear constraints

$$\pi(a \mid \omega) \geq 0, \qquad \sum_{a\in A} \pi(a \mid \omega) = 1, \qquad \pi(a \mid \omega) = 0 \quad \text{for } a \notin B.$$

For every fixed $(a, a', \mu) \in B \times A \times \mathcal{P}(\alpha)$, the robust-obedience constraint

$$\sum_{\omega\in\Omega} \pi(a \mid \omega)\mu(\omega)\big[u(a,\omega) - u(a',\omega)\big] \geq 0$$

defines a closed half-space in π. Thus, $\Pi_{RO}(\alpha)$, as an intersection of Π_B and closed half-spaces, is closed. It is therefore compact. It is convex because every robust-obedience inequality is linear in π. By assumption, it is nonempty.

By Lemma 2.1, $\mathcal{P}(\alpha)$ is nonempty, compact, and convex. Since $\pi \in \Pi_{RO}(\alpha)$, the Receiver follows every recommendation sent with positive probability under every $\mu \in \mathcal{P}(\alpha)$. Hence the Sender's expected payoff is

$$V(\pi \mid \mu) = \sum_{\omega\in\Omega}\sum_{a\in B} \mu(\omega)\pi(a \mid \omega)v(a,\omega).$$

This function is jointly continuous and bilinear in (π, μ). Therefore, for each fixed μ, $V(\cdot \mid \mu)$ is continuous and affine, hence concave, on $\Pi_{RO}(\alpha)$. For each fixed π, $V(\pi \mid \cdot)$ is continuous and affine, hence convex, on $\mathcal{P}(\alpha)$.

Sion's minimax theorem applies:

$$\max_{\pi\in\Pi_{RO}(\alpha)} \min_{\mu\in\mathcal{P}(\alpha)} V(\pi \mid \mu) = \min_{\mu\in\mathcal{P}(\alpha)} \max_{\pi\in\Pi_{RO}(\alpha)} V(\pi \mid \mu).$$

Compactness and continuity ensure that these extrema are attained. Let π^s attain the maxmin value and let μ^s attain the minimax value. Let their common value be w. Then

$$V(\pi^s \mid \mu^s) \geq \min_{\mu\in\mathcal{P}(\alpha)} V(\pi^s \mid \mu) = w,$$

and

$$V(\pi^s \mid \mu^s) \leq \max_{\pi\in\Pi_{RO}(\alpha)} V(\pi \mid \mu^s) = w.$$

Therefore,

$$V(\pi^s \mid \mu^s) = w.$$

It follows that

$$V(\pi \mid \mu^s) \leq V(\pi^s \mid \mu^s) \leq V(\pi^s \mid \mu)$$

for all $\pi \in \Pi_{RO}(\alpha)$ and $\mu \in \mathcal{P}(\alpha)$, proving the saddle-point inequalities.

It remains to prove vertex attainment. Fix any $\pi \in \Pi_{RO}(\alpha)$. Robust obedience implies that

$$V(\pi \mid \mu) = \sum_{\omega\in\Omega}\sum_{a\in B} \mu(\omega)\pi(a \mid \omega)v(a,\omega).$$

For fixed π, define

$$c_\omega(\pi) := \sum_{a\in B} \pi(a \mid \omega)v(a,\omega).$$

Then

$$V(\pi \mid \mu) = \sum_{\omega\in\Omega} c_\omega(\pi)\mu(\omega),$$

so $\mu \mapsto V(\pi \mid \mu)$ is a continuous linear functional. By Lemma 2.1, $\mathcal{P}(\alpha)$ is a nonempty compact polytope. Hence the minimum of $V(\pi \mid \cdot)$ over $\mathcal{P}(\alpha)$ is attained. Because a linear functional attains its minimum over a nonempty polytope at an extreme point, there exists

$$\widehat{\mu}_\pi \in \operatorname{ext}\big(\mathcal{P}(\alpha)\big)$$

such that

$$V(\pi \mid \widehat{\mu}_\pi) = \min_{\mu\in\mathcal{P}(\alpha)} V(\pi \mid \mu).$$

Applying this result to the maxmin-optimal policy π^s proves the final assertion. ■

E Proof of Observation 1

Since π is robustly obedient, the Receiver follows every recommendation sent with positive probability under every $\mu \in \mathcal{P}(\alpha)$. Thus,

$$V(\pi \mid \mu) = \sum_{\omega\in\Omega}\sum_{a\in B} \mu(\omega)\pi(a \mid \omega)v(a,\omega).$$

For fixed π, the coefficient

$$c_\omega(\pi) = \sum_{a\in B} \pi(a \mid \omega)v(a,\omega)$$

does not depend on μ. Therefore,

$$g_\pi(\mu) = \sum_{\omega\in\Omega} c_\omega(\pi)\mu(\omega)$$

is linear in μ. Since Ω is finite, g_π is finite-valued and continuous on $\mathbb{R}^{|\Omega|}$.

A finite-valued continuous linear functional is proper, convex, and lower semicontinuous. Therefore, by the Fenchel–Moreau theorem,

$$g_\pi = g_\pi^{**}.$$

For any $z \in \mathbb{R}^{|\Omega|}$,

$$\begin{aligned} g_\pi^*(z) &= \sup_{\mu\in\mathbb{R}^{|\Omega|}} \{\langle z, \mu\rangle - g_\pi(\mu)\} \\ &= \sup_{\mu\in\mathbb{R}^{|\Omega|}} \langle z - c(\pi), \mu\rangle. \end{aligned}$$

This supremum equals 0 when $z = c(\pi)$, and equals $+\infty$ otherwise. Hence,

$$g_\pi^*(z) = \begin{cases} 0, & z = c(\pi), \\ +\infty, & z \neq c(\pi). \end{cases}$$

■

F Proof of Claim 1

(⇐) If such an optimal direct policy $\pi^* \in \Pi_B$ exists and its recommended actions are Receiver-optimal for all $p \in P(\alpha)$, then $\pi^* \in \Pi_{\mathrm{RO}}(\alpha)$ by Definition 2.2. Thus $W_{\mathrm{RO}}(\alpha) \geq \min_{p\in P(\alpha)}\langle c(\pi^*), p\rangle = W(\alpha)$, which together with $W_{\mathrm{RO}}(\alpha) \leq W(\alpha)$ implies $W_{\mathrm{RO}}(\alpha) = W(\alpha)$.

(⇒) Conversely, if $W_{\mathrm{RO}}(\alpha) = W(\alpha)$, note that the robust problem attains its maximum over policies. Compactness of $\Pi_{\mathrm{RO}}(\alpha)$ holds because Π_B is a compact product of simplices and the robust obedience inequalities define closed linear half-spaces. Since $g_\pi(p) = \langle c(\pi), p\rangle$ is continuous in π, a robust maximizer $\pi^*_{\mathrm{RO}} \in \Pi_{\mathrm{RO}}(\alpha)$ exists. Since π^*_{RO} achieves value $W(\alpha)$ in the unrestricted class, $\pi^*_{\mathrm{RO}} \in \Pi_B$ is an unrestricted maxmin-optimal direct policy whose recommendations are uniformly Receiver-optimal across $P(\alpha)$. ■

G Proof of Theorem 4.3

Fix any $p \in P(\alpha)$, $a \in B$, and $a' \in A \setminus \{a\}$. By linearity of $D^{\pi^*}_{a,a'}(p)$ in p:

$$|D^{\pi^*}_{a,a'}(p) - D^{\pi^*}_{a,a'}(\bar{p})| = \left|\sum_{\omega\in\Omega} \pi^*(a \mid \omega)(p(\omega) - \bar{p}(\omega))[u(a,\omega) - u(a',\omega)]\right| \leq L_R\|p-$$

By the reverse triangle inequality:

$$D^{\pi^*}_{a,a'}(p) \geq D^{\pi^*}_{a,a'}(\bar{p}) - L_R r_1(P(\alpha), \bar{p}) \geq \gamma - L_R r_1(P(\alpha), \bar{p}).$$

Since $r_1(P(\alpha), \bar{p}) < \frac{\gamma}{L_R}$, we have $D^{\pi^*}_{a,a'}(p) > 0$ for all $a \in B, a' \neq a$ and all $p \in P(\alpha)$. Thus $\pi^* \in \Pi_{\mathrm{RO}}(\alpha)$, entailing $W_{\mathrm{RO}}(\alpha) = W(\alpha)$. ■

H Proof of Theorem 4.4

Proof of (i) [Policy Robustification]. Fix $a \in B$, $a' \neq a$, and $p \in P(\alpha)$. Note that $D^{\pi^*}_{a,a'}(p) - D^{\pi^*}_{a,a'}(\bar{p}) = \sum_{\omega\in\Omega} \pi^*(a \mid \omega)\big(p(\omega) - \bar{p}(\omega)\big)\big[u(a,\omega) - u(a',\omega)\big]$. Hence,

$$\left|D^{\pi^*}_{a,a'}(p) - D^{\pi^*}_{a,a'}(\bar{p})\right| \le \sum_{\omega\in\Omega} \pi^*(a \mid \omega)|p(\omega) - \bar{p}(\omega)| \cdot |u(a,\omega) - u(a',\omega)|$$
$$\le L_R \sum_{\omega\in\Omega} |p(\omega) - \bar{p}(\omega)| = L_R\|p - \bar{p}\|_1 \le L_R r_1.$$

Since π^* is obedient at $\bar{p}$, we have $D^{\pi^*}_{a,a'}(\bar{p}) \ge 0$, and therefore $D^{\pi^*}_{a,a'}(p) \ge -L_R r_1$. By linearity of $D^{\pi}_{a,a'}(p)$ in π, for $\pi_\delta = (1-\delta)\pi^* + \delta\pi_{\mathrm{int}}$, we have $D^{\pi_\delta}_{a,a'}(p) = (1-\delta)D^{\pi^*}_{a,a'}(p) + \delta D^{\pi_{\mathrm{int}}}_{a,a'}(p) \ge -(1-\delta)L_R r_1 + \delta\eta$. Setting $\delta = \frac{L_R r_1}{\eta + L_R r_1}$ yields $\delta\eta = (1-\delta)L_R r_1$, and hence $D^{\pi_\delta}_{a,a'}(p) \ge 0$ for all $a \in B, a' \neq a, p \in P(\alpha)$. Therefore, $\pi_\delta \in \Pi_{\mathrm{RO}}(\alpha)$. This proves part (i).

Proof of (ii) [Payoff Bound under Receiver Action Stability]. Under Receiver Action Stability, $V(\pi \mid p) = \sum_{\omega\in\Omega}\sum_{a\in B} p(\omega)\pi(a \mid \omega)v(a,\omega)$. Hence,

$$V(\pi_\delta \mid p) - V(\pi^* \mid p) = \sum_{\omega\in\Omega}\sum_{a\in B} p(\omega)\big[\pi_\delta(a \mid \omega) - \pi^*(a \mid \omega)\big]v(a,\omega)$$
$$= \delta \sum_{\omega\in\Omega}\sum_{a\in B} p(\omega)\big[\pi_{\mathrm{int}}(a \mid \omega) - \pi^*(a \mid \omega)\big]v(a,\omega).$$

Therefore,

$$|V(\pi_\delta \mid p) - V(\pi^* \mid p)| \le \delta M_v \sum_{\omega\in\Omega} p(\omega) \sum_{a\in B} |\pi_{\mathrm{int}}(a \mid \omega) - \pi^*(a \mid \omega)|.$$

For each ω, both $\pi_{\mathrm{int}}(\cdot \mid \omega)$ and $\pi^*(\cdot \mid \omega)$ are probability distributions on B, so $\sum_{a\in B} |\pi_{\mathrm{int}}(a \mid \omega) - \pi^*(a \mid \omega)| \le 2$. Thus, $|V(\pi_\delta \mid p) - V(\pi^* \mid p)| \le 2\delta M_v$, and consequently, $V(\pi_\delta \mid p) \ge V(\pi^* \mid p) - 2\delta M_v$. Since $\delta = \frac{L_R r_1}{\eta + L_R r_1} \le \frac{L_R r_1}{\eta}$, we obtain $2\delta M_v \le \frac{2M_v L_R}{\eta} r_1$. Using the general inequality $W_{\mathrm{RO}}(\alpha) \le W(\alpha)$, we conclude that

$$0 \le W(\alpha) - W_{\mathrm{RO}}(\alpha) \le \frac{2M_v L_R}{\eta} r_1(P(\alpha), \bar{p}).$$

This proves part (ii) and thus the full theorem as well. ■

I Proof of Claim 2

Let $\phi(x_1, \ldots, x_d)$ be a 3-CNF formula. Write S for its number of satisfying assignments.

State space. Add one dummy bit z to the assignment:

$$\Omega = \{0,1\}^{d+1}, \qquad N = |\Omega| = 2^{d+1}.$$

States with $z = 0$ represent assignments of ϕ; states with $z = 1$ are dummy states. Let $\bar{\mu}$ denote the uniform distribution, so $\bar{\mu}(\omega) = 1/N$, and set $\epsilon = 1/N$.

Actions. Use the receiver actions

$$A = \{a_0, a_1, c\} \cup \{b_s : s \in \Omega\},$$

with observed distribution $\alpha(a_0) = \alpha(a_1) = 1/2$. All auxiliary actions have zero observed probability. Define

$$h(z,x) = \begin{cases} 2\phi(x) - 1, & z = 0, \\ 1, & z = 1. \end{cases}$$

Receiver utilities. Set

$$u(a_0,\omega) = u(a_1,\omega) = 0,$$
$$u(c,\omega) = -h(\omega), \qquad u(b_s,\omega) = \epsilon - \mathbf{1}\{\omega = s\}.$$

These utilities are computable by polynomial-size circuits: the circuit evaluates ϕ, compares the state with the auxiliary action's index s, and performs rational arithmetic. The exponentially many auxiliary actions are represented by their $O(d)$-bit indices, not listed individually.

For either a_i, where $i \in \{0,1\}$, the best-response conditions against every b_s imply

$$\mathbb{E}_\mu[u(a_i,\omega) - u(b_s,\omega)] = \mu(s) - \epsilon \ge 0 \quad \forall s \in \Omega.$$

Thus every $\mu \in M_{a_i}$ satisfies $\mu(s) \ge 1/N$ for all s. Since the probabilities sum to one, these inequalities force

$$\mu(s) = 1/N \quad \forall s,$$

so $\mu = \bar{\mu}$. It remains to check that $\bar{\mu}$ satisfies the condition against c. We have

$$\mathbb{E}_{\bar{\mu}}[h] = \frac{1}{N}\left[(2S - 2^d) + 2^d\right]$$
$$= \frac{2S}{2^{d+1}} = \frac{S}{2^d} \ge 0.$$

The conditions against a_0 and a_1 hold with equality because their utilities coincide. Therefore,

$$M_{a_0} = M_{a_1} = \{\bar{\mu}\}.$$

Consequently, the identified set is exactly

$$\boxed{P(\alpha) = \tfrac{1}{2}M_{a_0} + \tfrac{1}{2}M_{a_1} = \{\bar{\mu}\}.}$$

This is derived from the original rationalization formula, not imposed independently. Moreover, neither observed action is statewise strictly dominated: they have identical utilities, every b_s is worse at state s, and c is worse at every dummy state.

Set

$$v(a_1,(z,x)) = \begin{cases} \phi(x), & z = 0, \\ 0, & z = 1, \end{cases} \qquad v(a_0,\omega) = 0,$$

and set the sender's utilities from auxiliary actions to zero. Choose the threshold

$$\tau = \frac{1}{N} = 2^{-(d+1)}.$$

Consider the policy that recommends a_1 at every state. Against a_0, obedience holds with equality. Against every b_s, its obedience margin is $\bar{\mu}(s) - \epsilon = 0$. Against c, its obedience margin is $\mathbb{E}_{\bar{\mu}}[h] = S/2^d \ge 0$. Thus the policy is robustly obedient. Its sender payoff is S/N. No policy can obtain more, because the sender receives zero from a_0 and the payoff from a_1 is at most one at satisfying assignment states and zero elsewhere. Therefore,

$$\boxed{W_{\mathrm{RO}}(\alpha) = \frac{S}{2^{d+1}}.}$$

It follows immediately that

$$\boxed{W_{\mathrm{RO}}(\alpha) \ge 2^{-(d+1)} \iff S \ge 1 \iff \phi \text{ is satisfiable.}}$$

The construction is polynomial-time in the 3-SAT input size. Hence 3-SAT reduces to the succinct robust-persuasion decision problem, proving NP-hardness. ■

J Proof of Proposition 5.1

Write $n = |\Omega| - 1$, the affine dimension of the simplex. Since μ° is interior, every feasible direction in the simplex's tangent space is locally available in both signs. For M_a to contain only μ°, the active best-response inequalities at μ° must have normals whose associated homogeneous inequalities admit only the zero direction. Equivalently, those normals must positively span the n-dimensional tangent space. A positive spanning set in dimension n requires at least $n + 1 = |\Omega|$ vectors. There are at most $|A| - 1$ comparisons against alternative actions, giving the claimed bound. ■

K Proof of Claim 3

Let $S = \mathrm{supp}(\mu^a)$, with $|S| = s$. The prior μ^a lies in the relative interior of the face $\Delta(S)$, whose tangent space has dimension $s - 1$.

For M_a to be a singleton, the active receiver-optimality inequalities, restricted to that tangent space, must have a feasible tangent cone consisting only of the zero vector. A collection of fewer than s homogeneous half-space inequalities cannot have this property in dimension $s-1$: their normals cannot positively span the entire dual space. Since there are only $|A| - 1$ comparisons against alternative actions,

$$s \le |A| - 1.$$

This proves the claim. ■

L Proof of Theorem 5.2

We give a polynomial-time many-one reduction from 3-UNSAT. Let

$$\phi(x_1, \dots, x_d) = \bigwedge_{j=1}^{m} C_j(x_1, \dots, x_d)$$

be a 3-CNF formula, where each clause C_j contains at most three literals. We may restrict attention to standard nondegenerate instances with $d \ge 1$ and nonempty clauses.

Let

$$\Omega = \{0,1\}^{d+1}.$$

Write each state as (z, x), where $z \in \{0,1\}$ and $x \in \{0,1\}^d$. Thus $|\Omega| = 2^{d+1}$, but a state can be encoded using only $d + 1$ bits. Take the action set

$$A = \{a_0, a_1, d_0, e_0\} \cup \{c_1, \dots, c_m\},$$

and let the observed distribution be

$$\alpha(a_0) = \alpha(a_1) = \tfrac{1}{2}, \qquad B = \mathrm{supp}(\alpha) = \{a_0, a_1\}.$$

Define the receiver's utility functions by

$$u(a_0, (z, x)) = u(a_1, (z, x)) = 0$$

for every state, and

$$u(d_0, (z, x)) = \begin{cases} -1, & z = 0, \\ 1, & z = 1, \end{cases}$$

$$u(e_0, (z, x)) = \begin{cases} 2^{-(d+1)}, & z = 0, \\ -1, & z = 1, \end{cases}$$

and, for each clause C_j,

$$u(c_j, (z, x)) = \begin{cases} 1 - C_j(x), & z = 0, \\ -1, & z = 1. \end{cases}$$

Here $C_j(x) \in \{0,1\}$ denotes the truth value of clause C_j under assignment x. Define the sender's utilities on the supported actions by

$$v(a_0, (z, x)) = 0$$

and

$$v(a_1, (z, x)) = \begin{cases} 1 - \phi(x), & z = 0, \\ 0, & z = 1. \end{cases}$$

Set the sender's utilities for all auxiliary actions to zero. Finally, set

$$\tau = 2^{-(d+1)}.$$

The construction has $m + 4$ actions and is polynomial in the length of ϕ: each utility is computable from the state bits and the input formula by a polynomial-size Boolean/arithmetic circuit, and the rational number $2^{-(d+1)}$ has a binary representation of $O(d)$ bits.

For each action a, let

$$M_a = \{\mu \in \Delta(\Omega) : a \in \arg\max_{a' \in A} \mathbb{E}_\mu[u(a', \omega)]\}.$$

Because a_0 and a_1 both have identically zero receiver utility,

$$M_{a_0} = M_{a_1} =: M.$$

Consequently,

$$P(\alpha) = \tfrac{1}{2} M_{a_0} + \tfrac{1}{2} M_{a_1} = M.$$

A distribution $\mu \in \Delta(\Omega)$ belongs to M precisely when no auxiliary action has strictly greater expected utility than a_0 (whose expected utility is zero). Write

$$\mu_z := \sum_{x \in \{0,1\}^d} \mu(z, x).$$

The constraint associated with d_0 is

$$-\mu_0 + \mu_1 \le 0, \quad \text{equivalently} \quad \mu_0 \ge \mu_1.$$

The constraint associated with e_0 is

$$2^{-(d+1)}\mu_0 - \mu_1 \le 0, \quad \text{equivalently} \quad \mu_1 \ge 2^{-(d+1)}\mu_0.$$

For clause action c_j, its expected utility is

$$\sum_{x : C_j(x) = 0} \mu(0, x) - \mu_1.$$

Its rationalization constraint is therefore

$$\mu_1 \ge \sum_{x : C_j(x) = 0} \mu(0, x), \qquad j = 1, \dots, m.$$

Thus

$$\boxed{\begin{aligned} P(\alpha) = \Big\{ \mu \in \Delta(\Omega) : \ & \mu_0 \ge \mu_1, \\ & \mu_1 \ge 2^{-(d+1)}\mu_0, \\ & \mu_1 \ge \sum_{x : C_j(x) = 0} \mu(0, x) \quad \forall j \Big\}. \end{aligned}} \tag{1}$$

If ϕ is unsatisfiable, then $W_{\mathrm{RO}}(\alpha) \ge \tau$. Consider the constant policy π^{a_1} that recommends a_1 with probability one in every state:

$$\pi^{a_1}(a_1 \mid \omega) = 1 \quad \text{for every } \omega \in \Omega.$$

This policy is robustly obedient. Indeed, the receiver is indifferent between a_0 and a_1 at every state, and its robust-obedience constraints against the auxiliary actions are precisely the inequalities

defining $P(\alpha)$ in (1). Hence the policy satisfies all obedience constraints for every $\mu \in P(\alpha)$. If ϕ is unsatisfiable, then $\phi(x) = 0$ for every $x \in \{0,1\}^d$. Therefore,

$$v(a_1, (0, x)) = 1, \qquad v(a_1, (1, x)) = 0.$$

The sender's expected utility under π^{a_1}, at any $\mu \in P(\alpha)$, is consequently

$$\sum_{\omega \in \Omega} \mu(\omega) v(a_1, \omega) = \mu_0.$$

By (1), $\mu_0 \geq \mu_1$, and since $\mu_0 + \mu_1 = 1$, it follows that $\mu_0 \geq \frac{1}{2}$. Thus

$$W_{\mathrm{RO}}(\alpha) \geq \min_{\mu \in P(\alpha)} \mu_0 \geq \frac{1}{2} \geq 2^{-(d+1)} = \tau. \tag{2}$$

If ϕ is satisfiable, then $W_{\mathrm{RO}}(\alpha) < \tau$. Suppose ϕ is satisfiable, and choose a satisfying assignment $x^* \in \{0,1\}^d$. Define a distribution μ^* by

$$\mu^*(0, x^*) = \frac{1}{2}, \qquad \mu^*(1, 0^d) = \frac{1}{2},$$

with zero probability assigned to every other state. We verify that $\mu^* \in P(\alpha)$. First,

$$\mu_0^* = \mu_1^* = \frac{1}{2},$$

so $\mu_0^* \geq \mu_1^*$. Second,

$$\frac{1}{2} \geq 2^{-(d+1)} \frac{1}{2}.$$

Finally, since x^* satisfies every clause, $C_j(x^*) = 1$ for each j. Hence

$$\sum_{x : C_j(x) = 0} \mu^*(0, x) = 0 \leq \frac{1}{2} = \mu_1^*.$$

All inequalities in (1) hold, establishing $\mu^* \in P(\alpha)$. At the state $(0, x^*)$, the sender receives zero from either supported action, because

$$v(a_0, (0, x^*)) = 0, \qquad v(a_1, (0, x^*)) = 1 - \phi(x^*) = 0.$$

At the state $(1, 0^d)$, the sender also receives zero from either supported action. Therefore, every policy supported on $B = \{a_0, a_1\}$ yields expected utility zero under μ^*. In particular, every robustly obedient policy has worst-case expected utility at most zero:

$$W_{\mathrm{RO}}(\alpha) \leq 0.$$

The constant policy π^{a_1} is robustly obedient, as established above, and the sender's utilities on the supported actions are nonnegative at every state. Hence $W_{\mathrm{RO}}(\alpha) \geq 0$, giving

$$W_{\mathrm{RO}}(\alpha) = 0 < 2^{-(d+1)} = \tau. \tag{3}$$

Combining (2) and (3), we obtain

$$\boxed{\phi \in \text{3-UNSAT} \iff W_{\mathrm{RO}}(\alpha) \geq \tau.}$$

The construction is computable in polynomial time, so this is a polynomial-time many-one reduction from 3-UNSAT to the decision problem $W_{\mathrm{RO}}(\alpha) \geq \tau$. Since 3-UNSAT is coNP-complete, the decision problem is coNP-hard.

It remains to verify the stated geometric property. Let $N = 2^{d+1}$, and consider the distribution $\bar{\mu}$ defined by

$$\bar{\mu}(0, x) = \frac{10}{19 \cdot 2^d}, \qquad \bar{\mu}(1, x) = \frac{9}{19 \cdot 2^d}$$

for every $x \in \{0,1\}^d$. Then

$$\bar{\mu}_0 = \frac{10}{19}, \qquad \bar{\mu}_1 = \frac{9}{19}.$$

The first inequality in (1) is strict because $10/19 > 9/19$. The second is strict because

$$\frac{9}{19} > 2^{-(d+1)} \frac{10}{19}.$$

For any nonempty clause C_j containing at most three literals, at most half of all assignments falsify the clause. Consequently,

$$\sum_{x : C_j(x) = 0} \bar{\mu}(0, x) \leq \frac{10}{38} = \frac{5}{19} < \frac{9}{19} = \bar{\mu}_1.$$

Thus every inequality defining $P(\alpha)$, apart from the simplex's normalization equality, holds strictly at $\bar{\mu}$. Moreover, $\bar{\mu}$ assigns positive probability to every state. It follows that $P(\alpha)$ contains a relatively open neighborhood of $\bar{\mu}$ in the probability simplex. Therefore,

$$\dim P(\alpha) = |\Omega| - 1 = 2^{d+1} - 1.$$

This establishes full dimensionality and completes the proof. ■

M Proof of Theorem 5.3

For each pair $a, a' \in A$, define the state-dependent utility difference

$$\Delta u_{a,a'}(\omega) := u(a, \omega) - u(a', \omega).$$

A policy π is robustly obedient precisely when

$$\sum_{\omega \in \Omega} \mu(\omega) \pi(a \mid \omega) \Delta u_{a,a'}(\omega) \geq 0 \tag{5}$$

for every $a, a' \in A$ and every $\mu \in P(\alpha)$. Its worst-case sender payoff is

$$\min_{\mu \in P(\alpha)} \sum_{\omega \in \Omega} \mu(\omega) \sum_{a \in B} \pi(a \mid \omega) v(a, \omega). \tag{6}$$

Consequently, $W_{\mathrm{RO}}(\alpha) \geq \tau$ holds exactly when there is a policy satisfying (5) and making (6) at least τ. Since $B = \{a_0, a_1\}$, write

$$p(\omega) := \pi(a_1 \mid \omega) \in [0,1], \qquad \pi(a_0 \mid \omega) = 1 - p(\omega).$$

Every coefficient in (5) and (6) is therefore affine in the statewise policy variable $p(\omega)$. For $b \in B$, the rationalization set is

$$M_b = \left\{ q \in \Delta(\Omega) : \sum_{\omega} q(\omega) \Delta u_{b,c}(\omega) \geq 0 \quad \forall c \in A \right\}. \tag{7}$$

By definition,

$$P(\alpha) = \sum_{b \in B} \alpha(b) M_b.$$

Fix any vector $c = (c_\omega)_{\omega \in \Omega}$. Minimizing a linear function over this Minkowski sum separates into the component problems:

$$\min_{\mu \in P(\alpha)} c^\top \mu = \sum_{b \in B} \alpha(b) \min_{q \in M_b} c^\top q. \tag{8}$$

For a fixed b, the linear program on the right is

$$\begin{aligned} \min_{q} \quad & \sum_{\omega} c_\omega q(\omega) \\ \text{s.t.} \quad & \sum_{\omega} q(\omega) = 1, \\ & \sum_{\omega} q(\omega) \Delta u_{b,c'}(\omega) \geq 0 \quad \forall c' \in A, \\ & q(\omega) \geq 0 \qquad\qquad\qquad\quad \forall \omega. \end{aligned} \tag{9}$$

Its dual is

$$\begin{aligned} \max_{\eta_b,\lambda_{b,c'}} \quad & \eta_b \\ \text{s.t.} \quad & \eta_b + \sum_{c'\in A} \lambda_{b,c'} \Delta u_{b,c'}(\omega) \le c_\omega \quad \forall \omega, \\ & \lambda_{b,c'} \ge 0 \qquad\qquad\qquad\quad \forall c' \in A, \end{aligned} \tag{10}$$

where η_b is unrestricted in sign. The primal is feasible and bounded because M_b is nonempty and compact. Strong LP duality therefore applies. Combining (8)–(10) yields

$$\min_{\mu\in P(\alpha)} c^\top \mu = \max_{\{\eta_b,\lambda_{b,c'}\}_{b\in B}} \sum_{b\in B} \alpha(b)\eta_b, \tag{11}$$

subject to

$$\eta_b + \sum_{c'\in A} \lambda_{b,c'} \Delta u_{b,c'}(\omega) \le c_\omega \quad \forall b \in B,\ \omega \in \Omega, \tag{12}$$

and $\lambda_{b,c'} \ge 0$. Crucially, there is only one collection of dual variables for each b, shared across all states. The state-indexed inequalities in (12) are the only constraints in this dual formulation that need to be enumerated over Ω.

Let

$$Q = \{(a, a') : a, a' \in A\}$$

be the set of obedience constraints, and let q_* denote the sender-payoff constraint. For each $q = (a, a') \in Q$, set

$$c^q_\omega(p) = \pi(a \mid \omega)\Delta u_{a,a'}(\omega), \qquad t_q = 0.$$

For the payoff constraint, set

$$c^{q_*}_\omega(p) = \sum_{a\in B} \pi(a \mid \omega) v(a, \omega), \qquad t_{q_*} = \tau.$$

Applying (11) to each of these constraints, we obtain the following exact equivalence:

$$W_{\mathrm{RO}}(\alpha) \ge \tau \iff \exists g \in G \quad \forall \omega \in \Omega \quad \exists p(\omega) \in [0,1] : \quad A_\omega g + B_\omega p(\omega) \le c_\omega. \tag{13}$$

Here g collects the variables $\eta_{q,b}$ and $\lambda_{q,b,c'}$ for all $q \in Q \cup \{q_*\}$ and $b \in B$. The set G imposes the global constraints

$$\sum_{b\in B} \alpha(b)\eta_{q,b} \ge t_q \quad \forall q \in Q \cup \{q_*\}, \tag{14}$$

together with $\lambda_{q,b,c'} \ge 0$. For each state, the inequalities in (13) are exactly the dual inequalities (12), with right-hand side $c^q_\omega(p)$. These are linear inequalities in g and $p(\omega)$: all utilities are fixed coefficients, so there are no products of decision variables. The number of global variables is polynomial in $|A|$. With $k = |A|$, there are $O(k^3)$ dual variables, up to an inessential constant factor. All coefficients can be computed in polynomial time from the input and the specified state.

For every state ω, define

$$S_\omega := \left\{ g \in \mathbb{R}^D : \exists p \in [0,1],\ A_\omega g + B_\omega p \le c_\omega \right\}, \tag{15}$$

where $D = \dim(g) = \mathrm{poly}(|A|)$. The set described jointly by g and p in (15) is a polyhedron. Its projection onto the g-coordinates is also a polyhedron, so S_ω is convex and closed. By (13),

$$W_{\mathrm{RO}}(\alpha) \ge \tau \iff G \cap \bigcap_{\omega\in\Omega} S_\omega \ne \varnothing. \tag{16}$$

Thus, a NO instance is precisely one for which this finite family of convex sets has empty intersection.

Suppose

$$G \cap \bigcap_{\omega\in\Omega} S_\omega = \varnothing. \tag{17}$$

If $G = \varnothing$, this can be detected by checking the polynomial-size rational LP defining G. Otherwise, Helly's theorem in $\mathbb{R}^D$ implies that some subfamily of at most $D + 1$ sets from

$$\{G\} \cup \{S_\omega : \omega \in \Omega\}$$

already has empty intersection. If this subfamily does not contain G, adjoining G preserves emptiness. Hence there is a subset $T \subseteq \Omega$, with

$$|T| \le D + 1, \tag{18}$$

such that

$$G \cap \bigcap_{\omega\in T} S_\omega = \varnothing. \tag{19}$$

Each state in T is encoded by d bits, so the list of states is a polynomial-size certificate because $D = \mathrm{poly}(|A|)$ and d is at most the input length.

A nondeterministic verifier for the complement language guesses the states in T. It constructs the rational LP consisting of: * the global constraints defining G; * the statewise dual inequalities for the guessed states; * one policy variable $p(\omega) \in [0, 1]$ for each $\omega \in T$.

This LP has polynomially many variables and constraints. Its coefficients can be evaluated in polynomial time, and their bit lengths are polynomially bounded. The verifier accepts if and only if the LP is infeasible, which can be checked in polynomial time by exact rational linear programming. For every NO instance, the Helly argument guarantees such a certificate. Conversely, if the restricted LP is infeasible, the full system is infeasible as well, so the instance is indeed a NO instance. Therefore,

$$\{\mathcal{I} : W_{\mathrm{RO}}(\alpha) < \tau\} \in \mathrm{NP},$$

and hence

$$\boxed{\text{DECISION-}W_{\mathrm{RO}} \in \mathrm{coNP}.}$$

■

N Proof of Theorem 5.4

Proving this result requires establishing two intermediate lemmas first.

Lemma 5.5: For each state $\omega \in \Omega$, define its utility type by

$$V(\omega) = ((u(a,\omega))_{a\in A}, (v(a,\omega))_{a\in B}) \in \mathbb{Z}^r, \qquad r = k + |B| \le 2k. \tag{21}$$

Let

$$\mathcal{V} = \{V(\omega) : \omega \in \Omega\}$$

be the set of attainable types. Then

$$|\mathcal{V}| \le (2dR + 1)^r \le (2dR + 1)^{2k}. \tag{22}$$

Moreover, $\mathcal{V}$ can be computed deterministically in time

$$2^{O(w)}(dR)^{O(k)}\,\mathrm{poly}(L). \tag{23}$$

Proof: Every coordinate of $V(\omega)$ is a sum of at most d integer-valued local components, each with magnitude at most R. Hence every coordinate belongs to

$$\{-dR, -dR + 1, \ldots, dR\}.$$

Since $V(\omega)$ has $r \le 2k$ coordinates, (22) follows. It remains to construct the dynamic program. For each local component, choose one bag containing its entire variable scope. Combine the components assigned to a given bag t into a vector-valued local function

$$\ell_t : \{0,1\}^{X_t} \longrightarrow \mathbb{Z}^r,$$

where X_t is the bag at node t. Its j-th coordinate is the sum of the contributions of the components assigned to t for the j-th utility in (21). Every local component is assigned to exactly one bag. Therefore, for every complete assignment ω,

$$V(\omega) = \sum_{t \in T} \ell_t(\omega|_{X_t}), \tag{24}$$

where T is the decomposition tree and $\omega|_{X_t}$ denotes the restriction of ω to bag X_t. We may convert the supplied decomposition to a rooted nice decomposition of width w with a polynomial increase in its number of nodes. The argument below also applies to a rooted decomposition with bounded branching.

Root the decomposition at t_0. Let U_t be the union of all variables appearing in bags in the subtree rooted at t. For a bag assignment $\sigma \in \{0,1\}^{X_t}$ and a vector $s \in \mathbb{Z}^r$, define

$$D_t(\sigma, s)$$

to be the number of assignments to $U_t \setminus X_t$ which, together with σ, produce total local contribution s from all bags in the subtree rooted at t. Only vectors in

$$\mathcal{S} = \{-dR, \dots, dR\}^r \tag{25}$$

need to be stored. Every partial sum is bounded coordinatewise by dR, because it contains only a subset of the local summands contributing to a complete utility.

Let $\mathrm{ch}(t)$ denote the children of t. For a fixed assignment σ to X_t, let

$$s_t = \ell_t(\sigma).$$

For each child c, sum over assignments ρ_c to X_c agreeing with σ on the separator $X_c \cap X_t$. The recurrence is

$$D_t(\sigma, s) = \sum_{\substack{(\rho_c)_{c \in \mathrm{ch}(t)} \\ \rho_c|_{X_c \cap X_t} = \sigma|_{X_c \cap X_t}}} \sum_{\substack{(s_c)_{c \in \mathrm{ch}(t)} \in \mathcal{S}^{|\mathrm{ch}(t)|} \\ s = s_t + \sum_{c \in \mathrm{ch}(t)} s_c}} \prod_{c \in \mathrm{ch}(t)} D_c(\rho_c, s_c). \tag{26}$$

At a leaf t, the recurrence reduces to

$$D_t(\sigma, s) = \begin{cases} 1, & s = \ell_t(\sigma), \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

The tree-decomposition property guarantees that, conditional on the assignment to X_t, assignments in different child subtrees are compatible and independent except for their shared separator variables. Thus (26) counts every assignment to $U_t \setminus X_t$ exactly once, with its correct total contribution vector. This proves the recurrence by induction from the leaves to the root.

At the root, every state assignment belongs to exactly one root-bag assignment and one total-sum vector. Consequently,

$$\mathcal{V} = \left\{ s \in \mathcal{S} : \sum_{\sigma \in \{0,1\}^{X_{t_0}}} D_{t_0}(\sigma, s) > 0 \right\}. \tag{28}$$

The counts can be stored in binary: no count exceeds 2^d, so at most $d+1$ bits are needed. Each bag has at most $w+1$ variables, giving 2^{w+1} bag assignments. The number of possible sum vectors is at most

$$(2dR+1)^r = (dR)^{O(k)}.$$

The convolution operations in (26) require at most a fixed power of this table size after conversion to a bounded-branching decomposition. Each operation uses polynomial-bit integers, and the number of bags and local components is polynomial in L. The resulting total time is

$$2^{O(w)} (dR)^{O(k)} \,\mathrm{poly}(L).$$

This proves the lemma. □

Lemma 5.6: Under the assumptions of Theorem 5.4, if there exists a robustly obedient policy achieving sender payoff at least τ, then there exists such a policy that is constant on every utility type:

$$V(\omega) = V(\omega') \quad \Longrightarrow \quad \pi(\cdot \mid \omega) = \pi(\cdot \mid \omega'). \tag{29}$$

Proof: The proof uses the dual representation from Part I, now for an arbitrary support B. For a recommendation $a \in B$ and a deviation $a' \in A$, define

$$c_\omega^{a,a'}(\pi) = \pi(a \mid \omega) \Delta u_{a,a'}(\omega).$$

For the sender-payoff constraint, define

$$c_\omega^*(\pi) = \sum_{a \in B} \pi(a \mid \omega) v(a, \omega).$$

By LP duality, the requirement

$$\min_{\mu \in P(\alpha)} \sum_\omega \mu(\omega) c_\omega(\pi) \ge t \tag{30}$$

is equivalent to the existence of variables

$$\eta_b \in \mathbb{R}, \qquad \lambda_{b,c} \ge 0 \quad (b \in B,\ c \in A)$$

satisfying

$$\eta_b + \sum_{c \in A} \lambda_{b,c} \Delta u_{b,c}(\omega) \le c_\omega(\pi) \quad \forall b \in B,\ \omega \in \Omega, \tag{31}$$

and

$$\sum_{b \in B} \alpha(b) \eta_b \ge t. \tag{32}$$

Apply this equivalence to every obedience constraint with $t = 0$ and to the sender-payoff constraint with $t = \tau$. The resulting global dual variables are shared across all states. Once these variables are fixed, the remaining conditions are linear constraints on the policy probabilities at each individual state. Now suppose π is a robustly obedient policy achieving sender payoff at least τ. Strong duality supplies global dual witnesses for all these constraints. Fix those witnesses. For each type $v \in \mathcal{V}$, choose one representative state ω_v with $V(\omega_v) = v$, and set

$$p_v(a) := \pi(a \mid \omega_v), \qquad a \in B.$$

Define a new policy by

$$\widehat{\pi}(a \mid \omega) = p_{V(\omega)}(a). \tag{33}$$

If $V(\omega) = V(\omega_v)$, then every receiver utility $u(c, \omega)$ and every relevant sender utility $v(a, \omega)$ equals its counterpart at ω_v. Thus every coefficient in the statewise dual inequalities (31) is identical at ω and ω_v. The policy probabilities p_v satisfy those inequalities at ω_v, so they also satisfy them at ω. The global constraints (32)

are unchanged. By the dual characterization, $\widehat{\pi}$ remains robustly obedient and achieves sender payoff at least τ. It is constant on each utility type by construction. □

This is not an assertion that arbitrary averaging of policies over states of the same type preserves robust obedience. It follows from the shared global dual witnesses and the fact that the resulting statewise inequalities depend only on the complete utility type.

Now we continue with the proof of Theorem 5.4: Let $\mathcal{V}$ be the attainable type set computed in Lemma 5.5. Introduce policy variables

$$p_v(a) \geq 0, \qquad \sum_{a \in B} p_v(a) = 1 \quad \forall v \in \mathcal{V}. \tag{34}$$

These represent the probability of recommending action a at every state of type v. For every obedience pair $q = (a, a') \in B \times A$ and every $b \in B$, introduce dual variables

$$\eta_{q,b} \in \mathbb{R}, \qquad \lambda_{q,b,c} \geq 0 \quad (c \in A).$$

Also introduce $\eta_{*,b}$ and $\lambda_{*,b,c} \geq 0$ for the sender-payoff constraint. For a type v, use the notation $u_v(c)$ for the coordinate $u(c, \omega)$ at any state ω of type v, and $s_v(a)$ for its sender-utility coordinate $v(a, \omega)$. These are well-defined because the type includes all relevant utility values. The reduced LP has the following constraints.

Obedience constraints. For every $q = (a, a') \in B \times A$, every $b \in B$, and every type $v \in \mathcal{V}$,

$$\eta_{q,b} + \sum_{c \in A} \lambda_{q,b,c}\big(u_v(b) - u_v(c)\big) \leq p_v(a)\big(u_v(a) - u_v(a')\big). \tag{35}$$

The corresponding global dual-objective constraint is

$$\sum_{b \in B} \alpha(b)\eta_{q,b} \geq 0 \qquad \forall q \in B \times A. \tag{36}$$

Sender-payoff constraint. For every $b \in B$ and every $v \in \mathcal{V}$,

$$\eta_{*,b} + \sum_{c \in A} \lambda_{*,b,c}\big(u_v(b) - u_v(c)\big) \leq \sum_{a \in B} p_v(a) s_v(a), \tag{37}$$

with global constraint

$$\sum_{b \in B} \alpha(b)\eta_{*,b} \geq \tau. \tag{38}$$

Together with (34) and nonnegativity of the λ-variables, these are all linear constraints.

Exactness of the reduced LP. If $W_{\mathrm{RO}}(\alpha) \geq \tau$, Lemma 5.6 supplies a type-constant policy. Strong duality for each robust-obedience constraint and the sender-payoff constraint supplies dual witnesses. These variables satisfy (34)–(38), so the reduced LP is feasible. Conversely, suppose the reduced LP is feasible. Define the policy

$$\pi(a \mid \omega) = p_{V(\omega)}(a).$$

For every state, (35) and (36) give a feasible dual certificate with objective at least zero for the corresponding obedience minimization. By weak duality, the associated minimum over $P(\alpha)$ is at least zero. Similarly, (37) and (38) give a feasible dual certificate with objective at least τ for the sender-payoff minimization. Its worst-case payoff is therefore at least τ. Hence

$$\boxed{W_{\mathrm{RO}}(\alpha) \geq \tau \iff \text{the reduced LP is feasible.}} \tag{39}$$

This argument uses the original M_b, $P(\alpha)$, and robust-obedience definition; no alternative economic model is introduced.

LP size and total running time. There are at most

$$|\mathcal{V}| \leq (2dR + 1)^{2k}$$

types. The LP has $O(k^3)$ global dual variables, $O(k^3)$ constraints per type, and at most $k|\mathcal{V}|$ policy variables. Its number of variables and constraints is therefore

$$(dR)^{O(k)}.$$

All coefficients have polynomial bit length: type coordinates are integers bounded by dR, and α, τ have polynomial-bit rational encodings. The LP can consequently be solved in deterministic polynomial time in its explicit encoding length, giving a running time of

$$(dR)^{O(k)} \operatorname{poly}(L). \tag{40}$$

The type-extraction dynamic program takes

$$2^{O(w)}(dR)^{O(k)} \operatorname{poly}(L)$$

time by Lemma 5.5. Combining it with the reduced LP proves

$$\boxed{T(d, w, k, R, L) = 2^{O(w)}(dR)^{O(k)} \operatorname{poly}(L).}$$

Since $d, R \leq \operatorname{poly}(L)$, this is of the form

$$f(w)\, L^{O(k)} \operatorname{poly}(L),$$

which establishes membership in XP parameterized by (w, k). ■

References

[CG19] André Chassein and Marc Goerigk. On the complexity of robust geometric programming with polyhedral uncertainty. *Operations Research Letters*, 47(1):21–24, 2019.

[DP22] Piotr Dworczak and Alessandro Pavan. Preparing for the worst but hoping for the best: Robust (bayesian) persuasion. *Econometrica*, 90(5):2017–2051, 2022.

[DX16] Shaddin Dughmi and Haifeng Xu. Algorithmic bayesian persuasion. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 412–425, 2016.

[FIKU08] Lance Fortnow, Russell Impagliazzo, Valentine Kabanets, and Christopher Umans. On the complexity of succinct zero-sum games. *Computational Complexity*, 17(3):353–376, 2008.

[Fre90] Eugene C Freuderl. Complexity of k-tree structurecl constraint satisfaction problems. 1990.

[FS22] Kaito Fujii and Shinsaku Sakaue. Algorithmic bayesian persuasion with combinatorial actions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 5016–5024, 2022.

[GHHS22] Ronen Gradwohl, Niklas Hahn, Martin Hoefer, and Rann Smorodinsky. Algorithms for persuasion with limited communication. *Mathematics of Operations Research*, 47(3):2520–2545, 2022.

[HF26] Yuwei Hu and Bryce L Ferguson. Robust information design with heterogeneous beliefs in bayesian congestion games. *arXiv preprint arXiv:2604.10831*, 2026.

[KG11] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[Kos22] Svetlana Kosterina. Persuasion with unknown beliefs. *Theoretical Economics*, 17(3):1075–1107, 2022.

[LR26] Marie Laclau and Ludovic Renou. Public persuasion. 2026.

[PCdG+19] Nathalie Peyrard, M-J Cros, Simon de Givry, Alain Franc, Stephane Robin, Regis Sabbadin, Thomas Schiex, and Matthieu Vignes. Exact or approximate inference in graphical models: why the choice is dictated by the treewidth, and how variable elimination can be exploited. *Australian & New Zealand Journal of Statistics*, 61(2):89–133, 2019.

[PS26] Alexei Parakhonyak and Anton Sobolev. Persuasion without priors: A. parakhonyak, a. sobolev. *Economic Theory*, 81(3):625–659, 2026.

[Ros26] Maxwell Rosenthal. Data-driven persuasion. *Journal of Mathematical Economics*, page 103268, 2026.

[SV12] Grant R Schoenebeck and Salil Vadhan. The computational complexity of nash equilibria in concisely represented games. *ACM Transactions on Computation Theory (TOCT)*, 4(2):1–50, 2012.